Dissertation

Average and Smoothed Complexity of Geometric Structures

Dipl.–Math. Valentina Damerow

Universität Paderborn
Fakultät für Elektrotechnik, Informatik und Mathematik
Institut für Informatik & Heinz Nixdorf Institut (HNI) &
Paderborn Institute for Scientific Computation (PaSCo)
Warburgerstraße 100, D – 33098 Paderborn

Reviewers: Prof. Dr. Friedhelm Meyer auf der Heide, Universität Paderborn
Prof. Dr. Peter Bürgisser, Universität Paderborn

Do not worry about your difficulties in mathematics;
I can assure you that mine are still greater.

Albert Einstein

Acknowledgments

During the last four years I had the golden opportunity to work for this Ph.D. thesis here at the University of Paderborn. I am especially grateful to Prof. Friedhelm Meyer auf der Heide who had the confidence in me to accomplish this goal. I experienced a lot of support from him and also from the members of my research group “Algorithms and Complexity” and many other colleagues in the Fürstenallee and on campus. It was sometimes a challenge but always a great gratification for me to work and live with them.

As my second advisor I would like to thank Prof. Peter Bürgisser for valuable comments on this thesis. My special thanks go also to my co-authors Marcin Bienkowski, Friedhelm Meyer auf der Heide, Harald Räcke, Christian Scheideler, and Christian Sohler, as well as to Marcel Rudolph Ackermann, Mirko Hessel-von Molo, Jarek Kutylowski, Harald Räcke, and Christian Sohler for carefully reading parts of this thesis.

I would like to thank also all members of the DFG graduate school No. 693 “Scientific Computation: Application-oriented Modelling and Development of Algorithms”, hosted by the Paderborn Institute for Scientific Computation (PaSCo). I enjoyed the interdisciplinary and international environment very much, it was a great source of inspiration for me.

Valentina Damerow
Paderborn, December 2005

Contents

1 Introduction

One of the main challenges in computer science is to determine the complexity of problems, structures, and algorithms. Since complexity and performance can vary drastically depending on the considered input instance, it is our goal to provide a profound theoretical analysis that gives reliable statements about the complexity and performance in applications, especially for algorithms. There are mainly two different traditional approaches to achieve this goal, which are the *worst case analysis* and the *average case analysis*.

Certainly, *worst case analysis* is the strongest and most reliable complexity measure we have. In worst case analysis, we ask for the maximum complexity of a particular problem over all input instances. The analysis is thus independent of input instances and holds for any input scenario and under any condition. But for many problems the worst case complexity bounds are rather pessimistic and for some problems and algorithms rarely or almost never encountered in applications. One famous example for this is the *simplex algorithm* for linear programming. The worst case complexity of the simplex algorithm is proven to be exponential, while at the same time it shows an extremely good performance in practice and is therefore widely used.

In order to provide a more 'realistic' analysis method researchers introduced the *average case analysis*. In the average case analysis a probability distribution on input instances is defined. The average case complexity of a problem or an algorithm is then the expected complexity measured on input instances from that distribution. For example, a low average case complexity provides at least some evidence that an algorithm might perform well in practice. But in most cases the average case analysis depends on the chosen probability distribution in the sense that we obtain different results for different probability distributions. Furthermore, we observe that the usually considered probability distributions provide in some cases a too optimistic complexity measure compared to the behavior encountered in applications. One reason for this is that in many applications the inputs have very special properties and that these properties cannot be captured by any probability distributions.

However, for many problems and algorithms we see a large discrepancy between their average case and worst case complexity. Moreover, the 'typical' complexity of a problem or algorithm one encounters in practice seems to lie usually somewhere between its average and worst case complexity. Thus there is need for a complexity measure that bridges this gap between average case and worst case.

In 2001, Spielman and Teng introduced the *smoothed case analysis* as a hybrid between average case and worst case analysis. In smoothed analysis, the input is object to slight random perturbations, e.g. modeled by adding a random vector from a fixed probability distribution to the input. The smoothed complexity of an algorithm or problem is then defined to be the worst case expected complexity over all perturbed inputs where the expectation is taken with respect to the random perturbation, e.g. the added random noise. The smoothed complexity is measured as a function of the input size and the magnitude of the perturbation. Indeed, later we will see that smoothed analysis is able to provide such a bridge between average case and worst case analysis.

Besides this appealing theoretical motivation the use of smoothed analysis is particularly well motivated in the context of *computational geometry*. In many areas we encounter applications from computational geometry, e.g. in computer-aided design, operations research, geographic information systems, computer graphics, and combinatorial optimization. In these applications of computational geometry the input data often come from experimental and physical measurements and are thus afflicted with some error since measuring devices have only limited accuracy. A standard assumption in physics is that this error is distributed according to the Gaussian normal distribution. Thus we can use smoothed analysis with Gaussian error to model inputs coming from physical measurements. By the assumptions that physical measurements are not precise and that the error is distributed according to the Gaussian normal distribution, smoothed analysis provides an expected worst case complexity measure for this class of inputs.

Another interesting motivation for smoothed analysis in the context of computational geometry is that the computations on a computer are carried out with limited accuracy. When we consider the case that the input points are computed with fixed precision arithmetic we observe that the rounded position of any point lies within a hypercube of a particular side length around its ‘real’ position where the side length of the hypercube depends on the rounding. We can now model this scenario by the assumption that every point is distributed uniformly in a hypercube centered at its ‘real’ position having an appropriate side length (depending on the considered rounding). We obtain thus an expected worst case complexity measure for computations under limited accuracy.

In this thesis, the combinatorial complexity of some fundamental geometric structures is considered. The main contribution is to introduce a formal model for smoothed analysis of the combinatorial complexity of geometric structures. Among other problems, this concept is applied to the number of *extreme points* of the *convex hull* of a point set in $\mathbb{R}^d$. In the following, some of the results for this problem are sketched.

In particular, for the convex hull we will assume that the input points are perturbed by Gaussian normal noise of variance σ^2. It is then shown that the worst case expected number of extreme points, or for short, the smoothed number of extreme points is polylogarithmic in the number of input points and polynomial in $1/\sigma$. For this result almost matching upper and lower bounds are given where the dimension d is considered as a constant. Interestingly, for uniform noise of variance σ^2 the smoothed number of extreme points is polynomial in the number of input points and $1/\sigma$. This reveals a significant discrepancy between the perturbation by Gaussian normal noise and by uniform noise. In other words, the assumptions of measurement errors and fixed precision arithmetic lead to quite different behavior and results.

Indeed, the analysis is much broader and provides actually bounds on the smoothed number of extreme points for a whole class of continuous probability distributions. An interesting open problem is to generally classify those probability distributions for which the smoothed complexity is low or high.

Smoothed analysis is also used to introduce a new complexity measure for motion, the *smoothed motion complexity*. Especially when considering applications on moving objects the influence of measurement errors is not negligible and of major importance. The use of smoothed analysis is thus very well motivated in the context of motion and motion complexity. The concept of smoothed motion complexity is then applied to the problem of maintaining a smallest axis-aligned bounding box of a moving point set.

Another fundamental geometric structure is the *Voronoi diagram* of a set of points in $\mathbb{R}^d$. The combinatorial complexity of Voronoi diagrams (= number of all faces) is assumed to be low in the average case but for most probability distributions explicit proofs are not published. In this thesis the case is considered that the points are chosen uniformly from a hypercube and it is shown that the expected complexity of the Voronoi diagram is then also linear in the number of points. Based on this average case analysis it seems possible to carry out a smoothed case analysis of the Voronoi diagram which is a quite interesting open problem.

1.1 Outline

In Chapter 2, an introduction to smoothed analysis is given with a formal definition and a broad overview on related work. In several papers, smoothed analysis and also its variants have been applied to problems from very distinct areas of research.

In the first technical Chapter 3, smoothed analysis is applied to the problem of counting the number of *left-to-right maxima* in a sequence of elements. The chapter starts with an average case analysis of the problem and presents then the smoothed case analysis in great detail where upper and lower bounds for various noise distributions are given. On the one hand, this problem and its analysis serves very well as an introductory example to the concepts and techniques used in this thesis. On the other hand, the left-to-right maxima problem represents somehow a 1-dimensional version of two other multi-dimensional problems that are considered in the next two chapters.

The problem to count the number of *extreme points* is a kind of canonical extension of the left-to-right maxima problem to higher dimensions, and it is treated in Chapter 4. The chapter starts with an average case analysis and presents then upper and lower bounds for the smoothed number of extreme points for various noise distributions.

In Chapter 5, motion and especially moving point sets are considered under the notion of *motion complexity* which is a complexity measure for movement of objects. It is proposed to extend this notion to *smoothed motion complexity*, which is as the name already indicates, a ‘smoothed’ version of motion complexity. This concept is then applied to the *bounding box* problem where the number of combinatorial changes to the description of the bounding box of a moving point set is considered. Upper and lower bounds on the smoothed motion complexity of the bounding box problem are given.

In the last technical Chapter 6, the *Voronoi diagram* of a random point set and particularly the expected number of *Voronoi vertices* is considered. It is shown to be only linear in the number of points for the case that the points are uniformly distributed in a hypercube. For the case that the points are chosen uniformly from inside a ball, Dwyer [Dwy91] gave already an explicit proof that the expected number of vertices is linear. He conjectured that similar results hold for other uniform distributions but proofs are lacking so far or not published in any form.

Each chapter starts with a short introduction and a brief overview of related work in order to help classify the considered problem and the obtained results, and ends with a summary and conclusion. Finally, in the last Chapter 7, a summary and conclusion of all results in this thesis is given together with a prospect of future work in this area.

1.2 Bibliographic Notes

Parts of the work presented here in this thesis have been published in preliminary form in the proceedings of the *European Symposium on Algorithms (ESA)* and the *European Workshop on Computation Geometry (EWCG)*. These publications are

- V. Damerow, H. Räcke, F. Meyer auf der Heide, C. Scheideler, and C. Sohler. Smoothed Motion Complexity. In *Proceedings of the 11th European Symposium on Algorithms (ESA)*, 2003. [DMR$^+$03]

- V. Damerow and C. Sohler. Smoothed Number of Extreme Points under Uniform Noise. In *Proceedings of the 20th European Workshop on Computational Geometry (EWCG)*, 2004. [DS04b]

- V. Damerow and C. Sohler. Extreme Points under Random Noise. In *Proceedings of the 12th European Symposium on Algorithms (ESA)*, 2004. [DS04a]

- M. Bienkowski, V. Damerow, F. Meyer auf der Heide, and C. Sohler. Average Case Complexity of Voronoi Diagrams of n Sites from the Unit Cube. In *Proceedings of the 21st European Workshop on Computational Geometry (EWCG)*, 2005. [BDMS05]

The results of Chapter 3 on the smoothed number of left-to-right maxima and of Chapter 5 on the smoothed motion complexity of the bounding box are based on the work in [DMR$^+$03]. The lower bounds for the smoothed number of extreme points are an extension of this work, too. The upper bounds for the smoothed number of extreme points are an extension of the results presented in the papers [DS04b] and [DS04a]. The average case analysis of the Voronoi diagram in Chapter 6 is based on the work presented in [BDMS05].

2 Smoothed Analysis

The general idea of smoothed analysis is to weaken the worst case complexity by adding small random noise to any input instance. The *smoothed complexity* of a problem is then the worst case expected complexity over all input instances, where the expectation is with respect to the random noise, and is given as a function of the input size and the relative magnitude of the perturbation. To see how this complexity measure compares to worst case and average case complexity we will consider these, too. In the following section, the worst case, average case and smoothed case complexity of *algorithms* is introduced and formally defined.

2.1 Smoothed Analysis of Algorithms

Let $\mathcal{X}_n$ denote the space of all input instances of length n to a particular algorithm $\mathcal{A}$, consider for example the space all linear programming problems of length n to the simplex algorithm. Let $\mathcal{T}_{\mathcal{A}}(x)$ denote the running time of algorithm $\mathcal{A}$ on input $x \in \mathcal{X}_n$. The *worst case complexity* of algorithm $\mathcal{A}$ is then the function

$$\mathcal{C}_{\text{worst}}(\mathcal{A}, n) \;:=\; \sup_{x \in \mathcal{X}_n} \mathcal{T}_{\mathcal{A}}(x) \;\; .$$

To consider the *average case complexity* of algorithm $\mathcal{A}$, a probability measure Δ on $\mathcal{X}_n$ is introduced. The average case complexity of algorithm $\mathcal{A}$ is then defined as

$$\mathcal{C}_{\text{ave}}(\mathcal{A}, n) \;:=\; \mathbf{E}_{\Delta}\big[\mathcal{T}_{\mathcal{A}}(x)\big] \;\; ,$$

where x is a random variable on $(\mathcal{X}_n, \Delta)$ and $\mathbf{E}_{\Delta}$ is the expectation with respect to Δ.

For the *smoothed complexity* of algorithm $\mathcal{A}$, in contrast, a probability measure Δ_x is considered for each input instance $x \in \mathcal{X}_n$, and the smoothed complexity of algorithm $\mathcal{A}$ is defined as

$$\mathcal{C}_{\text{smooth}}(\mathcal{A}, n) \;:=\; \sup_{x \in \mathcal{X}_n} \mathbf{E}_{\Delta_x}\big[\mathcal{T}_{\mathcal{A}}(y)\big] \;\; ,$$

where y is a random variable on $(\mathcal{X}_n, \Delta_x)$.

This definition of course heavily depends on the probability measures Δ_x that are used. Which measures to use depends on each individual problem, but normally one will consider measures which put much weight on inputs that are similar, or near, to x. This notion is best illustrated by the following examples.

1. If $\mathcal{X}_n$ is a finite discrete metric space with metric d, one can consider
$$\Delta_x(\{x'\}) = f(d(x,x'))$$
with some suitable function f that is e.g. supported on a bounded neighborhood of 0.

2. If $\mathcal{X}_n$ is a vector space and $\varphi : \mathcal{X}_n \to \mathbb{R}_{\geq 0}$ a suitable probability density function one can consider
$$\Delta_x(B) = \int_B \varphi(x - x')\,\mathrm{d}x'$$
where B is a subset of $\mathcal{X}_n$.

An important aspect of smoothed analysis is that the concept of smoothed complexity is a generalization of both average case and worst case complexity. While the former can be obtained by setting all distributions Δ_x equal to the same global distribution Δ, the latter is obtained in the case that the probability measures Δ_x are all concentrated in the point x itself. One could therefore also say that smoothed analysis actually interpolates between worst case and average case analysis.

Besides the general definition of smoothed complexity, in most papers that appeared on smoothed analysis, the following more concrete definition of smoothed complexity is used. Let $\mathcal{X}_n$ be a vector space, e.g. $\mathcal{X}_n = \mathbb{R}^n$. The smoothed complexity of algorithm $\mathcal{A}$ is then the function
$$\mathcal{C}_{\text{smooth}}(\mathcal{A}, n) \;:=\; \sup_{x \in \mathcal{X}_n} \mathbf{E}_\Delta\big[\mathcal{T}_\mathcal{A}(x + ||x|| \cdot \rho)\big] \;,$$
where Δ is again a probability measure on $\mathcal{X}_n$ and ρ a random vector on $(\mathcal{X}_n, \Delta)$.

The vector ρ is also denoted as the *random noise* by which the input instances to algorithm $\mathcal{A}$ are perturbed. By multiplying ρ with $||x||$, the magnitude of the perturbation is related to the magnitude of the input that is perturbed. This is important when the considered problems are invariant under scaling which is for example the case for linear programs. Otherwise it would happen that problems that are equivalent up to their ‘size’ obtain different results under the same perturbation. In the following we will denote this definition of smoothed analysis as the *additive perturbation scheme*.

The name “smoothed analysis” comes from the following observation. If we consider the complexity of a problem or an algorithm as a function from the input space to the combinatorial size of the problem or the running time of the algorithm or to any other complexity measure, and we plot this function, we obtain a complexity landscape, see also Figure 2.1. The smoothed complexity is then the highest peak in this landscape after it is convolved with a small random noise distribution. One could also interpret this as taking for each single input point the average complexity value over a small (weighted) neighborhood of this input point and assigning this value to the input point.

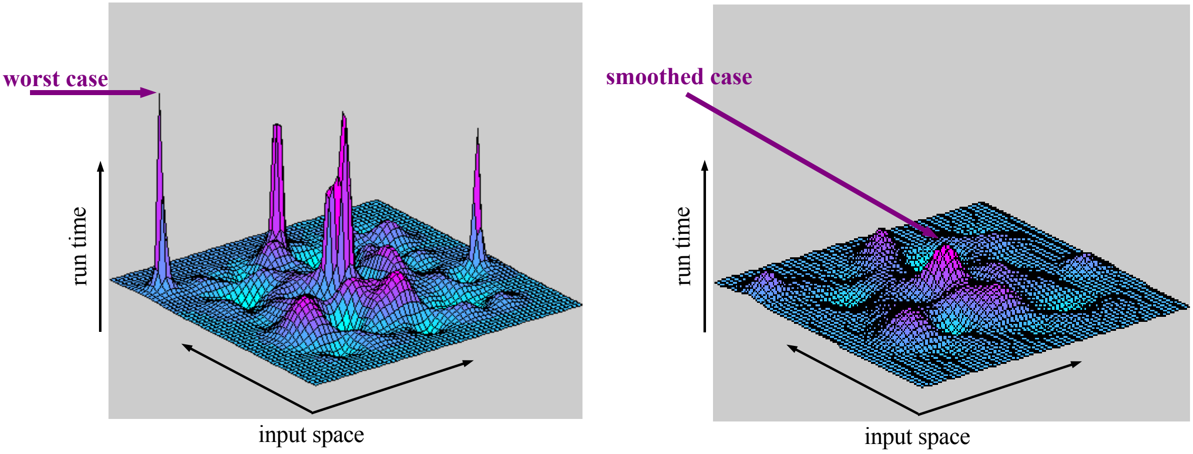

Figure 2.1: Complexity Landscape and Smoothed Complexity Landscape[1].

Smoothed analysis provides thus an insight into the topology of worst case instances in the input space. A low smoothed complexity shows whether bad inputs are pathological and isolated and can thus be avoided by already slight changes to the input instance. A high smoothed complexity reveals that bad inputs are somehow closely connected in terms of lying closely in the same neighborhood of the input space.

2.1.1 Smoothed Analysis of the Simplex Algorithm

Smoothed analysis was introduced by Spielman and Teng [ST04] to explain the typically good performance of the simplex algorithm in applications. For many variants of the simplex algorithm with different pivot rules, the worst case complexity is proven to be exponential, e.g. on the famous Klee-Minty cubes [KM72]. Nevertheless, the simplex algorithm shows an extraordinarily good performance on input instances from applications. The simplex algorithm is also known to have polynomial average case complexity under different notions of average case. The first average case analysis of the simplex algorithm was given by Borgwardt [Bor80], many other researchers followed him and investigated the average case complexity for other variants and pivot rules of the simplex algorithm. But still these results are not considered to give a satisfying explanation of the good behavior of the simplex algorithm encountered in practice.

Another intention for Spielman and Teng to introduce smoothed analysis as a new complexity measure was to model inputs that are encountered in applications. Besides the assumption that inputs are afflicted with measurement errors, another motivation for using randomly perturbed inputs comes from the observation, that usually inputs are formed in processes subject to chance, randomness, and arbitrary decisions. On the other hand

[1]Reproduction by courtesy of Daniel Spielman.

typical inputs can have very special properties such as being degenerated, which holds especially for the simplex algorithm. It is e.g. not unusual that typical input instances to the simplex algorithm contain many 0-entries in the constraints matrix. Indeed, this phenomenon is captured by taking the worst case over all perturbed input instances.

Spielman and Teng now consider a particular two-phase shadow-vertex simplex method on linear programs of the form

$$\begin{array}{ll} \text{maximize} & c^T \cdot x \\ \text{subject to} & A \cdot x \leq b \end{array}$$

where A is an $(m \times n)$-matrix, b is an m-vector, and c is an n-vector over the reals. Let $\mathcal{T}(A, b, c)$ denote the time complexity of this simplex method. They show that for every b and c, the smoothed complexity of this method,

$$\max_{A \in \mathbb{R}^{m \times n}} \mathbf{E}_G\big[\mathcal{T}(A + ||A|| \cdot G, b, c)\big] = \text{poly}(m, n, 1/\sigma)$$

is polynomial in $m, n,$ and $1/\sigma$, independent of b and c, where G is a Gaussian random $(m \times n)$-matrix of variance σ^2 centered at the origin.

Most remarkable about this result is, that the complexity of the simplex algorithm actually grows incredibly slow from the polynomial average case complexity to the exponential worst case complexity. This can be seen from the smoothed case bound, where the reciprocal of the standard deviation $1/\sigma$ goes only by a polynomial factor into the smoothed case bound.

The result of Spielman and Teng definitively marks a major step to an understanding of the behavior of the simplex algorithm in applications since it gives us a strong evidence that worst case instances are pathological and very isolated in the input space and therefore almost never encountered in practice.

2.1.2 Smoothed Analysis of Condition Numbers

Since 2001 several other papers on the smoothed analysis of algorithms have been written. The following results are all obtained for the additive perturbation scheme.

In numerical analysis, the *condition number* of a problem instance plays an important role. Generally, it is defined to indicate the sensitivity of the output to slight perturbations of the input instance, and the running time of an algorithm is often given in terms of the condition number of its input. Instead of bounding the condition number in the average case, it is proposed to consider the smoothed value of the condition number in order to show that already under small noise it is unlikely that a problem is ill-conditioned [SST03]. In this stream of research, the *Perceptron algorithm* [BD02] and *Renegar's algorithm* [DST03, ST03b] for linear programming, and most recently, *Gaussian elimination* [San04] are investigated.

In some sense connected to the analysis of smoothed condition numbers is the research work of Beier and Vöcking. They consider the sensitivity of perturbations to *discrete optimization problems* like the Knapsack problem [BV04a] and generally to the class of all binary optimization problems [BV04b]. The analyses are actually broader than in smoothed analysis since they can even handle the case that the elements of an input instance are from different probability measures. This enables them to study the effects of correlations between the random elements of input instances and gives also a nice framework to study semi-random input models where some elements of an input instance are adversarial and some stochastic. Maybe most remarkable is their result that a binary optimization problem has a polynomial smoothed complexity if and only if it has a pseudo-polynomial complexity. This result was recently extended to all integer linear programs [RV05].

2.1.3 Discrete Perturbation Models

We just saw how smoothed analysis with the additive perturbation scheme is applied to discrete algorithms. There are also attempts to define meaningful models for discrete perturbations in order to analyze other discrete algorithms. This makes sense in several settings, e.g. when investigating discrete structures like graphs and algorithms on graphs, but also for other problems and algorithms like sorting and scheduling.

Banderier et al. [BMB03] were the first to introduce discrete perturbations. They define *partial permutations* where the elements in a sequence are randomly permuted with probability p. In the partial permutations model they analyze then the *quicksort algorithm*[2].

Banderier et al. introduce also *partial bit randomization* where integers are perturbed by randomly choosing the last k bits. This model is extended by Becchetti et al. [BLMS$^+$03] and applied to clairvoyant scheduling. They introduce *smoothed competitive analysis* for online algorithms and present an analysis of the smoothed competitive ratio of the *multi-level feedback algorithm*. The usual competitive analysis gives often a too pessimistic estimation of the performance of an online algorithm since the online algorithm is measured against an optimal offline algorithm with full knowledge of the future. Sometimes the offline algorithm is thus simply too strong. In smoothed competitive analysis, the offline algorithm is somehow weaker since the input instances are randomly perturbed and it has not full knowledge about the future. In this sense, it seems quite natural to apply smoothed analysis to the analysis of online algorithms.

As already mentioned, Banderier et al. consider the number of comparisons of the *quicksort algorithm* in the partial permutations model. Eppstein [Spi] proposed to refine this model in such a way that the already sorted input and the reverse-sorted input stay unperturbed and that the perturbation of other inputs depends on their distance to these inputs. This perturbation would capture one significant property of the continuous perturbation scheme, that the zero (here the sorted and reverse-sorted inputs) stay unperturbed and that other inputs are perturbed in proportion to their distance to zero.

[2] They consider a deterministic variant of quicksort where the first element is always taken as pivot element.

Spielman and Teng [ST03a] try to further develop this concept and introduce *property-preserving perturbations*. The idea is to restrict a natural discrete perturbation model to preserve certain properties of the input. The notion of a 'natural' perturbation scheme in discrete settings is of course very vague since for most discrete problems is not nearly clear what natural means. E.g. for graphs Spielman and Teng propose to define perturbations by XORing an input graph with some random (sparse) graph such that some particular property of the input graph (such as having a ρ-clique) stays preserved by the perturbation. Then they measure the smoothed error probability of sub-linear time algorithms for this property. This idea is in some sense closely related to studies in the field of *property testing*.

In property testing one wants to decide whether an input instance has a certain property or differs significantly from all instances with that property. Property testing provides an alternative weakening of decision problems and is thus related to smoothed analysis. It seems therefore promising to join these two models. In this stream of research, Flaxman and Frieze [FF04] consider the diameter of randomly perturbed graphs and present an algorithm for recognizing strong connectivity of smoothed instances and also present a property tester for recognizing if a digraph is k-linked.

2.2 Smoothed Analysis of Geometric Structures

In this thesis the combinatorial complexity of (geometric) structures on point sets in $\mathbb{R}^d$ is considered, such as the number of *left-to-right maxima* in a sequence of elements or the number of *extreme points* of the *convex hull* of a point set. We investigate the smoothed case complexity of these structures under an additive perturbation scheme which is defined as follows.

Definition 1 (Perturbed Input Point) *For a fixed probability measure Δ defined on $\mathbb{R}^d$ consider an input point $p \in \mathbb{R}^d$ and a random noise vector ρ from Δ. Let $\widetilde{p}$ be the random vector that is given by*

$$\widetilde{p} \;:=\; p + \rho \;\; .$$

The random vector $\widetilde{p}$ is also denoted as the perturbed point to input point p, or for short as the ***perturbed input point****.*

For a set $\mathcal{P}$ of n input points $p_1, \ldots, p_n$ from $\mathbb{R}^d$ we denote by $\widetilde{\mathcal{P}}$ the set of perturbed input points under random noise Δ. It is given by

$$\widetilde{\mathcal{P}} \;:=\; \{p_1 + \rho_1, \ldots, p_n + \rho_n\} \;=\; \{\widetilde{p}_1, \ldots, \widetilde{p}_n\} \;\; ,$$

where $\rho_1, \ldots, \rho_n$ are independent and identically distributed random noise vectors from Δ.

The smoothed complexity of a geometrical structure is defined on input instances $\mathcal{P}$ where $p \in [0,1]^d$ for all $p \in \mathcal{P}$. The reason for this confindement is that the input point set needs to be normalized since the geometric structures considered here are invariant under scaling.

Definition 2 (Smoothed Complexity) *For a set $\mathcal{P}$ of n input points from $[0,1]^d$ consider some geometric structure on $\mathcal{P}$ and let $\mathcal{T}(\mathcal{P})$ denote a combinatorial complexity measure for this structure. For a fixed probability measure Δ on $\mathbb{R}^d$, the **smoothed complexity** of $\mathcal{T}(\mathcal{P})$ is defined as*

$$\max_{\mathcal{P}} \mathbf{E}_{\Delta}\big[\mathcal{T}(\widetilde{\mathcal{P}})\big]$$

where $\widetilde{\mathcal{P}}$ is the set of perturbed input points under random noise from Δ.

This perturbation scheme is also used by Bansal et al. who consider the labeling of smart dust [BMS04]. By smart dust a large set of small and very simple devices is meant, each consisting of a sensor and a sender that gathers sensor data and sends it to a central station. These devices are usually placed with low accuracy. It is thus very natural to model the imprecise information about the position of the sensor devices by random perturbations to the position information.

Banderier et al. [BMB03] consider also the smoothed complexity of geometric structures. Among other things they analyze the smoothed number of left-to-right maxima in the *partial permutations* model where each element is randomly permuted with probability p. This problem is also considered in this thesis, see Chapter 3 and especially page 14 where the result of Banderier et al. is discussed in more detail. However, the perturbation scheme of Banderier et al. significantly differs from the one used in this thesis.

2.2.1 Probability Distributions

In this thesis a very general class of probability measures is considered and the results hold for all measures from this class. In the following this class of probability measure is characterized by probability distribution functions.

In the 1-dimensional case we consider probability measures of the following kind. Let X be a 1-dimensional random vector taking values from some domain $R \subseteq \mathbb{R}$. Let the *probability distribution function* of X be given by

$$\Phi(x) \;:=\; \mathbf{Pr}[\,X \leq x\,] \;=\; \int_{-\infty}^{x} \varphi(z)\,\mathrm{d}z \;\;,$$

Here $\varphi : \mathbb{R} \to \mathbb{R}_{\geq 0}$ is a bounded, integrable function with $\int_{-\infty}^{\infty} \varphi(z)\;\mathrm{d}z = 1$ and is denoted as the *probability density function* of variable X. All probability distributions of the just described kind are denoted as *continuous probability distributions*.

In the d-dimensional case we consider probability measures that have as distribution function the d-fold product of a 1-dimensional continuous probability distribution function.

Let $X = (X_1, \ldots, X_d)$ be a d-dimensional random vector taking values from some area $R^d = \prod_{i=1}^{d} R \subseteq \mathbb{R}^d$ where $R \subseteq \mathbb{R}$. Let the *probability distribution function* of X be given by

$$\mathbf{Pr}[\, X_1 \leq x_1, \ldots, X_d \leq x_d \,] \;=\; \int_{-\infty}^{x_d} \cdots \int_{-\infty}^{x_1} \varphi(z_1) \cdots \varphi(z_d)\ \mathrm{d}z_1 \cdots\ \mathrm{d}z_d\ .$$

Again $\varphi : \mathbb{R} \to \mathbb{R}_{\geq 0}$ is a bounded, integrable function with $\int_{-\infty}^{\infty} \varphi(z)\ \mathrm{d}z = 1$ and is denoted as the 1-dimensional *components' density function* of X.

Note that all components of X are mutually independent and identically distributed. Probability distributions of the just described kind are denoted as *continuous d-dimensional product probability distributions*. Examples for such distributions are the d-dimensional Gaussian normal distribution or the uniform distribution in a d-dimensional hypercube.

Preliminaries. We consider the smoothed combinatorial complexity of geometric structures where the geometric structures are defined on point sets in $\mathbb{R}^d$. The smoothed complexity is usually given as a function of the input size, i.e. here the number of input points, and the reciprocal of the standard deviation of the 1-dimensional random noise distribution.

For uniform noise we consider the uniform distribution in a hypercube of side length 2ϵ centered at the origin. Note, that the standard deviation of the 1-dimensional uniform distribution (i.e. the uniform distribution in an interval $[-\epsilon, \epsilon]$) is then $\epsilon/\sqrt{3}$. Thus when considering uniform noise we do not explicitly state this but give the results only in terms of the side length of the hypercube. However, the results are still comparable to results for other noise distributions where the bounds are given in terms of the standard deviation.

The following other definitions hold throughout this thesis:

- The logarithm to basis 2 is denoted by $\log$, and the logarithm to basis e by $\ln$.
- The n-th harmonic number of first order is given by $\mathcal{H}(n) = \sum_{\ell=1}^{n} 1/\ell$, and of second order by $\mathcal{H}^{(2)}(n) = \sum_{\ell=1}^{n} 1/\ell^2$ for all $n \in \mathbb{N}$.
- It is $\mathcal{H}(n) = \ln(n) + \mathcal{O}(1)$.
- The Gamma function is given by $\Gamma(n+1) = n!$ and $\Gamma(n+1/2) = (2n)! \cdot \sqrt{\pi}/(n! \cdot 2^{2n})$ for all $n \in \mathbb{N}$.

3 Left-to-Right Maxima

In this first technical chapter we consider the problem to count the number of *left-to-right maxima* in a sequence of elements (numbers). We analyze the average and smoothed number of left-to-right maxima for all continuous probability distributions.

On the one hand, this problem is very basic and rather simple since it is one dimensional. It serves therefore as a good introductory example since proof ideas and basic technical steps will return later in the analysis of other problems. On the other hand, the results we obtain for this problem will be extended to higher dimensions in Chapter 4. There we consider the number of maximal points and extreme points of the convex hull. Maximal points are a kind of canonical extension of the left-to-right maxima problem to arbitrary dimensions. Chapter 5 deals with the number of combinatorial changes to the description of the bounding box of a moving point set. The left-to-right maxima problem serves there as an auxiliary problem to improve some of the results.

However, the presentation of techniques and proofs in this chapter is in great detail in order to make further reading more convenient. We will start now with the following formal problem definition.

Definition 3 (Left-to-Right Maxima) *Given is an arbitrary sequence $\mathcal{S}$ of n numbers (called elements) $\mathcal{S} = (s_1, \ldots, s_n)$ where $s_k \in \mathbb{R}$ for all $1 \leq k \leq n$. If all predecessors of s_k have smaller value than s_k, i.e. if $s_i < s_k$ for all $i < k$, element s_k is called a **left-to-right maximum** in $\mathcal{S}$. Let $\mathcal{L}(\mathcal{S})$ denote the number of left-to-right maxima in sequence $\mathcal{S}$.*

The main contribution of this chapter is a rather general lemma that can be used to obtain upper bounds on the smoothed number of left-to-right maxima for noise from continuous probability distributions. This lemma is applied to obtain explicit bounds for noise from the Gaussian normal distribution and the uniform distribution in a closed interval. The upper bounds are then complemented by lower bounds. Interestingly, these upper and lower bounds show that the smoothed number of left-to-right maxima differs significantly for random noise from the Gaussian normal and the uniform distribution. This is even more interesting since in the usual average case the expected number of left-to-right maxima is the same for all continuous probability distributions.

Related Work. In many textbooks about computer science and theory [Knu97, LL83, Kem84], the number of left-to-right maxima is considered in the context of permutations or in the analysis of basic algorithms, e.g. for finding a maximum in a sequence of numbers.

For an input sequence of n elements, the maximum number of left-to-right maxima is clearly n while it is 1 in the best case. The average number is known to be the n-th *harmonic number* $\mathcal{H}(n) = \sum_{\ell=1}^{n} 1/\ell$ for a wide variety of probability distributions. The standard deviation is known to be $(\mathcal{H}(n) - \mathcal{H}^{(2)}(n))^{1/2}$, where $\mathcal{H}^{(2)}(n) = \sum_{\ell=1}^{n} 1/\ell^2$ is the n-th harmonic number of second order.

The smoothed number of left-to-right maxima has already been analyzed but for a fundamentally different perturbation scheme than the one that is considered here. Banderier at al. [BMB03] introduce so called *partial permutations* where in a sequence of n elements each element is selected with some fixed probability p, and the selected elements are then randomly permuted where each permutation is equally likely. Under this model, the authors show that the smoothed number of left-to-right maxima is $\Omega(\sqrt{n/p})$ and $\mathcal{O}(\sqrt{n/p \cdot \log(n)})$. Interestingly, the worst case instance in the smoothed case under this model is the sequence $(n-k, n-k+1, \ldots, n, 1, 2, \ldots, n-k-1)$ for $k = \sqrt{n/p}$. For the perturbation scheme that is considered here, the worst case instance is the sequence $(1, 2, \ldots, n)$, see also Lemma 1.

Outline. In Section 3.1 the usual average case number of left-to-right maxima is considered. We will show how to derive the expected number of left-to-right maxima for a sequence of independent and identically distributed random elements chosen from a continuous probability distribution. The average case is already well known and a proof can actually be done by very simple considerations. But instead we present a more sophisticated way of showing the average case by use of integrals. This is done to better illustrate the approach for the smoothed case where the use of integrals is of major importance.

In Section 3.2 the smoothed case analysis is presented. The proofs are described in great detail since the basic steps of the analysis will return in Chapter 4 where we consider the number of *extreme points* in d dimensions. We derive a general lemma (Lemma 2) to upper bound the smoothed number of left-to-right maxima for all continuous noise distributions. This lemma is applied to obtain explicit bounds for noise from the Gaussian normal distribution $\mathrm{N}(0, \sigma)$ and the uniform distribution in an interval $[-\epsilon, \epsilon]$. For these noise distributions we get upper bounds of $\mathcal{O}(1/\sigma \cdot \log(n)^{3/2} + \log(n))$ and $\mathcal{O}(\sqrt{n \cdot \log(n)/\epsilon} + \log(n))$, respectively. Also some general class of unimodal probability distributions is considered which have monotonic density functions with one single 'peak'.

In Section 3.3 lower bounds for the smoothed number of left-to-right maxima are shown which prove that the upper bounds are almost tight for Gaussian normal and uniform noise.

In the last section the results of this chapter are shortly summarized. We also work out the property for continuous, unimodal probability distributions that are necessary to provide a low smoothed complexity that comes close to the average case complexity.

3.1 Average Case Analysis

We consider now the case when the elements $s_1, \ldots, s_n$ are independent and identically distributed random variables with probability distribution function $\Phi : \mathbb{R} \to [0,1]$. Let Φ have the integrable density function $\varphi : \mathbb{R} \to \mathbb{R}_{\geq 0}$ with $\int_{-\infty}^{\infty} \varphi(x) \, \mathrm{d}x = 1$. The probability distribution of input element s_i is then given by

$$\mathbf{Pr}[\, s_i \leq x \,] \;=\; \Phi(x) \;=\; \int_{-\infty}^{x} \varphi(z) \, \mathrm{d}z \;\; .$$

The following theorem holds for all distributions of the above kind.

Theorem 1 *The expected number of left-to-right maxima in a sequence $\mathcal{S}$ of n independent and identically distributed random variables chosen from a continuous probability distribution is*

$$\mathbf{E}\left[\mathcal{L}(\mathcal{S})\right] \;=\; \mathcal{H}(n) \;=\; \Theta(\log(n)) \;\; ,$$

where $\mathcal{H}(n)$ denotes the n-th harmonic number, i.e. $\mathcal{H}(n) = \sum_{k=1}^{n} 1/k$.

Proof. The probability for the k-th element s_k to be a left-to-right maximum in $\mathcal{S}$ is given by

$$\mathbf{Pr}[\, s_k \text{ is a left-to-right maximum in } \mathcal{S} \,] \;=\; \int_{-\infty}^{\infty} \varphi(x) \cdot \Phi(x)^{k-1} \, \mathrm{d}x \;\; . \tag{3.1}$$

This holds since $\Phi(x)$ is the probability that a random variable s_i is not greater than x, and since all variables are independent and identically distributed $\Phi(x)^{k-1}$ equals the probability that all $k-1$ predecessors of s_k are smaller than x. Consequently, the expression $\varphi(x) \cdot \Phi(x)^{k-1} \, \mathrm{d}x$ can be interpreted as the probability that the k-th element reaches x and is a left-to-right maximum. Hence, integration over x gives the probability that s_k is a left-to-right maximum.

In order to compute the integral in (3.1) we will use the substitution $z := \Phi(x)$, where $\mathrm{d}z = \varphi(x) \, \mathrm{d}x$. This yields

$$\int_{-\infty}^{\infty} \varphi(x) \cdot \Phi(x)^{k-1} \, \mathrm{d}x \;=\; \int_{0}^{1} z^{k-1} \, \mathrm{d}z \;=\; \frac{1}{k} \;\; .$$

Of course, this result only reveals the fact that the probability for the k-th element to be the largest among the first k elements is $1/k$. We can exploit linearity of expectation and sum up over the probabilities for all k which leads to

$$\mathbf{E}\left[\mathcal{L}(\mathcal{S})\right] \;=\; \sum_{k=1}^{n} \frac{1}{k} \;=\; \mathcal{H}(n) \;\; .$$

Consequently, the theorem is proven. □

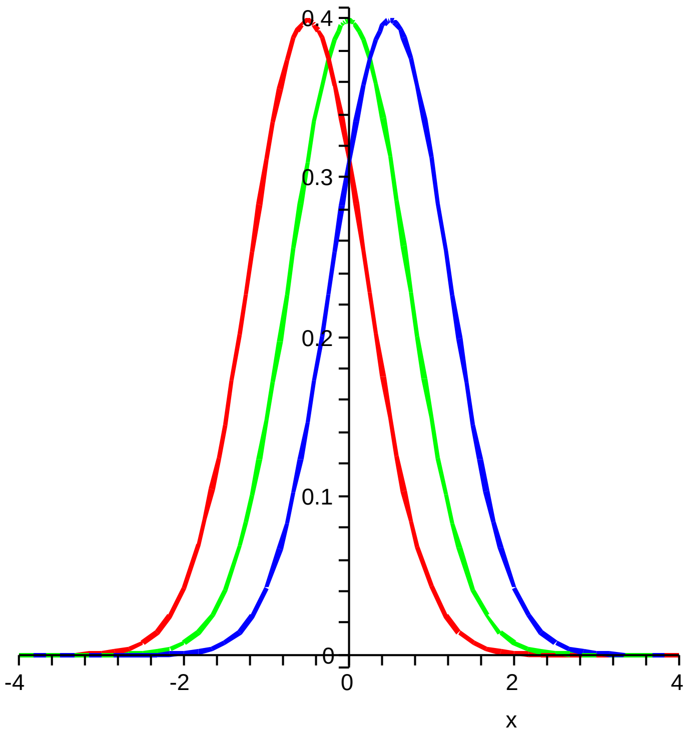

Figure 3.1: Shifted Gaussian bell curves – the density functions of $\mathrm{N}(0,1), \mathrm{N}(1/2,1)$ and $\mathrm{N}(-1/2,1)$.

3.2 Upper Bounds for the Smoothed Case

In the last section we saw how to express the probability that a random element is a left-to-right maximum by an integral expression. In order to investigate the smoothed number of left-to-right maxima we will express the probability that a perturbed element is a left-to-right maximum also by an integral expression in a similar way.

Let us first consider the random noise. Let $\rho_1, \ldots, \rho_n$ be independent and identically distributed random numbers from a continuous noise distribution with integrable density function φ and corresponding distribution function Φ. So for a given sequence of input elements $\mathcal{S} = (s_1, \ldots s_n)$ we consider the sequence of perturbed elements $\widetilde{\mathcal{S}} = (\widetilde{s}_1, \ldots, \widetilde{s}_n)$ where $\widetilde{s}_k = s_k + \rho_k$, for all $1 \leq k \leq n$. For reasons of normalization we assume that $s_1, \ldots, s_n \in (0,1]$.

First of all we make the following observation. When the noise distribution is fixed the sequence of perturbed elements also becomes a random distribution where an element $\widetilde{s}_k$ is a random number with density function $\varphi(x - s_k)$ and according distribution function $\mathbf{Pr}[\,\widetilde{s}_k \leq x\,] = \Phi(x - s_k)$. Contrary to the usual average case, the perturbed elements are not drawn from the same probability distribution and are thus not identically distributed. Instead each perturbed element has a density and distribution function that is a copy of the density and distribution function of the random noise with different parameters. The density and distribution function of $\widetilde{s}_k$ differ from the density and distribution function of the random noise only in the sense that the curves are shifted by the according amount of s_k. For example, in Figure 3.1 three shifted versions of the Gaussian density function are depicted.

Now we can write analogously to (3.1) that

$$\mathbf{Pr}\left[\,\widetilde{s}_k \text{ is a left-to-right maximum in } \widetilde{\mathcal{S}}\,\right] \;=\; \int_{-\infty}^{\infty} \varphi(x - s_k) \cdot \prod_{j=1}^{k-1} \Phi(x - s_j)\, \mathrm{d}x \;\; . \qquad (3.2)$$

In order to proceed we first need to show the following lemma.

Lemma 1 *The maximum expected number of left-to-right maxima in a sequence $\widetilde{\mathcal{S}}$ of perturbed elements is obtained for a sequence of input elements $\mathcal{S}$ that is monotonically increasing.*

Proof. Consider the two sequences of input elements $\mathcal{S}_1 = (s_1, \ldots, s_{k-2}, s_k, s_{k-1})$ and $S_2 = (s_1, \ldots, s_{k-2}, s_{k-1}, s_k)$, where $s_{k-1} < s_k$, and let the difference between the these two elements be $\gamma := s_k - s_{k-1} > 0$. For a fixed noise distribution we want now to show that

$$\beta \;:=\; \mathbf{E}\left[\mathcal{L}(\widetilde{\mathcal{S}}_2)\right] - \mathbf{E}\left[\mathcal{L}(\widetilde{\mathcal{S}}_1)\right] \;\geq\; 0 \;\; ,$$

where $\mathcal{L}(\widetilde{\mathcal{S}}_1)$ and $\mathcal{L}(\widetilde{\mathcal{S}}_2)$ denote the number of left-to-right maxima in the perturbed sequences $\widetilde{\mathcal{S}}_1$ and $\widetilde{\mathcal{S}}_2$ under noise from the fixed probability distribution, respectively.

To see this it suffices to consider for both sequences $\mathcal{S}_1$ and $\mathcal{S}_2$ the probability that s_{k-1} and that s_k become a left-to-right maximum in the perturbed sequences $\widetilde{\mathcal{S}}_1$ and $\widetilde{\mathcal{S}}_2$. Since all elements are independent the probabilities for the other elements to be a left-to-right maximum are equal for both sequences and we can neglect them. Hence we get that

$$\begin{aligned} \beta \;\;=\;\; & \Big(\mathbf{Pr}\Big[\,\widetilde{s}_{k-1} \text{ is left-to-right max. in } \widetilde{\mathcal{S}}_2\,\Big] + \mathbf{Pr}\Big[\,\widetilde{s}_k \text{ is left-to-right max. in } \widetilde{\mathcal{S}}_2\,\Big]\Big) \\ & - \Big(\mathbf{Pr}\Big[\,\widetilde{s}_k \text{ is left-to-right max. in } \widetilde{\mathcal{S}}_1\,\Big] + \mathbf{Pr}\Big[\,\widetilde{s}_{k-1} \text{ is left-to-right max. in } \widetilde{\mathcal{S}}_1\,\Big]\Big) \;\; . \end{aligned}$$

Let now φ and Φ denote the density function and distribution function of the fixed noise distribution, respectively. Further, for ease of notation let $\mathcal{F}(x) = \prod_{j=1}^{k-2} \Phi(x - s_j)$ denote the probability that the first $k-2$ elements in $\widetilde{S}_1$ and $\widetilde{S}_2$ are not greater than x. Then it follows analogously to (3.2) that

$$\begin{aligned}
\beta &= \left(\int_{-\infty}^{\infty} \varphi(x-s_{k-1}) \cdot \mathcal{F}(x)\,\mathrm{d}x + \int_{-\infty}^{\infty} \varphi(x-s_k) \cdot \Phi(x-s_{k-1}) \cdot \mathcal{F}(x)\,\mathrm{d}x\right) \\
&\quad - \left(\int_{-\infty}^{\infty} \varphi(x-s_k) \cdot \mathcal{F}(x)\,\mathrm{d}x + \int_{-\infty}^{\infty} \varphi(x-s_{k-1}) \cdot \Phi(x-s_k) \cdot \mathcal{F}(x)\,\mathrm{d}x\right) \\
&= \int_{-\infty}^{\infty} \varphi(x-s_{k-1}) \cdot (1-\Phi(x-s_k)) \cdot \mathcal{F}(x)\,\mathrm{d}x \qquad (3.3) \\
&\quad - \int_{-\infty}^{\infty} \varphi(x-s_k) \cdot (1-\Phi(x-s_{k-1})) \cdot \mathcal{F}(x)\,\mathrm{d}x \ . \qquad (3.4)
\end{aligned}$$

In the next step we exploit in (3.3) that $\Phi(x)$ is a positive and monotonically increasing function, and in (3.4) we substitute x by $x+\gamma$, which yields

$$\begin{aligned}
\beta &\geq \int_{-\infty}^{\infty} \varphi(x-s_{k-1}) \cdot (1-\Phi(x-s_{k-1})) \cdot \mathcal{F}(x)\,\mathrm{d}x \\
&\quad - \int_{-\infty}^{\infty} \varphi(x-s_{k-1}) \cdot (1-\Phi(x-s_{k-1}+\gamma)) \cdot \mathcal{F}(x)\,\mathrm{d}x \quad \geq \quad 0 \ .
\end{aligned}$$

The last inequality holds again since $\Phi(x)$ is a positive and monotonically increasing function. Thus Lemma 1 is shown. □

So from now on when considering the smoothed number of left-to-right maxima, we will assume that input sequence $\mathcal{S}$ is a sequence of monotonically increasing elements from $(0,1]$. The main idea for computing the integral (3.2) is now to subdivide the interval $(0,1]$ into $m = \lceil 1/\delta \rceil$ smaller intervals of length δ (the last one possibly shorter). Here δ is a small parameter that is specified later. Then the sequence $\mathcal{S}$ of unperturbed input elements is subdivided into m subsequences $\mathcal{S}_1, \ldots, \mathcal{S}_m$ where S_ℓ contains all elements $s \in \mathcal{S}$ that lie in the ℓ-th small interval, i.e.

$$\mathcal{S}_\ell = \big(s \in \mathcal{S} \bigm| (\ell-1) \cdot \delta < s \leq \ell \cdot \delta\big) \ .$$

Note that all subsequences $\mathcal{S}_1, \ldots, \mathcal{S}_m$ of $\mathcal{S}$ are also monotonically increasing, and that $\mathcal{S} = (\mathcal{S}_1, \ldots, \mathcal{S}_m)$ by Lemma 1. This enables us to utilize that

$$\mathcal{L}(\widetilde{\mathcal{S}}) \leq \sum_{\ell=1}^{m} \mathcal{L}(\widetilde{\mathcal{S}}_\ell) \ , \qquad (3.5)$$

where $\widetilde{\mathcal{S}}_\ell$ is the sequence of perturbed elements to input sequence $\mathcal{S}_\ell$ and $\mathcal{L}(\widetilde{\mathcal{S}}_\ell)$ is the number of left-to-right maxima in sequence $\widetilde{\mathcal{S}}_\ell$.

The advantage of this approach is that for small enough δ the elements of a subsequence $\widetilde{\mathcal{S}}_\ell$ behave almost as in the usual average case. Intuitively, the reason for this is that the unperturbed input elements lie very close together and that the density functions of the perturbed elements differ only very little.

Indeed, we will apply for each subsequence $\widetilde{\mathcal{S}}_\ell$ the average case bounds as if the elements were identically distributed random numbers. Of course we make here an error but this can easily be fixed. Then by (3.5) and a summation over all subsequences $\widetilde{\mathcal{S}}_\ell$, we obtain an upper bound on the smoothed number of left-to-right maxima for the whole sequence $\widetilde{\mathcal{S}}$.

In order to remedy the error between the smoothed number and the average number of left-to-right maxima in subsequence $\widetilde{\mathcal{S}}_\ell$ we proceed as follows. In a first step we cut off the "tail" of the considered noise distribution and treat it separately, i.e. we bound the tail probability (later denoted as $\mathcal{Z}$) and count the expected number of elements from the tail as left-to-right maxima. This gives us in the end an additive error depending on the considered noise distribution and on δ since the probability of the tail depends on these, too.

For the remaining part of the probability distribution, we observe then that the smoothed number and the average number of left-to-right maxima for each subsequence differ only by a multiplicative factor r which depends on the choice of δ. A multiplication of the average number of left-to-right maxima with r remedies then the error on the smoothed number of left-to-right maxima.

However, we have here a trade-off between δ and the number of subsequences. When we choose δ small in order to make the tail of the noise distribution not too heavy and thus the additive error small, the number of subsequences (which is $m = \lceil 1/\delta \rceil$) and therefore the smoothed number of left-to-right for the whole sequence becomes large.

In order to proceed with the analysis, we will now fix one of the subsequences and without loss of generality we consider the first subsequence S_1. Let n_1 denote the number of elements in subsequence $\mathcal{S}_1$, i.e. $\mathcal{S}_1 = (s_1, \ldots, s_{n_1})$. For an element s_k in $\mathcal{S}_1$ we know that

$$\begin{aligned}
\mathbf{Pr}\Big[\, \widetilde{s}_k \text{ is left-to-right max. in } \widetilde{\mathcal{S}}_1 \Big] \quad &= \quad \int_{-\infty}^{\infty} \varphi(x - s_k) \cdot \prod_{j=1}^{k-1} \Phi(x - s_j)\, \mathrm{d}x \\
&\le \quad \int_{-\infty}^{\infty} \varphi(x - s_{n_1}) \cdot \Phi(x - s_1)^{k-1}\, \mathrm{d}x \\
&\le \quad \int_{-\infty}^{\infty} \varphi(x) \cdot \Phi(x + \delta)^{k-1}\, \mathrm{d}x \ , \qquad (3.6)
\end{aligned}$$

where the last step follows by substituting x by $x+s_{n_1}$ and the observation that $s_{n_1} - s_1 \le \delta$.

We could easily solve this integral if instead of $\varphi(x)$ there would occur a $\varphi(x+\delta)$. Thus in a next step we will expand the integrand in (3.6) by a multiplication:

$$\varphi(x) = \varphi(x+\delta) \cdot \frac{\varphi(x)}{\varphi(x+\delta)} \ .$$

Where the ratio $\varphi(x)/\varphi(x+\delta)$ is bounded we can extract it from inside the integral and solve the remaining integral as in the average case analysis. Let

$$\mathcal{Z}^{\varphi}_{\delta,r} \;:=\; \left\{ x \in \mathbb{R} \;\middle|\; \frac{\varphi(x)}{\varphi(x+\delta)} > r \right\}$$

denote the subset of $\mathbb{R}$ that contains all elements x for which the ratio $\varphi(x)/\varphi(x+\delta)$ is larger than some positive constant r. To avoid difficulties with zeros of φ, we will use alternatively the definition $\mathcal{Z}^{\varphi}_{\delta,r} := \{x \in \mathbb{R} \mid \varphi(x) > r \cdot \varphi(x+\delta)\}$. The constant r is the same as seen earlier, i.e. it is the multiplicative error between the average case and smoothed case number of left-to-right maxima.

We can now formulate and prove the main lemma of this section as follows.

Lemma 2 *For a fixed continuous probability distribution with integrable density function φ and for positive parameters δ and r define the set $\mathcal{Z}^{\varphi}_{\delta,r} := \{x \in \mathbb{R} \mid \varphi(x) > r \cdot \varphi(x+\delta)\}$. Let $\mathcal{Z}$ be the probability of set $\mathcal{Z}^{\varphi}_{\delta,r}$, i.e.*

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

For random noise from the fixed probability distribution, the smoothed number of left-to-right maxima over all input sequences $\mathcal{S}$ of n elements from $(0,1]$ is

$$\max_{\mathcal{S}} \mathbf{E}\left[\mathcal{L}(\widetilde{\mathcal{S}})\right] \;\le\; \max\left\{ r \cdot \left\lceil \frac{1}{\delta} \right\rceil \cdot \mathcal{H}(n) + n \cdot \mathcal{Z}, \mathcal{H}(n) \right\} \ .$$

Proof. Let again $m := \lceil 1/\delta \rceil$. Without loss of generality we consider again the input subsequence $\mathcal{S}_1$. We saw already in (3.6) that for an element $s_k \in \mathcal{S}_1$ it is

$$\mathbf{Pr}\left[\, \widetilde{s_k} \text{ is left-to-right max. in } \widetilde{\mathcal{S}_1} \,\right] \le \int_{-\infty}^{\infty} \varphi(x) \cdot \Phi(x+\delta)^{k-1}\ \mathrm{d}x \ .$$

We can now compute this integral in the following way

$$\begin{aligned}
(3.6) \;&=\; \int_{\mathbb{R}} \varphi(x) \cdot \Phi(x+\delta)^{k-1}\ \mathrm{d}x \\
&=\; \int_{\mathbb{R}-\mathcal{Z}^{\varphi}_{\delta,r}} \varphi(x) \cdot \Phi(x+\delta)^{k-1}\ \mathrm{d}x \;+\; \int_{\mathcal{Z}^{\varphi}_{\delta,r}} \varphi(x) \cdot \Phi(x+\delta)^{k-1}\ \mathrm{d}x \\
&\le\; \int_{\mathbb{R}-\mathcal{Z}^{\varphi}_{\delta,r}} \varphi(x+\delta) \cdot \frac{\varphi(x)}{\varphi(x+\delta)} \cdot \Phi(x+\delta)^{k-1}\ \mathrm{d}x \;+\; \int_{\mathcal{Z}^{\varphi}_{\delta,r}} \varphi(x)\ \mathrm{d}x \\
&\le\; r \cdot \int_{\mathbb{R}-\mathcal{Z}^{\varphi}_{\delta,r}} \varphi(x+\delta) \cdot \Phi(x+\delta)^{k-1}\ \mathrm{d}x \;+\; \mathcal{Z} \\
&\le\; r \cdot \int_{\mathbb{R}} \varphi(x) \cdot \Phi(x)^{k-1}\ \mathrm{d}x \;+\; \mathcal{Z} \quad=\quad r \cdot \frac{1}{k} \;+\; \mathcal{Z} \ .
\end{aligned}$$

It follows for input subsequence $\mathcal{S}_1$ that

$$\mathbf{E}\left[\mathcal{L}(\widetilde{\mathcal{S}}_1)\right] \;\leq\; \sum_{k=1}^{n_1}(r \cdot \frac{1}{k} + \mathcal{Z}) \;=\; r \cdot \mathcal{H}(n_1) + n_1 \cdot \mathcal{Z} \;\;.$$

Analogous results hold also for subsequences $\mathcal{S}_2, \ldots, \mathcal{S}_m$, and by (3.5) it follows that

$$\mathbf{E}\left[\mathcal{L}(\widetilde{\mathcal{S}})\right] \;\leq\; \sum_{\ell=1}^{m} \mathbf{E}\left[\mathcal{L}(\widetilde{\mathcal{S}}_\ell)\right] \;\leq\; \sum_{\ell=1}^{m} r \cdot \mathcal{H}(n_\ell) + n_\ell \cdot \mathcal{Z} \;\leq\; r \cdot m \cdot \mathcal{H}(n) + n \cdot \mathcal{Z} \;\;.$$

Since the analysis is independent of the considered input sequence it holds for any monotonically increasing input sequence $\mathcal{S}$. To assure that the smoothed number of left-to-right maxima does not drop below the average number, we take the maximum over $r \cdot m \cdot \mathcal{H}(n) + n \cdot \mathcal{Z}$ and $\mathcal{H}(n)$. Thus Lemma 2 follows immediately. $\square$

3.2.1 Gaussian Normal Noise

In this subsection we show how to apply Lemma 2 to the case of normally distributed noise. Let now

$$\varphi(x) \;:=\; \frac{1}{\sqrt{2\pi}\,\sigma} \cdot e^{-\frac{x^2}{2\sigma^2}}$$

denote the density function for the standard Gaussian normal distribution with expectation 0 and variance σ^2, denoted by $\mathrm{N}(0, \sigma)$.

Theorem 2 *For random noise from the standard Gaussian normal distribution* $\mathrm{N}(0, \sigma)$*, the smoothed number of left-to-right maxima over all input sequences* $\mathcal{S}$ *of* n *elements from* $(0, 1]$ *is*

$$\begin{aligned}\max_{\mathcal{S}} \mathbf{E}\left[\mathcal{L}(\widetilde{\mathcal{S}})\right] \;&\leq\; \max\left\{e^3 \cdot \left\lceil \sqrt{\ln(n)}/\sigma \right\rceil \cdot \mathcal{H}(n) + 1, \mathcal{H}(n)\right\} \\ &=\; \mathcal{O}\left(\frac{1}{\sigma} \cdot \log(n)^{3/2} + \log(n)\right) \;\;.\end{aligned}$$

Proof. In order to utilize Lemma 2 we choose $\delta := \sigma/\sqrt{\ln(n)}$. For $x \leq \sigma\sqrt{2\ln(n)}$ it holds that

$$\varphi(x)/\varphi(x+\delta) \;=\; e^{(\delta/\sigma^2)\cdot x + \delta^2/(2\sigma^2)} \;=\; e^{x/(\sigma\sqrt{\ln(n)}) + 1/(2\ln(n))} \;\leq\; e^{\sqrt{2} + 1/(2\ln(n))} \;\leq\; e^3 \;\;.$$

Therefore, if we choose $r := e^3$ we can conclude that $\mathcal{Z}^{\varphi}_{\delta,r} \subset [\sigma\sqrt{2\ln(n)}, \infty)$. Now, we will derive a bound on $\mathcal{Z} = \int_{\mathcal{Z}^{\varphi}_{\delta,r}} \varphi(x)\;\mathrm{d}x$ by using the following claim.

Claim 1 *For the density function* $\varphi(x)$ *of the Gaussian normal distribution* $\mathrm{N}(0, \sigma)$ *it holds for any* $k \geq 1/\sqrt{2\pi}$ *that*

$$\int_{\sigma\cdot k}^{\infty} \varphi(x)\;\mathrm{d}x \;\leq\; e^{-k^2/2} \;\;.$$

The proof of this claim is deferred to the end of this subsection. With the claim it follows that

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

Altogether we can apply Lemma 2 with the parameters $\delta = \sigma/\sqrt{\ln n}$, $r = e^3$, and $\mathcal{Z} = 1/n$. It follows that for every input sequence $\mathcal{S}$ the expected number of left-to-right maxima in the perturbed sequence $\widetilde{\mathcal{S}}$ under Gaussian normal noise is at most

$$\mathbf{E}\left[\mathcal{L}(\widetilde{\mathcal{S}})\right] \;\le\; \max\left\{e^3 \cdot \left\lceil \sqrt{\ln(n)}/\sigma \right\rceil \cdot \mathcal{H}(n) + 1, \mathcal{H}(n)\right\}\ .$$

□

We observe that it does not make sense to apply Theorem 2 for arbitrary values of σ. If $\sigma \ge \Omega(\sqrt{\log(n)})$ we obtain the average case bound of $\mathcal{O}(\log(n))$, and thus we cannot distinguish in the analysis between the usual average case and the smoothed case. If $\sigma \le \mathcal{O}(\log(n)^{3/2}/n)$ we obtain an expected number of left-to-right maxima of $\mathcal{O}(n)$. This means that for variances this small the perturbation of (worst case input) instances does not show any effect in our analysis.

It remains now to prove Claim 1.

Proof of Claim 1. By a linear substitution $t = x^2/(2\sigma^2)$, $\mathrm{d}x = \sigma/\sqrt{2t}\,\mathrm{d}t$ we get that

$$\begin{aligned}
\int_{\sigma\cdot k}^{\infty} \varphi(x)\,\mathrm{d}x &= \int_{\sigma\cdot k}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma}\cdot e^{-\frac{x^2}{2\sigma^2}}\,\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}}\cdot\int_{k^2/2}^{\infty} e^{-t}\cdot\frac{1}{\sqrt{2t}}\,\mathrm{d}t \\
&\le \frac{1}{\sqrt{2\pi}}\cdot\frac{1}{k}\cdot e^{-k^2/2} \quad\le\quad e^{-k^2/2}\ .
\end{aligned}$$

$\square_{\text{Claim 1}}$

3.2.2 Uniform Noise

We consider now random noise that is uniformly distributed in the interval $[-\epsilon, \epsilon]$. The corresponding density function is given by

$$\varphi(x) := \begin{cases} \frac{1}{2\epsilon} & \text{if } x \in [-\epsilon,\epsilon] \\ 0 & \text{else}\ . \end{cases}$$

Theorem 3 *For random noise from the uniform distribution in* $[-\epsilon, \epsilon]$*, the smoothed number of left-to-right maxima over all input sequences* $\mathcal{S}$ *of* n *elements from* $(0, 1]$ *is*

$$\begin{aligned}\max_{\mathcal{S}} \mathbf{E}\left[\mathcal{L}(\widetilde{\mathcal{S}})\right] &\leq \max\left\{2 \cdot \left\lceil \sqrt{\frac{n \cdot \mathcal{H}(n)}{2\epsilon}} \right\rceil, \mathcal{H}(n)\right\} \\ &= \mathcal{O}\left(\sqrt{\frac{n \cdot \log(n)}{\epsilon}} + \log(n)\right) \quad .\end{aligned}$$

Proof. Again we want to utilize Lemma 2 . We choose $r = 1$, then it follows immediately that for $0 < \delta < 2\epsilon$ we have $\mathcal{Z}^{\varphi}_{\delta,r} = \left\{x \in \mathbb{R} \mid \varphi(x) > \varphi(x+\delta)\right\} = [\epsilon - \delta, \epsilon]$. We can now compute $\mathcal{Z}$ which is

$$\mathcal{Z} = \int_{\epsilon-\delta}^{\epsilon} \frac{1}{2\epsilon}\, \mathrm{d}x = \frac{\delta}{2\epsilon} \quad .$$

With Lemma 2 it follows that the smoothed number of left-to-right maxima is at most $\lceil 1/\delta \rceil \cdot \mathcal{H}(n) + n \cdot \delta/(2\epsilon)$. If we choose $\delta = \sqrt{2\epsilon \cdot \mathcal{H}(n)/n}$ we get that for every input sequence $\mathcal{S}$ the expected number of left-to-right maxima in the perturbed sequence $\widetilde{\mathcal{S}}$ under uniform noise is at most

$$\max_{\mathcal{S}} \mathbf{E}\left[\mathcal{L}(\widetilde{\mathcal{S}})\right] \leq \max\left\{2 \cdot \left\lceil \sqrt{\frac{n \cdot \mathcal{H}(n)}{2\epsilon}} \right\rceil, \mathcal{H}(n)\right\} \quad .$$

□

Again we observe that Theorem 3 is not applicable for uniform distributions with arbitrary values of ϵ. If $\epsilon \geq \Omega(n/\log(n))$ we obtain the average case bound of $\mathcal{O}(\log(n))$ and we cannot distinguish in our analysis between the usual average case and the smoothed case. If $\epsilon \leq \mathcal{O}(\log(n)/n)$ we obtain an expected number of left-to-right maxima of $\mathcal{O}(n)$ and we cannot distinguish if a (worst case) instance is perturbed or not.

3.2.3 Unimodal Noise Distributions

In this section we investigate upper bounds for general noise distributions that fulfill the following condition. We denote a continuous probability distribution as *unimodal* if the corresponding integrable density function is bounded and monotonically increasing on $\mathbb{R}_{\leq 0}$ and monotonically decreasing on $\mathbb{R}_{\geq 0}$. In other words, the density function has a single peak at $x = 0$. Note that the following analysis holds also for distributions that have a single peak not at $x = 0$ but at any other x. But for ease of notation we restrict the analysis to the first case. The following theorem gives now an upper bound on the number of left-to-right maxima for arbitrary unimodal noise distributions with peak at $x = 0$.

Theorem 4 *For random noise from a unimodal continuous probability distribution with density function φ, the smoothed number of left-to-right maxima over all input sequences $\mathcal{S}$ of n elements from $(0,1]$ is*

$$\begin{aligned}\max_{\mathcal{S}} \mathbf{E}\left[\mathcal{L}(\widetilde{\mathcal{S}})\right] &\leq \max\left\{7 \cdot \sqrt{n \cdot \mathcal{H}(n) \cdot \varphi(0)}, \mathcal{H}(n)\right\} \\ &= \mathcal{O}\left(\sqrt{n \cdot \log n \cdot \varphi(0)} + \log n\right) \ .\end{aligned}$$

Proof. Since we consider here monotonic probability distributions, $\varphi(0)$ denotes the (global) maximum of the density function. In order to utilize Lemma 2 we choose $r := 2$ whereas δ will be chosen later.

Again we need to derive a bound for $\mathcal{Z} = \int_{\mathcal{Z}^{\varphi}_{\delta,r}} \varphi(x) \, \mathrm{d}x$. We want now to find a covering for set $\mathcal{Z}^{\varphi}_{\delta,r}$ by defining sets $\mathcal{Z}_i, i \in \mathbb{N}$, such that $\bigcup_i \mathcal{Z}_i \supseteq \mathcal{Z}^{\varphi}_{\delta,r}$. Then we will estimate $\int_{\bigcup_i \mathcal{Z}_i} \varphi(x) \, \mathrm{d}x$ in order to derive an upper bound for $\mathcal{Z}$.

We observe that for $x + \delta \leq 0$ we have $\varphi(x) \leq \varphi(x+\delta)$ because of the monotonicity of φ. Hence, it is $\mathcal{Z}^{\varphi}_{\delta,r} \subseteq [-\delta, \infty)$. Now we partition $[-\delta, \infty)$ into intervals of the form $[(\ell-1)\cdot\delta, \ell\cdot\delta]$ for $\ell \in \mathbb{N}_0$. We define $\mathcal{Z}_i$ to be the i-th such interval that has a non-empty intersection with $\mathcal{Z}^{\varphi}_{\delta,r}$. If less than i intervals have a non-empty intersection then $\mathcal{Z}_i$ is the empty set. By this definition we obtained the wanted covering, and it is $\bigcup_i \mathcal{Z}_i \supseteq \mathcal{Z}^{\varphi}_{\delta,r}$ as desired.

We can now bound $\int_{\bigcup_i \mathcal{Z}_i} \varphi(x) \, \mathrm{d}x$ as follows. First suppose that all $\mathcal{Z}_i \subset \mathbb{R}_{\geq 0}$. Let z_i denote the start of interval $\mathcal{Z}_i$, i.e. $\mathcal{Z}_i = [z_i, z_i + \delta]$. Then we obtain that

$$\int_{\mathcal{Z}_i} \varphi(x) \, \mathrm{d}x \ \leq \ \delta \cdot \varphi(z_i) \ ,$$

because $\mathcal{Z}_i$ is of length δ and $\varphi(x)$ takes its maximum value within interval $\mathcal{Z}_i$ at $\varphi(z_i)$. If $\mathcal{Z}_1 = [-\delta, 0]$ it follows that $\int_{\mathcal{Z}_1} \varphi(x) \, \mathrm{d}x \leq \delta \cdot \varphi(0)$ for similar reasons.

Furthermore, it holds that $\varphi(z_{i+2}) \leq 1/2 \cdot \varphi(z_i)$ for all $i \in \mathbb{N}$. To see this consider another point in interval $\mathcal{Z}_i$ that belongs also to $\mathcal{Z}^{\varphi}_{\delta,r}$, e.g. $\hat{z}_i \in \mathcal{Z}_i \cap \mathcal{Z}^{\varphi}_{\delta,r}$. It is now

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

where we utilize that $\hat{z}_i \in \mathcal{Z}^{\varphi}_{\delta,r} = \{x \in \mathbb{R} \mid \varphi(x)/\varphi(x+\delta) > r\}$, that $r = 2$ and that $\hat{z}_i + \delta \leq z_{i+2}$.

Now we can combine everything and we obtain

$$\begin{aligned}\int_{\bigcup_i \mathcal{Z}_i} \varphi(x) \, \mathrm{d}x &\leq \int_{\mathcal{Z}_1} \varphi(x) \, \mathrm{d}x + \sum_{i \in \mathbb{N}} \int_{\mathcal{Z}_{2i-1}} \varphi(x) \, \mathrm{d}x + \sum_{i \in \mathbb{N}} \int_{\mathcal{Z}_{2i}} \varphi(x) \, \mathrm{d}x \\ &\leq \delta \cdot \varphi(0) + \sum_{i \in \mathbb{N}} \delta \cdot \varphi(z_{2i-1}) + \sum_{i \in \mathbb{N}} \delta \cdot \varphi(z_{2i}) \\ &\leq \delta \cdot \varphi(0) + \sum_{i \in \mathbb{N}} \frac{1}{2^{i-1}} \cdot \delta \cdot \varphi(z_1) + \sum_{i \in \mathbb{N}} \frac{1}{2^{i-1}} \cdot \delta \cdot \varphi(z_2) \\ &\leq \delta \cdot \varphi(0) + 2\delta \cdot \varphi(z_1) + 2\delta \cdot \varphi(z_2) \quad \leq \quad 5\delta \cdot \varphi(0) \ .\end{aligned}$$

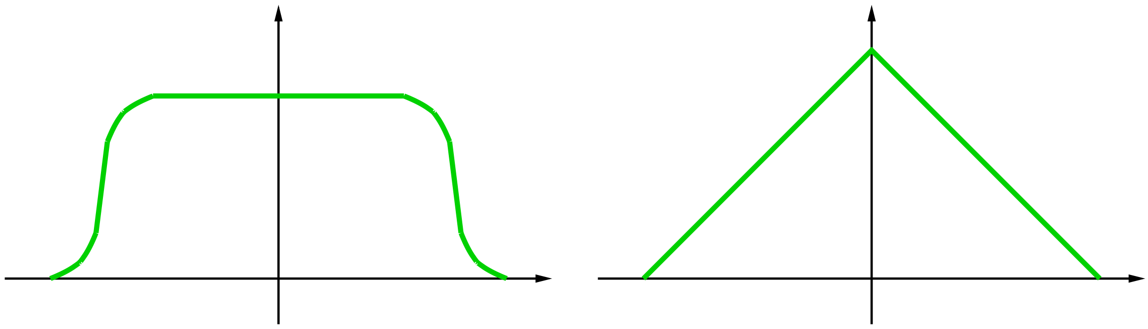

Figure 3.2: The curves of two density functions are depicted for probability distributions causing a high smoothed complexity.

It follows that $\mathcal{Z} \leq 5\delta \cdot \varphi(0)$ and Lemma 2 yields that the smoothed number of left-to-right maxima is at most $2 \cdot \lceil 1/\delta \rceil \cdot \mathcal{H}(n) + n \cdot 5\delta \cdot \varphi(0)$. Now, choosing $\delta := \sqrt{\mathcal{H}(n)/(n \cdot \varphi(0))}$ gives that for every input sequence $\mathcal{S}$ the expected number of left-to-right maxima in the perturbed sequence $\widetilde{\mathcal{S}}$ under noise from a unimodal distribution is at most

$$\mathbf{E}\left[\mathcal{L}(\widetilde{\mathcal{S}})\right] \leq \max\left\{7 \cdot \sqrt{n \cdot \mathcal{H}(n) \cdot \varphi(0)}, \mathcal{H}(n)\right\} .$$

□

It remains to mention that the Gaussian normal and uniform noise distributions are of course also unimodal distributions. For uniform noise, the Theorem 3 follows up to a constant factor also from this result since for the uniform distribution in an interval $[-\epsilon, \epsilon]$ the maximal density is clearly $1/(2\epsilon)$. If we consider here Gaussian normal noise, we obtain a much worse result than the one shown in Theorem 2 since the maximal density is $1/(\sqrt{2\pi}\,\sigma)$.

3.3 Lower Bounds for the Smoothed Case

To complete this chapter about the left-to-right maxima we will consider the tightness of the just seen upper smoothed case bounds. Of course, a general lower bound arises from the average case bound of $\Theta(\log(n))$. For the case of Gaussian normal noise this leaves only a gap of roughly $\sqrt{\log(n)}$ to the upper bound of $\mathcal{O}\left(1/\sigma \cdot \log(n)^{3/2} + \log(n)\right)$.

For the case of uniform noise this general lower bound reveals a much larger gap to the upper bound of $\mathcal{O}(\sqrt{n \cdot \log(n)/\epsilon} + \log(n))$. In the following we construct an explicit input sequence for which a large smoothed number of left-to-right maxima is achieved for the uniform noise distribution. This closes the gap to the upper bound up to a factor of roughly $\sqrt{\log(n)}$.

In fact, the following theorem holds not only for the uniform distribution but for all continuous probability distributions whose density functions have bounded support. Without

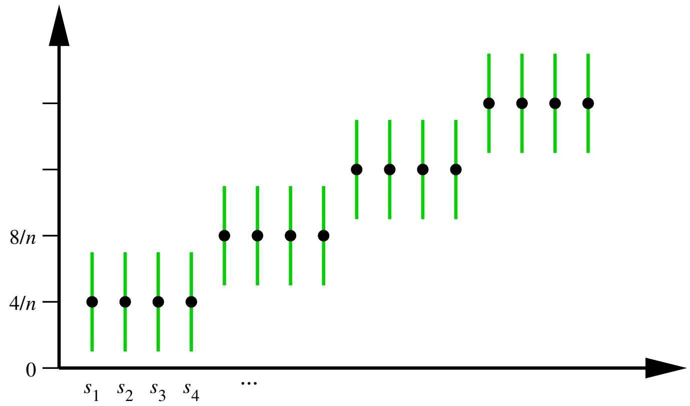

Figure 3.3: Input sequence for the lower bound construction for $\ell = 4$.

loss of generality we assume that the density functions are non-zero only in the interval $[-\epsilon, \epsilon]$ and zero everywhere else, see also Figure 3.2 which shows two examples of such density functions. Of course, the uniform distribution belongs to this class of distributions, too.

Theorem 5 *For all continuous probability distributions with density function φ of bounded support, i.e. φ is non-zero only in the interval $[-\epsilon, \epsilon]$, the smoothed number of left-to-right maxima over all input sequences $\mathcal{S}$ of n elements from $(0, 1]$ is*

$$\max_{\mathcal{S}} \mathbf{E}\left[\mathcal{L}(\widetilde{\mathcal{S}})\right] \geq \min\left\{\sqrt{n/\epsilon} \cdot \left(1 - \Phi\left(\epsilon - \sqrt{\epsilon/n}\right)^{\sqrt{n \cdot \epsilon}}\right), n\right\} ,$$

where Φ denotes the distribution function of φ. If the probability distribution is unimodal and $\varphi(\epsilon) \neq 0$ it is

$$\max_{\mathcal{S}} \mathbf{E}\left[\mathcal{L}(\widetilde{\mathcal{S}})\right] \geq \min\left\{(1 - 1/e) \cdot \left\lceil \sqrt{n \cdot \varphi(\epsilon)} \right\rceil, n\right\} .$$

Proof. Consider the following input sequence $\mathcal{S} = (s_1, \ldots, s_n)$ of n elements. For some $\ell \in \mathbb{N}, \ell \leq n$, subdivide $\mathcal{S}$ into $m := \lceil n/\ell \rceil$ subsequences $\mathcal{S}_1, \ldots, \mathcal{S}_m$ of length ℓ, the last subsequence possibly shorter. Let all elements in each particular subsequence have equal values such that the elements in a subsequence $\mathcal{S}_i$ have value $i \cdot \ell/n$, i.e. $s_{(i-1)\cdot\ell+1} = \cdots = s_{i\cdot\ell} = i \cdot \ell/n$ for all $1 \leq i \leq m$, see also Figure 3.3.

Let now $\rho_1, \ldots, \rho_n$ be n independent and identically distributed random variables from a continuous probability distribution whose density function φ has bounded support only in the interval $[-\epsilon, \epsilon]$. Again let $\widetilde{s}_k = s_k + \rho_k$ denote the perturbed element to input element s_k, for all $1 \leq k \leq n$.

If element s_k is from subsequence $\mathcal{S}_i$, i.e. $s_k = i \cdot \ell/n$, we observe that $\widetilde{s}_k$ is distributed in the interval $[i \cdot \ell/n - \epsilon, i \cdot \ell/n + \epsilon]$ after the perturbation. Consider now the case that $\widetilde{s}_k$

is large, i.e. $\widetilde{s}_k > (i-1) \cdot \ell/n + \epsilon$. It follows that $\widetilde{s}_k$ is also larger than all elements in the preceding subsequences $\widetilde{\mathcal{S}}_1, \ldots, \widetilde{\mathcal{S}}_{i-1}$. So if there is at least one element in subsequence $\widetilde{\mathcal{S}}_i$ that has a value larger than $(i-1) \cdot \ell/n + \epsilon$, then there is also at least an element in $\widetilde{\mathcal{S}}_i$ that is a left-to-right maximum.

For the input element s_k in subsequence $\mathcal{S}_i$, it is now

$$\begin{aligned} \mathbf{Pr}[\, \widetilde{s}_k \ \leq \ (i-1) \cdot \ell/n + \epsilon \,] &= \mathbf{Pr}[\, i \cdot \ell/n + \rho_k \ \leq \ (i-1) \cdot \ell/n + \epsilon \,] \\ &= \mathbf{Pr}[\, \rho_k \ \leq \ \epsilon - \ell/n \,] \quad =: \quad \Phi(\epsilon - \ell/n) \ . \end{aligned}$$

It follows then

$$\mathbf{Pr}\left[\text{no element in } \widetilde{\mathcal{S}}_i \text{ is larger than } (i-1) \cdot \ell/n + \epsilon \right] \ \leq \ \Phi(\epsilon - \ell/n)^{\ell} \ , \tag{3.7}$$

and thus

$$\mathbf{E}\left[\mathcal{L}(\widetilde{\mathcal{S}})\right] \ \geq \ m \cdot \left(1 - \Phi(\epsilon - \ell/n)^{\ell}\right) \ .$$

Choosing $\ell = \sqrt{n \cdot \epsilon}$ we get that $m = \lceil n/\ell \rceil = \sqrt{n/\epsilon}$. The first part of the theorem follows then immediately. Note that the smoothed number of left-to-right maxima cannot exceed n.

For unimodal noise distributions with density function $\varphi(\epsilon) \neq 0$ we can also conclude that $1 - \Phi(\epsilon - \ell/n) \geq \varphi(\epsilon) \cdot \ell/n$ and thus $\Phi(\epsilon - \ell/n) \leq (1 - \varphi(\epsilon) \cdot \ell/n)$. Choosing now $\ell = \left\lceil \sqrt{n/\varphi(\epsilon)} \right\rceil$ we get that $(3.7) \leq 1/e$ and therefore we have

$$\mathbf{E}\left[\mathcal{L}(\widetilde{\mathcal{S}})\right] \ \geq \ (1 - 1/e) \cdot \left\lceil \sqrt{n \cdot \varphi(\epsilon)} \right\rceil \ .$$

□

From the Theorem, the following Corollary follows immediately.

Corollary 1 *For random noise from the uniform distribution in the interval* $[-\epsilon, \epsilon]$*, the smoothed number of left-to-right maxima over all input sequences* $\mathcal{S}$ *of* n *elements from* $(0, 1]$ *is at least*

$$\begin{aligned} \max_{\mathcal{S}} \mathbf{E}\left[\mathcal{L}(\widetilde{\mathcal{S}})\right] &\geq \min\left\{(1 - 1/e) \cdot \left\lceil \sqrt{n/(2\epsilon)} \right\rceil, n\right\} \\ &= \Omega\left(\min\left\{\sqrt{n/\epsilon}, n\right\}\right) \ . \end{aligned}$$

3.4 Conclusion

The results of this chapter are summarized in the following tabular overview. The bounds are given in $\mathcal{O}$-notation.

	Upper Bounds	Lower Bounds
Gaussian $\mathrm{N}(0,\sigma)$	$\mathcal{O}\big((1/\sigma)\cdot\log(n)^{3/2}+\log(n)\big)$	$\Omega\big(\log(n)\big)$
uniform in $[-\epsilon,\epsilon]$	$\mathcal{O}\big(\sqrt{n\cdot\log(n)/\epsilon}+\log(n)\big)$	$\Omega\big(\min\big\{\sqrt{n/\epsilon},n\big\}\big)$

An interesting result is definitively that for different noise distributions we obtain a different smoothed complexity of the left-to-right maxima problem. We see that for Gaussian normal noise of variance σ^2, the smoothed number of left-to-right maxima is polylogarithmic in the number of elements and polynomial in $1/\sigma$. Interestingly, for uniform noise in an interval of length 2ϵ we obtain that the smoothed number of left-to-right maxima is polynomial in the number of elements and $1/\epsilon$.

Since both bounds are only a factor of roughly $\sqrt{\log(n)}$ away from corresponding lower bounds this discrepancy is really significant. This result is especially suprising since in the average case we obtain an expected number of left-to-right maxima of $\Theta(\log(n))$ for *all* continuous probability distributions.

4 Extreme Points

The convex hull of a point set in d-dimensional Euclidean space is one of the fundamental combinatorial structures in computational geometry. In this chapter we are interested in the number of extreme points of the convex hull of a point set.

Definition 4 (Convex Hull – Extreme Points) *Given is a set $\mathcal{P}$ of n points $p_1, \ldots, p_n$, where $p_k \in \mathbb{R}^d$ for all $1 \leq k \leq n$. The **convex hull** of $\mathcal{P}$ is the smallest convex set containing all points in $\mathcal{P}$ and is denoted by $\operatorname{conv}(\mathcal{P})$, i.e.*

$$\operatorname{conv}(\mathcal{P}) := \left\{ \sum_{i=1}^{n} \lambda_i \cdot p_i \;\middle|\; \lambda_i \geq 0, \sum_{i=1}^{n} \lambda_i = 1 \right\} .$$

*If the points in $\mathcal{P}$ are in general position (no $d+1$ points lie on a common hyperplane) the convex hull of $\mathcal{P}$ is a d-dimensional (simplicial) polytope. The faces of $\operatorname{conv}(\mathcal{P})$ of dimension 1 are called vertices or **extreme points** and their number is denoted by $\mathcal{V}(\mathcal{P})$.*

Analogously to Chapter 3, the main contribution in this chapter is a rather general lemma by which upper bounds on the smoothed number of extreme points can be obtained for noise from continuous d-dimensional product probability distributions. Again we apply this lemma explicitly to the cases that the random noise comes from the standard Gaussian normal distribution and from the uniform distribution in a hypercube.

Related Work. The convex hull of a point set in d-dimensional Euclidean space and its properties have been studied extensively in the last decades.

To compute the convex hull of a point set $\mathcal{P}$ means to compute a description of the polytope formed by $\operatorname{conv}(\mathcal{P})$. A convex polytope can be described in many ways where Seidel [Sei97] distinguishes between purely geometric and combinatorial descriptions. By purely geometric it is meant that the output consists only of the vertices (= extreme points) and/or the facets (= $(d-1)$-dimensional faces) of the polytope (given by coordinates and/or defining inequalities, respectively). A combinatorial description contains also further information about the facial structure, for example given by a Hasse diagram of the face lattice of the polytope. To compute this can be hard since McMullen [McM70] showed that the number of faces in a polytope is at worst $\Theta(n^{\lfloor d/2 \rfloor})$.

For geometric descriptions the case is different. To compute the extreme points of a convex hull is also known as the *irredundancy problem*. A point is extreme (or irredundant) if it cannot be represented as a convex combination of the remaining points in $\mathcal{P}$. To test whether a point is irredundant one has to solve for all n input points a linear programming problem in d variables with $n-1$ constraints.

For fixed dimension d this can be done in linear time e.g. by Megiddo's [Meg84] linear programming algorithm, which leads immediately to a an algorithm for the extreme point problem that runs in $\mathcal{O}(n^2)$ time. Matoušek could reduce this bound to $\mathcal{O}(n^{2-2/(\lfloor d/2\rfloor+1)} \cdot \log(n)^{\mathcal{O}(1)})$ by using a data structure for linear programming queries [Mat93] that exploits that the linear programs are closely related. One year later, Clarkson [Cla94] developed a simple output-sensitive algorithm of run-time $\mathcal{O}(n \cdot \mathcal{V}(\mathcal{P}))$ by reducing the number of involved constraints for every linear program from $n-1$ to $\mathcal{V}(\mathcal{P})$.

Chan [Cha96b] combined this idea with Matoušek's data structure and obtained for fixed $d > 3$ an $\mathcal{O}(n \cdot \log(\mathcal{V}(\mathcal{P}))^{d+2} + (n \cdot \mathcal{V}(\mathcal{P}))^{1-1/(\lfloor d/2\rfloor+1)} \cdot \log(n)^{\mathcal{O}(1)})$ time algorithm for the extreme point problem. Since an optimal output-sensitive algorithm is of time $\mathcal{O}(n \cdot \log(\mathcal{V}(\mathcal{P})))$ this algorithm is almost optimal (up to a factor of $\log(V(\mathcal{P}))^{\mathcal{O}(1)}$) when $\mathcal{V}(\mathcal{P}) = \mathcal{O}(n^{1/\lfloor d/2\rfloor} / \log(n)^K)$ for a sufficiently large constant K.

In dimensions 2 and 3, there is no need to distinguish between combinatorial and purely geometric polytope descriptions since they cannot differ much in terms of their sizes. Output sensitive algorithms with optimal time bound $\mathcal{O}(n \cdot \log(\mathcal{F}(\mathcal{P})))$ were given in the 2-dimensional case by Kirkpatrick and Seidel [KS86] and in the 3-dimensional case by Chazelle and Matoušek [CM95], where $\mathcal{F}(\mathcal{P})$ denotes the number of all faces of $\mathrm{conv}(\mathcal{P})$. Also Chan [Cha96a] obtained an algorithm of this time bound for dimensions 2 and 3 by similar techniques as described above.

Having optimal or almost optimal output-sensitive algorithms it remains to answer the question about the quantitative behavior of the extreme points. Several researchers have treated the combinatorial structure of the convex hull of n random points. In 1963/64, Rényi and Sulanke [RS63, RS64] were the first to consider the area and perimeter (length of the boundary) and the number of extreme points of the convex hull in expectation. For the latter they showed the following results. In the plane, let $\mathcal{P}$ be a set of n independent and identically distributed points uniformly chosen from a bounded convex set with continuously differentiable boundary (e.g. a sphere or ellipse). The expected number of vertices is then $\Theta(n^{1/3})$. If the points are uniformly chosen from a convex polygon the expected number of vertices is $\Theta(\log(n))$, and if the points are chosen from the 2-dimensional Gaussian normal distribution the expected number of vertices is $\Theta(\sqrt{\log(n)})$.

This work was continued by several other authors, e.g. by Efron [Efr65], Raynaud [Ray65, Ray70], Carnal [Car70], and Affentranger and Wieacker [AW91], and extended to higher dimensions and other probability distributions. For instance, Raynaud [Ray70] considered the case that the points in $\mathcal{P}$ are chosen uniformly from the d-dimensional unit ball and he showed that $\mathbf{E}[\mathcal{V}(\mathcal{P})] = \Theta(n^{(d-1)/(d+1)})$. For the case that the points are chosen uniformly from any d-dimensional polytope (e.g. the d-dimensional hypercube) Affentranger and Wieacker showed a bound of $\mathbf{E}[\mathcal{V}(\mathcal{P})] = \Theta(\log(n)^{d-1})$. If the points are chosen from the d-dimensional Gaussian normal distribution again Raynaud [Ray70] proved $\mathbf{E}[\mathcal{V}(\mathcal{P})] = \Theta(\log(n)^{(d-1)/2})$. In all these results d is considered to be a constant.

Generally, for continuous d-dimensional product probability distributions Bentley et al. [BKST78] and Buchta et al. [BTT85] derived an upper bound of $\mathcal{O}(\log(n)^{(d-1)})$ on the expected number of extreme points (this includes of course the Gaussian normal and the uniform distribution in a hypercube). Har-Peled [Har98] gave another nice proof of this result, and interesting for us is that these proofs are based on the computation of the expected number of *maximal points* (see below). Indeed, in our analysis we also exploit the close connection between extreme and maximal points.

Maximal Points. The problem of counting the number of *extreme points* is very closely related to the problem of counting the number of *maximal points* which we will exploit also for our analysis. in order to define maximal points we need to introduce also the notion of $orthants$ of a point.

Definition 5 (Orthant of a Point) *Consider a point $p = (p^{(1)}, \ldots, p^{(d)}) \in \mathbb{R}^d$. For any subset $\mathcal{I} \subseteq [d] = \{1, \ldots, d\}$ let*

$$\mathrm{o}_{\mathcal{I}}(p) \;:=\; \prod_{i \in \mathcal{I}} \left(-\infty, p^{(i)}\right] \times \prod_{i \notin \mathcal{I}} \left[p^{(i)}, \infty\right)$$

denote the ***orthant*** *centered at point p that is introduced by the index set $\mathcal{I}$.*

Definition 6 (Maximal Points) *Given is a set $\mathcal{P}$ of n points $p_1, \ldots, p_n$ in $\mathbb{R}^d$. A point $p_i \in \mathcal{P}$, $1 \leq i \leq n$, is denoted a* ***maximal point*** *of $\mathcal{P}$, if there is an index set $\mathcal{I} \subseteq [d]$ such that $\mathrm{o}_{\mathcal{I}}(p_i)$ is empty, i.e. no other point of $\mathcal{P} - \{p_i\}$ lies in $\mathrm{o}_{\mathcal{I}}(p_i)$.*

Let $\mathcal{D}(\mathcal{P})$ denote the number of maximal points of set $\mathcal{P}$.

The reason for the close relation between extreme points and maximal points lies now in the following observation. A point $p \in \mathcal{P}$ is not *maximal*, if each of the 2^d orthants centered at p contains at least one other point. In this case, p is not *extreme* either, see also Figure 4.1. It follows immediately that the number of maximal points is an upper bound on the number of extreme points, i.e.

$$\mathcal{V}(\mathcal{P}) \;\leq\; \mathcal{D}(\mathcal{P}) \;. \tag{4.1}$$

Buchta [Buc89] showed that the expected number of maximal points for a set $\mathcal{P}$ of n independent and identically distributed points chosen from a continuous d-dimensional product probability distribution is $\Theta(\log(n)^{d-1})$. This holds of course also for the uniform distribution in a d-dimensional hypercube and the d-dimensional Gaussian normal distribution. Dwyer [Dwy90] considered the case that the points are chosen from a d-dimensional ball and proved that $\mathbf{E}\left[\mathcal{D}(\mathcal{P})\right] = \Theta(n^{(d-1)/d})$.

The problem of counting the number of maximal points as it is treated here is essentially an extension of the one-dimensional left-to-right maxima problem (see the previous Chapter 3) to arbitrary dimensions. This will become very clear by a closer look at the analysis.

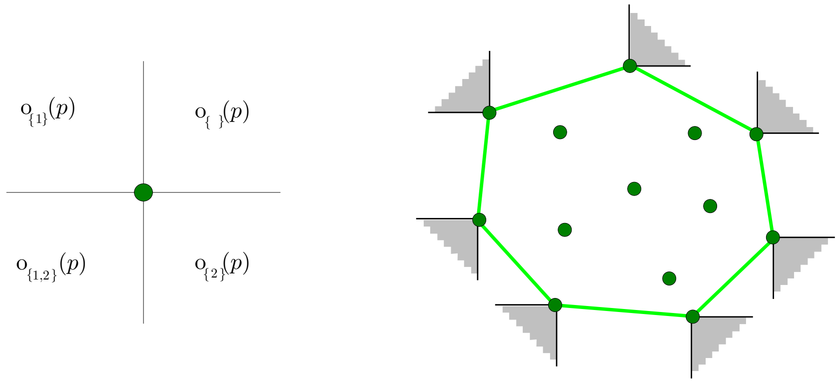

Figure 4.1: On the left side a point $p \in \mathbb{R}^2$ and its four orthants are depicted. In 2 dimensions, they are also called quadrants. On the right side we see, that extreme points are also maximal points.

Indeed, the way how we analyze the maximal points is almost analogous to the way how the left-to-right maxima were treated. Hence the chapters are almost equally structured.

Outline. In Section 4.1 we will consider the average case number of extreme points and present another proof that $\mathbf{E}[\mathcal{V}(\mathcal{P})] = \mathcal{O}(\log(n)^{d-1})$ for points chosen from continuous d-dimensional product probability distributions. Our proof makes extensive use of integrals, which seems to be very helpful for a better readability and understanding of the smoothed case analysis.

The smoothed case analysis is then presented in Section 4.2. The main contribution is a very general lemma which provides upper bounds for the smoothed number of extreme points when the random noise comes from a continuous d-dimensional product probability distribution. This lemma is a d-dimensional version of the main Lemma 2 in Section 3.2. The lemma is then explicitly applied for the cases that the random noise comes from the Gaussian normal distribution of variance[3] σ^2 and the uniform distribution in a hypercube of side-length 2ϵ. For these noise distributions we get upper bounds on the smoothed number of extreme points of $\mathcal{O}\left((1/\sigma)^d \cdot \log(n)^{(3/2)\cdot d-1}\right)$ and $\mathcal{O}\big((n \cdot \log(n)/\epsilon)^{d/(d+1)}\big)$, respectively.

In Section 4.3 the upper bounds are complemented by lower bounds. An explicit construction is presented for which the smoothed number of extreme points under uniform noise is large.

The chapter ends again with a brief summary and concluding remarks.

[3]Here the variance of the 1-dimensional Gaussian normal distribution of the components is meant.

4.1 Average Case Analysis

In this section we will consider the case that $\mathcal{P}$ is a set of n independent and identically distributed random points chosen from a continuous d-dimensional product probability distribution over $\mathbb{R}^d$. Let $X = (X_1, \ldots, X_d)$ be a random vector from such a distribution and let $\varphi : \mathbb{R} \to \mathbb{R}_{\geq 0}$ be the 1-dimensional density function of the components of X. The corresponding probability distribution function is then

$$\mathbf{Pr}[\, X_1 \leq x_1, \ldots, X_d \leq x_d \,] \;=\; \int_{-\infty}^{x_d} \cdots \int_{-\infty}^{x_1} \varphi(z_1) \cdots \varphi(z_d)\, \mathrm{d}z_1 \cdots \, \mathrm{d}z_d \;\; .$$

Note that all components of X are mutually independent and identically distributed. Examples for such continuous d-dimensional product probability distributions are the uniform distribution in a hypercube or the d-dimensional Gaussian normal distribution.

In this section we will show the following theorem.

Theorem 6 *The expected number of maximal points in a set $\mathcal{P}$ of n independent and identically distributed random points in $\mathbb{R}^d$ chosen from a continuous d-dimensional product probability distribution is*

$$\mathbf{E}\,[\mathcal{D}(\mathcal{P})] \;\leq\; 2^d \cdot \sum_{i_1=1}^{n} \frac{1}{i_1} \sum_{i_2=1}^{i_1} \frac{1}{i_2} \cdots \sum_{i_{d-1}=1}^{i_{d-2}} \frac{1}{i_{d-1}} \;=\; \Theta\left(\log(n)^{d-1}\right) \quad .$$

By (4.1) we can immediately conclude that also the following theorem holds.

Theorem 7 *The expected number of extreme points in a set $\mathcal{P}$ of n independent and identically distributed random points in $\mathbb{R}^d$ chosen from a continuous d-dimensional product probability distribution is*

$$\mathbf{E}\,[\mathcal{V}(\mathcal{P})] \;\leq\; \mathcal{O}\left(\log(n)^{d-1}\right) \quad .$$

We will now continue with the proof of Theorem 6.

Proof. First of all we recall that for an arbitrary point $p_k \in \mathcal{P}$ to be maximal it suffices to have at least one empty orthant. Without loss of generality let us now fix the orthant $\mathrm{o}_{[d]}(p_k)$ for further considerations. Since a point has 2^d orthants it follows by standard union bound that

$$\mathbf{Pr}[\, p_k \text{ is maximal} \,] \;=\; 2^d \cdot \mathbf{Pr}\big[\, \mathrm{o}_{[d]}(p_k) \text{ is empty} \,\big] \quad .$$

We will now show that

$$\mathbf{Pr}\big[\, \mathrm{o}_{[d]}(p_k) \text{ is empty} \,\big] \;\leq\; \mathcal{H}(n)^{d-1}/n \;\; . \tag{4.2}$$

By linearity of expectation the theorem follows then immediately.

For reasons of better understanding let us first establish the **2-dimensional case**. Consider a point $p_k \in \mathcal{P} \subset \mathrm{I\!R}^2$ having the coordinates $p_k = (x, y)$. It follows that p_k has four orthants which are also called quadrants. The probability for any other point $p_j \in \mathcal{P} - \{p_k\}$ to be not in the quadrant $\mathrm{o}_{\{1,2\}}(p_k) = (-\infty, x] \times (-\infty, y]$ is then equal to the probability that the point p_j lies in one of the three other quadrants of p_k. These three other quadrants are $\mathrm{o}_{\{\,\}}(p_k) := [x, \infty) \times [y, \infty)$, $\mathrm{o}_{\{1\}}(p_k) := (-\infty, x] \times [y, \infty)$, and $\mathrm{o}_{\{2\}}(p_k) := [x, \infty) \times (-\infty, y]$.

It follows

$$\begin{aligned}
\mathbf{Pr}\big[\, p_j \in \mathrm{o}_{\{\,\}}(p_k) \,\big] &= \mathbf{Pr}\Big[\, p_j^{(1)} \geq x \,\Big] \cdot \mathbf{Pr}\Big[\, p_j^{(2)} \geq y \,\Big] = (1 - \Phi(x)) \cdot (1 - \Phi(y)) \\
\mathbf{Pr}\big[\, p_j \in \mathrm{o}_{\{1\}}(p_k) \,\big] &= \mathbf{Pr}\Big[\, p_j^{(1)} \leq x \,\Big] \cdot \mathbf{Pr}\Big[\, p_j^{(2)} \geq y \,\Big] = \Phi(x) \cdot (1 - \Phi(y)) \\
\mathbf{Pr}\big[\, p_j \in \mathrm{o}_{\{2\}}(p_k) \,\big] &= \mathbf{Pr}\Big[\, p_j^{(1)} \geq x \,\Big] \cdot \mathbf{Pr}\Big[\, p_j^{(2)} \leq y \,\Big] = (1 - \Phi(x)) \cdot \Phi(y) \;.
\end{aligned}$$

The probability for any point $p_j \in \mathcal{P} - \{p_k\}$ to be not in $\mathrm{o}_{\{1,2\}}(p_k)$ is then given by

$$\begin{aligned}
&\mathbf{Pr}\big[\, p_j \notin \mathrm{o}_{\{1,2\}}(p_k) \,\big] \\
&\quad = (1 - \Phi(x)) \cdot \Phi(y) + (1 - \Phi(x)) \cdot (1 - \Phi(y)) + \Phi(x) \cdot (1 - \Phi(y)) \\
&\quad = 1 - \Phi(x) \cdot \Phi(y) \;.
\end{aligned}$$

Since there are $n - 1$ other points in $\mathcal{P} - \{p_k\}$ the probability that $\mathrm{o}_{\{1,2\}}(p_k)$ is empty is

$$\mathbf{Pr}\big[\, \mathrm{o}_{\{1,2\}}(p_k) \text{ is empty} \,\big] = \int_{\mathrm{I\!R}^2} \varphi(x) \cdot \varphi(y) \cdot (\underbrace{1 - \Phi(x) \cdot \Phi(y)}_{=:\, z \,=\, z(x)})^{n-1} \,\mathrm{d}x \,\mathrm{d}y \;. \tag{4.3}$$

This integral can be solved by two linear substitutions. In a first step we will substitute $1 - \Phi(x) \cdot \Phi(y) =: z = z(x)$, as indicated. Then it is $\mathrm{d}z = -\varphi(x) \cdot \Phi(y) \cdot \mathrm{d}x$ and this yields the integral

$$(4.3) = \int_{\mathrm{I\!R}} \int_{1-\Phi(y)}^{1} \frac{\varphi(y)}{\Phi(y)} \cdot z^{n-1} \,\mathrm{d}z \,\mathrm{d}y = \frac{1}{n} \int_{\mathrm{I\!R}} \frac{\varphi(y)}{\Phi(y)} \cdot (1 - (\underbrace{1 - \Phi(y)}_{=:\, z})^{n}) \,\mathrm{d}y \;. \tag{4.4}$$

By the indicated second substitution $1 - \Phi(y) =: z = z(y)$ where $\mathrm{d}z = -\varphi(y) \cdot \mathrm{d}y$, we get

$$\begin{aligned}
(4.4) &= \frac{1}{n} \int_0^1 \frac{1 - z^n}{1 - z} \,\mathrm{d}z = \frac{1}{n} \int_0^1 \sum_{i=0}^{n-1} z^i \,\mathrm{d}z = \frac{1}{n} \sum_{i=0}^{n-1} \int_0^1 z^i \,\mathrm{d}z \\
&= \frac{1}{n} \sum_{i=1}^{n} \frac{1}{i} = \frac{\mathcal{H}(n)}{n} \;,
\end{aligned}$$

and thus (4.2) follows immediately for $d = 2$. The theorem is therefore shown for the planar case.

We consider now the ***d*-dimensional case**. All the ideas and concepts needed are already introduced in the 2-dimensional case. The presentation is thus rather short.

Again, we consider a particular point $p_k = (x^{(1)}, \ldots, x^{(d)})$ and fix one of its orthants, again $\mathrm{o}_{[d]}(p_k) := \prod_{i=1}^{d}(-\infty, x^{(i)}]$. The probability for any other point $p_j \in \mathcal{P} - \{p_k\}$ to be not in $\mathrm{o}_{[d]}(p_k)$ is given by

$$\mathbf{Pr}\big[\, p_j \notin \mathrm{o}_{[d]}(p_k) \,\big] \;=\; \sum_{\mathcal{I}\subset[d]} \mathbf{Pr}[\, p_j \in \mathrm{o}_{\mathcal{I}}(p_k) \,] \;=\; \sum_{\mathcal{I}\subset[d]} \prod_{i\in\mathcal{I}} \Phi(x^{(i)}) \prod_{i\notin\mathcal{I}} (1-\Phi(x^{(i)})) \;\; .$$

This expression can be simplified by using the following claim.

Claim 2

$$\sum_{\mathcal{I}\subset[d]} \prod_{i\in\mathcal{I}} \Phi(x^{(i)}) \prod_{i\notin\mathcal{I}} (1-\Phi(x^{(i)})) \;=\; 1-\Phi(x^{(1)})\cdots\Phi(x^{(d)})$$

The proof of claim 2 is done by an induction on d and is deferred to the end of this section.

It follows that

$$\begin{aligned} &\mathbf{Pr}\big[\, \mathrm{o}_{[d]}(p_k) \text{ is empty} \,\big] = \\ &\quad \int_{\mathbb{R}^d} \varphi(x^{(1)})\cdots\varphi(x^{(d)}) \cdot (\underbrace{1-\Phi(x^{(1)})\cdots\Phi(x^{(d)})}_{=:\, z \,=\, z(x^{(1)})})^{n-1} \,\mathrm{d}x^{(1)}\cdots\,\mathrm{d}x^{(d)} \end{aligned} \qquad (4.5)$$

which is the d-dimensional analogue to (4.3).

Analogously to the 2-dimensional case, this integral can be solved by d repeated linear substitutions. We will present here the first three substitutions for a better illustration of the concept although the depiction is a little bit difficult.

The first substitution is already indicated, i.e. $1-\Phi(x^{(1)})\cdots\Phi(x^{(d)}) =: z = z(x^{(1)})$.

We get that $\mathrm{d}z = -\varphi(x^{(1)})\Phi(x^{(2)})\cdots\Phi(x^{(d)})\,\mathrm{d}x^{(1)}$ and it follows

$$
\begin{aligned}
(4.5) \;=\; & \int_{\mathbb{R}^{d-1}} \int_{1-\Phi(x^{(2)})\cdots\Phi(x^{(d)})}^{1} \frac{\varphi(x^{(2)})\cdots\varphi(x^{(d)})}{\Phi(x^{(2)})\cdots\Phi(x^{(d)})} \cdot z^{n-1}\,\mathrm{d}z\,\mathrm{d}x^{(2)}\cdots\,\mathrm{d}x^{(d)} \\
= \;& \frac{1}{n}\int_{\mathbb{R}^{d-1}} \frac{\varphi(x^{(2)})\cdots\varphi(x^{(d)})}{\Phi(x^{(2)})\cdots\Phi(x^{(d)})} \cdot (1-(\underbrace{1-\Phi(x^{(2)})\cdots\Phi(x^{(d)})}_{=:\, z = z(x^{(2)})})^{n})\,\mathrm{d}x^{(2)}\cdots\,\mathrm{d}x^{(d)} \\
= \;& \frac{1}{n}\int_{\mathbb{R}^{d-2}} \int_{1-\Phi(x^{(3)})\cdots\Phi(x^{(d)})}^{1} \frac{\varphi(x^{(3)})\cdots\varphi(x^{(d)})}{\Phi(x^{(3)})\cdots\Phi(x^{(d)})} \cdot \frac{1-z^{n}}{1-z}\,\mathrm{d}z\,\mathrm{d}x^{(3)}\cdots\,\mathrm{d}x^{(d)} \\
= \;& \frac{1}{n}\sum_{i_1=1}^{n}\frac{1}{i_1}\int_{\mathbb{R}^{d-2}} \frac{\varphi(x^{(3)})\cdots\varphi(x^{(d)})}{\Phi(x^{(3)})\cdots\Phi(x^{(d)})} \cdot (1-(\underbrace{1-\Phi(x^{(3)})\cdots\Phi(x^{(d)})}_{=:\, z = z(x^{(3)})})^{i_1})\,\mathrm{d}x^{(3)}\cdots\,\mathrm{d}x^{(d)} \\
= \;& \frac{1}{n}\sum_{i_1=1}^{n}\frac{1}{i_1}\int_{\mathbb{R}^{d-3}} \int_{1-\Phi(x^{(4)})\cdots\Phi(x^{(d)})}^{1} \frac{\varphi(x^{(4)})\cdots\varphi(x^{(d)})}{\Phi(x^{(4)})\cdots\Phi(x^{(d)})} \cdot \frac{1-z^{i_1}}{1-z}\,\mathrm{d}z\,\mathrm{d}x^{(4)}\cdots\,\mathrm{d}x^{(d)} \\
= \;& \frac{1}{n}\sum_{i_1=1}^{n}\frac{1}{i_1}\sum_{i_2=1}^{i_1}\frac{1}{i_2}\int_{\mathbb{R}^{d-3}} \frac{\varphi(x^{(4)})\cdots\varphi(x^{(d)})}{\Phi(x^{(4)})\cdots\Phi(x^{(d)})} \cdot (1-(\underbrace{1-\Phi(x^{(4)})\cdots\Phi(x^{(d)})}_{=:\, z = z(x^{(4)})})^{i_2})\,\mathrm{d}x^{(4)}\cdots\,\mathrm{d}x^{(d)} \\
= \;& \cdots \;=\; \frac{1}{n}\sum_{i_1=1}^{n}\frac{1}{i_1}\sum_{i_2=1}^{i_1}\frac{1}{i_2}\cdots\sum_{i_{d-1}=1}^{i_{d-2}}\frac{1}{i_{d-1}} \;\leq\; \frac{\mathcal{H}(n)^{d-1}}{n} \;.
\end{aligned}
$$

Therefore, (4.2) is shown and the theorem follows immediately. □

It remains to show Claim 2 which will be done now.

Claim 2

$$\sum_{\mathcal{I}\subset[d]}\prod_{i\in\mathcal{I}}\Phi(x^{(i)})\prod_{i\notin\mathcal{I}}(1-\Phi(x^{(i)})) \;=\; 1-\prod_{i=1}^{d}\Phi(x^{(i)})$$

Proof of Claim 2. By induction on d.

For $d = 2$, it follows immediately that the claimed equality holds. The induction step $d-1 \to d$ follows also easily. For ease of notation, we use the following abbreviation. Let

$$f(d) \;:=\; \sum_{\mathcal{I}\subset[d]}\prod_{i\in\mathcal{I}}\Phi(x^{(i)})\prod_{i\notin\mathcal{I}}(1-\Phi(x^{(i)})) \;.$$

It is then

$$\begin{aligned} f(d) &= f(d-1)\cdot\Phi(x^{(d)}) + f(d-1)\cdot(1-\Phi(x^{(d)})) + \prod_{i=1}^{d-1}\Phi(x^{(i)})\cdot(1-\Phi(x^{(d)})) \\ &= (1-\prod_{i=1}^{d-1}\Phi(x^{(i)}))\cdot(\Phi(x^{(d)})+(1-\Phi(x^{(d)}))) + \prod_{i=1}^{d-1}\Phi(x^{(i)}) - \prod_{i=1}^{d}\Phi(x^{(i)}) \\ &= 1-\prod_{i=1}^{d}\Phi(x^{(i)}) \ , \end{aligned}$$

which concludes the proof. $\square_{\text{Claim 2}}$

4.2 Upper Bounds for the Smoothed Case

In this section we consider the smoothed number of extreme points. A general lemma for random noise from continuous d-dimensional product probability distributions is derived and explicitly applied to noise from the Gaussian normal and the uniform distribution in a hypercube.

We start with a formal description of the perturbation. Consider a set $\mathcal{P}$ of *input points* $p_1, \ldots, p_d$ where all input points come from the unit hypercube $[0,1]^d$ for reasons of normalization. Let $r_1, \ldots, r_d$ be independent and identically distributed random d-vectors from a fixed continuous d-dimensional product probability distribution. We denote the random vectors as *noise vectors* and their distribution as the *noise distribution*. The set of *perturbed points* $\widetilde{\mathcal{P}} = \{\widetilde{p}_1, \ldots, \widetilde{p}_d\}$ is then given by

$$\widetilde{p}_k = p_k + r_k, \qquad \text{for } 1 \le k \le d \ .$$

We observe that each perturbed point is itself a random vector from a continuous d-dimensional probability distribution of the following kind. Consider the input point p_k and its corresponding perturbed point $\widetilde{p}_k = p_k + r_k$ where r_k is a random vector from a fixed noise distribution. The noise distribution is the d-fold product of a 1-dimensional probability distribution and all the components of r_k are from this same 1-dimensional distribution. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be the 1-dimensional integrable density function of the 1-dimensional distribution of the components of the random noise and let $\Phi : \mathbb{R} \to [0,1]$ be the corresponding distribution function.

The components of $\widetilde{p}_k$ are not anymore from an identical but from slightly different probability distributions. The distributions of the components of $\widetilde{p}_k$ depend on the components of the input point p_k, e.g., the 1-dimensional density function of the i-th component of $\widetilde{p}_k$ is $\varphi(x-p_k^{(i)})$ and the corresponding distribution function is $\Phi(x-p_k^{(i)})$, for all $1 \le i \le d$. The density and distribution functions of the components of $\widetilde{p}_k$ are thus slightly shifted copies of the density and distribution function of the 1-dimensional noise distribution, namely of φ and Φ, respectively.

In this sense, also $\widetilde{\mathcal{P}}$ is a set of independent but *not* identically distributed random points and the distribution of a particular point $\widetilde{p}_k$ is the d-fold product of similar but slightly shifted 1-dimensional distributions.

Now we will start with the actual analysis of the smoothed number of extreme points. Indeed, the approach is the same as for the average case analysis in the previous Section 4.1. Instead of extreme points we will again consider maximal points to upper bound the number of extreme points. As before we will bound the probability that for a fixed input point p_k the perturbed point $\widetilde{p}_k$ is maximal by considering a fixed orthant of $\widetilde{p}_k$ and computing the probability that this orthant is empty.

Let the perturbed point have the coordinates $\widetilde{p}_k := (x^{(1)}, \ldots, x^{(d)})$. Without loss of generality we fix again orthant $\mathrm{o}_{[d]}(\widetilde{p}_k) = \prod_{i=1}^{d}(-\infty, x^{(i)}]$. We will now derive an integral expression for the probability that $\mathrm{o}_{[d]}(\widetilde{p}_k)$ is empty. For any other input point $p_j \in \mathcal{P} - \{p_k\}$ it holds that

$$\begin{aligned}
\mathbf{Pr}\left[\,\widetilde{p}_j \notin \mathrm{o}_{[d]}(\widetilde{p}_k)\,\right] &= \sum_{\mathcal{I}\subset[d]} \mathbf{Pr}[\,\widetilde{p}_j \in \mathrm{o}_{\mathcal{I}}(\widetilde{p}_k)\,] \\
&= \sum_{\mathcal{I}\subset[d]} \prod_{i\in\mathcal{I}} \Phi(x^{(i)} - p_j^{(i)}) \prod_{i\notin\mathcal{I}} (1 - \Phi(x^{(i)} - p_j^{(i)})) \\
&= 1 - \Phi(x^{(1)} - p_j^{(1)}) \cdots \Phi(x^{(d)} - p_j^{(d)})
\end{aligned}$$

where the last step follows by Claim 2. This yields

$$\begin{aligned}
\mathbf{Pr}\left[\,\mathrm{o}_{[d]}(\widetilde{p}_k) \text{ is empty}\,\right] = \int_{\mathbb{R}^d} & \varphi(x^{(1)} - p_k^{(1)}) \cdots \varphi(x^{(d)} - p_k^{(d)}) \cdot \\
& \prod_{j\neq k} \Big(1 - \Phi(x^{(1)} - p_j^{(1)}) \cdots \Phi(x^{(d)} - p_j^{(d)})\Big)\ \mathrm{d}x^{(1)} \cdots\ \mathrm{d}x^{(d)} \ ,
\end{aligned} \tag{4.6}$$

which is the “smoothed” analogue to (4.5).

The approach to solve this integral is very similar to the approach in Section 3.2 where we considered the smoothed number of left-to-right maxima. The main idea is to subdivide the unit hyper-cube into $m := \lceil 1/\delta \rceil^d$ smaller axis-aligned hypercubes of side length δ. Then the input set $\mathcal{P}$ is divided into sets $\mathcal{P}_1, \ldots, \mathcal{P}_m$ where $\mathcal{P}_\ell$ is the subset of $\mathcal{P}$ that is located in the ℓ-th small hypercube, where some ordering among the small hypercubes is assumed. Now we can compute for all subsets $\mathcal{P}_\ell$ the expected number of maximal points and exploit that

$$\mathbf{E}\left[\mathcal{V}(\widetilde{\mathcal{P}})\right] \le \mathbf{E}\left[\mathcal{D}(\widetilde{\mathcal{P}})\right] \le \sum_{\ell=1}^{m} \mathbf{E}\left[\mathcal{D}(\widetilde{\mathcal{P}}_\ell)\right] \ . \tag{4.7}$$

The motivation for this approach is the same as in the previous chapter. Intuitively, the advantage is that for small enough δ the input points of a subset $\mathcal{P}_\ell$ lie so close together that after perturbation the points behave almost as in the random average case.

Note that differently to Section 3.2 it is not necessary to establish an analogue to lemma 1. The reason is that the number of left-right-maxima in a sequence is depending on the ordering of the sequence. This is of course not the case for the number of extreme/maximal points since this number is independent of the ordering in which the input points are considered.

Without loss of generality assume now that the input subset $\mathcal{P}_\ell$ lies in the small hypercube $[0,\delta]^d$ and that $\mathcal{P}_\ell$ is of magnitude n_ℓ. Consider $p_k \in \mathcal{P}_\ell$, and let $\bar{\delta} = (\delta, \ldots, \delta)$ and $\bar{0} = (0, \ldots, 0)$. The integral in (4.6) can be simplified by the following observation. The probability that for any other point $p_j \in \mathcal{P}_\ell - \{p_k\}$ the perturbed point $\widetilde{p_j}$ lies not in $\mathrm{o}_{[d]}(\widetilde{p_k})$ is maximized if $p_k = \bar{0}$ and $p_j = \bar{\delta}$. This yields

$$\mathbf{Pr}\left[\,\mathrm{o}_{[d]}(\widetilde{p_k}) \text{ is empty}\,\right] \;\leq\; \int_{\mathrm{I\!R}^d} \varphi(x^{(1)}) \cdots \varphi(x^{(d)}) \cdot \\ (1 - \Phi(x^{(1)} - \delta) \cdots \Phi(x^{(d)} - \delta))^{n_\ell - 1} \,\mathrm{d}x^{(1)} \cdots \,\mathrm{d}x^{(d)} \;. \tag{4.8}$$

In order to solve the integral (4.8) we will expand the product of density functions by a multiplication in the following way

$$\varphi(x^{(1)}) \cdots \varphi(x^{(d)}) \;=\; \varphi(x^{(1)} - \delta) \cdots \varphi(x^{(d)} - \delta) \cdot \frac{\varphi(x^{(1)})}{\varphi(x^{(1)} - \delta)} \cdots \frac{\varphi(x^{(d)})}{\varphi(x^{(d)} - \delta)} \;. \tag{4.9}$$

If now the ratios $\varphi(x)/\varphi(x-\delta)$ were bounded by some parameter r we could replace in the integral (4.8) the ratios by the parameter r and solve the remaining integral very easily as seen in the average case analysis in the previous section. In the following lemma this is actually done, namely to identify the regions where the ratios of densities are bounded by r and where they exceed this bound, and to solve the integral (4.8) for both regions.

The bounded ratios of density functions formalize what was meant when writing “behave almost as in the average case”. The parameters δ and r enable us to trade between the number of subsets m and the size of the regions where the ratios are not bounded by r.

Now we can state the main lemma of this section which is also a d-dimensional analogue to Lemma 2, the main lemma about the smoothed number of the left-to-right maxima.

Lemma 3 *For a fixed continuous d-dimensional product probability distribution with one dimensional density function φ and for positive parameters δ and r define the set*

$$\mathcal{Z}^{\varphi}_{\delta,r} := \{x \in \mathrm{I\!R} \mid \varphi(x) > r \cdot \varphi(x - \delta)\} \;.$$

Let $\mathcal{Z}$ be the probability of set $\mathcal{Z}^{\varphi}_{\delta,r}$, i.e., let

$$\mathcal{Z} := \sum_{i=0}^{d-1} \binom{d}{i} \underbrace{\int_{\mathcal{Z}^{\varphi}_{\delta,r}} \cdots \int_{\mathcal{Z}^{\varphi}_{\delta,r}}}_{d-i} \underbrace{\int_{\mathrm{I\!R} - \mathcal{Z}^{\varphi}_{\delta,r}} \cdots \int_{\mathrm{I\!R} - \mathcal{Z}^{\varphi}_{\delta,r}}}_{i} \varphi(x^{(1)}) \cdots \varphi(x^{(d)}) \,\mathrm{d}x^{(1)} \cdots \,\mathrm{d}x^{(d)} \;.$$

For random noise from the fixed probability distribution, the smoothed number of extreme points over all input sets $\mathcal{P}$ of n points from $[0,1]^d$ is

$$\max_{\mathcal{P}} \mathbf{E}\left[\mathcal{V}(\widetilde{\mathcal{P}})\right] \;\leq\; \max\left\{2^d \cdot \left(r^d \cdot \lceil 1/\delta\rceil^d \cdot \mathcal{H}(n)^{d-1} + n \cdot \mathcal{Z}\right), \mathcal{H}(n)^{d-1}\right\} \quad .$$

Proof. As already described earlier we plan to exploit inequality (4.7) and consider therefore the set $\mathcal{P}_\ell \subset \mathcal{P}$ of n_ℓ points lying in $[0,\delta]^d$. For a point $p_k \in \mathcal{P}_\ell$ we saw already in (4.8) an expression for the probability that $\widetilde{p}_k$ has a fixed empty orthant. We will now further transform this integral by subdividing the domain $\mathbb{R}^d$ of the integral into 2^d subdomains in the following way

$$\begin{aligned}
&\mathbf{Pr}\left[\, \mathrm{o}_{[d]}(\widetilde{p}_k) \text{ is empty} \,\right] \\
&\leq \int_{\mathbb{R}^d} \varphi(x^{(1)}) \cdots \varphi(x^{(d)}) \cdot (1 - \Phi(x^{(1)} - \delta) \cdots \Phi(x^{(d)} - \delta))^{n_\ell - 1} \,\mathrm{d}x^{(1)} \cdots \,\mathrm{d}x^{(d)} \\
&= \sum_{i=0}^{d} \binom{d}{i} \underbrace{\int_{\mathcal{Z}^{\varphi}_{\delta,r}} \cdots \int_{\mathcal{Z}^{\varphi}_{\delta,r}}}_{d-i} \underbrace{\int_{\mathbb{R}-\mathcal{Z}^{\varphi}_{\delta,r}} \cdots \int_{\mathbb{R}-\mathcal{Z}^{\varphi}_{\delta,r}}}_{i} \varphi(x^{(1)}) \cdots \varphi(x^{(d)}) \cdot \\
&\qquad\qquad (1 - \Phi(x^{(1)} - \delta) \cdots \Phi(x^{(d)} - \delta))^{n_\ell - 1} \,\mathrm{d}x^{(1)} \cdots \,\mathrm{d}x^{(d)} \\
&\leq \mathcal{Z} + \int_{\mathbb{R}-\mathcal{Z}^{\varphi}_{\delta,r}} \cdots \int_{\mathbb{R}-\mathcal{Z}^{\varphi}_{\delta,r}} \varphi(x^{(1)}) \cdots \varphi(x^{(d)}) \cdot \\
&\qquad\qquad (1 - \Phi(x^{(1)} - \delta) \cdots \Phi(x^{(d)} - \delta))^{n_\ell - 1} \,\mathrm{d}x^{(1)} \cdots \,\mathrm{d}x^{(d)} \quad .
\end{aligned}$$

The last step follows by the observation that $(1 - \Phi(x^{(1)} - \delta) \cdots \Phi(x^{(d)} - \delta)) \leq 1$. Thus the first d summands can be bounded by $\mathcal{Z}$ and it remains to treat the last summand. Indeed, for the last summand we can expand the product of density functions as described before in (4.9) and then bound the ratios $\varphi(x)/\varphi(x-\delta)$ by r because $\mathcal{Z}^{\varphi}_{\delta,r}$ was defined in this way. This gives us then for the last summand

$$\begin{aligned}
&\int_{\mathbb{R}-\mathcal{Z}^{\varphi}_{\delta,r}} \cdots \int_{\mathbb{R}-\mathcal{Z}^{\varphi}_{\delta,r}} \varphi(x^{(1)}) \cdots \varphi(x^{(d)}) \cdot \\
&\qquad\qquad (1 - \Phi(x^{(1)} - \delta) \cdots \Phi(x^{(d)} - \delta))^{n_\ell - 1} \,\mathrm{d}x^{(1)} \cdots \,\mathrm{d}x^{(d)} \\
&\leq r^d \cdot \int_{\mathbb{R}-\mathcal{Z}^{\varphi}_{\delta,r}} \cdots \int_{\mathbb{R}-\mathcal{Z}^{\varphi}_{\delta,r}} \varphi(x^{(1)} - \delta) \cdots \varphi(x^{(d)} - \delta) \cdot \\
&\qquad\qquad (1 - \Phi(x^{(1)} - \delta) \cdots \Phi(x^{(d)} - \delta))^{n_\ell - 1} \,\mathrm{d}x^{(1)} \cdots \,\mathrm{d}x^{(d)} \\
&\leq r^d \cdot \int_{\mathbb{R}^d} \varphi(x^{(1)}) \cdots \varphi(x^{(d)}) \cdot (1 - \Phi(x^{(1)}) \cdots \Phi(x^{(d)}))^{n_\ell - 1} \,\mathrm{d}x^{(1)} \cdots \,\mathrm{d}x^{(d)} \\
&= r^d \cdot \frac{1}{n_\ell} \sum_{i_1=1}^{n_\ell} \frac{1}{i_1} \sum_{i_2=1}^{i_1} \frac{1}{i_2} \cdots \sum_{i_{d-1}=1}^{i_{d-2}} \frac{1}{i_{d-1}} \quad \leq \quad r^d \cdot \frac{\mathcal{H}(n_\ell)^{d-1}}{n_\ell} \quad .
\end{aligned}$$

To see the second last step we refer to the proof of Theorem 6, the average case analysis. It remains to conclude that

$$\mathbf{E}\left[\mathcal{D}(\widetilde{\mathcal{P}}_\ell)\right] \;\leq\; 2^d \cdot (r^d \cdot \mathcal{H}(n_\ell)^{d-1} + \mathcal{Z} \cdot n_\ell) \;\; .$$

Remember that we have $m = \lceil 1/\delta \rceil^d$ subsets $\mathcal{P}_\ell$. From (4.7) it follows that

$$\mathbf{E}\left[\mathcal{D}(\widetilde{\mathcal{P}})\right] \;\leq\; \sum_{\ell=1}^{m} \mathbf{E}\left[\mathcal{D}(\widetilde{\mathcal{P}}_\ell)\right] \;\leq\; 2^d \cdot (r^d \cdot \lceil 1/\delta \rceil^d \cdot \mathcal{H}(n_\ell)^{d-1} + \mathcal{Z} \cdot n) \;\; .$$

Since this result holds for all sets $\mathcal{P}$ of input points from $[0,1]^d$, Lemma 3 is proven.

□

4.2.1 Normal Gaussian Noise

We will apply now Lemma 3 to the case that the random noise comes from the Gaussian normal distribution with expectation 0 and variance σ^2, also denoted as $\mathrm{N}(0,\sigma)$. The 1-dimensional density function of the Gaussian normal distribution is

$$\varphi(x) \;:=\; \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{x^2}{2\sigma^2}} \;\; .$$

Theorem 8 *For random noise from the standard Gaussian normal distribution* $\mathrm{N}(0,\sigma)$*, the smoothed number of extreme points over all input sets* $\mathcal{P}$ *of* n *points from* $[0,1]^d$ *is*

$$\begin{aligned} \max_{\mathcal{P}} \mathbf{E}\left[\mathcal{V}(\widetilde{\mathcal{P}})\right] \;&\leq\; \max\left\{ 2^{6d+2} \cdot d^{d/2} \cdot \left(\frac{1}{\sigma}\right)^d \cdot \ln(n)^{d/2} \cdot \mathcal{H}(n)^{d-1}, \mathcal{H}(n)^{d-1} \right\} \\ &=\; \mathcal{O}\left(\left(\frac{1}{\sigma}\right)^d \cdot \log(n)^{\frac{3}{2}\cdot d - 1} + \log(n)^{d-1} \right) \;\; . \end{aligned}$$

Proof. In order to utilize Lemma 3 we need to choose the two parameters r and δ. Let $\delta := \sigma/\beta$ where $\beta := \sqrt{\ln\left(1/\left(\sqrt[d]{1+1/n} - 1\right)\right)}$. For $x \leq -\sqrt{2}\sigma\beta$ it holds that

$$\frac{\varphi(x)}{\varphi(x-\delta)} \;=\; e^{-\frac{\delta}{\sigma^2}\cdot x + \frac{\delta^2}{2\sigma^2}} \;=\; e^{-\frac{x}{\sigma\cdot\beta} + \frac{1}{2\beta^2}} \;\leq\; e^{\sqrt{2} + \frac{1}{2\beta^2}} \;\leq\; e^3 \;\; .$$

Therefore, if we choose $r := e^3$, we can conclude that $\mathcal{Z}^{\varphi}_{\delta,r} \subset \left(-\infty, -\sqrt{2}\sigma\beta\right]$. With this approximation we can take the next crucial step, namely to bound $\mathcal{Z}$. It is

$$\begin{aligned}
\mathcal{Z} &= \sum_{i=0}^{d-1} \binom{d}{i} \underbrace{\int_{\mathcal{Z}^{\varphi}_{\delta,r}} \cdots \int_{\mathcal{Z}^{\varphi}_{\delta,r}}}_{d-i} \underbrace{\int_{\mathbb{R}-\mathcal{Z}^{\varphi}_{\delta,r}} \cdots \int_{\mathbb{R}-\mathcal{Z}^{\varphi}_{\delta,r}}}_{i} \varphi(x^{(1)}) \cdots \varphi(x^{(d)}) \, \mathrm{d}x^{(1)} \cdots \, \mathrm{d}x^{(d)} \\
&\leq \sum_{i=0}^{d-1} \binom{d}{i} \underbrace{\int_{-\infty}^{-\sqrt{2}\sigma\beta} \cdots \int_{-\infty}^{-\sqrt{2}\sigma\beta}}_{d-i} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{i} \varphi(x^{(1)}) \cdots \varphi(x^{(d)}) \, \mathrm{d}x^{(1)} \cdots \, \mathrm{d}x^{(d)} \\
&= \sum_{i=0}^{d-1} \binom{d}{i} \left(\int_{-\infty}^{-\sqrt{2}\sigma\beta} \varphi(x) \, \mathrm{d}x \right)^{d-i} .
\end{aligned} \tag{4.10}$$

From Claim 1 on page 21 it follows, that we can estimate the tail of the standard Gaussian normal probability distribution $\mathrm{N}(0,\sigma)$ in the following way. It is $\int_{-\infty}^{-k\sigma} \varphi(x) \, \mathrm{d}x \leq e^{-k^2/2}$ for any $k \geq 1/\sqrt{2\pi}$. Hence we have

$$\int_{-\infty}^{-\sqrt{2}\sigma\beta} \varphi(x) \, \mathrm{d}x \leq e^{-\beta^2} = \sqrt[d]{1 + \frac{1}{n}} - 1 . \tag{4.11}$$

Combining (4.10) and (4.11) we get

$$\mathcal{Z} \leq \sum_{i=0}^{d-1} \binom{d}{i} \left(\sqrt[d]{1 + \frac{1}{n}} - 1 \right)^{d-i} = \left(1 + \sqrt[d]{1 + \frac{1}{n}} - 1 \right)^{d} - 1 = \frac{1}{n} .$$

We can now apply Lemma 3 with $r = e^3$ and $\mathcal{Z} \leq 1/n$ and it follows

$$\mathbf{E}\left[\mathcal{V}(\widetilde{\mathcal{P}})\right] \leq \max\left\{ 2^d \cdot \left(e^{3d} \cdot \lceil 1/\delta \rceil^d \cdot \mathcal{H}(n)^{d-1} + 1 \right), \mathcal{H}(n)^{d-1} \right\} .$$

It remains to consider δ which was earlier chosen to be σ/β. We will exploit the following claim.

Claim 3

$$\sqrt[d]{1+b} \geq 1 + \frac{b}{2^d} \qquad \forall\, 0 < b < 1$$

Proof of Claim 3.

$$\left(1 + \frac{b}{2^d}\right)^d = \sum_{i=0}^{d} \binom{d}{i} \left(\frac{b}{2^d}\right)^i \leq 1 + \sum_{i=0}^{d} \binom{d}{i} \cdot \frac{b}{2^d} = 1 + b$$

$\square_{\text{Claim 3}}$

It follows that

$$\beta \;=\; \sqrt{\ln\left(1/\left(\sqrt[d]{1+1/n}-1\right)\right)} \;\leq\; \sqrt{\ln(2^d\cdot n)} \;\leq\; \sqrt{d\cdot\ln(n)}\;\;,$$

and thus we get

$$\lceil 1/\delta\rceil \;\leq\; (1/\sigma)^d\cdot\beta^d+1 \;\leq\; 2\cdot d^{d/2}\cdot(1/\sigma)^d\cdot\ln(n)^{d/2}\;\;.$$

Now we combine the results and conclude that for every input set $\mathcal{P}$ the expected number of extreme points under Gaussian normal noise is at most

$$\mathbf{E}\left[\mathcal{V}(\widetilde{\mathcal{P}})\right] \;\leq\; \max\left\{2^{6d+2}\cdot d^{d/2}\cdot(1/\sigma)^d\cdot\ln(n)^{d/2}\cdot\mathcal{H}(n)^{d-1},\mathcal{H}(n)^{d-1}\right\}\;\;,$$

which proves Theorem 8.

□

4.2.2 Uniform Noise

In this section, we will now consider random noise that is uniformly distributed in a hypercube of side length 2ϵ centered at the origin. This distribution has for its components the 1-dimensional density function

$$\varphi(x) = \begin{cases} \frac{1}{2\epsilon} & \text{if } x\in[-\epsilon,\epsilon] \\ 0 & \text{else}\;\;. \end{cases}$$

Theorem 9 *For random noise from the uniform distribution from a hyper-cube of side length 2ϵ, the smoothed number of extreme points over all input sets $\mathcal{P}$ of n points from $[0,1]^d$ is*

$$\begin{aligned}\max_{\mathcal{P}}\mathbf{E}\left[\mathcal{V}(\widetilde{\mathcal{P}})\right] \;&\leq\; \max\left\{2^{d+1}\cdot\left(\frac{d\cdot n}{2\epsilon}\right)^{\frac{d}{d+1}}\cdot\mathcal{H}(n)^{\frac{d-1}{d+1}},\mathcal{H}(n)^{d-1}\right\}\\ &=\; \mathcal{O}\left(\left(\frac{n\cdot\log(n)}{\epsilon}\right)^{\frac{d}{d+1}}+\log(n)^{d-1}\right)\;\;.\end{aligned}$$

Proof. Again we want to utilize Lemma 3. We choose $r=1$, then it follows that for $0<\delta<2\epsilon$ we have $\mathcal{Z}^{\varphi}_{\delta,r}=\{x\in\mathbb{R}\mid\varphi(x)>\varphi(x-\delta)\}=[-\epsilon,-\epsilon+\delta]$. In the next step we need to compute $\mathcal{Z}$ which is given by

$$\begin{aligned}
\mathcal{Z} &= \sum_{i=0}^{d-1} \binom{d}{i} \underbrace{\int_{-\epsilon}^{\delta-\epsilon} \cdots \int_{-\epsilon}^{\delta-\epsilon}}_{d-i} \underbrace{\int_{\delta-\epsilon}^{\epsilon} \cdots \int_{\delta-\epsilon}^{\epsilon}}_{i} \left(\frac{1}{2\epsilon}\right)^{d} \ \mathrm{d}x^{(1)} \cdots \ \mathrm{d}x^{(d)} \\
&= \sum_{i=0}^{d-1} \binom{d}{i} \cdot \left(\frac{1}{2\epsilon}\right)^{d} \cdot \delta^{d-i} \cdot (2\epsilon-\delta)^{i} \quad = \quad \left(\frac{1}{2\epsilon}\right)^{d} \left((2\epsilon)^{d} - (2\epsilon-\delta)^{d}\right) \\
&= 1 - \left(1 - \frac{\delta}{2\epsilon}\right)^{d} \quad \leq \quad 1 - \left(1 - d \cdot \frac{\delta}{2\epsilon}\right) \quad = \quad d \cdot \frac{\delta}{2\epsilon} \ .
\end{aligned}$$

We can now apply Lemma 3 and it follows that for every input set $\mathcal{P}$ the expected number of extreme points under uniform noise is at most

$$\mathbf{E}\left[\mathcal{V}(\widetilde{\mathcal{P}})\right] \ \leq \ \max\left\{2^{d} \cdot \left(\lceil 1/\delta \rceil^{d} \cdot \mathcal{H}(n)^{d-1} n + n \cdot d \cdot \frac{\delta}{2\epsilon}\right), \mathcal{H}(n)^{d-1}\right\} \ .$$

If we choose $\delta = (2\epsilon \cdot \mathcal{H}(n)^{d-1}/(d \cdot n))^{1/(d+1)}$ we obtain the theorem. □

4.3 Lower Bounds for the Smoothed Case

In this section we consider lower bounds for the smoothed number of extreme and maximal points.

For random noise from the Gaussian normal distribution we can refer to the average case bounds which directly imply lower bounds for the smoothed case. In Theorem 6 we saw that the expected number of maximal points is $\Theta(\log(n)^{d-1})$. Raynaud showed that the expected number of extreme points is $\Theta(\log(n)^{(d-1)/2})$ [Ray70]. We recall that in Theorem 8 the smoothed number of maximal and extreme points under Gaussian normal noise was shown to be at most $\mathcal{O}((1/\sigma)^{d} \cdot \log(n)^{3/2 \cdot d-1} + \log(n)^{d-1})$. For the smoothed number of maximal points, the gap between this upper bound and the average case bound is thus roughly a factor of $\log(n)^{d/2}$. For the smoothed number of extreme points, the gap is roughly a factor of $\log(n)^{d}$.

We will now show that the smoothed number of extreme points for uniform noise from a hyper-cube is significantly larger than for the case of normally distributed noise. We do this by constructing an input set of points such that the set of perturbed points has a large expected number of extreme points.

Before we start with this construction we will introduce the notion of *spherical caps* on the unit sphere. Let

$$\Omega_d \ := \ \left\{(x_1, \ldots, x_d) \in \mathbb{R}^d \mid \sum_{i=1}^{d} x_i^2 = 1\right\}$$

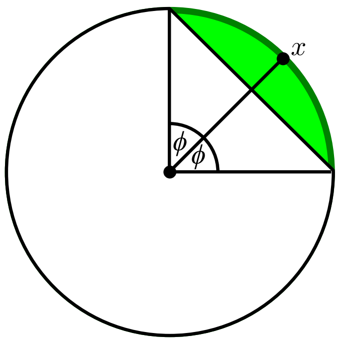

Figure 4.2: Spherical cap $\mathrm{cap}_d(x, \phi)$ and the corresponding region (shaded).

denote the d-dimensional Euclidean sphere of unit radius and let

$$\mathcal{S}_d = \frac{d \cdot \pi^{d/2}}{\Gamma(d/2+1)}$$

be the $(d-1)$-dimensional content (surface area) of Ω_d.

The *angular separation* between two points $x, y \in \Omega_d$ is the angle between the line segment joining the origin with x and the line segment joining the origin with y. The angular separation is thus $\arccos(x \cdot y)$ where $x \cdot y$ is the inner (dot) product[4] of x and y. The set of points on Ω_d whose angular separation from a fixed point $x \in \Omega_d$ is at most ϕ is called a *spherical cap centered at* x *with angular radius* ϕ and is denoted by $\mathrm{cap}_d(x, \phi)$, i.e.

$$\mathrm{cap}_d(x, \phi) := \left\{ y \in \Omega_d \;\middle|\; x \cdot y > \cos(\phi) \right\} .$$

When the center of a spherical cap is not relevant, the notation may be abbreviated as $\mathrm{cap}_d(\phi)$. Furthermore, we will consider the convex closure of $\mathrm{cap}_d(\phi)$. Let us denote

$$\begin{aligned} \overline{\mathrm{cap}}_d(x, \phi) &:= \mathrm{conv}(\mathrm{cap}_d(x, \phi)) \\ &= \left\{ \sum_{i=1}^{d} \lambda_i \cdot z_i \;\middle|\; z_1, \ldots, z_d \in \mathrm{cap}_d(\phi), \lambda_i \geq 0, \sum_{i=1}^{d} \lambda_i = 1 \right\} \end{aligned}$$

as the *region of spherical cap* $\mathrm{cap}_d(x, \phi)$, see also Figure 4.2.

An important property of spherical caps is expressed in the following observation.

[4]For $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$ the inner dot product is defined as $x \cdot y = \sum_{i=1}^{d} x_i \cdot y_i$.

Observation 1 *Consider a set $\mathcal{P}$ of points lying inside the sphere Ω_d and an arbitrary $x \in \Omega_d$. If the region $\overline{\mathrm{cap}}_d(x, \phi)$ is non-empty it contains at least one point of $\mathcal{P}$ that is an extreme point of* $\mathrm{conv}(\mathcal{P})$.

Our plan is now the following. First we try to place a large number of non-intersecting spherical caps on Ω_d. Then we select the position of the input points in set $\mathcal{P}$ such that it is guaranteed, that after the perturbation no point of $\widetilde{\mathcal{P}}$ lies outside the unit sphere Ω_d and that the number of spherical caps with non-empty region is large.

So in a first step we will investigate how many spherical caps of fixed angular radius ϕ can be placed on the unit sphere Ω_d such that they are non-intersecting. This problem is also studied in the context of *spherical codes*. A spherical code is a so-called *channel code* (or *error-correcting code*) and consists of a finite set of points on Ω_d. Spherical codes have important applications to transmission over the Gaussian channel and to many other areas [CS88]. The minimum angular separation of a spherical code $\mathcal{C} \subset \Omega_d$ is the minimum over all pairwise angular separations and is denoted by $\mathrm{sep}(\mathcal{C})$, i.e.

$$\mathrm{sep}(\mathcal{C}) \;:=\; \min\left\{\arccos(c_1 \cdot c_2) \;\middle|\; c_1, c_2 \in \mathcal{C} \subset \Omega_d\right\} \quad .$$

A desirable property of a spherical code is to have a large minimum angular separation. The maximum number of points of a spherical code in d dimensions having a minimum angular separation greater than or equal to γ is commonly denoted by $\mathcal{M}(d, \gamma)$, i.e.

$$\mathcal{M}(d, \gamma) \;:=\; \max\left\{|\mathcal{C}| \;\middle|\; \mathcal{C} \subset \Omega_d \quad \text{and} \quad \mathrm{sep}(\mathcal{C}) \geq \gamma\right\} \quad .$$

This is also interesting for us since a lower bound on $\mathcal{M}(d, 2\phi)$ provides also a lower bound on the number of non-intersecting spherical caps of angular radius ϕ that can be placed on Ω_d. The reason is that for any two points in $\mathrm{cap}_d(\phi)$, the angular separation is at most 2ϕ. In other words, we can place in each point of a spherical code of minimum angular separation 2ϕ the center of a spherical cap with angular radius ϕ such that all spherical caps are non-intersecting.

We obtain a lower bound on $\mathcal{M}(d, \gamma)$ by the following simple observation. Consider an optimal spherical code $\mathcal{C} = \{c_1, \ldots, c_m\} \subset \Omega_d$ such that $|\mathcal{C}| = m = \mathcal{M}(d, \gamma)$ and $\mathrm{sep}(\mathcal{C}) = \gamma$. Now we place at each point c_i of the code $\mathcal{C}$ the center of a spherical cap of angular radius γ and consider their union, i.e. $\cup_{i=1}^{m} \mathrm{cap}_d(c_i, \gamma)$.

If there is now a point $x \in \Omega_d$ and $x \notin \cup_{i=1}^{m} \mathrm{cap}_d(c_i, \gamma)$, it follows immediately that $\arccos(x \cdot c_i) \geq \gamma$. Thus $\mathcal{C} \cup \{x\}$ is also a spherical code of minimum angular separation γ with $m + 1$ points, a contradiction to $m = \mathcal{M}(d, \gamma)$ being the maximum. Therefore we have $\Omega_d = \cup_{i=1}^{m} \mathrm{cap}_d(c_i, \gamma)$ and $\mathcal{M}(d, \gamma) \cdot \mathcal{S}(\mathrm{cap}_d(\gamma)) \geq \mathcal{S}_d$, where $\mathcal{S}(\mathrm{cap}_d(\gamma))$ denotes the $(d-1)$-dimensional content (surface area) of $\mathrm{cap}_d(\gamma)$.

We conclude that

$$\mathcal{M}(d, \gamma) \;\geq\; \frac{\mathcal{S}_d}{\mathcal{S}(\mathrm{cap}_d(\gamma))} \quad .$$

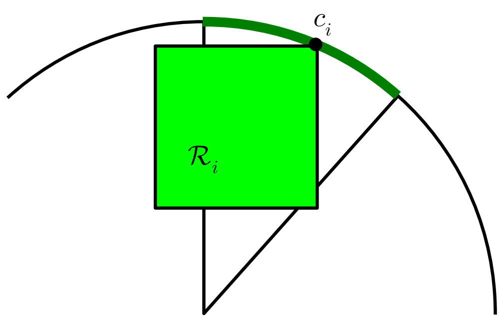

Figure 4.3: Spherical cap $\mathrm{cap}_d(c_i, \phi)$ and the range cube $\mathcal{R}_i$.

Lemma 4 *The $(d-1)$-dimensional content of a spherical cap of angular radius γ is*

$$\begin{aligned}\mathcal{S}(\mathrm{cap}_d(\gamma)) &= \mathcal{S}_{d-1} \cdot \int_0^\gamma \sin(\vartheta)^{d-2} \, \mathrm{d}\vartheta \\ &= \mathcal{S}_{d-1} \cdot \left(\frac{1}{d-1} \cdot \gamma^{d-1} - \frac{d-2}{6(d+1)} \cdot \gamma^{d+1} + \mathcal{O}(\gamma^{d+3}) \right) \quad .\end{aligned}$$

It follows immediately that $\mathcal{M}(d, \gamma) = \Omega(\gamma^{-(d-1)})$. The proof of this lemma is deferred to the end of this section.

For a given vector ϕ let now $\ell = \ell(\phi)$ be the number of non-intersecting spherical caps on Ω_d and let them be centered at points $c_1, \ldots, c_\ell \in \Omega_d$, i.e. the spherical caps are given by $\mathrm{cap}_d(c_1, \phi), \ldots, \mathrm{cap}_d(c_\ell, \phi)$ where $c_i \cdot c_j > \cos(2\phi)$ for all $1 \leq i < j \leq \ell$.

In the next step we will now consider the positions of the input points. Let again $\mathcal{P} = (p_1, \ldots, p_n) \subset \mathbb{R}^d$ be the set of input points and consider independently and identically distributed random noise vectors $r_1, \ldots, r_n$ chosen from the d-dimensional uniform distribution in the hyper-cube $[-\epsilon, \epsilon]^d$. The perturbed point $\widetilde{p}_k = p_k + r_k$ is then uniformly distributed in the hyper-cube $\prod_{i=1}^d [p_k^{(i)} - \epsilon, p_k^{(i)} + \epsilon]$ which we will denote as the *range cube* for input point p_k.

For every spherical cap $\mathrm{cap}_d(c_i, \phi)$ we try to place a bunch of at least $\lfloor n/\ell \rfloor$ input points at exactly the same position such that one vertex of their common range cube lies in c_i, for $1 \leq i \leq \ell$. Let $\mathcal{R}_i$ denote the common range cube for the points placed in such a way at spherical cap $\mathrm{cap}_d(c_i, \phi)$, see also Figure 4.3.

Furthermore, if for a spherical cap $\mathrm{cap}_d(c_i, \phi)$ the input points can be placed such that their common range cube $\mathcal{R}_i$ lies completely inside Ω_d it can be shown that the intersection volume between $\mathcal{R}_i$ and the region $\overline{\mathrm{cap}}_d(c_i, \phi)$ is large. Therefore it will be likely that one

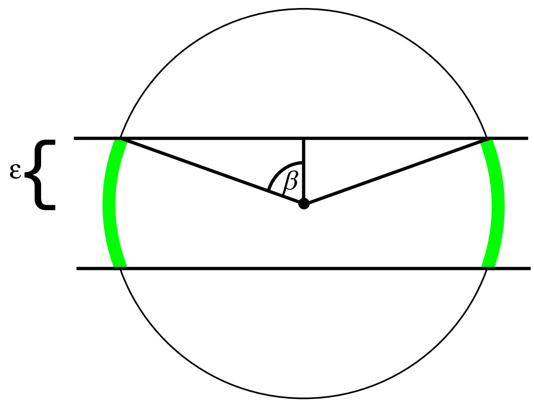

Figure 4.4: In the 2-dimensional case, the spherical segment $\mathrm{seg}_2^{(1)}(\epsilon)$ consists of the two indicated parts of Ω_d. What remains when they are removed from Ω_d are two spherical caps each of angular radius $\beta = \arccos(\epsilon)$.

of the points from the range cube $\mathcal{R}_i$ lies in $\overline{\mathrm{cap}}_d(c_i, \phi)$ after perturbation. By exploiting Observation 1 we can then derive a lower bound on the smoothed number of extreme points.

We call the spherical cap $\mathrm{cap}_d(c_i, \phi)$ *valid* if we can place input points as just described, such that their common range cube $\mathcal{R}_i$ is contained inside Ω_d and a vertex of it lies in c_i. In the next lemma we will investigate the number of valid spherical caps which will be denoted by $\ell_v(\phi) = \ell_v$.

Lemma 5 *Let $\epsilon \leq 1/\sqrt{2}$. Choose ϕ such that $\epsilon = \sin(\phi)$. The number of valid spherical caps of angular radius ϕ is then*

$$\ell_v(\phi) \;\geq\; \frac{2d \cdot \mathcal{S}(\mathrm{cap}_d(\pi/2 - \phi)) - (d-1) \cdot \mathcal{S}_d}{\mathcal{S}(\mathrm{cap}_d(2\phi))} \;=\; \Omega\left(\phi^{-(d-1)}\right) \quad .$$

Proof. A spherical cap $\mathrm{cap}_d(c_i, \phi)$ is non-valid if the corresponding range cube $\mathcal{R}_i$ lies not completely inside Ω_d. We define now d spherical segments which will be denoted by $\mathrm{seg}_d^{(1)}(\epsilon), \ldots, \mathrm{seg}_d^{(d)}(\epsilon)$, where

$$\mathrm{seg}_d^{(i)}(\epsilon) = \left\{ x = (x_1, \ldots, x_d) \in \Omega_d \;\middle|\; -\epsilon \leq x_i \leq \epsilon \right\}$$

for $1 \leq i \leq d$, see also Figure 4.4. From Figure 4.5 it follows immediately that $\mathrm{cap}_d(c_i, \phi)$ is non-valid if and only if its center c_i is from one of these spherical segments since then the corresponding range cube juts out of the unit sphere.

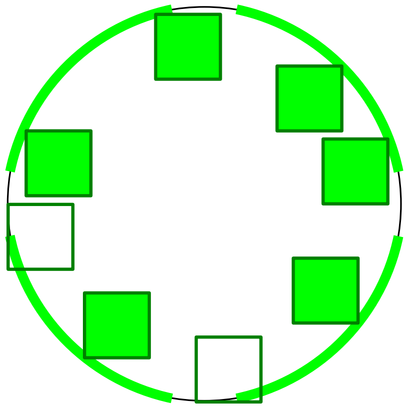

Figure 4.5: Range cubes of valid spherical caps (filled) and of non-valid spherical caps (non-filled).

Let us now consider a fixed spherical segment $\mathrm{seg}_d^{(i)}(\epsilon)$. Indeed, what remains from the sphere Ω_d when $\mathrm{seg}_d^{(i)}(\epsilon)$ is removed are two spherical caps each of angular radius $\beta := \arccos(\epsilon) = \pi/2 - \phi$. The $(d-1)$-dimensional content (surface area) of the spherical segment $\mathrm{seg}_d^{(i)}(\epsilon)$ is then given by $\mathcal{S}_d - 2 \cdot \mathcal{S}(\mathrm{cap}_d(\beta)))$. We have now

$$\begin{aligned} \ell_v(\phi) &\geq \frac{\mathcal{S}_d - d \cdot (\mathcal{S}_d - 2 \cdot \mathcal{S}(\mathrm{cap}_d(\beta))))}{\mathcal{S}(\mathrm{cap}_d(2\phi))} \\ &= \frac{2d \cdot \mathcal{S}(\mathrm{cap}_d(\pi/2 - \phi)) - (d-1) \cdot \mathcal{S}_d}{\mathcal{S}(\mathrm{cap}_d(2\phi))} \\ &= \Omega(\phi^{-(d-1)}) \ , \end{aligned}$$

where the last step follows with Lemma 4. □

In the next lemma we consider the intersection between the region of a valid spherical cap and the corresponding range cube.

Lemma 6 *Let $\mathrm{cap}_d(c_i, \phi)$ be a valid spherical cap and let $\mathcal{R}_i$ be the corresponding range cube of side-length 2ϵ, i.e. a vertex of $\mathcal{R}_i$ lies in c_i and $\mathcal{R}_i$ is completely inside Ω_d. The d-dimensional volume of intersection between the region $\overline{\mathrm{cap}}_d(c_i, \phi)$ and $\mathcal{R}_i$ is at least*

$$\min \left\{ \left(\frac{11}{24} \phi^2 \right)^d \cdot \epsilon^{d-1} \cdot \sqrt{1 - (d-1) \cdot \epsilon^2}, (2\epsilon)^d \right\} \ .$$

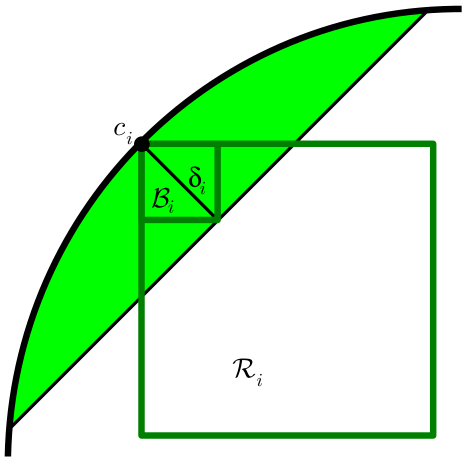

Figure 4.6: The intersection volume between $\overline{\text{cap}}_d(c_i, \phi)$ and $\mathcal{R}_i$.

Proof. Consider the line segment that joins the origin and point c_i and let us denote that part of it that lies inside the region $\overline{\text{cap}}_d(c_i, \phi)$ by δ_i, see Figure 4.6. If $\phi \leq 1$ the length of δ_i is then given by

$$||\delta_i||_2 \;=\; 1 - \cos(\phi) \;\geq\; 1 - (1 - \frac{1}{2}\phi^2 + \frac{1}{24}\phi^4) \;\geq\; \frac{11}{24}\phi^2 \;\; .$$

Now we need to find a lower bound on the intersection volume between $\overline{\text{cap}}_d(c_i, \phi)$ and $\mathcal{R}_i$. Therefore consider the axis-aligned box $\mathcal{B}_i$ that has one vertex lying in c_i, is completely contained in $\overline{\text{cap}}_d(c_i, \phi)$, and has δ_i as the diagonal that joins the vertex at c_i with the 'opposite' vertex of $\mathcal{B}_i$. It follows immediately that $\mathcal{B}_i$ is also completely contained in $\mathcal{R}_i$. We will now approximate the intersection volume between $\overline{\text{cap}}_d(c_i, \phi)$ and $\mathcal{R}_i$ by the volume of box $\mathcal{B}_i$.

Let $x_1, \ldots, x_d$ denote now the side-lengths of $\mathcal{B}_i$. The volume of $\mathcal{B}_i$ is then given by $\text{vol}(\mathcal{B}) = \prod_{i=1}^{d} x_i$. Since δ_i is a diagonal of $\mathcal{B}_i$, it follows that

$$x_d \;=\; \sqrt{||\delta_i||_2^2 - x_1^2 - \ldots - x_{d-1}^2} \;\; . \tag{4.12}$$

Since $\text{cap}_d(c_i, \phi)$ is a valid spherical cap we know about the components of its center $c_i = (c_i^{(1)}, \ldots, c_i^{(d)})$ that $c_i^{(i)} \notin [-\epsilon, \epsilon]$ for all $1 \leq i \leq d$, where again $\epsilon = \sin(\phi)$. We can conclude that $x_i \geq \epsilon \cdot ||\delta_i||_2$ for all $1 \leq i \leq d$. It remains to find out about the minimum volume of box $\mathcal{B}_i$, which is done with the following claim.

Claim 4 *The volume of $\mathcal{B}_i$ is minimal if $d-1$ of the sides of $\mathcal{B}_i$ are of length $\epsilon \cdot ||\delta_i||_2$.*

The proof of this claim is deferred for now.

Without loss of generality we can assume now that $x_1 = \cdots = x_{d-1} = \epsilon \cdot ||\delta_i||_2$, and it follows by (4.12) that $x_d = ||\delta_i||_2 \cdot \sqrt{1-(d-1)\cdot\epsilon^2}$. The volume of box $\mathcal{B}_i$ is then

$$\mathrm{vol}(\mathcal{B}_i) \;=\; ||\delta_i||_2^d \cdot \epsilon^{d-1} \cdot \sqrt{1-(d-1)\cdot\epsilon^2} \;\; .$$

Since the intersection volume between $\overline{\mathrm{cap}}_d(c_i,\phi)$ and $\mathcal{R}_i$ should not drop below the volume of $\mathcal{R}_i$ which is $(2\epsilon)^d$, Lemma 6 follows immediately. □

It remains to show that Claim 4 holds.

Proof of Claim 4. For ease of notation we abbreviate $b := \epsilon \cdot ||\delta_i||_2$. Let again $x_1,\ldots,x_d$ denote the side-lengths of $\mathcal{B}_i$. Without loss of generality we assume that the x_j's are in increasing order and that x_k is the smallest one not equal to b, i.e. $b = x_1 = \ldots = x_{k-1} < x_k \le x_{k+1} \le \ldots \le x_d$. Let also $c := ||\delta_i||_2^2 - x_1^2 - \ldots - x_{k-1}^2 - x_{k+1}^2 - \ldots x_{d-1}^2$. It follows then by (4.12) that $x_d = \sqrt{c - x_k^2}$ and therefore

$$\mathrm{vol}(\mathcal{B}_i) \;=\; x_1 \cdots x_{d-1} \cdot \sqrt{c - x_k^2} \;\; .$$

The plan is now to decrease the side-length x_k and to increase x_d accordingly such that the diagonal has still a length of $||\delta_i||_2$. Therefore, consider now the axis-aligned box $\widehat{\mathcal{B}}_i$ with side-lengths $y_1,\ldots,y_d$ such that $y_1 = \ldots = y_k = b$, and $y_j = x_j$ for $k+1 \le j \le d-1$. Again by (4.12) it follows then that $y_d = \sqrt{c-b^2}$. Now we have that

$$\mathrm{vol}(\mathcal{B}_i) - \mathrm{vol}(\widehat{\mathcal{B}}_i) \;=\; x_1 \cdots x_{k-1} \cdot x_{k+1} \cdots x_{d-1} \cdot \left(x_k \cdot \sqrt{c - x_k^2} - b \cdot \sqrt{c - b^2} \right) \;>\; 0 \;\; .$$

The last step follows since $x_k > b$ and $\sqrt{c - x_k^2} = x_d > b$. □$_{\text{Claim 4}}$

In the next lemma we are ready to find out about the probability that the region of a valid spherical cap contains at least one point and we thus have an extreme point after perturbation for every valid spherical cap. Recall that the number of valid spherical caps is denoted by $\ell_v(\phi)$.

Lemma 7 *For k a sufficiently large constant, if $\phi = k \cdot (\epsilon/n)^{1/(3d-1)}$, the region of every valid spherical cap is nonempty with probability at least $1 - 1/e$, after perturbation.*

Proof. Let $\ell_v(\phi)$, the number of valid spherical caps, be as in Lemma 5. For every valid spherical cap $\mathrm{cap}(c_i,\phi)$ we place a bunch of at least $\lfloor n/\ell_v(\phi) \rfloor = \mathcal{O}(n \cdot \phi^{d-1})$ input points in the described way, such that a vertex of their common range cube $\mathcal{R}_i$ lies in c_i and that $\mathcal{R}_i$ is completely contained inside Ω_d.

From Lemma 6 it follows immediately that the probability that none of these points lies in the region $\overline{\mathrm{cap}}(c_i,\phi)$ after perturbation is

$$\begin{aligned}&\mathbf{Pr}[\,\overline{\mathrm{cap}}(c_i,\phi) \text{ is empty}\,]\\ &\le \left(1 - \frac{\min\left\{\left(\frac{11}{24}\phi^2\right)^d \epsilon^{d-1} \cdot \sqrt{1-(d-1)\cdot\epsilon^2}, (2\epsilon)^d\right\}}{(2\epsilon)^d}\right)^{\lfloor n/\ell_v(\phi)\rfloor} \;\; .\end{aligned}$$

For k a sufficiently large constant, we choose $\phi = k \cdot (\epsilon/n)^{1/(3d-1)}$ such that

$$\frac{(2\epsilon)^d}{(\frac{11}{24}\phi^2)^d \cdot \epsilon^{d-1} \cdot \sqrt{1-(d-1)\cdot\epsilon^2}} \;\le\; \frac{n}{\ell_v(\phi)} \;.$$

Lemma 7 is thus proven.

□

By combining Lemmas 5 and 7 the next theorem follows immediately.

Theorem 10 *For random noise from the uniform distribution in a d-dimensional hypercube of side-length $2\epsilon \le \sqrt{2}$, the smoothed number of extreme points over all input sets $\mathcal{P}$ of n points is*

$$\max_{\mathcal{P}} \mathbf{E}\left[\mathcal{V}(\widetilde{\mathcal{P}})\right] \;=\; \Omega\left(\min\left\{\left(\frac{n}{\epsilon}\right)^{\frac{d-1}{3d-1}}, n\right\}\right) \;.$$

It remains to prove Lemma 4.

Lemma 4 *The $(d-1)$-dimensional content of a spherical cap of angular radius γ is*

$$\begin{aligned}\mathcal{S}(\mathrm{cap}_d(\gamma)) \;&=\; \mathcal{S}_{d-1}\cdot\int_0^\gamma \sin(\vartheta)^{d-2}\,\mathrm{d}\vartheta\\ &=\; \mathcal{S}_{d-1}\cdot\left(\frac{1}{d-1}\cdot\gamma^{d-1} - \frac{d-2}{6(d+1)}\cdot\gamma^{d+1} + \mathcal{O}(\gamma^{d+3})\right) \;.\end{aligned}$$

Proof. Consider the spherical cap of angular radius γ center at $e_1 := (1,0,\ldots,0) \in \Omega_d$ which is given by $\mathrm{cap}_d(e_1,\gamma) = \{x \in \Omega_d \mid e_1 \cdot x > \cos(\gamma)\}$.

In 3-dimensional Euclidean space we can use *Pappus' Centroid Theorem* [Wei] which is also known as *Guldin's First Rule* to compute the surface area of a spherical cap as follows. The theorem gives a general formula to compute the 2-dimensional content of a *surface of revolution*. Consider the curve of the integrable function $f : [a,b] \to \mathbb{R}$ that does not intersect the x_1-axis. This curve is called the *generating curve* and its length is given by $\int_a^b \sqrt{1+f'(x)^2}\,\mathrm{d}x$ where $f'(x) = \mathrm{d}f(x)/\mathrm{d}x$. The content $\mathcal{S}_f(a,b)$ of the surface generated by the revolution of the generating curve about the x_1-axis is then given by

$$\mathcal{S}_f(a,b) \;=\; \int_a^b \mathcal{S}_2\cdot f(x)\sqrt{1+f'(x)^2}\,\mathrm{d}x \;=\; 2\pi\int_a^b f(x)\sqrt{1+f'(x)^2}\,\mathrm{d}x \;.$$

Note that $\mathcal{S}_2 \cdot f(x) = 2\pi\cdot f(x)$ denotes the 1-dimensional content of a 2-dimensional sphere of radius $f(x)$.

In our case, the generating function is that of a semi-circle, namely $f(x) = \sqrt{1-x^2}$ and $f'(x) = -x/\sqrt{1-x^2}$, and $a = \cos(\gamma)$ and $b = 1$, so we get that

$$\mathrm{cap}_3(\gamma) \;=\; 2\pi\int_{\cos(\gamma)}^1 1\,\mathrm{d}x \;=\; 2\pi\cdot(1-\cos(\gamma)) \;.$$

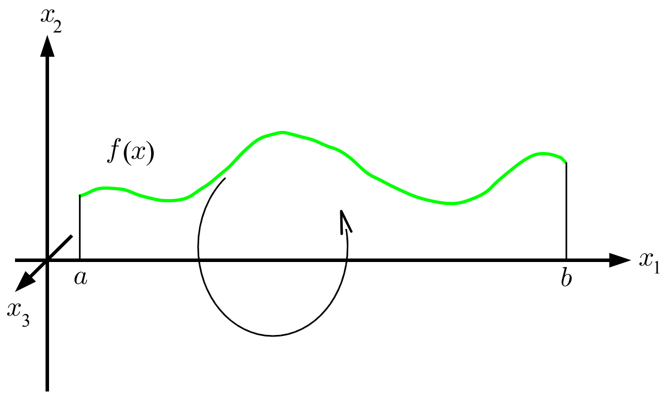

Figure 4.7: Generating curve of function $f(x)$ revolves around the x_1-axis.

The Guldin-Pappus' Theorem has been extended to higher dimensions [Kur53]. Consider a $(d-1)$-dimensional surface whose intersection with a hyperplane perpendicular to the x_1-axis is a $(d-1)$-dimensional hypersphere of radius $f(x)$, where $f : [a, b] \to \mathbb{R}$ is an integrable function that does not intersect the x_1-axis. The $(d-1)$-dimensional content $\mathcal{S}_f(a, b)$ of this surface is then given by

$$\mathcal{S}_f(a,b) \;=\; \int_a^b \mathcal{S}_{d-1} \cdot f(x)^{d-2} \sqrt{1 + f'(x)^2}\,\mathrm{d}x \quad .$$

Note that $\mathcal{S}_{d-1} \cdot f(x)^{d-2}$ denotes the $(d-2)$-dimensional content of a $(d-1)$-dimensional sphere of radius $f(x)$.

To compute the $(d-1)$-dimensional content of the surface of $\mathrm{cap}_d(e_1, \gamma)$ we use again that $f(x) = \sqrt{1-x^2}$ and $f'(x) = -x/\sqrt{1-x^2}$, and $a = \cos(\gamma)$ and $b = 1$, so we get that

$$\begin{aligned} \mathcal{S}(\mathrm{cap}_d(\gamma)) &= \mathcal{S}_{d-1} \cdot \int_{\cos(\gamma)}^1 \sqrt{1-x^2}^{\,d-2} \cdot \sqrt{1 + \frac{x^2}{1-x^2}}\,\mathrm{d}x \\ &= \mathcal{S}_{d-1} \cdot \int_{\cos(\gamma)}^1 \sqrt{1-x^2}^{\,d-3}\,\mathrm{d}x \quad . \end{aligned}$$

By a linear substitution $x = \cos(\vartheta),\ \mathrm{d}x = -\sin(\vartheta)\,\mathrm{d}\vartheta$ and by the series expansion of

$\sin(\vartheta)$ we get that

$$\begin{aligned}\mathcal{S}(\mathrm{cap}_d(\gamma)) &= \mathcal{S}_{d-1} \cdot \int_0^\gamma \sin(\vartheta)^{d-2}\,\mathrm{d}\vartheta \\ &= \mathcal{S}_{d-1} \cdot \int_0^\gamma \left(\sum_{i=0}^{\infty}(-1)^i \frac{\vartheta^{2i+1}}{(2i+1)!}\right)^{d-2} \mathrm{d}\vartheta \\ &= \mathcal{S}_{d-1} \cdot \int_0^\gamma \vartheta^{d-2} - \frac{d-2}{6} \cdot \vartheta^d + \mathcal{O}(\vartheta^{d+2})\,\mathrm{d}\vartheta \\ &= \mathcal{S}_{d-1} \cdot \left(\frac{1}{d-1} \cdot \gamma^{d-1} - \frac{d-2}{6(d+1)} \cdot \gamma^{d+1} + \mathcal{O}(\gamma^{d+3})\right) \quad ,\end{aligned}$$

which concludes the proof of Lemma 4.

□

4.4 Conclusion

The following table gives an overview on the results of this chapter. Depicted are upper and lower bounds for the smoothed number of extreme/maximal points for noise from the Gaussian normal distribution and the uniform distribution. The bounds are given in $\mathcal{O}$-notation, the dimension d is considered as a constant.

	Upper Bounds	Lower Bounds
Gaussian $\mathrm{N}(0,\sigma)$	$\mathcal{O}\big((1/\sigma)^d \cdot \log(n)^{\frac{3}{2}\cdot d-1} + \log(n)^{d-1}\big)$	$\Omega\big(\log(n)^{(d-1)/2}\big)$ $(\star)$ $\Omega\big(\log(n)^{d-1}\big)$ $(\star\star)$
uniform in $[-\epsilon,\epsilon]^d$	$\mathcal{O}\big((n \cdot \log(n)/\epsilon)^{\frac{d}{d+1}} + \log(n)^{d-1}\big)$	$\Omega\big(\min\{(n/\epsilon)^{\frac{d-1}{3d-1}}, n\}\big)$

The lower smoothed case bounds for random noise from the Gaussian normal distribution are obtained from the average case bounds on the number of extreme points $(\star)$ [Ray70] and on the number of maximal points $(\star\star)$, see also Theorem 6.

We observe that for Gaussian normal noise the lower and upper smoothed bounds leave a gap of roughly $\log(n)^{d/2}$ for the number of maximal points and roughly $\log(n)^d$ for the number of extreme points. The gap for the uniform noise is much smaller, the bounds differ only by a factor of roughly $\log(n)^{d/(d+1)}$. This might be due to the fact that for Gaussian normal noise we do not have an explicit lower bound constructions and rely on the average case bounds.

However, we observe analogously to the results of the previous chapter, that there is a significant difference between the behavior under Gaussian normal noise and uniform noise.

From the upper bound results we observe again that our analysis method cannot be applied for Gaussian normal noise of arbitrary deviation. For $\sigma \leq \mathcal{O}(\sqrt{\log(n)})$ we obtain the average case bound of $\mathcal{O}(\log(n)^{d-1})$. If $\sigma \geq \Omega(\log(n)^{3/2-1/d}/n^{1/d})$ we obtain $\mathcal{O}(n)$ many extreme points which means that our analysis cannot distinguish between the perturbed and unperturbed case. For uniform noise we observe the same. Here the upper bound are meaningless if $\epsilon \leq \mathcal{O}(n/\log(n)^{d-1-1/d})$ or $\epsilon \geq \Omega(\log(n)/n^{1/d})$.

5 Bounding Box of a Moving Point Set

The goal of this chapter is to present an interesting application of smoothed analysis in the area of analysing motion and the complexity of motion. When talking about motion, we consider a non-static scenario in $\mathbb{R}^d$ where given (input) points move. The movement of each point is predictable for some time in the future. The motivation to consider motion comes from the fact that many applications are based on algorithms dealing with moving objects. There has been diverse research on moving objects and the question how to deal with motion computationally and algorithmically.

This chapter gives a brief introduction into this area of research where we focus on a complexity measure for movement of objects, the *motion complexity*. We then introduce *smoothed motion complexity* which is an extension of motion complexity that is based, as the name points out, on smoothed analysis. The motivation for using smoothed analysis in this context comes from the observation that in applications usually the data about positions of moving objects are inherently noisy due to measurement errors. Again we will model this measurement error by Gaussian normal noise.

To illustrate the concept of *motion complexity* and *smoothed motion complexity*, we consider as an example the problem to maintain a smallest orthogonal bounding box of a moving point set under linear motion.

Outline. In Section 5.1, a brief introduction into the area of analysing motion with a focus on *kinetic data structures* is given. We introduce also *motion complexity* and *smoothed motion complexity*.

In Section 5.2, the concept of motion complexity is illustrated by an example. We consider the problem to maintain the smallest orthogonal bounding of a linearly moving point set in $\mathbb{R}^d$. We show that the motion complexity of the bounding box is closely related to the number of extreme points of a set of points. Thus we obtain bounds on the motion complexity of the bounding box applying the bounds on the number of *extreme points* from the previous Chapter 4. At the end of this section, we see how the upper bounds on the smoothed motion complexity can be improved. This is done by applying results from Chapter 3 on the number of *left-to-right maxima*.

The last Section 5.3 briefly summarizes this chapter and gives some conclusions.

5.1 Analysing Motion

The task to process a set of continuously moving objects arises in a broad variety of applications, e.g. in mobile ad-hoc networks, traffic control systems, and computer graphics (rendering moving objects). Therefore, researchers investigated data structures for certain attributes of moving point sets that can be efficiently maintained under continuous motion, e.g. to answer proximity queries [BGZ97], maintain a clustering [Har04], a convex hull [BGH99], or some connectivity information of the moving point set [HS01].

5.1.1 Kinetic Data Structures

Basch et al. [BGH99] introduce *kinetic data structures* which as a framework for data structures for moving objects. 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, i.e. linear functions or low-degree polynomials. This specification is called a *flight plan*. The flight plan may change from time to time and these updates are reported to the kinetic data structure. The goal is to maintain the description of a combinatorial structure as the objects move according to the flight plan.

For a kinetic data structure, the number of combinatorial changes in the description of the maintained attribute that occur during linear (or low degree algebraic) motion are called *external events*. Events that are processed by the data structure because of internal needs are called *internal events*.

The efficiency of a kinetic data structures is then analyzed by comparing the worst case number of internal events and external events it processes against the worst case number of external events. Using this framework many interesting and efficient kinetic data structures have been developed, e.g. for connectivity of discs [GHSZ01] and rectangles [HS01], convex hulls [BGH99], proximity problems [BGZ97], and collision detection for simple polygons [KSS02].

Basch et al. [BGH99] develop also a kinetic data structure to maintain a bounding box of a moving point set in $\mathbb{R}^d$. The number of events these data structures process is $O(n \log n)$ which is close to the worst case motion complexity of $\Theta(n)$, as we will see later. Agarwal and Har-Peled [AH01] show that it is possible to maintain an $(1+\epsilon)$-approximation of such a bounding box efficiently. The advantage of this approach is that the motion complexity of this approximation is only $O(1/\sqrt{\epsilon})$.

5.1.2 Motion Complexity

We will use now the concept of external events in order to introduce *motion complexity* as a complexity measure for motion. Thus, motion complexity is already implicitly contained in the framework of kinetic data structures. We consider here moving point sets in $\mathbb{R}^d$ under linear motion that are defined as follows.

Definition 7 (Moving Point Set) *Given is a set $\mathcal{P}$ of n points in $\mathbb{R}^d$. We call $\mathcal{P}$ a* ***linearly moving point set*** *or a* ***points set under linear motion*** *if the following holds.*

The position $\mathrm{pos}_i(t)$ of the ith point of $\mathcal{P}$ at time t is given by a linear function of t. Thus we have $\mathrm{pos}_i(t) = p_i + s_i \cdot t$ where p_i is the initial position and s_i the speed vector of the ith point.

For linearly moving point sets we consider certain geometric structures such as the convex hull or the Voronoi diagram of the point sets. The worst case number of external events with respect to the maintainance of such a structure over time is denoted as the *worst case motion complexity*. Analogously, we denote the average case number of external events as the *average case motion complexity* of that structure.

The worst case motion complexity is the maximum number of external events over all choices of speed values and initial positions. The average motion complexity is the expected number of external events, where all speed values and initial positions are independent and identically distributed random vectors chosen from a fixed probability distribution.

The average case motion complexity has already been considered in the past. If n particles are drawn independently from the unit square then it has been shown that the expected number of combinatorial changes to the description of the convex hull is $\Theta(\log(n)^2)$ and of the Voronoi diagram $\Theta(n^{3/2})$, and to the closest pair problem $\Theta(n)$ [ZDBI97].

5.1.3 Smoothed Motion Complexity

Many applications are based on algorithms dealing with moving objects, but usually data about positions of moving objects are inherently noisy due to measurement errors. Therefore we introduce *smoothed motion complexity* that considers this imprecise information and uses smoothed analysis to model noisy data.

In the context of mobile data this means that both the speed vector and the starting position of an input point are slightly perturbed by random noise from a fixed noise distribution. The smoothed motion complexity is then the worst case expected motion complexity over all input instances perturbed in such a way. The speed vectors and initial positions are normalized such that $p_i, s_i \in [-1, 1]^d$.

5.2 Motion Complexity of the Bounding Box

To illustrate the concept of *motion complexity* and *smoothed motion complexity* we consider the problem of maintaining a bounding box of a moving point set under linear motion. This problem is formally described in the following definition.

Definition 8 (Bounding Box) *Given is a set $\mathcal{P}$ of n linearly moving points in $\mathbb{R}^d$.*

At a particular point of time t_0, the ***bounding box*** *of set $\mathcal{P}$ with given initial positions $p_1, \ldots, p_n$ and speed vectors $s_1, \ldots, s_n$ is then the smallest orthogonal box containing* *all points of $\mathcal{P}$. At any time the bounding box is uniquely defined and combinatorially described by at most $2d$ **bounding points**, i.e. the points that attain the maximum and minimum value in each of the d dimensions.*

For example, in the one dimensional case, the bounding box problem can be interpreted as a race where all participants (drivers) have slightly different starting positions. (Without loss of generality, we assume that all drivers have the same direction.) The radio reporter has to tell the audience each time when the leading position changes, i.e. another driver becomes the leader. At a particular point of time, no further changes will occur, namely when the fastest driver has become the leader. The number of times that the radio reporter announces a change in the leading position is the number we are interested in.

More generally, for a set $\mathcal{P}$ of linearly moving points, if the combinatorial description of the bounding box of $\mathcal{P}$ changes, i.e. any bounding point of $\mathcal{P}$ changes, then an *external event* occurs. The *motion complexity of the bounding box* is thus the number of combinatorial changes over time to the set of at most $2d$ bounding points defining the bounding box.

Clearly the worst case motion complexity of the bounding box is at most $2d \cdot n = \Theta(n)$. In the best case, the motion complexity of the bounding box is 0, while the average case motion complexity is $\mathcal{O}(\log(n))$, as we will see later.

When we consider the smoothed motion complexity of the bounding box we add to each coordinate of the speed vector and each coordinate of the initial position of every input point an i.i.d. random vector from a fixed probability distribution over $\mathbb{R}^d$, e.g. the d-dimensional Gaussian normal distribution. The smoothed motion complexity is then the worst case expected motion complexity over all choices of speed vectors and initial positions.

In the following we will see, that the 1-dimensional bounding box problem for linearly moving point sets is dual to the 2-dimensional convex hull of a point set. The results from the previous chapter carry thus immediately over to the bounding box problem.

5.2.1 Duality between Bounding Box and Convex Hull

We make the following simplifications. We will consider only the 1-dimensional problem, i.e. all points move along a line such that their ordering changes only when they overtake each other. Since all dimensions are independent from each other, a bound for the 1-dimensional problem can be multiplied by d to yield a bound for the problem in d dimensions.

We map now each point with initial position p_i and speed value s_i to a point $P_i = (s_i, -p_i)$ in $\mathbb{R}^2$. Then we can utilize that the number of combinatorial changes to the description of the 1-dimensional bounding box is equal to the number of extreme points of the convex hull of the P_i's. This relation is easily seen by the following considerations.

In a first step we consider the following scenario. For all $1 \leq i \leq n$, we consider for the ith point with initial position p_i and speed value s_i the function $\mathrm{pos}_i(t) = p_i + s_i \cdot t$

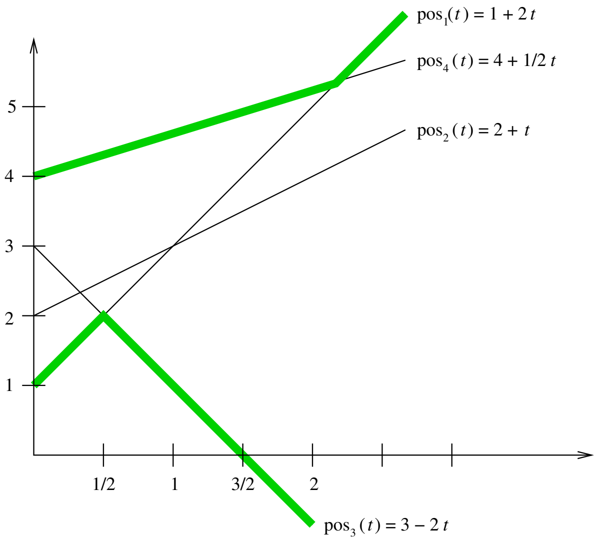

Figure 5.1: The upper and lower envelope of a line arrangement and the corresponding dual set of points with the upper and lower convex hull are depicted.

which gives us the position of the ith point at any point of time. Considering the graphs of the position functions of all points in the set $\mathcal{P}$, we obtain a line arrangement which we will denote as $\ell(\mathcal{P})$. The *upper and lower envelope* of the line arrangement $\ell(\mathcal{P})$ is the boundary of the top and bottom cell of $\ell(\mathcal{P})$, which is a chain of edges defined as the maximum and minimum of the linear functions whose graphs are the lines in $\ell(\mathcal{P})$, respectively. We observe now the following.

The edges of the upper envelope of $\ell(\mathcal{P})$ give us exactly the points that are rightmost and thus the boundary points on the right side of the bounding box over time. By symmetry, this holds also for the lower envelope of $\ell(\mathcal{P})$, which gives us the points that are leftmost, i.e. the boundary points on the left side of the bounding box. So by the number of edges on the upper and lower envelope of $\ell(\mathcal{P})$ we obtain immediately the number of different bounding points for 'both' boundaries and thus the motion complexity of the bounding box of $\mathcal{P}$.

In a second step we will consider the following simple *duality transform* [dvOS00]. For each line $\ell : y := b + m \cdot x$ in $\mathbb{R}^2$ we consider the point $\ell^\star = (m, -b)$ in $\mathbb{R}^2$ called its dual. The dual of a point $p = (p_x, p_y)$ is then the line $p^\star : y := -p_y + p_x \cdot x$. We observe that this duality transform preserves ordering, i.e. a point p lies above a line ℓ if and only if the point $\ell^\star$ lies above line $p^\star$.

From this observation it follows immediately that the line to function $\mathrm{pos}_i(t) = p_i + s_i \cdot t$ belongs to the lower envelope of $\ell(\mathcal{P})$ if and only if its dual, the point $P_i = (s_i, -p_i)$ is a vertex of the upper convex hull of $P_1, \ldots, P_n$ which consists of the convex hull edges that have all remaining points below their supporing line. Of course, the same holds for the lower envelope, the point P_i is then a vertex of the lower convex hull, see also Figure 5.1.

By this method it follows immediately that the number of extreme points in 2 dimensions is equal to the number of edges of the lower and upper envelope of the dual line arrangement and thus equal to the motion complexity of the bounding box in 1 dimension. The bounds on the average and smoothed number of extreme points in 2 dimensions from Chapter 4 carry thus over to the bounding box problem.

5.2.2 Average Motion Complexity of the Bounding Box

The result of Renyi and Sulanke [RS63] on the average number of extreme points of the convex hull when points are independent and identically distributed random vectors chosen from the 2-dimensional Gaussian normal distribution implies the following theorem.

Corollary 2 *The average motion complexity of the bounding box of a set of n linearly moving points in d-dimensional space, where initial positions and speed vectors are independent and identically distributed random vectors chosen from the d-dimensional Gaussian normal distribution, is*

$$\mathcal{O}\left(\sqrt{\log(n)}\right) \quad .$$

Another bound on the average motion complexity of the bounding box follows from the 2-dimensional version of Theorem 7 for all continuous probability distributions.

Corollary 3 *The average motion complexity of the bounding box of a set of n linearly moving points in d-dimensional space, where initial positions and speed vectors are independent and identically distributed random vectors chosen from a d-dimensional continuous probability distribution, is*

$$\mathcal{O}\left(\log(n)\right) \quad .$$

5.2.3 Upper Bounds on the Smoothed Motion Complexity

From the 2-dimensional version of Theorem 8 which upper bounds the smoothed number of extreme points under Gaussian normal noise we get an upper bound on the smoothed motion complexity of the bounding box under Gaussian normal noise.

Corollary 4 *The smoothed motion complexity of the bounding box over all sets of n linearly moving points in d-dimensional space, with initial positions and speed vectors from $[-1,1]^d$ perturbed by random noise from the Gaussian normal distribution of deviation σ, is*

$$\mathcal{O}\left(\left(\frac{1}{\sigma}\right)^2 \cdot \log(n)^2 + \log(n)\right) \quad .$$

By Theorem 9, the smoothed number of extreme points under uniform noise is covered. This gives us the following corollary.

Corollary 5 *The smoothed motion complexity of the bounding box over all sets of n linearly moving points in d-dimensional space, with initial positions and speed vectors from $[-1,1]^d$ perturbed by random noise from the uniform distribution in a hypercube of side length 2ϵ, is*

$$\mathcal{O}\left(\left(\frac{n \cdot \log(n)}{\epsilon}\right)^{2/3} + \log(n)\right) .$$

5.2.4 Lower Bounds on the Smoothed Motion Complexity

Also the lower bounds on the smoothed number of extreme points carry over to the motion complexity of the bounding box. By the average case bound of Renyi and Sulanke [RS63] we get the following lower bound on the smoothed motion complexity under Gaussian normal noise.

Corollary 6 *The smoothed motion complexity of the bounding box problem over all sets of n linearly moving points in d-dimensional space, with initial positions and speed vectors from $[-1,1]^d$ perturbed by random noise from the Gaussian normal distribution of deviation σ, is*

$$\Omega\left(\sqrt{\log(n)}\right) .$$

From the 2-dimensional version of Theorem 10 follows the next corollary.

Corollary 7 *The smoothed motion complexity of the bounding box problem over all sets of n linearly moving points in d-dimensional space, with initial positions and speed vectors from $[-1,1]^d$ perturbed by random noise from the uniform probability distribution in a hypercube of side length 2ϵ, is*

$$\Omega\left(\min\left\{\sqrt[5]{\frac{n}{\epsilon^2}}, n\right\}\right) .$$

5.2.5 Improved Upper Bounds on the Smoothed Motion Complexity

In this subsection we show how the upper bounds on the smoothed motion complexity can easily be improved by a simple consideration. Again we consider only the 1-dimensional problem and exploit that the problem is dimension-wise independent. Results hold thus also for the d-dimensional case.

We observe that adding a constant to all initial positions and speed values, or multiplying these values by a constant does not change the motion complexity of the bounding box. Thus we assume that the points are ordered by their increasing initial positions and that they are all moving to the left with absolute speed values between 0 and 1. We count thus only events that occur because the leftmost point of the 1-dimensional bounding box changes.

A necessary condition for the jth point to cause an external event is that all its "preceding" points have smaller absolute speed values, i.e. that $s_i < s_j$, for all $i < j$. If this is the case, then s_j is clearly a *left-to-right maximum* as seen in Chapter 3. Since we are interested in upper bounds we can neglect the initial positions of the points and need only to focus on the sequence of absolute speed values $(s_1, \ldots, s_n)$ and count the left-to-right maxima in this sequence.

It follows immediately that the results on the number of left-to-right maxima carry over to the bounding box problem. Theorem 2 in Chapter 3 implies thus directly the following corollary.

Corollary 8 *The smoothed motion complexity of the bounding box problem over all sets of n linearly moving points in d-dimensional space, with initial positions in $\mathbb{R}^d$ and speed vectors from $[-1,1]^d$ perturbed by random noise from the Gaussian normal distribution of deviation σ, is*

$$\mathcal{O}\left(\frac{1}{\sigma} \cdot \log(n)^{3/2} + \log(n)\right) \quad .$$

We can also use Theorem 4 for unimodal noise distributions immediately. Recall, that a random noise distribution is *unimodal* if the corresponding density function is monotonically increasing on $\mathbb{R}_{\leq 0}$ and monotonically decreasing on $\mathbb{R}_{\geq 0}$.

Corollary 9 *The smoothed motion complexity of the bounding box problem over all sets of n linearly moving points in d-dimensional space, with initial positions in $\mathbb{R}^d$ and speed vectors from $[-1,1]^d$ perturbed by random noise from a continuous unimodal probability distribution with 1-dimensional density function φ, is*

$$\mathcal{O}\left(\sqrt{n \cdot \log(n) \cdot \varphi(0)} + \log(n)\right) \quad .$$

These improved bounds on the smoothed motion complexity of the bounding box imply something else. The bounds on the smoothed number of extreme points as derived in Chapter 4 are not exactly tight, at least for the 2-dimensional case.

5.3 Conclusion

In this chapter an introduction to the analysis of motion was given. We introduced *motion complexity* and *smoothed motion complexity* as a measure for the complexity of maintaining combinatorial structures of moving data, i.e. points.

We saw that for the problem of maintaining the bounding box of a set of linearly moving points, the smoothed motion complexity differs significantly from the worst case motion complexity. This makes it unlikely that the worst case is attained in typical applications.

Therefore it seems promising to reconsider the use of worst case analysis for algorithms dealing with moving objects. Especially in the development of kinetic data structures this might lead to interesting new results.

6 Voronoi Diagram and Delaunay Triangulation

The *Voronoi diagram* together with its dual structure the *Delaunay triangulation* are two very important and fundamental structures in computer science. Especially in computational geometry they constitute a central topic in research [Aur91, For97] with many applications. Both structures are also well known and widely used in several other fields of (natural) science. Besides mathematics and computer science, Voronoi diagrams and Delaunay triangulations can be found in physics, geology, agriculture, geography, and many other disciplines.

Voronoi diagrams have the great advantage to be a rather simple but quite elegant structure. There are also many extensions of the basic Voronoi diagrams which are obtained by varying metric, sites, environment, and constraints. In computer science they are widely used in clustering, mesh generation, graphics, curve and surface reconstruction, and other applications [OBS92].

Voronoi diagrams are named after the Russian mathematician Voronoi [Vor08] who generalized an original idea of Gauss [Gau40] to higher dimensions. Gauss' work was motivated by the study of quadratic forms and was also exploited and further developed by Dirichlet [Dir50]. Voronoi diagrams have been 'reinvented' by other researchers, e.g. by the physicists Wigner and Seits [WS33], the meteorologist Thiessen [Thi11] and the biologist Blum [Blu73]. In these fields of science, the Voronoi diagram is thus known by other names such as Wigner-Seitz diagram, Thiessen diagram, and Blum transform. In mathematics the Voronoi diagram is usually also known as Dirichlet tessellation.

Definition 9 (Voronoi diagram) *Given is a set $\mathcal{P}$ of n points – also called **sites** – in $\mathbb{R}^d$.*

*The **Voronoi cell** of a site p consists of all points that are strictly closer to p than to any other site in $\mathcal{P} - \{p\}$.*

*The **Voronoi face** of a nonempty subset $\mathcal{T}$ of $\mathcal{P}$ consists of all points that are equidistant to all sites of $\mathcal{T}$ and closer to any site of $\mathcal{T}$ than to any other site in $\mathcal{P} - \mathcal{T}$.*

*The **Voronoi diagram** of $\mathcal{P}$ – denoted $\mathcal{VD}(\mathcal{P})$ – is the collection of all nonempty Voronoi faces and forms a cell complex partitioning $\mathbb{R}^d$.*

The Voronoi cell of a site p is always a nonempty, open, convex, full-dimensional subset of $\mathbb{R}^d$. The one-dimensional and two-dimensional Voronoi faces are also called *Voronoi vertices* and *Voronoi edges*, respectively.

In his work, Voronoi [Vor08] introduced also the Delaunay triangulation for sites that form a lattice. Later Delaunay [Del34] extended this definition to irregularly placed sites

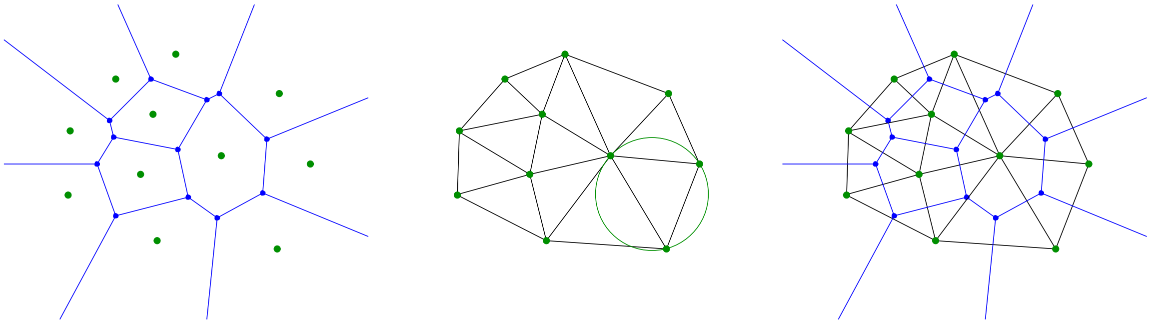

Figure 6.1: The Voronoi diagram and Delaunay triangulation of a set of points in $\mathbb{R}^2$.

and the structure was then named after him.

Definition 10 (Delaunay triangulation) *Given is a set $\mathcal{P}$ of n points – also called* ***sites*** *– in $\mathbb{R}^d$.*

Let $\mathcal{T}$ be a subset of $\mathcal{P}$ such that a sphere through all the sites of $\mathcal{T}$ is empty, i.e. all other sites in $\mathcal{P} - \mathcal{T}$ are lying exterior of this sphere. The ***Delaunay face*** *of $\mathcal{T}$ is the (relative) interior of the convex hull of $\mathcal{T}$.*

The ***Delaunay triangulation*** *of $\mathcal{P}$ – denoted $\mathcal{DT}(\mathcal{P})$ – is the collection of all Delaunay faces and forms a cell complex partitioning the convex hull of $\mathcal{P}$.*

Throughout this chapter we assume that the set $\mathcal{P}$ is in *general position*, i.e. no $d+2$ sites lie on a common d-sphere and no $k+2$ sites lie on a common k-flat, for $k < d$. It follows that the Delaunay triangulation of $\mathcal{P}$ is a simplicial cell complex, i.e. all d-dimensional Delaunay faces are simplices. Therefore we will consider from now on only *Delaunay simplices*.

Duality. From Figure 6.1 we see that there is an obvious one-to-one correspondence between the Voronoi diagram and the Delaunay triangulation of a set $\mathcal{P}$. By mapping the Voronoi face of a set $\mathcal{T} \subseteq \mathcal{P}$ to the corresponding Delaunay face of $\mathcal{T}$ we obtain a *duality* between cell complexes that reverses face ordering. The dimensions of the Voronoi and Delaunay face of a particular set $\mathcal{T} \subseteq \mathcal{P}$ sum up to the dimension d.

Relation to Convexity. There is a close connection between Delaunay triangulations in $\mathbb{R}^d$ and convex hulls in $\mathbb{R}^{d+1}$, and between Voronoi diagrams in $\mathbb{R}^d$ and half-space intersections in $\mathbb{R}^{d+1}$. Brown [Bro79, Bro80] was the first to give a transform that relates the dual of Voronoi diagrams (i.e. Delaunay triangulations) in $\mathbb{R}^2$ to convex hulls in $\mathbb{R}^3$ via a stereographical projection that maps the sites into points lying on a 3-sphere. Brown's result and further extensions to higher dimensions, enabled the complete analysis of the combinatorial complexity of Voronoi diagrams and Delaunay triangulations [Kle80, Sei87] in arbitrary dimensions since known results on the size of polytopes carry over.

Related Work. The Voronoi diagram can be computed in linear time from the Delaunay triangulation, using the one-to-one correspondence between their faces. A vast variety of basic and (relatively) simple algorithms exists for the construction of Delaunay diagrams such as the plane sweep [For87] and the divide-and-conquer [Sha78] algorithm in $\mathbb{R}^2$, and the (randomized) incremental [GKS92, Cha91], and the gift-wrapping [CK70] algorithm in arbitrary dimensions.

In fact, most of these algorithms are actually specialized convex hull algorithms (except the plane sweep algorithm) due to the close relation to convexity. Any $(d+1)$-dimensional convex hull algorithm can be used to compute a d-dimensional Delaunay triangulation. All these algorithms depend in their run time on the number of faces of the Delaunay triangulation. Unfortunately, in d dimensions this number is $\Theta(n^{\lceil d/2 \rceil})$ in the worst case [Kle80, Sei87] (for the 'general' diagram with the Euclidean metric).

Recent research attempts to quantify situations when the complexity (= number of faces) of the Voronoi diagram and Delaunay triangulation is low or when it is high [Eri01]. For sets $\mathcal{P}$ in general position, the number of all faces of either structure is asymptotically equal to the number of Delaunay simplices or Voronoi vertices, respectively. In this chapter we will thus concentrate on these two.

Average Case Complexity. The average case complexity was considered by Dwyer [Dwy91] who showed that for n independent and identically distributed random point sites chosen uniformly from the unit d-ball the expected number of Delaunay simplices is $\Theta(n)$. It has been conjectured that this bound also holds for any uniform distribution in a convex domain but until now no proofs were given [Dwy91, GN03].

In this chapter we consider the case that the point sites are chosen uniformly from inside an axis-aligned hypercube. Our contribution is the first published proof that shows that the expected complexity of Voronoi diagram and Delaunay triangulation is then linear. The proof is based on a rather technical lemma (Section 6.2) that bounds the intersection volume between the unit hypercube and a randomly chosen ball. How this lemma helps to bound the expected number of Delaunay simplices is shown in the following section.

6.1 Average Case Analysis

In this section we will consider the case that $\mathcal{P}$ is a set of n independent and identically distributed random points chosen uniformly from inside the unit d-hypercube $[0,1]^d$. We will derive a bound on the expected number of Delaunay simplices in $\mathcal{DT}(\mathcal{P})$ by exploiting the following observation. Any simplex that is the convex hull of $d+1$ sites from $\mathcal{P}$ is a Delaunay simplex if and only if its circumball is empty, i.e. contains no other site from $\mathcal{P}$.

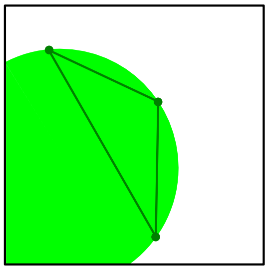

Figure 6.2: For a triangle Δ in $[0,1]^2$, $\mathrm{circumball}(\Delta) \cap [0,1]^2$ is depicted.

We will now consider all subsets of $\mathcal{P}$ with $d+1$ elements and without loss of generality we assume some ordering among these subsets. For all $1 \leq i \leq \binom{n}{d+1}$, let X_i be a random $\{0,1\}$ variable such that $X_i = 1$ if and only if the ith subset of $\mathcal{P}$ has the following property: the convex hull of the ith subset is a Delaunay simplex, i.e. no other point is contained inside this simplex.

Generally, we can then use that

$$\mathbf{E}\,[\text{number of Delaunay simplices in } \mathcal{DT}(\mathcal{P})] \;=\; \mathbf{E}\left[\sum_i X_i\right]$$
$$=\; \sum_i \mathbf{Pr}[\, X_i = 1\,] \;=\; \binom{n}{d+1} \cdot \mathbf{Pr}[\,\mathrm{circumball}(\Delta) \text{ is empty}\,] \quad ,$$

where Δ is the convex hull of $d+1$ independent and identically distributed random points chosen uniformly from $[0,1]^d$. By $\mathrm{circumball}(\Delta)$ we denote the circumball of Δ which is the smallest d-dimensional ball enclosing Δ.

Let $\mathrm{vol}(\mathrm{circumball}(\Delta))$ denote the d-dimensional volume (content) of $\mathrm{circumball}(\Delta)$. Unfortunately, in general it is

$$\mathbf{Pr}[\,\mathrm{circumball}(\Delta) \text{ is empty}\,] \;\neq\; \big(1 - \mathrm{vol}(\mathrm{circumball}(\Delta))\big)^{n-(d+1)}$$

for the following reason. All random point sites are chosen from inside $[0,1]^d$, but some part of $\mathrm{circumball}(\Delta)$ might lie outside of $[0,1]^d$. Of course, the probability for a random point site to be in a part of $\mathrm{circumball}(\Delta)$ that is not in $[0,1]^d$ is equal to 0 and therefore we must not consider these ‘outer’ parts of $\mathrm{circumball}(\Delta)$. Therefore we have to bound the volume of $\mathrm{circumball}(\Delta) \cap [0,1]^d$, which causes the main difficulty in our analysis, see also Figure 6.2.

Fortunately, we can show the following lemma that is crucial for the further analysis. The rather technical proof of this lemma is deferred to Section 6.2.

Lemma 8 *Let Δ be a random d-simplex, i.e. Δ is the convex hull of $d+1$ independent and identically distributed random points chosen uniformly from $[0,1]^d$. For any constant $a \in [0,1]$ it holds that*

$$\mathbf{Pr}\big[\operatorname{vol}(\operatorname{circumball}(\Delta) \cap [0,1]^d) \le a\,\big] \;\le\; \operatorname{const}_d \cdot a^d \;\;,$$

where

$$\operatorname{const}_d \;=\; d^{3d+1} \cdot 2^d \cdot \left(\max\left\{\frac{2^{2d}\cdot\Gamma(1+d/2)}{\pi^{d/2}}, d!\right\}\right)^d \;\le\; (c\cdot d)^{d^2} \;\;,$$

for c some constant factor. Here Γ denotes the Gamma-function where $\Gamma(1/2)=\sqrt{\pi}$, and for all $x \in \mathbb{N}$ it is $\Gamma(x+1)=x!$ and $\Gamma(x+1/2)=(2x)!\cdot\sqrt{\pi}/(x!\cdot 2^{2x})$.

Based on this lemma we will now establish the main theorem of this section.

Theorem 11 *For n independent and identically distributed random points chosen uniformly from $[0,1]^d$ it is*

$$\mathbf{E}\left[\textit{number of Delaunay simplices}\right] \;\le\; n\cdot\operatorname{const}_d\cdot\big((d+4)\cdot 2^{(d+4)\cdot d}+2\big) \;=\; \mathcal{O}(n)$$

where $\operatorname{const}_d \le (c\cdot d)^{d^2}$ is the same constant as in Lemma 8.

Proof. The main idea of this proof is to consider in a first step (classes of) simplices with a ‘large’ circumball. Then it is more likely that another point site lies in the circumball of these simplices and that they are therefore not Delaunay simplices. In a second step we show that the remaining simplices with a ‘small’ circumball are only very few.

Without loss of generality we assume that n is a power of 2. Let us now consider the $\binom{n}{d+1}$ possible simplices that have $d+1$ of the given n random point sites as vertices. For the simplices with ‘large’ circumball we define classes $\mathcal{C}_0,\ldots,\mathcal{C}_{\log(n)-1}$ such that for a simplex Δ it holds that

$$\Delta \in \mathcal{C}_i \quad\iff\quad \frac{1}{2^{i+1}} < \operatorname{vol}(\operatorname{circumball}(\Delta)\cap[0,1]^d) \le \frac{1}{2^i}\;\;.$$

From Lemma 8 it follows immediately that

$$\mathbf{Pr}[\,\Delta\in\mathcal{C}_i\,] \;\le\; \mathbf{Pr}\left[\operatorname{vol}(\operatorname{circumball}(\Delta)\cap[0,1]^d) \le \frac{1}{2^i}\right] \;\le\; \operatorname{const}_d\cdot\left(\frac{1}{2^i}\right)^d \;\;.$$

The probability for a simplex $\Delta\in\mathcal{C}_i$ to be a Delaunay simplex is

$$\begin{aligned}\mathbf{Pr}\big[\operatorname{circumball}(\Delta)\text{ is empty}\;\big|\;\Delta\in\mathcal{C}_i\,\big] \;&\le\; \left(1-\frac{1}{2^{i+1}}\right)^{n-(d+1)}\\ &\le\; \left(\frac{1}{e}\right)^{\frac{n-(d+1)}{2^{i+1}}} \;\le\; \left(\frac{1}{2}\right)^{\frac{n-(d+1)}{2^{i+1}}}\;\;.\end{aligned}$$

Now we can bound the expected number of Delaunay simplices for each class $\mathcal{C}_i$. For $0 \le i \le \log(n) - 1$ we get that

$$\begin{aligned}
&\mathbf{E}\left[\text{number of Delaunay simplices} \in \mathcal{C}_i\right] \\
\le\; & \binom{n}{d+1} \cdot \mathbf{Pr}[\, \Delta \in \mathcal{C}_i \,] \cdot \mathbf{Pr}\left[\, \text{circumball}(\Delta) \text{ is empty} \;\middle|\; \Delta \in \mathcal{C}_i \,\right] \\
\le\; & \binom{n}{d+1} \cdot \text{const}_d \cdot \left(\frac{1}{2^i}\right)^d \cdot \left(\frac{1}{2}\right)^{\frac{n-(d+1)}{2^{i+1}}} \quad .
\end{aligned}$$

The expected number of Delaunay simplices for all classes $\mathcal{C}_0, \dots, \mathcal{C}_{\log(n)-1}$ is then

$$\begin{aligned}
&\sum_{i=0}^{\log(n)-1} \mathbf{E}\left[\text{number of Delaunay simplices} \in \mathcal{C}_i\right] \\
\le\;& \binom{n}{d+1} \cdot \text{const}_d \sum_{i=0}^{\log(n)-1} \left(\frac{1}{2}\right)^{i\cdot d + \frac{n-(d+1)}{2^{i+1}}} \\
=\;& \binom{n}{d+1} \cdot \text{const}_d \sum_{i=0}^{\log(n)-1} \left(\frac{1}{2}\right)^{(\log(n)-(i+1))\cdot d + \frac{n-(d+1)}{2^{\log(n)-(i+1)+1}}} \\
=\;& \binom{n}{d+1} \cdot \text{const}_d \sum_{i=0}^{\log(n)-1} \left(\frac{1}{2}\right)^{\log(n)\cdot d + 2^i \cdot \frac{n-(d+1)}{n} - (i+1)\cdot d} \\
=\;& \binom{n}{d+1} \cdot \text{const}_d \cdot \frac{1}{n^d} \sum_{i=0}^{\log(n)-1} \left(\frac{1}{2}\right)^{2^i \cdot \left(1 - \frac{d+1}{n}\right) - (i+1)\cdot d} \\
\le\;& n \cdot \text{const}_d \cdot ((d+4) \cdot 2^{(d+4)\cdot d} + 1) \quad .
\end{aligned} \tag{6.1}$$

The last step follows immediately if $d + 3 \ge \log(n) - 1$ since

$$\sum_{i=0}^{d+3} \left(\frac{1}{2}\right)^{2^i \cdot \left(1 - \frac{d+1}{n}\right) - (i+1)\cdot d} \quad \le \quad \sum_{i=0}^{d+3} 2^{(i+1)\cdot d} \quad \le \quad (d+4) \cdot 2^{(d+4)\cdot d} \quad .$$

In the case that $d + 3 < \log(n) - 1$ we can bound the rest of the sum as follows

$$\sum_{i=d+4}^{\log(n)-1} \left(\frac{1}{2}\right)^{2^i \cdot \left(1 - \frac{d+1}{n}\right) - (i+1)\cdot d} \quad \le \quad \sum_{i=d+4}^{\log(n)-1} \left(\frac{1}{2}\right)^i \quad \le \quad 1 \quad .$$

Here we exploit that $n \ge 2 \cdot (d+1)$ and $i \ge d + 4$, since then it holds for all d that

$$2^i \cdot \left(1 - \frac{d+1}{n}\right) - (i+1) \cdot d \quad \ge \quad 2^{i-1} - (i+1) \cdot d \quad \ge \quad i \quad .$$

The expected number of remaining simplices with 'small' circumball can be bounded using Lemma 8, too. Let $\mathcal{C}_{re}$ denote the set of simplices such that for a simplex Δ it holds that

$$\Delta \in \mathcal{C}_{re} \quad \iff \quad \mathrm{vol}(\mathrm{circumball}(\Delta) \cap [0,1]^d) \le \frac{1}{n} \ .$$

The expected cardinality of $\mathcal{C}_{re}$ is then

$$\begin{aligned}\mathbf{E}\left[\text{number of simplices } \in \mathcal{C}_{re}\right] &\le \binom{n}{d+1} \cdot \mathbf{Pr}\left[\mathrm{vol}(\mathrm{circumball}(\Delta) \cap [0,1]^d) \le \frac{1}{n}\right] \\ &\le n^{d+1} \cdot \mathrm{const}_d \cdot \frac{1}{n^d} \quad = \quad n \cdot \mathrm{const}_d \ . \end{aligned} \tag{6.2}$$

Since there are so few simplices with small circumball in expectation we do not need to find out how many of these are possibly Delaunay simplices, and we can rather count them all.

Now we can combine the results (6.1) and (6.2) and by linearity of expectation it follows that

$$\begin{aligned}\mathbf{E}\left[\text{number of Delaunay cells}\right] &\le \sum_{i=0}^{\log(n)-1} \mathbf{E}\left[\text{number of Delaunay simplices} \in \mathcal{C}_i\right] \\ &\qquad + \mathbf{E}\left[\text{number of simplices } \in \mathcal{C}_{re}\right] \\ &\le n \cdot \mathrm{const}_d \cdot \left((d+4) \cdot 2^{(d+4)\cdot d} + 2\right) \ ,\end{aligned}$$

which concludes the proof of Theorem 11. □

6.2 Proof of Lemma 8

Let $p_1, \ldots, p_{d+1} \in [0,1]^d$ be $d+1$ independent and identically distributed random points chosen uniformly from the d-dimensional unit hypercube. Let $\Delta = \Delta(p_1, \ldots, p_{d+1})$ be the convex hull of the random points $p_1, \ldots, p_{d+1}$. For abbreviation we denote Δ also as the random simplex.

The volume of $\mathrm{circumball}(\Delta)$ is given by $\mathcal{V}_d \cdot r^d$ where $r = r(\Delta)$ is the radius of $\mathrm{circumball}(\Delta)$ and $\mathcal{V}_d = \pi^{d/2}/\Gamma(1+d/2)$ is the volume of the unit d-ball. We can approximate the radius $r(\Delta)$ and thus the volume of $\mathrm{circumball}(\Delta)$ by the following observation.

Observation 2 *With the just made definitions it holds that*

$$\begin{aligned} 2 \cdot r(\Delta) &\ge \max_{1 \le i < j \le d+1} \|p_i - p_j\|_2 \\ &\ge \max_{1 \le i < j \le d+1} \|p_i - p_j\|_\infty \quad =: \quad \mathrm{maxwidth}(\Delta) \ , \end{aligned}$$

and therefore it is

$$\mathrm{vol}(\mathrm{circumball}(\Delta)) \;\geq\; \frac{1}{2^d} \cdot \mathcal{V}_d \cdot \mathrm{maxwidth}(\Delta)^d \;\; .$$

The term $\mathrm{maxwidth}(\Delta)^d$ can also be interpreted as the volume of a smallest hypercube that contains all the points $p_1, \ldots, p_{d+1}$. The smallest hypercube is of course not uniquely defined but we know that its side-length is exactly $\mathrm{maxwidth}(\Delta)$ by definition. In other words, we approximate the volume of $\mathrm{circumball}(\Delta)$ by the volume of a smallest hypercube containing all the vertices of simplex Δ.

It is convenient to reformulate the random process under which this average case analysis is carried out in the following way. Instead of considering $d+1$ many d-dimensional random variables (= random point sites) we combine the elements of the random variables coordinate-wise, which leads to d sets of $d+1$ random numbers each.

Formally speaking, let us consider the random point sites $p_1, \ldots, p_{d+1} \in [0,1]^d$ where $p_i = (p_i^{(1)}, \ldots, p_i^{(d)})$ for $1 \leq i \leq d+1$. Let then $\mathcal{P}_1, \ldots, \mathcal{P}_d$ be the sets such that $\mathcal{P}_j = \{p_1^{(j)}, \ldots, p_{d+1}^{(j)}\}$ for $1 \leq j \leq d$. Note that both random processes are equivalent.

Furthermore, let

$$\mathrm{width}(\mathcal{P}_j) \;:=\; \max \mathcal{P}_j - \min \mathcal{P}_j$$

denote the maximal distance between two elements in $\mathcal{P}_j$. With the just made definition we can now redefine the variable $\mathrm{maxwidth}$ in the following way as

$$\mathrm{maxwidth}(\mathcal{P}_1, \ldots, \mathcal{P}_d) \;:=\; \max_{1 \leq j \leq d} \mathrm{width}(\mathcal{P}_j) \;\; ,$$

which is consistent with the earlier definition. Indeed, for a set of point sites $p_1, \ldots, p_{d+1}$ it is $\mathrm{maxwidth}(\Delta) = \mathrm{maxwidth}(\mathcal{P}_1, \ldots, \mathcal{P}_d)$ where Δ is the convex hull of the point sites. When the set of point sites is clear from the context we use only $\mathrm{maxwidth}$.

Since we actually want to find a lower bound on the volume of $\mathrm{circumball}(\Delta) \cap [0,1]^d$ we consider the (smallest) hypercube containing all point sites that has minimal volume when intersected with $[0,1]^d$. In order to have a minimal intersection volume the hypercube might jut out of $[0,1]^d$ in some dimensions. In Figure 6.3 two examples are given.

Therefore, we introduce the variable value that indicates how much each dimension contributes to the volume of the minimal intersection between a smallest hypercube containing all the point sites and $[0,1]^d$. If for a fixed dimension the coordinates of all point sites are close to 0 (or 1) then this dimension might contribute less than $\mathrm{maxwidth}$ to the volume, namely only the distance of the maximal coordinate to 0 (or the minimal coordinate to 1).

Formally, we define now the value of set $\mathcal{P}_j$ to be

$$\mathrm{value}(\mathcal{P}_j) := \begin{cases} \mathrm{maxwidth}(\mathcal{P}_1, \ldots, \mathcal{P}_d) & \textbf{if } \max \mathcal{P}_j - \mathrm{maxwidth} \geq 0 \\ & \text{and } \min \mathcal{P}_j + \mathrm{maxwidth} \leq 1 \\ \min \{\max \mathcal{P}_j, 1 - \min \mathcal{P}_j\} & \textbf{else} \end{cases} \;\; .$$

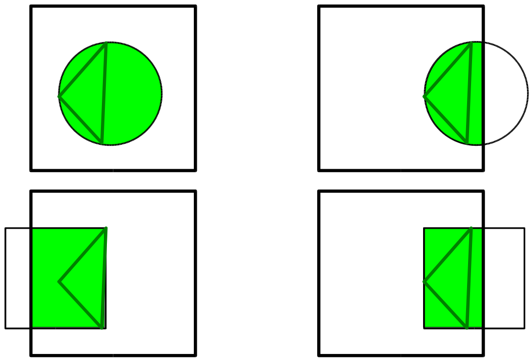

Figure 6.3: Two random triangles in $[0,1]^2$ and the smallest enclosing sphere and smallest enclosing square with minimal intersection area are depicted. In the first example the x_1-coordinates of the three triangle's vertices lie close to 0, in the second example close to 1, and so the squares jut out of $[0,1]^2$ accordingly.

With these definitions we can formulate the following lemma.

Lemma 9 *With the just made definition it holds that*

$$\mathrm{vol}(\mathrm{circumball}(\Delta) \cap [0,1]^d) \;\geq\; \min\left\{ \left(\frac{1}{2}\right)^{2d} \cdot \mathcal{V}_d, \frac{1}{d!} \right\} \cdot \prod_{j=1}^{d} \mathrm{value}(\mathcal{P}_j) \enspace .$$

Proof. In order to prove the lemma we will consider two cases, namely that the center of $\mathrm{circumball}(\Delta)$ lies inside $[0,1]^d$ (case i) and that it does not (case ii).

(i) If the center of $\mathrm{circumball}(\Delta)$ lies inside $[0,1]^d$ it follows immediately that at least a fraction of $(1/2)^d$ of $\mathrm{circumball}(\Delta)$ lies inside $[0,1]^d$. Together with Observation 2 we get that

$$\begin{aligned}
\mathrm{vol}(\mathrm{circumball}(\Delta) \cap [0,1]^d) \;&\geq\; \left(\frac{1}{2}\right)^{d} \cdot \mathrm{vol}\big(\mathrm{circumball}(\Delta)\big) \\
&\geq\; \left(\frac{1}{2}\right)^{d} \cdot \mathcal{V}_d \cdot \left(\frac{\mathrm{maxwidth}(\Delta)}{2}\right)^{d} \\
&\geq\; \left(\frac{1}{2}\right)^{2d} \cdot \mathcal{V}_d \cdot \prod_{j=1}^{d} \mathrm{value}(\mathcal{P}_j) \enspace .
\end{aligned}$$

(ii) Let $\mathcal{B} = \mathcal{B}(\text{circumball}(\Delta) \cap [0,1]^d)$ be the smallest axis parallel box containing $\text{circumball}(\Delta) \cap [0,1]^d$ and let $b_1, \ldots, b_d$ be the width of box $\mathcal{B}$ in each dimension, i.e. $\text{vol}(\mathcal{B}) = \prod_{i=1}^d b_i$. In a first step we want to show that

$$d! \cdot \text{vol}(\text{circumball}(\Delta) \cap [0,1]^d) \;\geq\; \text{vol}(\mathcal{B}) \;. \tag{6.3}$$

We will now distinguish two cases, namely that $\text{circumball}(\Delta) \cap [0,1]^d$ contains at least $d+1$ of the 2^d vertices of $\mathcal{B}$ and that $\text{circumball}(\Delta) \cap [0,1]^d$ contains less than $d+1$ vertices of $\mathcal{B}$.

So, consider the case that $\text{circumball}(\Delta) \cap [0,1]^d$ contains at least $d+1$ vertices of $\mathcal{B}$. Let $e_1, \ldots, e_d$ be the unit vectors, i.e. the j-th entry in e_j is a one and all other entries are 0. Without loss of generality we assume that $\bar{0} = (0, \ldots, 0)$ is a vertex of $\mathcal{B}$ and that $\bar{0}, b_1 \cdot e_1, \ldots, b_d \cdot e_d$ are the vertices of $\mathcal{B}$ that lie also in $\text{circumball}(\Delta) \cap [0,1]^d$. The convex hull of these vertices, namely the simplex $\mathcal{S} := \text{conv}(\bar{0}, b_1 \cdot e_1, \ldots, b_d \cdot e_d)$ is then also completely contained in $\text{circumball}(\Delta) \cap [0,1]^d$ and therefore it is $\text{vol}(\text{circumball}(\Delta) \cap [0,1]^d) \geq \text{vol}(\mathcal{S})$. From geometry [HRZ97] it is known that

$$\text{vol}(\mathcal{S}) \;=\; \frac{1}{d!} \cdot \text{vol}(\mathcal{B})$$

and therefore (6.3) follows.

Let us now consider the case that $\text{circumball}(\Delta) \cap [0,1]^d$ contains less than $d+1$ vertices of $\mathcal{B}$. Our goal is now to place a collection of simplices inside $\text{circumball}(\Delta) \cap [0,1]^d$ such that their sum of volumes is also equal to $1/d! \cdot \text{vol}(\mathcal{B})$.
Let $\mathbf{c} := (c_1, \ldots, c_d)$ be the center of $\text{circumball}(\Delta)$ and consider the d hyper-planes $x_1 = c_1, \ldots, x_d = c_d$ that subdivide $\text{circumball}(\Delta)$ into 2^d equal parts. These hyper-planes subdivide also $\mathcal{B}$ into at most 2^{d-1} smaller boxes. (Remember, the center $\mathbf{c}$ is not contained in $[0,1]^d$, therefore we cannot get more than 2^{d-1} of them.) For each of the smaller boxes we will find at least $d+1$ vertices that lie also in $\text{circumball}(\Delta) \cap [0,1]^d$. As before we can take their convex hull to construct the desired simplices, see also Figure 6.4. It follows immediately that their sum of volumes is equal to $1/d! \cdot \text{vol}(\mathcal{B})$. Thus (6.3) is shown.

In the next step we want to show that

$$\text{vol}(\mathcal{B}) \;=\; \prod_{i=1}^{d} b_i \;\geq\; \prod_{j=1}^{d} \text{value}(\mathcal{P}_j) \tag{6.4}$$

which follows immediately with the following Claim 5, though we defer its proof to the end of the proof of Lemma 9.

Claim 5 *It holds that* $\quad \text{value}(\mathcal{P}_j) \leq b_j \quad$ *for all* $1 \leq j \leq d$.

Figure 6.4: Depicted are the two possible cases in $\mathbb{R}^2$. In the first case three vertices of box $\mathcal{B}$ are contained in $\mathrm{circumball}(\Delta) \cap [0,1]^2$ and their convex hull, the simplex $\mathcal{S}$ lies also completely in $\mathrm{circumball}(\Delta) \cap [0,1]^2$. In the second case only two vertices of box $\mathcal{B}$ lie in $\mathrm{circumball}(\Delta) \cap [0,1]^2$. The hyperplane $x_1 = c_1$ subdivides $\mathcal{B}$ into two smaller boxes that have both three vertices contained in $\mathrm{circumball}(\Delta) \cap [0,1]^2$. The convex hull of these give the two simplices $\mathcal{S}_1$ and $\mathcal{S}_2$.

By combining now (6.3) and (6.4) we see that

$$d! \cdot \mathrm{vol}(\mathrm{circumball}(\Delta) \cap [0,1]^d) \quad \geq \quad \mathrm{vol}(\mathcal{B}) \quad \geq \quad \prod_{j=1}^{d} \mathrm{value}(\mathcal{P}_j)$$

which concludes the case (ii) and therefore the proof of Lemma 9.

$\square$

It remains to show that Claim 5 holds.

Proof of Claim 5. Again we will consider two cases, namely that $\mathrm{circumball}(\Delta) \cap [0,1]^d$ and therefore $\mathcal{B}$ contains at least one vertex of $[0,1]^d$ (case a) and that it does not contain a vertex of $[0,1]^d$ (case b).

(a) Let us assume that $\mathrm{circumball}(\Delta) \cap [0,1]^d$ contains a vertex of $[0,1]^d$, say $\mathbf{v} = (v_1, \ldots, v_d)$ where $v_j \in \{0,1\}$ for $1 \leq j \leq d$. Since $\mathcal{B}$ contains all vertices of the simplex Δ we can conclude that either $\max \mathcal{P}_j \leq b_j$ if $v_j = 0$ or that $1 - \min \mathcal{P}_j \leq b_j$ if $v_j = 1$ for $1 \leq j \leq d$.
If $\mathrm{value}(\mathcal{P}_j) = \min\{\max \mathcal{P}_j, 1 - \min \mathcal{P}_j\}$ then the claim follows immediately. If $\mathrm{value}(\mathcal{P}_j) = \mathrm{maxwidth}$ then $\max \mathcal{P}_j - \mathrm{maxwidth} \geq 0$ and $\min \mathcal{P}_j + \mathrm{maxwidth} \leq 1$. It follows that $\mathrm{maxwidth} \leq b_j$ and therefore $\mathrm{value}(\mathcal{P}_j) \leq b_j$ for $1 \leq j \leq d$.

(b) Let us now assume that $\mathrm{circumball}(\Delta) \cap [0,1]^d$ contains no vertex of $[0,1]^d$. Since the center of $\mathrm{circumball}(\Delta) \notin [0,1]^d$ we can assume that there is some number k, $1 \le k < d$ such that up to ordering

$$b_1 = \cdots = b_k > b_{k+1} > \cdots > b_d \enspace .$$

In other words, there is a k-dimensional face $\mathcal{F}$ of $[0,1]^d$ such that the intersection of $[0,1]^d$ and $\mathcal{F}$ is a k-ball. (If we would 'add' the next dimension we would obtain a $(k+1)$-dimensional spherical cap.)

Since the box $\mathcal{B}$ contains all vertices of the simplex Δ it follows immediately that $b_1 = \cdots = b_k \ge \mathrm{maxwidth}$. Therefore we can conclude that $\mathrm{value}(\mathcal{P}_j) \le b_j$ for $1 \le j \le k$.

Now consider some number $j > k$. Let $\mathcal{F}_{x_j=0}$ be the facet (= $(d-1)$-dimensional face) of $[0,1]^d$ that is contained in the hyperplane $x_j = 0$ and let $\mathcal{F}_{x_j=1}$ be the facet of $[0,1]^d$ that is contained in the hyperplane $x_j = 1$. Since $b_j < b_1$ the box $\mathcal{B}$ 'touches' either $\mathcal{F}_{x_j=1}$ or $\mathcal{F}_{x_j=0}$, i.e. facets of $\mathcal{B}$ are either contained in $\mathcal{F}_{x_j=1}$ or in $\mathcal{F}_{x_j=0}$. And again, since $\mathcal{B}$ also contains all vertices of the simplex Δ it follows that either $\max \mathcal{P}_j \le b_j$ or that $1 - \min \mathcal{P}_j \le b_j$. Analogously to case (a) it follows that $\mathrm{maxwidth} \le b_j$ and therefore $\mathrm{value}(\mathcal{P}_j) \le b_j$ for $k < j \le d$.

$\square_{\text{Claim 5}}$

We can utilize Lemma 9 in the following way. We will show that for any value $a \in [0,1]$ it is

$$\mathbf{Pr}\left[\prod_{j=1}^{d} \mathrm{value}(\mathcal{P}_j) \le a\right] \;\le\; d^{3d+1} \cdot 2^d \cdot a^d \enspace . \tag{6.5}$$

We know that $\mathrm{vol}(\mathrm{circumball}(\Delta) \cap [0,1]^d) \ge \min\left\{(1/2^{2d}) \cdot \mathcal{V}_d, \frac{1}{d!}\right\} \cdot \prod_{j=1}^{d} \mathrm{value}(\mathcal{P}_j)$ by Lemma 9. So we can conclude that for any value $a \in [0,1]$ it is

$$\mathbf{Pr}\left[\,\mathrm{vol}(\mathrm{circumball}(\Delta) \cap [0,1]^d) \le a\,\right] \le \mathrm{const}_d \cdot a^d \enspace ,$$

where $\mathrm{const}_d = d^{3d+1} \cdot 2^d \cdot \left(\max\left\{2^{2d}/\mathcal{V}_d, d!\right\}\right)^d$. Thus Lemma 8 is also shown.

In order to show (6.5) we will now establish two lemmas. Lemma 10 will cover the case that $\mathrm{maxwidth}(\mathcal{P}_1, \ldots, \mathcal{P}_d)$ is at most $\sqrt[d]{a}$, while Lemma 11 will cover the case that $\mathrm{maxwidth}(\mathcal{P}_1, \ldots, \mathcal{P}_d)$ is larger than $\sqrt[d]{a}$.

Before we proceed let us briefly recall the (reformulated) random process. Instead of random point sites $p_1, \ldots, p_{d+1}$ we consider d sets $\mathcal{P}_1, \ldots, \mathcal{P}_d$ each of $d+1$ independent and identically distributed random numbers chosen uniformly from the interval $[0,1]$.

Lemma 10 *For any value $a \in [0,1]$ it holds that*

$$\mathbf{Pr}\left[\prod_{j=1}^{d} \mathrm{value}(\mathcal{P}_j) \le a \;\wedge\; \mathrm{maxwidth}(\mathcal{P}_1, \ldots, \mathcal{P}_d) \le \sqrt[d]{a}\right] \;\le\; (d+1)^d \cdot a^d \enspace .$$

Proof. First of all we notice that from $\mathrm{maxwidth}(\mathcal{P}_1, \ldots, \mathcal{P}_d) \leq \sqrt[d]{a}$ it follows immediately that $\prod_{j=1}^{d} \mathrm{value}(\mathcal{P}_j) \leq a$ and therefore it is

$$\mathbf{Pr}\left[\prod_{j=1}^{d} \mathrm{value}(\mathcal{P}_j) \leq a \;\wedge\; \mathrm{maxwidth}(\mathcal{P}_1, \ldots, \mathcal{P}_d) \leq \sqrt[d]{a}\right]$$
$$= \mathbf{Pr}\big[\,\mathrm{maxwidth}(\mathcal{P}_1, \ldots, \mathcal{P}_d) \leq \sqrt[d]{a}\,\big] \quad .$$

Furthermore, it suffices to bound only the probability that $\mathrm{width}(\mathcal{P}_j) \leq \sqrt[d]{a}$, because it also holds that

$$\mathbf{Pr}\big[\,\mathrm{maxwidth}(\mathcal{P}_1, \ldots, \mathcal{P}_d) \leq \sqrt[d]{a}\,\big] = \prod_{j=1}^{d} \mathbf{Pr}\big[\,\mathrm{width}(\mathcal{P}_j) \leq \sqrt[d]{a}\,\big] \quad .$$

The idea is now to fix for set $\mathcal{P}_j$ the two elements[4] that attain the maximal distance, i.e. the elements $\max \mathcal{P}_j$ and $\min \mathcal{P}_j$. Now we can bound the probability that their distance does not exceed $\sqrt[d]{a}$ and that the remaining $d-1$ elements in $\mathcal{P}_j$ have values between $\max \mathcal{P}_j$ and $\min \mathcal{P}_j$. It is

$$\mathbf{Pr}\big[\,\mathrm{width}(\mathcal{P}_j) \leq \sqrt[d]{a}\,\big] = (d+1) \cdot d \cdot \int_0^1 \int_{\max\{0, y - \sqrt[d]{a}\}}^{y} (y-x)^{d-1}\, \mathrm{d}x\, \mathrm{d}y \tag{6.6}$$

where the outer integral denotes the range of element $\max \mathcal{P}_j (= y)$ and the inner integral the range of element $\min \mathcal{P}_j (= x)$. The integration boundaries assure that the distance of x and y is at most $\sqrt[d]{a}$. The integrand $(y-x)^{d-1}$ denotes exactly the probability that all remaining $d-1$ elements of $\mathcal{P}_j$ lie between y and x. The fore-factor $(d+1) \cdot d$ comes from fixing the maximal and minimal element in $\mathcal{P}_j$.

In order to solve integral (6.6) we will split it up in the following way to remove the maximum expression from the integration boundary of the inner integral

$$\begin{aligned}
&\int_0^1 \int_{\max\{0, y - \sqrt[d]{a}\}}^{y} (y-x)^{d-1}\, \mathrm{d}x\, \mathrm{d}y \\
&= \int_0^{\sqrt[d]{a}} \int_0^{y} (y-x)^{d-1}\, \mathrm{d}x\, \mathrm{d}y + \int_{\sqrt[d]{a}}^{1} \int_{y-\sqrt[d]{a}}^{y} (y-x)^{d-1}\, \mathrm{d}x\, \mathrm{d}y \\
&= \frac{1}{d} \cdot \left(\int_0^{\sqrt[d]{a}} y^d\, \mathrm{d}y + \int_{\sqrt[d]{a}}^{1} a\, \mathrm{d}y \right) = \frac{1}{d} \cdot \left(\frac{1}{d+1} \cdot a^{\frac{d+1}{d}} + a - a^{\frac{d+1}{d}} \right) \\
&= \frac{1}{d} \cdot a \cdot \left(1 - \left(1 - \frac{1}{d+1} \right) \cdot a^{1/d} \right) \leq \frac{1}{d} \cdot a \quad .
\end{aligned}$$

[4]Note that all elements are distinct with probability 1.

It follows that

$$\mathbf{Pr}\big[\,\mathrm{width}(\mathcal{P}_j) \le \sqrt[d]{a}\,\big] \le (d+1)\cdot a \quad \Rightarrow \quad \mathbf{Pr}\big[\,\mathrm{maxwidth} \le \sqrt[d]{a}\,\big] \le (d+1)^d \cdot a^d \ ,$$

which concludes the proof of Lemma 10.

□

Lemma 11 *For any value of* $a \in [0,1]$ *it holds that*

$$\mathbf{Pr}\left[\prod_{j=1}^{d} \mathrm{value}(\mathcal{P}_j) \le a \ \wedge \ \mathrm{maxwidth}(\mathcal{P}_1,\ldots,\mathcal{P}_d) > \sqrt[d]{a}\right] \ \le \ d^{3d+1}\cdot 2^d \cdot a^d \ .$$

Proof. Without loss of generality, we assume some ordering on the sets $\mathcal{P}_1,\ldots,\mathcal{P}_d$ as described now. Let the first set $\mathcal{P}_1$ attain the maximal width, i.e. $\mathrm{width}(\mathcal{P}_1) \ge \mathrm{width}(\mathcal{P}_j)$ for $2 \le j \le d$. It follows that $\mathrm{maxwidth}(\mathcal{P}_1,\ldots,\mathcal{P}_d) := \mathrm{width}(\mathcal{P}_1)$.

Before we proceed let us briefly recall the definition of the function value, i.e.

$$\mathrm{value}(\mathcal{P}_j) := \begin{cases} \mathrm{maxwidth} & \textbf{if } \max \mathcal{P}_j - \mathrm{maxwidth} \ge 0 \\ & \text{and } \min \mathcal{P}_j + \mathrm{maxwidth} \le 1 \\ \min\{\max \mathcal{P}_j, 1 - \min \mathcal{P}_j\} & \textbf{else} \ . \end{cases}$$

For the remaining sets $\mathcal{P}_2,\ldots,\mathcal{P}_d$ let d_f denote the number of sets for which the **if**-case is true (the elements of these sets lie *far away* from 0 or 1) and let d_c denote the number of sets for which the **else**-case is true (the elements of these sets lie *close* to 0 or 1), such that $d = 1 + d_f + d_c$. Furthermore, let the sets be ordered such that the **if**-case is true for the sets $\mathcal{P}_2,\ldots,\mathcal{P}_{d_f+1}$ and that the **else**-case is true for the sets $\mathcal{P}_{d_f+2},\ldots,\mathcal{P}_d$. Note that for given d_f and d_c there are exactly $d!/(1!\cdot d_f!\cdot d_c!)$ ways to fix the described ordering for the sets $\mathcal{P}_1,\ldots,\mathcal{P}_d$.

We will summarize the considerations just made in the following definition of three conditions. Let

$$\mathrm{A} \ : \ \mathrm{width}(\mathcal{P}_1) > \sqrt[d]{a}$$

$$\mathrm{B}(d_f) \ : \ \bigwedge_{2\le j\le d_f+1} \Big(\mathrm{width}(\mathcal{P}_j) \le \mathrm{width}(\mathcal{P}_1) \ \wedge \ \big(\max \mathcal{P}_j \ge \mathrm{width}(\mathcal{P}_1) \ \wedge \ 1 - \min \mathcal{P}_j \ge \mathrm{width}(\mathcal{P}_1)\big)\Big)$$

$$\mathrm{C}(d_f) \ : \ \bigwedge_{d_f+2\le j\le d} \Big(\mathrm{width}(\mathcal{P}_j) < \mathrm{width}(\mathcal{P}_1) \ \wedge \ \big(\max \mathcal{P}_j < \mathrm{width}(\mathcal{P}_1) \ \vee \ 1 - \min \mathcal{P}_j < \mathrm{width}(\mathcal{P}_1)\big)\Big) \ .$$

For condition $\mathrm{C}(d_f)$, if $(\max \mathcal{P}_j < \mathrm{width}(\mathcal{P}_1) \ \vee \ \min \mathcal{P}_j < 1 - \mathrm{width}(\mathcal{P}_1))$ is fulfilled it follows immediately that $(\mathrm{width}(\mathcal{P}_j) < \mathrm{width}(\mathcal{P}_1))$ is also fulfilled.

We can now rewrite

$$\begin{aligned} &\mathbf{Pr}\left[\prod_{j=1}^{d} \mathrm{value}(\mathcal{P}_j) \leq a \ \wedge \ \mathrm{maxwidth}(\mathcal{P}_1, \ldots, \mathcal{P}_d) > \sqrt[d]{a}\right] \\ &= \sum_{d_f=0}^{d-2} \frac{d!}{1! \cdot d_f! \cdot d_c!} \cdot \mathbf{Pr}\left[\prod_{j=1}^{d} \mathrm{value}(\mathcal{P}_j) \leq a \ \wedge \ \mathrm{A} \wedge \mathrm{B}(d_f) \wedge \mathrm{C}(d_f)\right] \ , \end{aligned}$$

where $d_c = d - d_f - 1$. The sum goes only up to $d_f = d - 2$ since d_c has to be greater than 0 in order to make it possible that $\prod_{j=1}^{d} \mathrm{value}(\mathcal{P}_j) \leq a$, since otherwise it is $\mathrm{value}(\mathcal{P}_j) > \sqrt[d]{a}$ for all j. Furthermore, by condition $\mathrm{B}(d_f)$ we know that $\mathrm{value}(\mathcal{P}_j) = \mathrm{width}(\mathcal{P}_1)$ for $1 \leq j \leq d_f + 1$ and we can conclude that

$$\begin{aligned} &\mathbf{Pr}\left[\prod_{j=1}^{d} \mathrm{value}(\mathcal{P}_j) \leq a \ \wedge \ \mathrm{A} \wedge \mathrm{B}(d_f) \wedge \mathrm{C}(d_f)\right] \\ &= \mathbf{Pr}\left[\prod_{j=d_f+2}^{d} \mathrm{value}(\mathcal{P}_j) \leq \frac{a}{\mathrm{width}(\mathcal{P}_1)^{d_f+1}} \ \wedge \ \mathrm{A} \wedge \mathrm{B}(d_f) \wedge \mathrm{C}(d_f)\right] \ . \end{aligned} \tag{6.7}$$

Our goal is now to find an integral expression for (6.7). It seems useful to consider first the three conditions $\mathrm{A}, \mathrm{B}(d_f)$ and $\mathrm{C}(d_f)$ separately. While conditions $\mathrm{B}(d_f)$ and $\mathrm{C}(d_f)$ are mutually independent, they both depend on condition A.

So in a first step we will find an expression for the probability that condition A holds, i.e. $\mathrm{width}(\mathcal{P}_1) > \sqrt[d]{a}$. As in the proof of Lemma 10 we fix the two elements in set $\mathcal{P}_1$ that have maximal distance, i.e. $\max \mathcal{P}_1 (= y)$ and $\min \mathcal{P}_1 (= x)$. Then we can write

$$\mathbf{Pr}[\,\mathrm{A}\,] \ = \ (d+1) \cdot d \cdot \int_{\sqrt[d]{a}}^{1} \int_{0}^{y - \sqrt[d]{a}} (y-x)^{d-1} \, \mathrm{d}x \, \mathrm{d}y \ . \tag{6.8}$$

In a second next step we will bound the probability that condition $\mathrm{B}(d_f)$ is true. By (6.8), we assume that $\mathrm{width}(\mathcal{P}_1) = (y - x)$. It is then

$$\begin{aligned} \mathbf{Pr}[\,\mathrm{B}(d_f)\,] \ &\leq \ \prod_{j=2}^{d_f+1} \mathbf{Pr}[\,\mathrm{width}(\mathcal{P}_j) \leq (y-x)\,] \\ &\leq \ \big((d+1) \cdot (y-x)^d\big)^{d_f} \ , \end{aligned}$$

where the last step follows analogously to the proof of Lemma 10 on page 79.

In a third step we will consider the probability for condition $\mathrm{C}(d_f)$ to be true. The condition $\mathrm{C}(d_f)$ holds if for all sets $\mathcal{P}_j$, $j \in \{d_f + 2, \ldots, d\}$, either $\max \mathcal{P}_j < \mathrm{width}(\mathcal{P}_1)$ or $1 - \min \mathcal{P}_j < \mathrm{width}(\mathcal{P}_1)$ is true.

We define now index sets $\mathcal{I}$ and $\mathcal{I}'$, where $\mathcal{I} \cup \mathcal{I}' = \{d_f + 2, \ldots, d\}$ and $\mathcal{I} \cap \mathcal{I}' = \emptyset$, such that for $\mathcal{P}_j$, $j \in \mathcal{I}$ the first case and for $\mathcal{P}_j$, $j \in \mathcal{I}'$ the second case is true. Then we can write

$$\begin{aligned}
\mathbf{Pr}[\,\mathrm{C}(d_f)\,] &= \prod_{j=d_f+2}^{d} \mathbf{Pr}[\,\max \mathcal{P}_j < \mathrm{width}(\mathcal{P}_1) \ \vee \ 1 - \min \mathcal{P}_j < \mathrm{width}(\mathcal{P}_1)\,] \\
&\leq \sum_{\mathcal{I},\mathcal{I}'} \left(\prod_{j\in\mathcal{I}} \mathbf{Pr}[\,\max \mathcal{P}_j < (y-x)\,] \cdot \prod_{j\in\mathcal{I}'} \mathbf{Pr}[\,1 - \min \mathcal{P}_j < (y-x)\,] \right) \\
&= 2^{d_c} \cdot \prod_{j=d_f+2}^{d} \mathbf{Pr}[\,\max \mathcal{P}_j < (y-x)\,] \quad ,
\end{aligned}$$

where the last step follows since $\mathbf{Pr}[\,1 - \min \mathcal{P}_j < (y-x)\,] = \mathbf{Pr}[\,\max \mathcal{P}_j < (y-x)\,]$.

We can now conclude for (6.7) that

$$\begin{aligned}
&\mathbf{Pr}\left[\prod_{j=d_f+2}^{d} \mathrm{value}(\mathcal{P}_j) \leq \frac{a}{\mathrm{width}(\mathcal{P}_1)^{d_f+1}} \ \wedge \ \mathrm{A} \wedge \mathrm{B}(d_f) \wedge \mathrm{C}(d_f) \right] \\
&= (d+1) \cdot d \cdot \int_{\sqrt[d]{a}}^{1} \int_{0}^{y-\sqrt[d]{a}} (y-x)^{d-1} \cdot \mathbf{Pr}[\,\mathrm{B}(d_f)\,] \cdot \\
&\qquad \mathbf{Pr}\left[\prod_{j=d_f+2}^{d} \mathrm{value}(\mathcal{P}_j) \leq \frac{a}{(y-x)^{d_f+1}} \ \wedge \ \mathrm{C}(d_f) \right] \mathrm{d}x \, \mathrm{d}y \\
&\leq (d+1)^{d_f+1} \cdot d \cdot 2^{d_c} \cdot \int_{\sqrt[d]{a}}^{1} \int_{0}^{y-\sqrt[d]{a}} (y-x)^{d-1+d_f\cdot d} \cdot \\
&\qquad \mathbf{Pr}\left[\prod_{j=d_f+2}^{d} \max \mathcal{P}_j \leq \bar{a} \ \wedge \bigwedge_{d_f+2\leq j\leq d} \max \mathcal{P}_j < (y-x) \right] \mathrm{d}x \, \mathrm{d}y
\end{aligned}$$

where $\bar{a} := a/(y-x)^{d_f+1}$. What remains to be shown is captured by the following claim.

Claim 6 *For any value of* $\bar{a} \in [0,1]$ *it holds that*

$$\begin{aligned}
&\mathbf{Pr}\left[\prod_{j=d_f+2}^{d} \max \mathcal{P}_j \leq \bar{a} \ \wedge \bigwedge_{d_f+2\leq j\leq d} \max \mathcal{P}_j < (y-x) \right] \\
&\leq \ 2 \cdot (d+1)^{d_c-1} \cdot (d_c-1)^{d_c-1} \cdot \bar{a}^d \cdot (y-x)^{d_c} \quad \leq \quad \mathcal{O}(d) \cdot \bar{a}^d \cdot (y-x)^{d_c} \quad .
\end{aligned}$$

The proof of Claim 6 is deferred to the end of this subsection.

From Claim 6 it follows that

$$\begin{aligned}
(6.7) \quad \le \quad & (d+1)^{d_f+1} \cdot d \cdot 2^{d_c} \cdot 2 \cdot (d+1)^{d_c-1} \cdot (d_c-1)^{d_c-1} \cdot \\
& \qquad \int_{\sqrt[d]{a}}^{1} \int_{0}^{y-\sqrt[d]{a}} \bar{a}^d \cdot (y-x)^{d-1+d_f \cdot d+d_c} \, \mathrm{d}x \, \mathrm{d}y \\
= \quad & (d+1)^{d-1} \cdot d \cdot 2^{d_c+1} \cdot (d_c-1)^{d_c-1} \cdot a^d \cdot \int_{\sqrt[d]{a}}^{1} \int_{0}^{y-\sqrt[d]{a}} (y-x)^{d_c-1} \, \mathrm{d}x \, \mathrm{d}y \ .
\end{aligned}$$

Now for the integral in this last expression we get

$$\begin{aligned}
\int_{\sqrt[d]{a}}^{1} \int_{0}^{y-\sqrt[d]{a}} (y-x)^{d_c-1} \, \mathrm{d}x \, \mathrm{d}y \quad &= \quad \frac{1}{d_c} \cdot \int_{\sqrt[d]{a}}^{1} y^{d_c} - \left(\sqrt[d]{a}\right)^{d_c} \, \mathrm{d}y \\
= \ \frac{1}{d_c} \cdot \left(\frac{1}{d_c+1} - \left(\sqrt[d]{a}\right)^{d_c} - \left(\frac{1}{d_c+1} \cdot \left(\sqrt[d]{a}\right)^{d_c+1} - \left(\sqrt[d]{a}\right)^{d_c+1} \right) \right) & \\
= \ \frac{1}{d_c} \cdot \left(\frac{1}{d_c+1} - \left(\sqrt[d]{a}\right)^{d_c} \cdot \underbrace{\left(1 - \left(1 - \frac{1}{d_c+1} \right) \cdot \sqrt[d]{a} \right)}_{>0} \right) \quad &\le \quad \frac{1}{d_c \cdot (d_c+1)} \ .
\end{aligned}$$

And finally it follows that

$$\begin{aligned}
& \mathbf{Pr}\left[\prod_{j=1}^{d} \mathrm{value}(\mathcal{P}_j) \le a \ \wedge \ \mathrm{maxwidth}(\mathcal{P}_1, \ldots, \mathcal{P}_d) > \sqrt[d]{a} \right] \\
& \quad \le \quad \sum_{d_f=0}^{d-2} \frac{d!}{1! \cdot d_f! \cdot d_c!} \cdot (d+1)^{d-1} \cdot d \cdot 2^{d_c+1} \cdot (d_c-1)^{d_c-1} \cdot \frac{1}{d_c \cdot (d_c+1)} \cdot a^d \\
& \quad \le \quad d^{3d+1} \cdot 2^d \cdot a^d
\end{aligned}$$

which concludes the proof of Lemma 11. □

From Lemma 10 and Lemma 11 it follows that (6.2) holds for

$$\mathrm{const}_d \ = \ d^{3d+1} \cdot 2^d \cdot \left(\max\left\{ \frac{2^{2d}}{\mathcal{V}_d}, d! \right\} \right)^d \ \le \ (c \cdot d)^{d^2} \ ,$$

for some constant factor c, and thus Lemma 8 is shown.

It remains to prove Claim 6 from the previous page. Recall that $\bar{a} = a/(y-x)^{d_f+1}$ and that $d = d_f + d_c + 1$.

Claim 6 *For any value of $\bar{a} \in [0,1]$ it holds that*

$$\mathbf{Pr}\left[\prod_{j=d_f+2}^{d} \max \mathcal{P}_j \leq \bar{a} \ \wedge \bigwedge_{d_f+2 \leq j \leq d} \max \mathcal{P}_j < (y-x)\right]$$
$$\leq \ 2 \cdot (d+1)^{d_c-1} \cdot (d_c-1)^{d_c-1} \cdot \bar{a}^d \cdot (y-x)^{d_c} \quad \leq \quad \mathcal{O}(d) \cdot \bar{a}^d \cdot (y-x)^{d_c} \ .$$

Proof of Claim 6. For ease of notation the enumeration of the sets $\mathcal{P}_{d_f+2}, \ldots, \mathcal{P}_d$ is changed to $\mathcal{P}_1, \ldots, \mathcal{P}_{d_c}$. For $1 \leq j \leq d_c - 1$ it is then

$$\mathbf{Pr}[\max \mathcal{P}_j < (y-x)] \ = \ (d+1) \cdot \int_0^{(y-x)} z_j^d \, \mathrm{d}z_j$$

where the fore-factor $(d+1)$ comes from fixing the maximal element $\max \mathcal{P}_j (= z_j)$. The integrant z_j^d denotes exactly the probability that the remaining d elements in $\mathcal{P}_j$ are at most z_j. By the integration boundaries it follows that all elements in $\mathcal{P}_j$ are smaller than $(y-x)$.

Now for the set $\mathcal{P}_{d_c}$, in order to guarantee that $\prod_{j=1}^{d_c} \max \mathcal{P}_j \leq \bar{a}$ it is necessary that $\max \mathcal{P}_{d_c}$ does not exceed $\bar{a}/z_1 \cdots z_{d_c-1}$ (and also not $(y-x)$). It is

$$\begin{aligned}
&\mathbf{Pr}\left[\max \mathcal{P}_{d_c} < \min\left\{(y-x), \frac{\bar{a}}{z_1 \cdots z_{d_c-1}}\right\}\right] \\
&= \ (d+1) \cdot \int_0^{\min\left\{(y-x), \frac{\bar{a}}{z_1 \cdots z_{d_c-1}}\right\}} z_{d_c}^d \, \mathrm{d}z_{d_c} \\
&= \ \left(\min\left\{(y-x), \frac{\bar{a}}{z_1 \cdots z_{d_c-1}}\right\}\right)^{d+1} \quad =: \quad \mathcal{K} \ .
\end{aligned}$$

From the independence of the events $\max \mathcal{P}_j \leq \bar{a}$ it follows that

$$\begin{aligned}
&\mathbf{Pr}\left[\prod_{j=1}^{d_c} \max \mathcal{P}_j \leq \bar{a} \;\middle|\; \max \mathcal{P}_j < (y-x), 1 \leq j \leq d_c\right] \\
&= \ (d+1)^{d_c-1} \cdot \int_0^{(y-x)} \cdots \int_0^{(y-x)} z_1^d \cdots z_{d_c-1}^d \cdot \mathcal{K} \, \mathrm{d}z_1 \cdots \, \mathrm{d}z_{d_c-1} \qquad (6.9)
\end{aligned}$$

where the fore-factor $(d+1)^{d_c-1}$ comes from fixing the maximal element in the sets $\mathcal{P}_1, \ldots, \mathcal{P}_{d_c-1}$.

In order to solve the integral in (6.9) we start with some preliminary observation. For all $k \in \mathbb{N}$ let

$$\mathcal{R}(k) := \frac{\bar{a}}{(y-x)^k} \ .$$

Observation 3 *Since $\bar{a} = a/(y-x)^{d-d_c}$ and $(y-x) > \sqrt[d]{a}$ it follows that*

$$\mathcal{R}(d_c - 1) \; < \; (y-x) \; .$$

Now we will split up the outermost integral in (6.9) into two integrals, one going from 0 to $\mathcal{R}(d_c-1)$ and the other from $\mathcal{R}(d_c-1)$ to $(y-x)$. The first integral can be solved in a straightforward way. To solve the second integral we split it up again into two integrals with appropriate integration boundaries. This process continues and leads finally to a sum of solvable integrals.

It is now

$$\begin{aligned}
&\int_0^{(y-x)} \cdots \int_0^{(y-x)} z_1^d \cdots z_{d_c-1}^d \cdot \mathcal{K} \, \mathrm{d}z_1 \cdots \, \mathrm{d}z_{d_c-1} \\
&= \int_0^{\mathcal{R}(d_c-1)} \int_0^{(y-x)} \cdots \int_0^{(y-x)} z_1^d \cdots z_{d_c-1}^d \cdot \mathcal{K} \, \mathrm{d}z_1 \cdots \, \mathrm{d}z_{d_c-1} && (6.10) \\
&+ \int_{\mathcal{R}(d_c-1)}^{(y-x)} \int_0^{(y-x)} \cdots \int_0^{(y-x)} z_1^d \cdots z_{d_c-1}^d \cdot \mathcal{K} \, \mathrm{d}z_1 \cdots \, \mathrm{d}z_{d_c-1} \; . && (6.11)
\end{aligned}$$

In integral (6.10) the variable z_{d_c-1} is bounded by $\mathcal{R}(d_c-1)$ and the variables $z_1, \ldots, z_{d_c-2}$ are bounded by $(y-x)$. Therefore, we can conclude that

$$\frac{\bar{a}}{z_1 \cdots z_{d_c-1}} \; \geq \; \frac{\bar{a}}{(y-x)^{d_c-2} \cdot \mathcal{R}(d_c-1)} \; = \; (y-x) \; .$$

It follows then that $\mathcal{K} = (y-x)^{d+1}$ and therefore it is

$$\begin{aligned}
(6.10) \; &= \int_0^{\mathcal{R}(d_c-1)} \int_0^{(y-x)} \cdots \int_0^{(y-x)} z_1^d \cdots z_{d_c-1}^d \cdot (y-x)^{d+1} \, \mathrm{d}z_1 \cdots \, \mathrm{d}z_{d_c-1} \\
&= \int_0^{\mathcal{R}(d_c-1)} \left(\frac{1}{d+1}\right)^{d_c-2} \cdot (y-x)^{(d+1)\cdot(d_c-1)} \cdot z_{d_c-1}^d \, \mathrm{d}z_{d_c-1} \\
&= \left(\frac{1}{d+1}\right)^{d_c-1} \cdot (y-x)^{(d+1)\cdot(d_c-1)} \cdot \bar{a}^{d+1} \cdot \left(\frac{1}{(y-x)}\right)^{(d+1)\cdot(d_c-1)} \\
&= \left(\frac{1}{d+1}\right)^{d_c-1} \cdot \bar{a}^{d+1} \; .
\end{aligned}$$

In order to solve integral (6.11) we will split up the second outermost integral into two integrals, one going from 0 to $\mathcal{R}(d_c-2)/z_{d_c-1}$ and the other from $\mathcal{R}(d_c-2)/z_{d_c-1}$ to $(y-x)$.

$$\begin{aligned}
(6.11) \; &= \int_{\mathcal{R}(d_c-1)}^{(y-x)} \int_0^{\mathcal{R}(d_c-2)/z_{d_c-1}} \cdots \int_0^{(y-x)} z_1^d \cdots z_{d_c-1}^d \cdot \mathcal{K} \, \mathrm{d}z_1 \cdots \, \mathrm{d}z_{d_c-1} && (6.12) \\
&+ \int_{\mathcal{R}(d_c-1)}^{(y-x)} \int_{\mathcal{R}(d_c-2)/z_{d_c-1}}^{(y-x)} \cdots \int_0^{(y-x)} z_1^d \cdots z_{d_c-1}^d \cdot \mathcal{K} \, \mathrm{d}z_1 \cdots \, \mathrm{d}z_{d_c-1} && (6.13)
\end{aligned}$$

The integral (6.12) can be solved in a straightforward way. Since the variable z_{d_c-2} is bounded by $\mathcal{R}(d_c-2)/z_{d_c-1}$ we can conclude that

$$\frac{\bar{a}}{z_1 \cdots z_{d_c-1}} \;\geq\; \frac{\bar{a} \cdot z_{d_c-1}}{(y-x)^{d_c-3} \cdot \mathcal{R}(d_c-2) \cdot z_{d_c-1}} \;=\; (y-x) \;.$$

It follows again that $\mathcal{K} = (y-x)^{d+1}$ and therefore it is

$$\begin{aligned}
(6.12) \;&=\; \int_{\mathcal{R}(d_c-1)}^{(y-x)} \int_0^{\mathcal{R}(d_c-2)/z_{d_c-1}} \cdots \int_0^{(y-x)} z_1^d \cdots z_{d_c-1}^d \cdot (y-x)^{d+1} \,\mathrm{d}z_1 \cdots \,\mathrm{d}z_{d_c-1} \\
&=\; \int_{\mathcal{R}(d_c-1)}^{(y-x)} \int_0^{\mathcal{R}(d_c-2)/z_{d_c-1}} \left(\frac{1}{d+1}\right)^{d_c-3} \cdot z_{d_c-2}^d \cdot z_{d_c-1}^d \cdot (y-x)^{(d+1)\cdot(d_c-2)} \,\mathrm{d}z_{d_c-2} \,\mathrm{d}z_{d_c-1} \\
&=\; \int_{\mathcal{R}(d_c-1)}^{(y-x)} \left(\frac{1}{d+1}\right)^{d_c-2} \cdot \bar{a}^{d+1} \cdot \frac{1}{z_{d_c-1}} \,\mathrm{d}z_{d_c-1} \\
&=\; \left(\frac{1}{d+1}\right)^{d_c-2} \cdot \bar{a}^{d+1} \cdot \ln\left(\frac{(y-x)^{d_c}}{\bar{a}}\right) \;.
\end{aligned}$$

In order to solve integral (6.13) we will split up the third outermost integral into two integrals, one going from 0 to $\mathcal{R}(d_c-3)/(z_{d_c-2} \cdot z_{d_c-1})$ and the other from $\mathcal{R}_{d_c-3}/(z_{d_c-2} \cdot z_{d_c-1})$ to $(y-x)$. The whole process continues now analogously. The computation of the next integral is briefly outlined for reasons of better understanding although the depiction becomes more and more uncomfortable. It is

$$(6.13) \;=\; \int_{\mathcal{R}(d_c-1)}^{(y-x)} \int_{\mathcal{R}(d_c-2)/z_{d_c-1}}^{(y-x)} \int_0^{\mathcal{R}(d_c-3)/(z_{d_c-2}\cdot z_{d_c-1})} \cdots \int_0^{(y-x)} \quad \cdots \qquad (6.14)$$

$$+\; \int_{\mathcal{R}(d_c-1)}^{(y-x)} \int_{\mathcal{R}(d_c-2)/z_{d_c-1}}^{(y-x)} \int_{\mathcal{R}(d_c-3)/(z_{d_c-2}\cdot z_{d_c-1})}^{(y-x)} \cdots \int_0^{(y-x)} \quad \cdots \qquad (6.15)$$

As before we can conclude from the ranges of the variables $z_1, \ldots, z_{d_c-1}$ that again $\mathcal{K} = (y-x)^{d+1}$. After solving the inner integrals up to the two outermost ones it remains that

$$\begin{aligned}
(6.14) \;&=\; \int_{\mathcal{R}(d_c-1)}^{(y-x)} \int_{\mathcal{R}(d_c-2)/z_{d_c-1}}^{(y-x)} \left(\frac{1}{d+1}\right)^{d_c-3} \cdot \bar{a}^{d+1} \cdot \frac{1}{z_{d_c-2} \cdot z_{d_c-1}} \,\mathrm{d}z_{d_c-2} \,\mathrm{d}z_{d_c-1} \\
&=\; \left(\frac{1}{d+1}\right)^{d_c-3} \cdot \bar{a}^{d+1} \cdot \int_{\mathcal{R}(d_c-1)}^{(y-x)} \ln\left(z_{d_c-1} \cdot \frac{(y-x)^{d_c-1}}{\bar{a}}\right) \cdot \frac{1}{z_{d_c-1}} \,\mathrm{d}z_{d_c-1} \\
&\leq\; \left(\frac{1}{d+1}\right)^{d_c-3} \cdot \bar{a}^{d+1} \cdot \ln\left(\frac{(y-x)^{d_c}}{\bar{a}}\right) \cdot \int_{\mathcal{R}(d_c-1)}^{(y-x)} \frac{1}{z_{d_c-1}} \,\mathrm{d}z_{d_c-1} \\
&=\; \left(\frac{1}{d+1}\right)^{d_c-3} \cdot \bar{a}^{d+1} \cdot \ln\left(\frac{(y-x)^{d_c}}{\bar{a}}\right)^2 \;.
\end{aligned}$$

With integral (6.15) we proceed analogously. To summarize the results so far we conclude that

$$(6.9) \;\leq\; (d+1)^{d_c-1} \cdot \bar{a}^{d+1} \cdot \sum_{i=1}^{d_c} \left(\frac{1}{d+1}\right)^{d_c-i} \cdot \ln\Bigg(\underbrace{\frac{(y-x)^{d_c}}{\bar{a}}}_{>1}\Bigg)^{i-1} \quad . \tag{6.16}$$

This can be simplified by the following observation.

Observation 4 *For all $x > 1$ and all $k \geq 1$ it is* $\ln(x)^k \leq (k/e)^k \cdot x$.

It follows that

$$\begin{aligned}(6.16) \;&\leq\; (d+1)^{d_c-1} \cdot \bar{a}^{d+1} \cdot \sum_{i=1}^{d_c} \left(\frac{1}{d+1}\right)^{d_c-i} \cdot \left(\frac{i-1}{e}\right)^{i-1} \frac{(y-x)^{d_c}}{\bar{a}} \\ &\leq\; 2 \cdot (d+1)^{d_c-1} \cdot (d_c-1)^{d_c-1} \cdot \bar{a}^d \cdot (y-x)^{d_c}\end{aligned}$$

and thus Claim 6 is shown. $\square_{\text{Claim 6}}$

Finally, let us have a brief look at Observation 4 which follows also easily. Consider the following function together with its first and second derivative, namely

$$\begin{aligned} f_k(x) \;&:=\; \ln(x)^k/x \\ f_k'(x) \;&=\; (k \cdot \ln(x)^{k-1} - \ln(x)^k)/x^2 \\ f_k''(x) \;&=\; (k \cdot (k-1) \cdot \ln(x)^{k-2} - 3k \cdot \ln(x)^{k-1} + 2 \cdot \ln(x)^k)/x^3 \quad . \end{aligned}$$

It holds now that $f_k'(x) = 0 \iff k \cdot \ln(x)^{k-1} = \ln(x)^k \iff x = e^k$. Since $f_k''(e^k) = -k^{k-1}/e^{3k} < 0$ it follows that for $x = e^k$ the function $f(x)$ attains a maximum. Therefore it is

$$f_k(x) \;\leq\; f_k(e^k) \;=\; \left(\frac{k}{e}\right)^k \qquad \text{for } x > 1 \text{ and } k \geq 1 \quad .$$

7 Summary and Open Problems

In this thesis, the concept of smoothed analysis is applied to the area of computational geometry. In the past, the use of smoothed analysis had a basically complexity theoretical motivation which holds of course also for problems in computational geometry. But besides this, we identified two other reasons that motivate the use of smoothed analysis particularly well in the field of computational geometry.

In many applications, concepts and methods from computational geometry are applied to data coming from physical measurements which are imprecise and thus afflicted with some noise. By the assumption that such measurement errors are distributed according to the Gaussian normal distribution, smoothed analysis provides a new complexity measure for this class of inputs. Another motivation lies in the fact that computers use only fixed precision arithmetic. This error can be modeled by the assumption that an input point is uniformly distributed in a hypercube around its real position.

In this thesis, the complexity of a fundamental geometric structure is considered and analysed in the smoothed case, namely the number of extreme points of the convex hull of a point set. It seems very promising to continue this work and to consider the smoothed complexity of other geometric structures such as the Voronoi diagram or Delaunay triangulation of point sets.

A first step toward the smoothed analysis of the Voronoi diagram is already done. The average case analysis for points chosen uniformly from a hypercube might be a good starting point. Since the Voronoi diagram is a rather relevant geometric structure that can be found in a large variety of applications, a smoothed analysis of this structure is definitively very interesting.

Another interesting and surprising result is surely that different probability distributions lead to a different smoothed complexity. E.g. for the number of extreme points of the convex hull, there is a significant gap between the smoothed complexity under Gaussian normal and uniform noise. The smoothed number of extreme points under Gaussian normal noise is only poly-logarithmic in the number of input points, while it is polynomial under uniform noise. This result implies directly the following questions and open problems.

- For which probability distributions is the smoothed complexity low or high?
- What is the decisive property of these distributions to cause either low or high smoothed complexity?
- Is the smoothed complexity under Gaussian normal noise the 'best' we can get?

- Is the smoothed complexity under uniform noise the 'worst' we can get?

It remains to mention that smoothed analysis is also very well motivated in the context of analysing motion. In this thesis, a new complexity measure for motion is introduced, the smoothed motion complexity. Especially in the development of algorithms and data structures smoothed analysis might lead to very interesting new results. A first step might be to develop a kinetic data structure for maintaining the bounding box of a linearly moving point set that is efficient compared to the smoothed motion complexity instead of the worst case motion complexity.

Bibliography

[AH01] P. Agarwal and S. 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.

[Aur91] F. Aurenhammer. Voronoi diagrams – a survey of a fundamental geometric data structure. *ACM Computing Surveys*, 23:345–405, 1991.

[AW91] F. Affentranger and J. A. Wieacker. On the convex hull of uniform randowm points in a simple d-polytope. *Discrete & Computational Geometry*, 6:291–305, 1991.

[BD02] A. Blum and J. Dunagan. Smoothed analysis of the perceptron algorithm. In *Proceedings of the 13th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 905–914, 2002.

[BDMS05] M. Bienkowski, V. Damerow, F. Meyer auf der Heide, and C. Sohler. Average case complexity of voronoi digrams of n sites from the unit cube. In *Proceedings of the 21st European Workshop on Computational Geometry (EWCG)*, pages 167–170, 2005.

[BGH99] J. Basch, L. J. Guibas, and J. Hershberger. Data structures for mobile data. *Journal of Algorithms*, 31(1):1–28, 1999.

[BGZ97] J. Basch, L. J. Guibas, and L. Zhang. Proximity problems on moving points. In *Proceedings of the 13th Annual ACM Symposium on Computational Geometry (SoCG)*, pages 344–351, 1997.

[BKST78] J. L. Bentley, H. T. Kung, M. Schkolnick, and C. D. Thompson. On the average number of maxima in a set of vectors. *Journal of the Association for Computing Machinery*, 25:536–543, 1978.

[BLMS$^+$03] L. Becchetti, S. Leonardi, A. Marchetti-Spaccamela, G. Schäfer, and T. Vredeveld. Average and smoothed competitive analysis of the multi-level feedback algorithm. In *Proceedings of the 44th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 462–471, 2003.

[Blu73] H. Blum. Biological shape and visual science (Part i). *Journal of Theoretical Biology*, 38:205–287, 1973.

[BMB03] C. Banderier, K. Mehlhorn, and R. Beier. Smoothed analysis of three combinatorial problems. In *Proceedings of the 28th International Symposium on Mathematical Foundations of Computer Science (MFCS)*, pages 198–207, 2003.

[BMS04] V. Bansal, F. Meyer auf der Heide, and C. Sohler. Labeling smart dust. In *Proceedings of the 12th European Symposium on Algorithms (ESA)*, pages 77–88, 2004.

[Bor80] K. H. Borgwardt. *The simplex method: a probabilistic analysis*. Springer, 1980.

[Bro79] K. Q. Brown. Voronoi diagrams from convex hulls. *Information Processing Letters*, 9:223–228, 1979.

[Bro80] K. Q. Brown. *Geometric transforms for fast geometric algorithms*. PhD thesis, Carnegie-Mellon University, Pittsburgh, Pennsylvania, USA, 1980.

[BTT85] C. Buchta, F. Tichy, and R. F. Tichy. Stochastical approximation of convex bodies. *Mathematische Annalen*, 271(2):225–235, 1985.

[Buc89] C. Buchta. On the average number of maxima in a set of vectors. *Information Processing Letters*, 33:63–65, 1989.

[BV04a] R. Beier and B. Vöcking. Random knapsack in expected polynomial time. *Journal of Computer and System Sciences*, 69(3):306–329, 2004. Also in *Proceedings of the 35th Annual ACM Symposium on Theory of Computing (STOC)*, pages 232–241, 2003.

[BV04b] R. Beier and B. Vöcking. Typical properties of winners and losers in discrete optimization. In *Proceedings of the 36th ACM Symposium on Theory of Computing (STOC)*, pages 343–352, 2004.

[Car70] H. Carnal. Die konvexe Hülle von n rotationssymmetrisch verteilten Punkten. *Zentralblatt für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 15:168–176, 1970.

[Cha91] B. Chazelle. An optimal convex hull algorithm and new results on cuttings. In *Proceedings of the 32th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 29–38, 1991.

[Cha96a] T. M. Chan. Optimal output-sensitive convex hull algorithms in two and three dimensions. *Discrete & Computational Geometry*, 16:361–368, 1996.

[Cha96b] T. M. Chan. Output-sensitive results on convex hulls, extreme points, and related problems. *Discrete & Computational Geometry*, 16:369–378, 1996.

[CK70] D. R. Chand and S. S. Kapur. An algorithm for convex polytopes. *Journal of the ACM*, 17(1):78–86, 1970.

[Cla94] K. L. Clarkson. More output-sensitive geometric algorithms. In *Proceedings of the 35th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 695–702, 1994.

[CM95] B. Chazelle and J. Matoušek. Derandomizing an output-sensitive convex hull algorithm in three dimenisons. *Comput. Geom. Theory Appl.*, 5:27–32, 1995.

[CS88] J. H. Conway and N. J. A. Sloane. *Sphere Packing, Lattices and Groups*. Springer-Verlag, New York, 1988.

[Del34] B. N. Delaunay. Sur la sphere vide. *Bulletin of Academy of Sciences of the USSR*, pages 793–800, 1934.

[Dir50] G. L. Dirichlet. Über die Reduction der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen. *Journal für die Reine und Angewandte Mathematik*, 40:209–227, 1850.

[DMR+03] V. Damerow, F. Meyer auf der Heide, H. Räcke, C. Scheideler, and C. Sohler. Smoothed motion complexity. In *Proceedings of the 11th European Symposium on Algorithms (ESA)*, pages 161–171, 2003.

[DS04a] V. Damerow and C. Sohler. Extreme points under random noise. In *Proceedings of the 12th European Symposium on Algorithms (ESA)*, pages 264–274, 2004.

[DS04b] V. Damerow and C. Sohler. Smoothed number of extreme points under uniform noise. In *Proceedings of the 20th European Workshop on Computational Geometry (EWCG)*, pages 93–96, 2004.

[DST03] J. Dunagan, D. A. Spielman, and S.-H. Teng. Smoothed analysis of interior-point algorithms: Condition number. *Computing Research Repository*, cs.DS/0302011, 2003.

[dvOS00] M. de Berg, M. van Kreveld, M. Overmars, and O. Schwarzkopf. *Computational Geometry*. Springer, 2 edition, 1997, 2000.

[Dwy90] R. A. Dwyer. Kinder, gentler average-case analysis for convex hulls and maximal vectors. *SIGACT News*, 21:64–71, 1990.

[Dwy91] R. A. Dwyer. Higher-dimensional Voronoi diagrams in linear expected time. *Discrete & Computational Geometry*, 6:343–367, 1991. Also in *Proceedings of the 5th Annual ACM Symposium on Computational Geometry (SoCG)*, pages 326–333, 1989.

[Efr65] B. Efron. The convex hull of a random set of points. *Biometrika*, 52:331–343, 1965.

[Eri01] J. Erickson. Nice point sets can have nasty Delaunay triangulations. In *Proceedings of the 17th Annual ACM Symposium on Computational Geometry (SoCG)*, pages 96–105, 2001.

[FF04] A. D. Flaxman and A. M. Frieze. The diameter of randomly perturbed digraphs and some applications. In *Proceedings of the 7th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems and 8th International Workshop on Randomization and Computation (APPROX-RANDOM)*, pages 345–356, 2004.

[For87] S. Fortune. A sweepline algorithm for Vorono diagrams. *Algorithmica*, 2:153–174, 1987.

[For97] S. Fortune. *Voronoi diagrams and Delaunay triangulations*, chapter 20 in J. E. Goodman and J. O'Rourke, editors, *Handbook of discrete and computational geometry*, pages 377–388. CRC Press, 1997.

[Gau40] C. F. Gauss. Recursion der Untersuchungen über die Eigenschaften der positiven ternären quadratischen Formen von Ludwig August Seeber. *Journal für die Reine und Angewandte Mathematik*, 20:312–320, 1840.

[GHSZ01] L. J. Guibas, J. Hershberger, S. Suri, and L. Zhang. Kinetic connectivity for unit disks. *Discrete & Computational Geometry*, 25(4):591–610, 2001.

[GKS92] L. J. Guibas, D. E. Knuth, and M. Sharir. Randomized incremental construction of Delaunay and Voronoi diagrams. *Algorithmica*, 7(4):381–413, 1992. Also in *Proceedings of the 17th International Colloquium on Automata, Languages and Programming (ICALP)*, pages 414–431, 1990.

[GN03] M. J. Golin and H.-S. Na. On the average complexity of 3D-Voronoi diagrams of random points on convex polytopes. *Computational Geometry*, 25:197–231, 2003.

[Har98] S. Har-Peled. On the expected complexity of random convex hulls. Technical Report 330, School of Mathematical Science, Tel-Aviv University, Tel-Aviv, Israel, 1998.

[Har04] S. Har-Peled. Clustering motion. *Discrete & Computational Geometry*, 31(4):545–565, 2004. Also in *Proceedings of the 42th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 84–93, 2001.

[HRZ97] M. Henk, J. Richter-Gebert, and G. M. Ziegler. *Basic Properties of Convex Polytopes*, chapter 13 in J. E. Goodman and J. O'Rourke, editors, *Handbook of discrete and computational geometry*, pages 243–270. CRC Press, 1997.

[HS01] J. Hershberger and S. Suri. Simplified kinetic connectivity for rectangles and hypercubes. In *Proceedings of the 12th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 158–167, 2001.

[Kem84] R. Kemp. *Fundamentals of the Average Case Analysis of Particular Algorithms*. B. G. Teubner, John Wiley & Sons, 1984.

[Kle80] V. Klee. On the complexity of d-dimensional Voronoi diagrams. *Archiv der Mathematik*, 34:75–80, 1980.

[KM72] V. Klee and G. J. Minty. How good is the simplex algorithm? *In Shisha, O., editor, Inequalities – III*, pages 159–175, Academic Press, 1972.

[Knu97] D. E. Knuth. *The Art of Computer Programming*, volume 1. Addison Wesley Longman, 3 edition, 1997.

[KS86] D. G. Kirkpatrick and R. Seidel. The ultimate planar convex hull algorithm? *SIAM Journal on Computing*, 15:287–299, 1986.

[KSS02] D. Kirkpatrick, J. Snoeyink, and B. Speckmann. Kinetic collision detection for simple polygons. *International Journal of Computational Geometry and Applications*, 12(1-2):3–27, 2002.

[Kur53] M. Kurita. On some formulas about volume and surface area. *Nagoya Mathematical Journal*, 6:109–117, 1953.

[LL83] G. Louchard and G. Latouche, editors. *Probability Theory and Computer Science*. Academic Press, 1983.

[Mat93] J. Matoušek. Linear optimization queries. *Journal of Algorithms*, 14:432–448, 1993. Also with O. Schwartzkopf in *Proceedings of the 8th Annual ACM Symposium on Computation Geometry (SoCG)*, pages 16–25, 1992.

[McM70] P. McMullen. The number of faces of a convex polytope. *Mathematika*, 17:179–184, 1970.

[Meg84] N. Megiddo. Linear programming in linear time when the dimension is fixed. *Journal of the Association for Computing Machinery*, 31:114–127, 1984.

[OBS92] A. Okabe, B. Boots, and K. Sugihara. *Spatial Tesselation: Concepts and Applications of Voronoi Diagrams*. Wiley, Chichester, 1992.

[Ray65] H. Raynaud. Sur le comportement asymptotique de l'enveloppe convexe d'un nuage de points tirés au hasard dans R^n. *Comptes Rendus de l'Acadèmie des Sciences Paris*, 261:627–629, 1965.

[Ray70] H. Raynaud. Sur l'enveloppe convexe des nuages de points aléatoires dans R^n. *Journal of Applied Probability*, 7:35–48, 1970.

[RS63] A. Rényi and R. Sulanke. Über die konvexe Hülle von n zufällig gewählten Punkten I. *Zentralblatt für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 2:75–84, 1963.

[RS64] A. Rényi and R. Sulanke. Über die konvexe Hülle von n zufällig gewählten Punkten II. *Zentralblatt für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 3:138–147, 1964.

[RV05] H. Röglin and B. Vöcking. Smoothed analysis of integer programming. In *Proceedings of the 11th Conference on Integer Programming and Combinatorial Optimization*, pages 276–290, 2005.

[San04] A. Sankar. *Smoothed Analysis of Gaussian elimination*. PhD thesis, Massachussetts Institute of Technology, Boston, Massachusetts, USA, 2004.

[Sei87] R. Seidel. On the number of faces in higher-dimensional Voronoi diagrams. In *Annual ACM Symposium on Computational Geometry (SoCG)*, pages 181–185, 1987.

[Sei97] R. Seidel. *Convex hull computations*, chapter 19 in J. E. Goodman and J. O'Rourke, editors, *Handbook of discrete and computational geometry*, pages 361–375. CRC Press, 1997.

[Sha78] M. I. Shamos. *Computational Geometry*. PhD thesis, Yale University, New Haven, Connecticut, USA, 1978.

[Spi] D. A. Spielman. `http://www-math.mit.edu/~spielman/SmoothedAnalysis/mfcs03comment.html`.

[SST03] A. Sankar, D. A. Spielman, and S.-H. Teng. Smoothed analysis of the condition numbers and growth factors of matrices. *Computing Research Repository*, cs.NA/0310022, 2003.

[ST03a] D. A. Spielman and S.-H. Teng. Smoothed analysis: motivation and discrete models. In *Proceedings of the 8th Workshop on Algorithms and Data Structures (WADS)*, pages 256–270, 2003.

[ST03b] D. A. Spielman and S.-H. Teng. Smoothed analysis of interior-point algorithms: termination. *Computing Research Repository*, cs.NA/0310022, 2003.

[ST04] D. A. Spielman and S.-H. Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *Journal of the ACM*, 51(3):385–463, 2004. Also in *Proceedings of the 33th Annual ACM Symposium on Theory of Computing (STOC)*, pages 296–305, 2001.

[Thi11] A. H. Thiessen. Percipitation average for large area. *Monthly Weather Review*, 39:1082–1084, 1911.

[Vor08] M. G. Voronoi. Nouvelles applications des parametres continus à la theorie des formes quadratique, deuxième mémoire: recherches sur les parallélloèdres primitifs. *Journal für die Reine und Angewandte Mathematik*, 134:198–287, 1908.

[Wei] E. W. Weisstein. Pappus's centroid theorem. From MathWorld–A Wolfram Web Resource. `http://mathworld.wolfram.com/PappussCentroidTheorem.html`.

[WS33] E. Wigner and F. Seitz. On the constitution of metallic sodium. *Physical Reviews*, 43:804–810, 1933.

[ZDBI97] L. Zhang, H. Devarajan, J. Basch, and P. Indyk. Probabilistic analysis for combinatorial functions of moving points. In *Proceedings of the 13th Annual ACM Symposium on Computational Geometry (SoCG)*, pages 442–444, 1997.