Berechnung konvexer Hüllen in erwarteter Linearzeit

Jung-Bae Son

Diplomarbeit nach einem Thema von
Prof. Dr. G. Hotz

Universität des Saarlandes
Fachbereich Informatik
Lehrstuhl für angewandte Mathematik II und Informatik

31. Dezember 1998

Zusammenfassung

In dieser Arbeit wird ein Reduktions-Algorithmus zur Berechnung der konvexen Hülle einer Punktmenge im reellen, d-dimensionalen, euklidischen Raum angeben und seine erwartete Laufzeit ermittelt.

Der Algorithmus geht folgendermaßen vor: Zunächst wird in einem Aussiebschritt die Anzahl der Eingabepunkte reduziert. Von den übriggebliebenen Punkten wird anschließend mit einem Algorithmus der worst-case Laufzeit $O(n \log n)$ im $\mathbb{R}^d$, $d \leq 3$, und $O(n^{\lfloor \frac{d}{2} \rfloor + 1})$ im $\mathbb{R}^d$, $d > 3$, die konvexe Hülle berechnet. Das Aussiebverfahren ist derart, daß die konvexe Hülle der übriggebliebenen Punkte mit der konvexen Hülle der Eingabepunkte übereinstimmt.

Bei Gleichverteilung auf einem d-dimensionalen Hyperquader beträgt die erwartete Laufzeit des Verfahrens $O(n)$. Liegt eine Wahrscheinlichkeitsdichte $f : \mathbb{R}^d \to \mathbb{R}$ vor, bei der für alle $x \in \mathbb{R}^d$ der Funktionswert $f(x) > 0$ ist, z.B. die Dichte der d-dimensionalen Standard-Normalverteilung, dann betragen die erwarteten Kosten ebenfalls $O(n)$. Bei bestimmten rotationssymmetrischen Dichten im $\mathbb{R}^d$ läßt sich die konvexe Hülle ebenfalls asymptotisch in erwarteter Linearzeit berechnen. Ist $d \geq 4$, so berechnet das Verfahren für bestimmte komponenten-unabhängige Dichten eine Obermenge der konvexen Hülle asymptotisch in linearer erwarteter Laufzeit. Bei Dichten $f : \mathbb{R}^d \to \mathbb{R}$, deren gesamte Wahrscheinlichkeitsmasse in einem hyperquaderförmigen Gebiet H versammelt ist, so daß

$$\int_{\mathbb{R}^d} f(x)\, dx = \int_H f(x)\, dx = 1$$

und $f(x) > 0$ für alle $x \in H$ ist, liegt die erwartete Laufzeit mit $O(n) + o(n \log n)$ für $d \leq 3$ und $O(n) + o(n^{\lfloor \frac{d}{2} \rfloor})$ – wenn man einen worst-case $O(n^{\lfloor \frac{d}{2} \rfloor})$ statt $O(n^{\lfloor \frac{d}{2} \rfloor + 1})$ Algorithmus benutzt – für $d > 3$ deutlich unter der worst-case Laufzeit gängiger Algorithmen zur Berechnung konvexer Hüllen.

Danksagung

An dieser Stelle möchte ich mich zunächst bei Herrn Hotz für eine äußerst interessante Aufgabenstellung und zahlreiche, wertvolle Denkanstöße bedanken. Weiterhin gilt mein Dank Frank Schulz, der immer offen und bereit für kritische Diskussionen war und es immer wieder verstand, seinen Zeigefinger auf den Kern des Problemes zu legen. Genauso bin ich Elmar Schömer zu Dank verpflichtet, ohne den die Betrachtung des höherdimensionalen Falles nicht in dieser Form vorliegen würde. Ebenfalls möchte ich mich bei Thomas Chadzelek bedanken, der vehement darauf hinwies, den Ansatz unter dem Aspekt der Berechenbarkeit und formalen Korrektheit zu überprüfen. Für das Korrekturlesen und die vielen hilfreichen Verbesserungsvorschläge, durch die diese Arbeit erst ihre jetzige Form annahm, stehe ich bei Thomas Chadzelek, Barbara Ruppert und Frank Schulz tief in der Schuld. Mein besonderer Dank gebührt meinen Eltern, die all dies erst ermöglicht haben.

Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, daß ich die Arbeit selbstständig verfaßt und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Inhaltsverzeichnis

1. Einleitung

Die Berechnung von konvexen Hüllen ist ein Thema, das auf dem Gebiet der *geometrischen Algorithmen* anzusiedeln ist. Konvexe Hüllen sind dabei nicht nur wegen ihrer Anwendungsvielfalt in der Praxis sondern auch als Hilfsmittel zur Lösung anderer – scheinbar völlig andersartiger – Fragestellungen von Bedeutung. Ein guter Überblick wird von Preparata und Shamos in [21] auf den Seiten 89-177 gegeben. Hier seien lediglich einige Beispiele genannt: in der Praxis werden konvexe Hüllen in der Bildverarbeitung, Mustererkennung, Kollisionserkennung und bei Aufteilungs- und Allokierungsproblemen verwendet. Konvexe Hüllen können auch zur Lösung von Problemen aus der Statistik eingesetzt werden, z.B. beim Clustering oder zur Ermittlung von robusten Schätzern. Eine Aufgabenstellung aus dem Gebiet der geometrischen Algorithmen, zu deren Lösung man konvexe Hüllen benutzt, ist die Delauney Triangulierung. Ein anderes Problem aus den geometrischen Algorithmen, das mit der Berechnung von konvexen Hüllen verwandt ist, ist die Frage nach der Bestimmung der Maxima einer Punktmenge. Erstaunlich ist vielleicht die Tatsache, daß sich die Berechnung der konvexen Hülle im Zweidimensionalen im Wesentlichen auf einen Sortieralgorithmus reduziert.

Der Berechnung konvexer Hüllen ist in den letzten Jahren ein vermehrtes Interesse zuteil geworden. Zur Zeit gibt es viele Verfahren zur Berechnung der konvexen Hüllen sowohl für den $\mathbb{R}^2$, als auch für den $\mathbb{R}^d$, $d > 2$. Eines der ersten effizienten Verfahren zur Berechnung konvexer Hüllen in der Ebene ist der sogenannte *Graham's Scan* [10]. Das Verfahren weist eine worst-case Laufzeit von $O(n \log n)$ auf, wobei n die Anzahl der Punkte sei, deren konvexe Hülle zu bestimmen ist. Anstatt wie bei Graham's Scan die Punkte der konvexen Hülle zu berechnen, gibt es Verfahren, z.B. *Jarvis's March* [15], bei denen die Seiten der konvexen Hülle bestimmt werden. Jarvis's March berechnet die konvexe Hülle von n Punkten in $O(n \cdot h)$ Schritten, wobei h die Anzahl der Seiten der konvexen Hülle ist. Ein weiterer Algorithmus zur Berechnung der konvexen Hülle im $\mathbb{R}^2$ mit worst-case Laufzeit $O(n \log n)$ wird im dritten Band von Mehlhorns „Data Structures and Algorithms“[19] angegeben. Hla Min und Si-Qing Zheng [20] haben einen Zeit-Platz optimalen Algorithmus entwickelt, mit dem sich in der Ebene die konvexe Hülle im ungünstigsten Fall mit Kosten $\Theta(n \log n)$ ausrechnen läßt. Von Kirkpatrick und Seidel [17] wurde ein Algorithmus mit worst-case Laufzeit $O(n \log h)$ entdeckt, wobei h die Anzahl der Hüllenpunkte ist, den sie als den „ultimativen“ bezeichnet haben. Sie zeigten in ihrer Arbeit darüber hinaus, daß ihr Verfahren asymptotisch optimal ist. Macht man Einschränkungen bezüglich der Eingabepunkte, so lassen sich deutlich geringere worst-case Laufzeiten erzielen. Unter der Annahme, daß es sich bei den Eingabepunkten um die Eckpunkte eines zweidimensionalen, einfachen Polygons (d.h. ohne sich schneidende Kanten) handelt, haben Graham und Yao [11] einen Linearzeit Algorithmus angegeben.

Bei der Berechnung konvexer Hüllen im Höherdimensionalen werden Verallgemeinerungen der oben genannten Verfahren eingesetzt. So handelt es sich z.B. bei Jarvis's

March um die zweidimensionale Variante einer Methode, die als *Gift-Wrapping* bezeichnet wird. Bei Gift-Wrapping Verfahren beträgt die Laufzeit $O(n \cdot h)$, wobei h die Anzahl der Flächen der konvexen Hülle ist. Eine andere Methode, der *Beneath-Beyond* Ansatz, stammt aus dem Bereich der dynamischen Algorithmen. Preparata und Shamos stellen in [21] auf den Seiten 125-134 beide Ansätze ausführlich vor und geben in beiden Fällen eine worst-case Laufzeit von $O(n^{\lfloor \frac{d}{2} \rfloor + 1})$ an. Seidel [22] benutzt einen zum Beneath-Beyond analogen Ansatz und gibt ein Verfahren mit worst-case Laufzeit $O(n^{\frac{\lfloor d+1 \rfloor}{2}})$ an, das für gerade Dimensionen nachweislich optimal ist. Weitere Verbesserungen sind der Algorithmus von Clarkson, Mehlhorn und Seidel [7] mit einer erwarteten Laufzeit von $O(n \log n + n^{\lfloor \frac{d}{2} \rfloor})$ und der Algorithmus von Chazelle [6] mit worst-case Laufzeit $O(n \log n + n^{\lfloor \frac{d}{2} \rfloor})$.

Bei fast allen obigen Laufzeitangaben handelt es sich um worst-case Angaben. Die Laufzeit von Jarvis's March – einem Gift-Wrapping Verfahren – oder dem Algorithmus von Kirkpatrick und Seidel hängt von der Anzahl der Flächen bzw. Punkte der konvexen Hülle ab. Diese kann im Mittel jedoch deutlich geringer sein als im ungünstigsten Fall. Daher kann es sich durchaus lohnen, die erwartete Laufzeit von konvexe Hülle Algorithmen zu betrachten. Bei den beiden eben genannten Algorithmen ist diese Analyse besonders einfach, da man lediglich die erwartete Anzahl der Punkte bzw. Flächen der konvexen Hülle betrachten muß. Die erwartete Anzahl der Hüllenpunkte bzw. -flächen hängt von der Wahrscheinlichkeitsverteilung ab, gemäß der die Punkte im Raum verteilt sind. Dwyer gibt in [9] eine ausführliche Übersicht über die erwartete Anzahl von Hüllenpunkten bei verschiedenen Wahrscheinlichkeitsverteilungen an. Bei einer Gleichverteilung auf einem Quadrat besitzt die konvexe Hülle im Mittel $\Theta(\log n)$ Punkte. Die Anzahl der Hüllenpunkte kann daher drastisch kleiner sein als die Anzahl der Eingabepunkte. Einige Algorithmen machen von dieser Beobachtung Gebrauch und bestimmen durch eine Vorauswahl eine Untermenge U der Eingabepunkte M, die mit großer Wahrscheinlichkeit alle Hüllenpunkte enthält, und berechnen anschließend die Hülle von U anstatt von M. Da U im Mittel wesentlich weniger Punkte enthalten wird als M, wird die erwartete Laufzeit geringer sein. Arne Maus [18] hat einen Algorithmus angegeben, der im $\mathbb{R}^2$ bei Gleichverteilung auf einem Rechteck im Mittel die konvexe Hülle in Linearzeit berechnet. Akl und Toussaints Algorithmus [1] berechnet ebenso wie Devroye und Toussaints Algorithmus [8] bei Gleichverteilung auf dem Einheitsquadrat und Einheitskreis die konvexe Hülle in erwarteter Linearzeit. Jünger, Borgwardt, Gaffke und Reinelt [16] haben ein Verfahren entwickelt, das zunächst einen „throw-away"-Schritt ausführt und dann unter Verwendung eines Algorithmus zur Berechnung der konvexen Hülle mit worst-case Laufzeit $O(n \log n)$ bei Gleichverteilung auf dem Einheitsquadrat bzw. -kreis die konvexe Hülle in erwarteter Zeit $O(n)$ berechnet. Ähnlich verfährt Golin und Sedgewick's Verfahren [12]. Es führt zunächst einen Sieb-Schritt aus und berechnet anschließend eine Obermenge der konvexen Hülle. Bei Gleichverteilung auf einem d-dimensionalen Hyperquader ist die erwartete Laufzeit ebenfalls linear. Bentley und Shamos [3] haben einen divide-and-conquer Algorithmus angegeben, mit dem sich im $\mathbb{R}^2$ und $\mathbb{R}^3$ bei N^p-Verteilungen die konvexe Hülle in erwarteter Linearzeit berechnen läßt. N^p-Verteilungen sind Wahrscheinlichkeitsverteilungen, bei denen eine Punktmenge der Größe N, für eine Konstante $p < 1$, $O(N^p)$ Extrem-Punkte besitzt. N^p-Verteilungen sind typischerweise gleichverteilt und/oder komponenten-unabhängig. Für komponenten-unabhängige Verteilungen haben Bentley, Clarkson und Levine [2] ein divide-and-conquer Verfahren beschrieben, das im $\mathbb{R}^2$ und $\mathbb{R}^3$ die konvexe Hülle in erwarteter Linearzeit und im $\mathbb{R}^d$, $d > 3$, eine Obermenge der konvexen Hülle in erwarteter Linearzeit berechnet. Borgwardt [4] erhält durch Verbindung eines „throw-away"-Schritts und Gift-Wrapping Algorithmus ein Verfahren, das im $\mathbb{R}^d$ für rotationssymmetrische Verteilungen eine erwartete Laufzeit zwischen $O(n)$ und $O(n^2)$ aufweist.

Auch in der vorliegenden Arbeit soll ein Verfahren zur Berechnung von konvexen Hüllen bezüglich seiner erwarteten Laufzeit untersucht werden. Es handelt sich dabei um die Verallgemeinerung eines Algorithmus zur Berechnung der konvexen Hülle einer Punktmenge im $\mathbb{R}^2$, den G. Hotz bereits in seiner Vorlesung „Algorithmische Informationstheorie"([14], S. 54–59) vorgestellt hat. Die Idee des Algorithmus ist dabei folgende: Zunächst reduziert man die Ausgangspunktmenge durch ein Aussiebverfahren. Dabei soll jedoch garantiert sein, daß trotz der Reduktion die konvexe Hülle der reduzierten Punktmenge identisch ist mit der konvexen Hülle der Ausgangspunktmenge. Anschließend berechnet man von dieser reduzierten Punktmenge mit einem Algorithmus mit worst-case Laufzeit $O(n \log n)$ (im $\mathbb{R}^d$, $d > 3$, mit worst-case Laufzeit $O(n^{\lfloor \frac{d}{2} \rfloor + 1})$), den wir im folgenden als *Standardalgorithmus* bezeichnen, die konvexe Hülle.

Abhängig von der Güte des Aussiebverfahrens wird man im Mittel eine mehr oder weniger deutliche Laufzeitverringerung erwarten dürfen. Im ersten Abschnitt werden wir ausführlich ein Reduktionsverfahren vorstellen, das obige Eigenschaft besitzt. An dieser Stelle sei deshalb lediglich die Grundidee skizziert.

Die Grundidee basiert auf folgender Beobachtung: Weiß man von einem konvexen Körper K, daß er gänzlich von der konvexen Hülle eingeschlossen wird, dann benötigt man keinen der Punkte, die in K liegen, zur Berechnung der konvexen Hülle, d.h. nur noch die Punkte außerhalb von K tragen zur konvexen Hülle bei. Wir überlagern daher den $\mathbb{R}^2$ mit seitenparallelen, ineinander geschachtelten Rechtecken. Indem man die Seitenlinien eines Rechtecks verlängert, generiert man im Rechteck, das darum liegt, vier Eckfelder. Sei $M = \{a_1, \ldots, a_n\} \subset \mathbb{R}^2$ die Menge der Eingabepunkte. Sobald nun alle vier Eckfelder eines Rechtecks R mit einem Punkt aus M besetzt sind, werden alle Punkte in den von R umfaßten Rechtecken nicht mehr zur Berechnung der konvexen Hülle benötigt.

Die Angabe dieser Aussiebkonstruktion alleine reicht nicht aus, um aussagen zu können, wie gut die Reduktion im Mittel sein wird. Dafür muß man zuerst wissen, gemäß welcher Wahrscheinlichkeitsdichte die Punkte im Raum verteilt sind. In Abschnitt 3 wird untersucht, welche Eigenschaften eine Wahrscheinlichkeitsdichte besitzen muß, damit das Verfahren die konvexe Hülle einer Punktmenge im Mittel asymptotisch in Linearzeit berechnet. So fordern wir, daß durch die Schachtelungs-

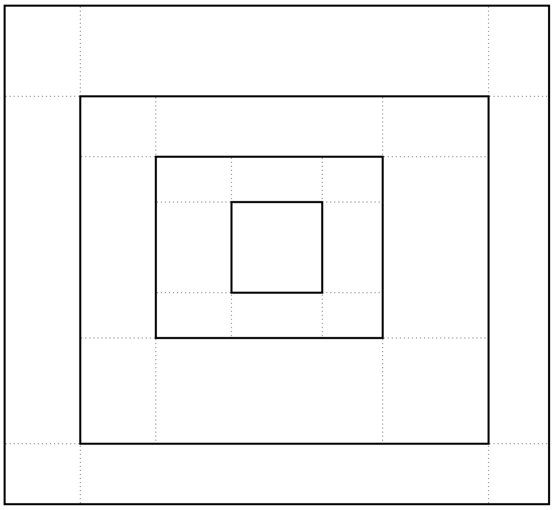

Abbildung 1: Reduktions-Konstruktion

konstruktion eine Wahrscheinlichkeitsverteilung auf dem $\mathbb{R}^2$ definiert wird, deren Wahrscheinlichkeitsmasse in den Schachtelungsgebieten exponentiell abnimmt. Das bedeutet genauer gesagt: Wenn a ein Punkt des $\mathbb{R}^2$ ist, dann beträgt die Wahrscheinlichkeit, daß a in einem Schachtelungsgebiet der Schachtelungstiefe l liegt, $p \cdot (1-p)^l$. Der Wert $p \in (0,1)$ sei die Wahrscheinlichkeit, daß ein Punkt im zentralen Rechteck der Schachtelung liegt. Das zentrale Rechteck besitzt Schachtelungstiefe 0. Außerdem soll folgende Fairneß-Bedingung gelten: ist k die Anzahl der Eckfelder des Rechtecks in Schachtelungstiefe l, dann sollen alle Eckfelder eine Wahrscheinlichkeit größer als $\frac{p(1-p)^l}{2kl}$ haben.

Man kann dieses Reduktionsverfahren auch leicht verallgemeinern, indem man die Schachtelung nicht mit Rechtecken sondern mit regulären Polygonen durchführt. Wir werden jedoch sehen, daß die konkrete Form der Aussiebkonstruktion keinen Einfluß auf die asymptotische Linearität der erwarteten Laufzeit des Algorithmus hat.

Die Idee läßt sich auch auf höherdimensionale Räume ausweiten. Dazu muß man von einer Schachtelung von Rechtecken zu einer Schachtelung von Hyper-Quadern übergehen. Gehen wir wiederum von einer exponentiellen Abnahme der Wahrscheinlichkeitsmasse in der Schachtelungstiefe aus, dann ermöglicht das Verfahren auch im $\mathbb{R}^d$ asymptotisch eine lineare erwartete Laufzeit.

Schwächt man die Fairneß-Bedingung ab, so kann man zeigen, daß bei einer Vielzahl von Riemann-Dichten $f : \mathbb{R}^d \to \mathbb{R}$ lineare erwartete Laufzeit erreicht wird. Dazu gehören folgende Dichten:

- Dichten von Gleichverteilungen auf einem d-dimensionalen Hyperquader,

- Dichten bestimmter rotationssymmetrischer Verteilungen im $\mathbb{R}^d$,
- Dichten mit $f(x) > 0$ für alle $x \in \mathbb{R}^d$, z.B. die Dichte der d-dimensionalen Standard-Normalverteilung.

Bei bestimmten komponenten-unabhängigen Dichten läßt sich eine Obermenge der konvexen Hülle in erwarteter Linearzeit berechnen. Geben wir die Fairneß-Bedingung ganz auf, so garantiert das Verfahren bei einer Klasse von Dichten immer noch eine erwartete Laufzeit von

$$\begin{array}{rl} & O(n) + o(n \log n) \text{ für } d \leq 3 \\ \text{und} & O(n) + o(n^{\lfloor \frac{d}{2} \rfloor}) \text{ für } d > 3. \end{array}$$

Solches sind Dichten $f : \mathbb{R}^d \to \mathbb{R}$, für die ein Hyperquader H existiert, der die gesamte Wahrscheinlichkeitsmasse enthält, d.h.

$$\int_{\mathbb{R}^d} f(x)\, dx = \int_H f(x)\, dx = 1$$

und $f(x) > 0$ für alle $x \in H$ ist.

2. Der Algorithmus im $\mathbb{R}^2$

Wir überdecken den $\mathbb{R}^2$ mit seitenparallelen, ineinander geschachtelten Rechtecken. Das zentrale Rechteck, $D^{(0)}$, sei durch die Geraden l_1, l_2, l_3 und l_4 begrenzt. Die Punkte $a \in \mathbb{R}^2$, die in $D^{(0)}$ liegen, können bei geeigneter Orientierung der Geraden durch die vier Abfragen

$$l_i < a \text{ bzw. } l_i > a, \quad \text{für } i = 1, \ldots, 4$$

lokalisiert werden. (Die Abfrage mit „$\leq$" oder „$<$" ist unerheblich, da wir annehmen wollen, daß die Geraden selber masselos sind. Die Abfrage $a < l_i$ bedeutet, daß je nach Orientierung der Geraden a links oder rechts bzw. ober- oder unterhalb der Geraden liegt.)

Als nächstes legen wir um $D^{(0)}$ ein weiteres Rechteck, welches durch die Begrenzungsgeraden l_5, l_6, l_7 und l_8 gegeben sei. Dabei gelte $l_i \parallel l_{i+4}$ für $i = 1, \ldots, 4$. Der entartete Fall $l_i = l_{i+4}$ sei nicht erlaubt. Auf diese Weise haben wir nun um $D^{(0)}$ einen *Ring* von acht weiteren Feldern, $D^{(1)}, \ldots, D^{(8)}$, gelegt, deren Seiten von Geradenabschnitten der Geraden $l_1, \ldots, l_8$ gebildet werden. Umgeben wir das ganze mit einem weiteren Rechteck mit den Begrenzungsgeraden l_9, l_{10}, l_{11} und l_{12}, dann erzeugen wir um die bisherige Konstruktion einen Ring aus 16 zusätzlichen Feldern, $D^{(9)}, \ldots, D^{(24)}$, und so weiter. Es sei:

$$\begin{aligned} Ring^{(0)} &= D^{(0)}, \\ Ring^{(l)} &= D^{((2l-1)^2)}, \ldots, D^{((2l+1)^2-1)} \end{aligned}$$

für $l = 1, \ldots, \infty$.

Definition 2.1 (Rechteck-Ring-Schachtelung) *Die oben beschriebene Konstruktion aus geschachtelten, seitenparallelen Rechtecken heiße* Rechteck-Ring-Schachtelung.
Ein Punkt in $Ring^{(l)}$ liegt in Schachtelungstiefe l.

In Abbildung 2 ist eine Rechteck-Ring-Schachtelung bis Tiefe 3 abgebildet.

Bemerkung 2.1 *Es ergibt sich konstruktionsbedingt, daß auf $Ring^{(l)}$ $8 \cdot l$ Felder liegen. Führt man diese Konstruktion bis ins Unendliche fort, so entsteht ein Gitter, das den $\mathbb{R}^2$ überdeckt. Die Felder des Gitters sind die Rechtecke $D^{(i)}$, $i = 0, \ldots, \infty$.*

Abbildung 3 gibt durch die Eintragungen in den Rechtecken die Anzahl der Entscheidungen an, die maximal erforderlich sind, um einen Punkt dem jeweiligen $D^{(i)}$ zuzuordnen.

Wir sehen, daß jeder Punkt in $Ring^{(l)}$ zur Lokalisierung in eines der Felder von $Ring^{(l)}$ höchstens $4 + 2l$ Linienvergleiche benötigt. Man kann einen Suchbaum angeben, bei dem die Werte aus Abbildung 3 eine obere Schranke für die Anzahl der

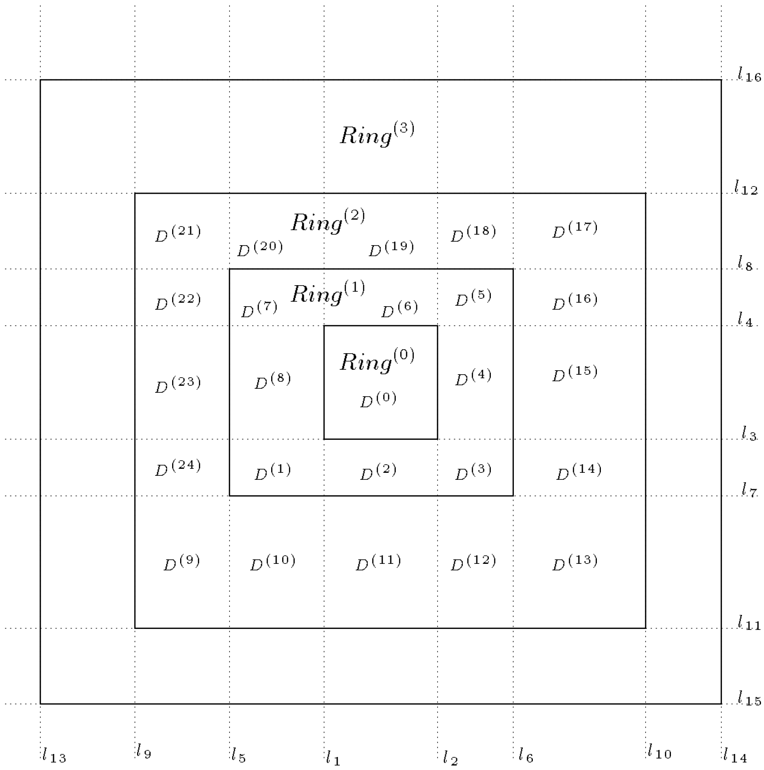

Abbildung 2: Ringkonstruktion für die konvexe Hülle

Abfragen bilden. In Abbildung 4 auf Seite 18 ist illustriert, wie der obere Teil dieses Suchbaums aussieht. Im Baum entsprechen die Kanten den jeweiligen Abfragen (z.B. $> l_1$ oder $< l_7$, etc.) und führen so zu semialgebraischen Mengen D_{Index}. Der Index korrespondiert mit den Abfragen, die zu dieser Menge führen. Sind Abfragen für eine semialgebraische Menge irrelevant, so werden sie im Index auch nicht aufgeführt. $D^{(1)}$ ist beispielsweise anhand der vier Abfragen $< l_1$, $< l_3$, $> l_5$ und $> l_7$ lokalisierbar; die Abfragen nach l_2, l_4 und l_6 bringen keinen Erkenntnisgewinn und sind daher für $D^{(1)}$ überflüssig.

Wie man sieht, wird die Anzahl der erforderlichen Abfragen, wie sie in Abbildung 3 vorgegeben wurde, nicht überschritten. Führt man diese Konstruktion bis ins Unendliche fort, dann erhält man den Suchbaum $\mathcal{B}$ eines Gitters, das den gesamten $\mathbb{R}^2$ überdeckt.

Man wähle nun n Punkte, $a_1, \ldots, a_n$, beliebig aus dem $\mathbb{R}^2$. Es soll nun die Frage geklärt werden, wie man aus der Lokalisation der Punkte $a_1, \ldots, a_n$ in dem Baum $\mathcal{B}$ Vorteile zur Konstruktion der konvexen Hülle von $\{a_1, \ldots, a_n\}$ ziehen kann.

Zuerst wird für jeden der n Punkte bestimmt, in welchem Feld $D^{(i)}$ er sich befindet. Danach machen wir Gebrauch von folgender Beobachtung: sobald in einem

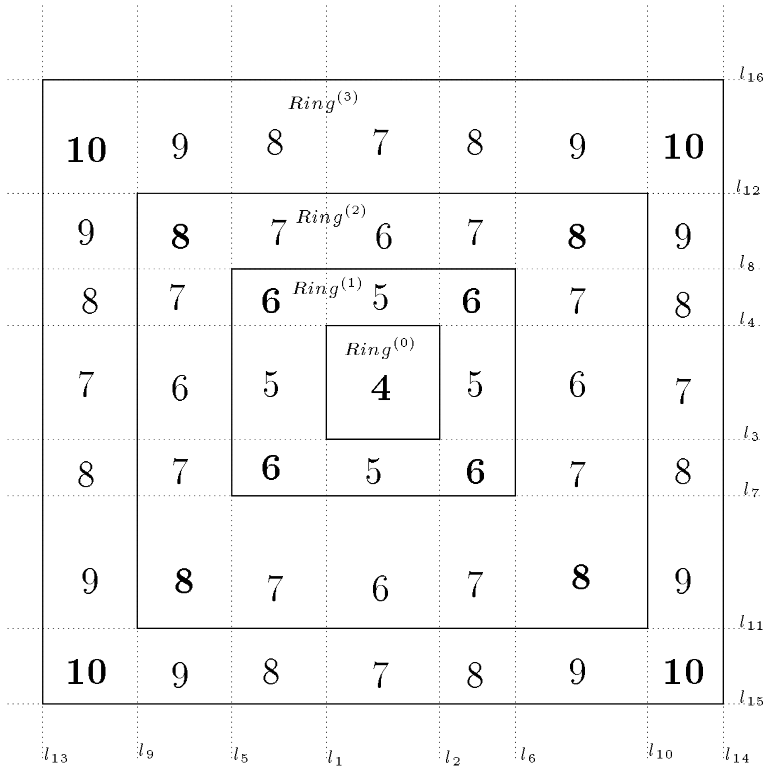

Abbildung 3: Suchaufwand in Ringkonstruktion

Ring, o.B.d.A. $Ring^{(l)}$, alle vier Eckfelder mit Punkten aus $\{a_1, \ldots, a_n\}$ besetzt sind, werden alle Ringe, die von $Ring^{(l)}$ umfaßt werden, nicht mehr zur Berechnung der konvexen Hülle gebraucht. D.h. $Ring^{(0)}, \ldots, Ring^{(l-1)}$ sind für die Berechnung der konvexen Hülle von $\{a_1, \ldots, a_n\}$ irrelevant.

Im folgenden werden wir einige Begriffe in besonderer Bedeutung verwenden.

Definition 2.2 (jenseits) *Liegt ein Punkt in oder außerhalb eines Ringes $Ring^{(l)}$, so befindet er sich* jenseits *von $Ring^{(l)}$.*

Seien $a_1, \ldots, a_n$ Punkte des $\mathbb{R}^2$.

Definition 2.3 (vollbesetzt) *Ringe, bei denen jedes Eckfeld mindestens einen der n Punkte $a_1, \ldots, a_n$ enthält, heißen* vollbesetzt.

Definition 2.4 (aktiv) *Wenn $Ring^{(l)}$ vollbesetzt ist, heißt jeder Punkt $a \in \{a_1, \ldots, a_n\}$ jenseits von $Ring^{(l)}$* aktiver *Punkt von $Ring^{(l)}$.*

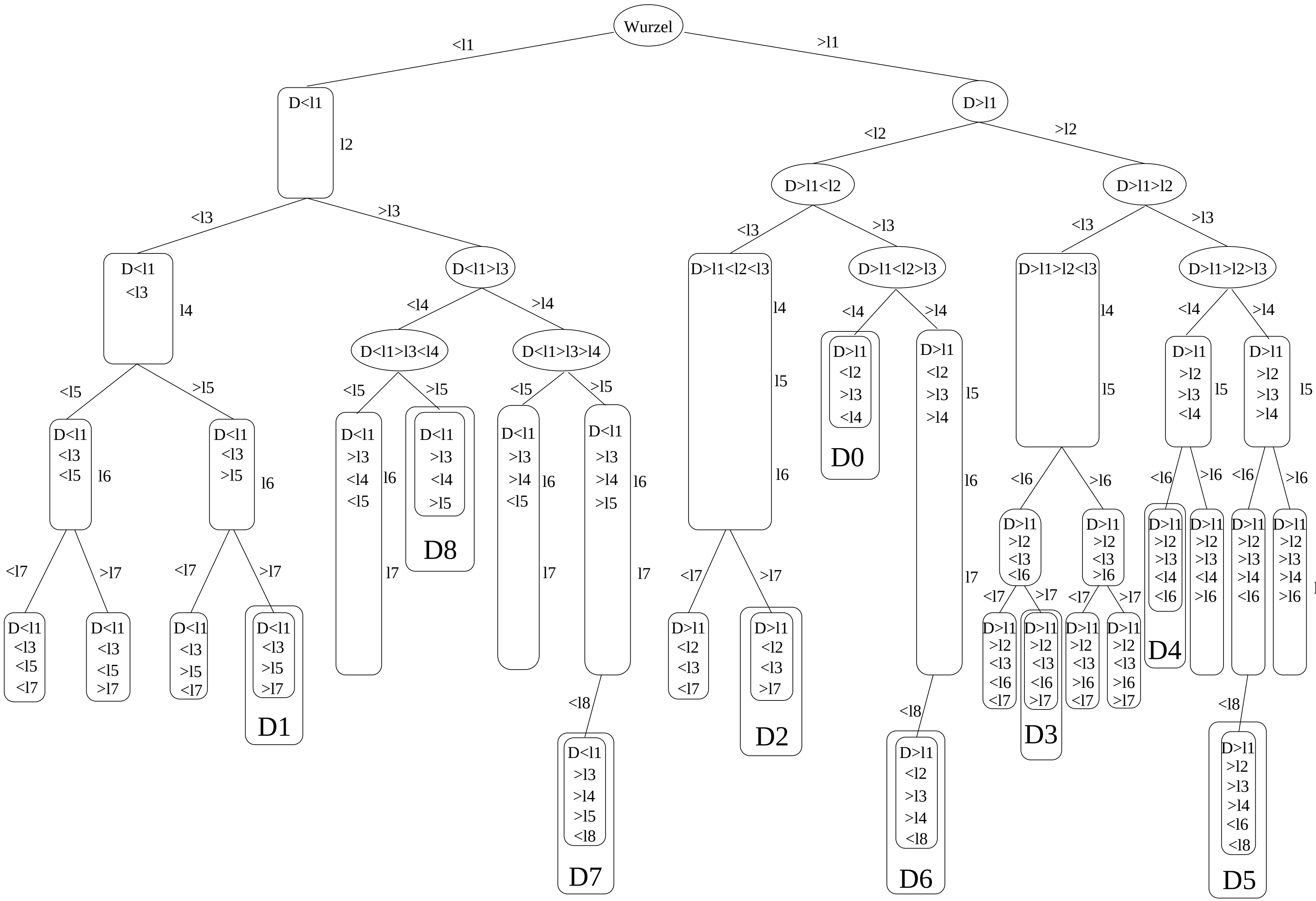

Abbildung 4: Suchbaum

Aufgrund der Beschaffenheit der Ringkonstruktion ist klar, daß die konvexe Hülle von $\{a_1, \ldots, a_n\}$ den äußersten vollbesetzten Ring enthält oder in ihm verläuft. Die Idee ist also, den äußersten vollbesetzten Ring zu suchen, der durch die Punkte $a_1, \ldots, a_n$ erzeugt wird, und die konvexe Hülle seiner aktiven Punkte zu ermitteln. Sei eine Rechteck-Ring-Schachtelung und der dazugehörige Suchbaum bereits gegeben, dann lautet der Algorithmus zur Berechnung der konvexen Hülle von $\{a_1, \ldots, a_n\}$ wie folgt:

Algorithmus A

1. lokalisiere die n Punkte und bestimme den äußersten ***vollbesetzten*** Ring $Ring^{(l)}$
2. benutze einen Standardalgorithmus zur Berechnung der konvexen Hülle der ***aktiven*** Punkte von $Ring^{(l)}$

Im weiteren Verlauf soll die erwartete Laufzeit dieses Algorithmus abgeschätzt werden. Als Kostenmaß des ersten Schrittes nehmen wir die Anzahl der Linienvergleiche und eine konstante Anzahl von Speicheranweisungen bzw. -vergleichen. Im zweiten Schritt sei das Kostenmaß dasjenige, welches auch in der vorhandenen Literatur zur Bestimmung der Kosten des Standardalgorithmus verwendet wird.

Zur Bestimmung der erwarteten Laufzeit brauchen wir eine Wahrscheinlichkeitsdichte bzw. Wahrscheinlichkeitsverteilung auf dem $\mathbb{R}^2$. Welcher Art diese Wahrscheinlichkeitsverteilung ist, wird später erläutert.

Die Lokalisierung der Punkte in $Ring^{(l)}$ erfordert, wie bereits erwähnt, $4 + 2l$ Linienvergleiche. Die Bestimmung des äußersten *vollbesetzten* Ringes läßt sich mit der Lokalisierung der Punkte koppeln. Diesen Mehraufwand (*overhead*) pro Punktlokalisation wollen wir durch eine Konstante c_{oh} abschätzen[1]. Somit kommt man pro Punkt mit $4 + 2l + c_{oh}$ Abfragen für Schritt 1 des Algorithmus aus.

In den folgenden Abschnitten soll die erwartete Berechnungskomplexität der Schritte 1 und 2 des Algorithmus detaillierter untersucht werden.

[1] Gehen wir davon aus, daß der Suchbaum $\mathcal{B}$ zur Punktlokalisation vorhanden ist. Jeder Knoten im Baum, der einem Eckfeld der Rechteck-Ring-Schachtelung entspricht, enthalte drei Zeiger auf die Speicherzellen, in denen die Punkte der anderen Eckfelder des Ringes gespeichert werden. Fällt ein Punkt nun in ein Eckfeld, so kann über die drei Zeiger in konstanter Zeit der Inhalt der anderen Speicherzellen ausgelesen werden. So kann man in konstanter Zeit prüfen, ob ein Ring vollbesetzt ist. Benutzt man zum Abspeichern einen Assoziativspeicher, mit dem sich in einem Durchgang mehrere Speicherzellen auslesen lassen, ließe sich mit einer Abfrage prüfen, ob die anderen Eckfelder leer sind. Mit einer zweiten Abfrage kann man dann den Zähler für den äußersten vollbesetzten Ring auffrischen. Im Falle der größtmöglichen Idealisierung ist also $c_{oh} = 2$.

3. Konvexe Hüllen im $\mathbb{R}^2$ bei einer Rechteck-Ring-Schachtelung

Damit wir die erwartete Laufzeit von Algorithmus `A` ermitteln können, müssen wir wissen, gemäß welcher Wahrscheinlichkeitsdichte die Punkte des $\mathbb{R}^2$ verteilt sind. Aus der Stochastik wissen wir, daß man durch Integration von Wahrscheinlichkeitsdichten auf $\mathbb{R}^2$ Wahrscheinlichkeitsverteilungen auf $\mathbb{R}^2$ erhält.

Eine Wahrscheinlichkeitsverteilung P auf $\mathbb{R}^2$ ist eine Funktion, die einem Gebiet G des $\mathbb{R}^2$ eine Wahrscheinlichkeit $P(G)$ – genauer einen Wert zwischen 0 und 1 – zuweist. Dabei bezeichnet $P(G)$ den Betrag der Wahrscheinlichkeit, daß ein Punkt des $\mathbb{R}^2$ in dem Gebiet G liegt. Sei $f : \mathbb{R}^2 \to \mathbb{R}$ eine Wahrscheinlichkeitsdichte. Aufgrund des oben genannten Zusammenhangs zwischen Wahrscheinlichkeitsdichte und Wahrscheinlichkeitsverteilung ist

$$P(G) = \int_G f(x)\, dx.$$

Wir werden zeigen, daß Algorithmus `A` für eine bestimmte Klasse von Wahrscheinlichkeitsdichten eine lineare erwartete Laufzeit besitzt. Die Wahrscheinlichkeitsdichten dieser Klasse seien dadurch charakterisiert, daß es für jede dieser Dichten eine Rechteck-Ring-Schachtelung gibt, so daß folgende beiden Eigenschaften erfüllt sind:

E1 Die Wahrscheinlichkeit, in $Ring^{(l)}$ zu liegen, nimmt mit zunehmender Schachtelungstiefe exponentiell ab.

E2 Alle Eckfelder eines Ringes $Ring^{(l)}$ haben die gleiche Wahrscheinlichkeit $q(l)$ und sind gegenüber den anderen Feldern von $Ring^{(l)}$ stochastisch nicht benachteiligt.

Formal bedeutet dies folgendes:
Ad **E1**) Sei die Wahrscheinlichkeit von $Ring^{(0)}$

$$P(Ring^{(0)}) = p \qquad \text{für } 0 < p < 1.$$

Dann betrage die Wahrscheinlichkeit, in $Ring^{(l+1)}$ zu liegen,

$$\begin{aligned} P(Ring^{(l+1)}) &= (1-p) \cdot P(Ring^{(l)}) \\ &= p(1-p)^l \qquad \text{für } l = 0, \ldots, \infty. \end{aligned}$$

Ad **E2**) Wie wir in Bemerkung 2.1 bereits feststellten, besteht $Ring^{(l)}$ aus $8{\cdot}l$ Feldern. Die Eckfelder von $Ring^{(l)}$ sind daher stochastisch nicht benachteiligt falls

$$q(l) \geq \frac{p(1-p)^l}{8l}$$

ist.

3.1. Verringerung der Beobachtungstiefe

Jede Berechnung von konvexen Hüllen im $\mathbb{R}^2$ beginnt mit einer Vorgabe von o.B.d.A. n Punkten aus dem $\mathbb{R}^2$. Da wir an der erwarteten Laufzeit des Algorithmus zur Berechnung der konvexen Hülle interessiert sind, betrachten wir zufällige Auswahlen von n Punkten. Die zufällige Auswahl von n Punkten aus dem $\mathbb{R}^2$ stellt ein Wahrscheinlichkeitsexperiment dar. Der Einfachheit halber nehmen wir an, daß die n Punkte unabhängig voneinander ausgewählt werden[2]; Punkte können daher auch mehrfach vorkommen. Der Ausgang dieses Experiments ist ein n-Tupel $a = (a_1, \ldots, a_n) \in (\mathbb{R}^2)^n$ und die Gesamtheit aller möglichen Ausgänge ist gleich $(\mathbb{R}^2)^n$.

Sei $D = \{D^{(0)}, D^{(1)}, D^{(2)}, \ldots\}$ die Menge der Felder des Gitters, das durch die Rechteck-Ring-Schachtelung erzeugt wird. Das n-Tupel $\omega = (\omega_1, \ldots, \omega_n) \in D^n$ sei eine Auswahl von n Gitterfeldern. Jede Auswahl $a = (a_1, \ldots, a_n)$ von n Punkten des $(\mathbb{R}^2)^n$ wird surjektiv auf eine Auswahl $\omega = (\omega_1, \ldots, \omega_n)$ von n Gitterfeldern abgebildet. Betrachtet man die Vorgehensweise von Algorithmus `A`, dann sieht man, daß alle n-Tupel aus $(\mathbb{R}^2)^n$, die auf dasselbe n-Tupel ω abgebildet werden, dieselbe Laufzeit haben. Es ist daher dasselbe, ob man die Laufzeit des Algorithmus für eine Auswahl $a = (a_1, \ldots, a_n)$ von n Punkten aus $(\mathbb{R}^2)^n$ oder für die dazugehörige Auswahl $\omega = (\omega_1, \ldots, \omega_n)$ von n Feldern des Gitters D betrachtet.

Da jede Auswahl von n Punkten des $\mathbb{R}^2$ surjektiv auf eine Auswahl von n Gitterfeldern aus D abgebildet wird, unterteilt die Rechteck-Ring-Schachtelung den $(\mathbb{R}^2)^n$ in $|D^n|$ Äquivalenzklassen. Die Menge aller Äquivalenzklassen ist gerade D^n. Zwei n-Tupel $a, a' \in (\mathbb{R}^2)^n$ liegen in derselben Äquivalenzklasse $\omega \in D^n$, falls a und a' auf ω abgebildet werden.

Dadurch erreichen wir eine Verringerung der Beobachtungstiefe. Wir fassen alle n-Tupel $a \in (\mathbb{R}^2)^n$ in Äquivalenzklassen $\omega \in D^n$ zusammen und betrachten statt einem Wahrscheinlichkeitsexperiment mit kontinuierlichen Ausgängen in $(\mathbb{R}^2)^n$ ein diskretes Wahrscheinlichkeitsexperiment mit Ausgängen in D^n, das für die Laufzeit des Algorithmus dieselben Ergebnisse liefert[3].

---

[2] Dies ist eine Annahme, die in der Literatur häufig vorkommt, vgl.[16],[9],[4],[2]

[3] Da die n Punkte aus $\mathbb{R}^2$ unabhängig voneinander gewählt werden, handelt es sich um ein n-stufiges Wahrscheinlichkeitsexperiment, bei dem die Ausgänge jeder Stufe unabhängig voneinander sind und ein Ergebnis aus demselben Ergebnisraum $\mathbb{R}^2$ liefern. In anderen Worten, das gleiche Experiment wird n-mal wiederholt. Der Wahrscheinlichkeitsraum ist daher $\left((\mathbb{R}^2)^n, (\mathcal{B}^2)^n\right)$, der n-fache Produktraum von $(\mathbb{R}^2, \mathcal{B}^2)$, wobei $\mathcal{B}^2$ die Borel-σ-Algebra auf $\mathbb{R}^2$ ist. Die surjektive Abbildung zur Verringerung der Beobachtungstiefe ist eine Zufallsvariable, die das auf $\left((\mathbb{R}^2)^n, (\mathcal{B}^2)^n\right)$ gegebene Wahrscheinlichkeitsmaß auf den Wahrscheinlichkeitsraum $(D^n, Pot(D^n))$ des diskreten Wahrscheinlichkeitsexperiments überträgt. Somit erhalten wir auch eine Wahrscheinlichkeitsverteilung für das Experiment in Abschnitt 3.2.

3.2. Das Wahrscheinlichkeitsexperiment

Betrachten wir nun als Wahrscheinlichkeitsexperiment die Auswahl von n Feldern des Gitters D. Aufgrund der Annahmen des vorherigen Abschnitts werden die Felder unabhängig voneinander gewählt; folglich kann man eine Auswahl von n Feldern des Gitters auch als n-fache Ziehung von Kugeln aus einer Urne betrachten, wobei die Kugeln in der Urne den Feldern des Gitters entsprechen. Da jedes Feld des Gitters beliebig oft gewählt werden kann, entspricht dies im Urnenmodell einer Ziehung, wobei nach jeder Ziehung die Kugel wieder in die Urne zurückgelegt wird, d.h. Wiederholungen sind erlaubt. Die n Urnenziehungen sind also stochastisch unabhängig.

Wir können daher die Wahl von n Feldern des Gitters durch eine n-fache Urnenziehung modellieren, bei der die einzelnen Ziehungen jeweils stochastisch unabhängig sind. Für dieses Wahrscheinlichkeitsexperiment läßt sich leicht ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ angeben.

Betrachten wir zunächst den Wahrscheinlichkeitsraum $(\Omega_e, \mathcal{A}_e, P_e)$ für die 1-fache Urnenziehung. Der Stichprobenraum ist

$$\Omega_e = \{D^{(0)}, D^{(1)}, D^{(2)}, \ldots\} = D,$$

wobei das Ereignis $\{D^{(i)}\}$ bedeutet: „`Feld` $D^{(i)}$ `wurde gewählt`". Der Ereignisraum

$$\mathcal{A}_e = Pot(\Omega_e)$$

ist die Potenzmenge des Stichprobenraumes. Es sei f eine Wahrscheinlichkeitsdichte mit den Eigenschaften **E1** und **E2**. Das n-Tupel $a = (a_1, \ldots, a_n)$ sei eine Auswahl von n Punkten des $(\mathbb{R}^2)^n$. Da das Ereignis „`Feld` $D^{(i)}$ `wurde gewählt`" nichts anderes bedeutet als, daß einer der Punkte aus a in Feld $D^{(i)}$ liegt, ist

$$P_e(\{D^{(i)}\}) = \int_{D^{(i)}} f(x)\,dx.$$

Mit dieser Definition ist P_e eine korrekte Wahrscheinlichkeitsdichte, denn dann erfüllt P_e die Normierungsbedingung $P_e(\Omega_e) = 1$:

$$P_e(\Omega_e) = P_e(\{D^{(0)}, D^{(1)}, D^{(2)}, \ldots\}),$$

da Elementarereignisse paarweise disjunkt sind, ist

$$\begin{aligned}
P_e(\{D^{(0)}, D^{(1)}, D^{(2)}, \ldots\}) &= \sum_{i=0}^{\infty} P_e(\{D^{(i)}\}) \\
&= \sum_{l=0}^{\infty} P_e(Ring^{(l)}) \\
&= p \cdot \sum_{l=0}^{\infty} (1-p)^l,
\end{aligned}$$

$\sum_{l=0}^{\infty}(1-p)^l$ ist eine geometrische Reihe, so daß

$$\begin{aligned} P_e(\Omega_e) &= p \cdot \frac{1}{1-(1-p)} \\ &= 1 \end{aligned}$$

ist.

Da sich in unserem Falle die n-fache Urnenziehung durch n stochastisch unabhängige 1-fache Urnenziehungen modellieren läßt, ist der Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ der n-fachen Urnenziehung gleich dem n-fachen Produktraum von $(\Omega_e, \mathcal{A}_e, P_e)$. Damit ist der Stichprobenraum

$$\Omega = \{\omega = (\omega_{e_1}, \ldots, \omega_{e_n}) \mid \omega_{e_i} \in \Omega_e\},$$

der Ereignisraum

$$\mathcal{A} = Pot(\Omega)$$

und die Wahrscheinlichkeitsverteilung

$$P(\{\omega\}) = \prod_{i=1}^{n} P_e\left(\{\omega_{e_i}\}\right) \quad , \text{für } \omega \in \Omega.$$

Definition 3.1 (n-Auswahl) *Die Elemente $\omega \in \Omega$, d.h. eine Auswahl von n Gitterfeldern werden wir künftig kurz* n-Auswahl *nennen.*

3.3. erwartete Kosten von Algorithmus A

Bestimmen wir nun die erwartete Laufzeit des auf Seite 19 angegebenen Verfahrens zur Berechnung der konvexen Hülle von n Punkten $a_1, \ldots, a_n$ des $\mathbb{R}^2$ bzw. einer n-Auswahl.

3.3.1. erwartete Kosten von Schritt 1

Betrachten wir zunächst die erwartete Laufzeit von Schritt 1. Wir wollen die erwarteten Kosten zur Lokalisation von n Punkten in der Rechteck-Ring-Schachtelung und der gleichzeitigen Bestimmung des äußersten vollbesetzten Ringes ermitteln.

Wir definieren dazu Zufallsvariablen $X_1, \ldots, X_n$ auf $(\Omega_e, \mathcal{A}_e, P_e)$. Dabei sei X_i die Anzahl der Abfragen, die zur Lokalisierung von Punkt a_i im Gitter benötigt werden. Die Zufallsvariable $X = \sum_{i=0}^{n} X_i$ auf $(\Omega, \mathcal{A}, P)$ sei die Anzahl der Abfragen, um bei einer Auswahl von n Punkten alle n Punkte zu lokalisieren.

Die erwarteten Kosten zur Lokalisierung eines Punktes in der Rechteck-Ring-Schachtelung betragen

$$E[X_i] = \sum_{j=0}^{\infty} P(X_i = j) \cdot j.$$

Aus Abschnitt 2 wissen wir bereits, daß man maximal $4 + 2l + c_{oh}$ Abfragen braucht, um zu entscheiden, ob ein Punkt in $Ring^{(l)}$ liegt und $Ring^{(l)}$ der äußerste vollbesetzte Ring ist. Daher läßt sich $E[X_i]$ durch

$$\begin{aligned}
E[X_i] &\leq \sum_{l=0}^{\infty} P(Ring^{(l)}) \cdot (4 + 2l + c_{oh}) \\
&= \sum_{l=0}^{\infty} p \cdot (1-p)^l \cdot (4 + 2l + c_{oh}) \\
&= (4 + c_{oh})\, p \cdot \sum_{l=0}^{\infty} (1-p)^l + 2p \cdot \sum_{l=0}^{\infty} l \cdot (1-p)^l \\
&= 4 + c_{oh} + 2p \cdot \frac{1-p}{(1-(1-p))^2} \\
&= 2 + \frac{2}{p} + c_{oh}
\end{aligned}$$

abschätzen. Da die Zufallsvariablen $X_1, \ldots, X_n$ auf demselben Wahrscheinlichkeitsraum definiert sind, ist

$$\begin{aligned}
E[X] &= E[X_1 + X_2 + \ldots + X_n] \\
&= E[X_1] + E[X_2] + \ldots + E[X_n].
\end{aligned}$$

Die erwartete Laufzeit von Schritt 1 ist demnach kleiner als

$$E[X] \leq (2 + \frac{2}{p} + c_{oh}) \cdot n.$$

3.3.2. erwartete Kosten von Schritt 2

Bevor wir jetzt das mittlere Laufzeitverhalten des zweiten Schrittes analysieren, führen wir vereinfachende Notationen ein und machen einige Vorbetrachtungen.

Als erstes definieren wir Zufallsvariablen Y, Y_C und Y_l auf $(\Omega, \mathcal{A}, P)$. Gegeben sei eine n-Auswahl. Die Zufallsvariable Y ist dann die Anzahl der Punkte, für die nach dem Aussieb-Schritt mit einem Standardverfahren die konvexe Hülle berechnet wird und Y_C gebe die Kosten dieser Berechnung an. Die Zufallsvariable Y_l bezeichne die Anzahl der Punkt, die *jenseits* von $Ring^{(l)}$ liegen.

Es folgen nun Begriffs- und Mengendefinitionen.

Definition 3.2 (erzeugt) *Eine n-Auswahl $(w_1, \ldots, w_n)$ von n Feldern des Gitters* erzeugt *einen Rechteck-Ring $Ring^{(l)}$, falls bei dieser Auswahl $Ring^{(l)}$ der äußerste vollbesetzte Ring ist.*

Eine n-Auswahl entspricht einer Menge von n-Tupeln $a \in (\mathbb{R}^2)^n$. Erzeugt eine n-Auswahl einen Rechteck-Ring $Ring^{(l)}$, dann erzeugt auch jedes dazugehörige n-Tupel $a = (a_1, \ldots, a_n)$ diesen Ring, d.h. es besitzt jeder Ring außerhalb von $Ring^{(l)}$ mindestens ein Eckfeld, in dem keiner der Punkte $a_1, \ldots, a_n$ liegt.

Definition 3.3 **($\mathcal{W}_l$)** $\mathcal{W}_l = \left\{\omega \in \Omega \mid \omega \textit{ erzeugt } Ring^{(l)}\right\}$.

$\mathcal{W}_l$ ist die Menge aller n-Auswahlen ω, die $Ring^{(l)}$ erzeugen. Es ist $\mathcal{W}_l \cap \mathcal{W}_{l'} = \emptyset$, falls $l \neq l'$, und $\Omega = \bigcup\limits_{l=0}^{\infty} \mathcal{W}_l$ ist.

Da Y eine Zufallsvariable ist, entspricht die Menge $\{Y = i\}$ der Menge aller n-Auswahlen, bei denen nach dem Aussiebschritt i Punkte an das Standardverfahren weitergegeben werden.

$\{Y_l = i\}$ ist die Menge aller n-Auswahlen, bei denen i Punkte jenseits von $Ring^{(l)}$ liegen.

$\{Y = i\} \cap \mathcal{W}_l$ ist daher die Menge aller n-Auswahlen, bei denen $Ring^{(l)}$ erzeugt wird und i Punkte an das Standardverfahren weitergegeben werden. D.h. es ist die Menge aller n-Auswahlen, bei denen i die Anzahl der aktiven Punkte von $Ring^{(l)}$ ist. Somit ist

$$\{Y = i\} \cap \mathcal{W}_l \quad = \quad \{Y_l = i\} \cap \mathcal{W}_l.$$

Damit läßt sich $\{Y = i\}$ auch als

$$\begin{aligned} \{Y = i\} &= \sum_{l=0}^{\infty} \{Y = i\} \cap \mathcal{W}_l \\ &= \sum_{l=0}^{\infty} \{Y_l = i\} \cap \mathcal{W}_l \end{aligned}$$

schreiben.

Bemerkung 3.1 *Bei der Vereinigung von disjunkten Mengen wird der Deutlichkeit halber das Summenzeichen benutzt.*

Wir besitzen nun alle formalen Angaben, um die erwartete Anzahl der Punkte anzugeben, die zur Berechnung der konvexen Hülle benötigt werden.

$$\begin{aligned} E[Y] &= \sum_{i=0}^{n} P(Y = i) \cdot i \\ &= \sum_{i=0}^{n} \sum_{l=0}^{\infty} P(\{Y_l = i\} \cap \mathcal{W}_l) \cdot i \end{aligned}$$

Wenn der Erwartungswert existiert, konvergiert die Reihe absolut und jede Umordnung konvergiert daher gegen denselben Grenzwert

$$\begin{aligned} E[Y] &= \sum_{l=0}^{\infty} P(\mathcal{W}_l) \cdot \sum_{i=0}^{n} \frac{P(\{Y_l = i\} \cap \mathcal{W}_l)}{P(\mathcal{W}_l)} \cdot i \\ &= \sum_{l=0}^{\infty} P(\mathcal{W}_l) \cdot E[Y_l | \mathcal{W}_l]. \end{aligned}$$

Bemerkung 3.2 *$E[Y_l|\mathcal{W}_l]$ ist die erwartete Anzahl der Punkte jenseits von Ring$^{(l)}$ unter der Bedingung, daß Ring$^{(l)}$ der äußerste vollbesetzte Ring ist.*

Unglücklicherweise ist es sehr schwierig, einen Term für $P(\mathcal{W}_l)$ oder $E[Y_l|\mathcal{W}_l]$ anzugeben. Numerische Simulationen und der Vergleich mit einem ähnlichen aber sehr viel einfacheren Wahrscheinlichkeitsexperiment legen nahe, daß folgende Vermutungen gelten:

Vermutung 3.1 *Gegeben eine Rechteck-Ring-Schachtelung. Für große Werte n und festes $p \in (0,1)$ gibt es eine maximale Schachtelungstiefe l_0, so daß für alle Schachtelungstiefen $l \leq l_0$ die Ungleichung*

$$E[Y_l|\mathcal{W}_l] \leq E[Y_l]$$

gilt.

Vermutung 3.2 *Sei f eine Wahrscheinlichkeitsdichte, für die die Eigenschaften* **E1** *und* **E2** *erfüllt sind. Dann existiert ein Wert $b > 1$, so daß für große Werte n, $0 < p < 1$ und alle Schachtelungstiefen $l \leq l_n = \lfloor \alpha \log_b n \rfloor$ mit $0 < \alpha < -\frac{1}{\log_b(1-p)}$ die Abschätzung*

$$E[Y_l|\mathcal{W}_l] \leq E[Y_l]$$

erfüllt ist.

Bemerkung 3.3 *Im folgenden werden wir, soweit nicht anders angegeben, statt $\log_b x$ die einfachere Notation $\log x$ verwenden.*

Ein exakter Beweis für die Gültigkeit beider Vermutungen ist schwierig. Ein Satz aus der Stochastik, der allgemeine Aussagen über Ungleichungen dieser Art macht, ist anscheinend nicht bekannt. Es bleibt daher keine andere Wahl, die Gültigkeit dieser Ungleichung zu zeigen, als die Erwartungswerte selbst auszurechnen. Dazu müssen wir die involvierten Wahrscheinlichkeiten ausrechnen. Dies können wir aber nur, indem wir für jedes Ereignis angeben, aus welchen Elementarereignissen es zusammengesetzt ist. Bei der Kompliziertheit des zugrundeliegenden Wahrscheinlichkeitsexperiments ist dies jedoch ein äußerst aufwendiger Vorgang. Es sei daher auf numerische Simulationen[4] und die Analyse eines vereinfachten Wahrscheinlichkeits-experiments[5] verwiesen, die die Aussagen beider Lemmata empirisch bestätigen.

Für Vermutung 3.1 können wir jedoch eine Plausibilitätsbetrachtung angeben. Betrachten wir für große n eine beliebige, aber feste Schachtelungstiefe l_0.Es ist $E[Y_{l_0}|\mathcal{W}_{l_0}]$ der Erwartungswert der Punkte jenseits $Ring^{(l_0)}$ unter der Bedingung, daß $Ring^{(l_0)}$ der äußerste vollbesetzte Ring ist. Da wegen der Bedingung immer *mindestens* vier Punkte jenseits von $Ring^{(l_0)}$ sind, beträgt $E[Y_{l_0}|\mathcal{W}_{l_0}]$ mindestens vier. Wenn nur wenige Punkte jenseits von $Ring^{(l_0)}$ liegen, wird deshalb $E[Y_{l_0}|\mathcal{W}_{l_0}]$ größer als der *unbedingte* Erwartungswert, $E[Y_{l_0}]$, sein. Für große Werte von n werden jedoch viele Punkte jenseits von $Ring^{(l_0)}$ liegen. Liegen so viele Punkte jenseits von $Ring^{(l_0)}$, daß jedesmal, wenn wir n Punkte aus $\mathbb{R}^2$ auswählen, $Ring^{(l_0)}$ vollbesetzt ist, dann verliert die Bedingung „$Ring^{(l_0)}$ `muß vollbesetzt sein`" ihre Bedeutung. Daher verliert der bedingte Erwartungswert $E[Y_{l_0}|\mathcal{W}_{l_0}]$ für große Werte von n seinen *Vorteil* gegenüber dem unbedingten Erwartungswert $E[Y_{l_0}]$.

Bemerkung 3.4 *Es sei darauf hingewiesen, daß l_0 in Vermutung 3.1 abhängig von n ist. Aufgrund der Plausibilitätsbetrachtung ist es leicht nachvollziehbar, daß l_0 monoton steigend in n ist.*

Zur Abschätzung von $P(\mathcal{W}_l)$ verwenden wir die beiden Mengen

Definition 3.4 ($R^{(l)}$) $R^{(l)} = \{\omega \in \Omega \mid \textit{bei } \omega \textit{ ist } Ring^{(l)} \textit{vollbesetzt}\}$
$R^{(l)}$ sei die Menge aller n-Auswahlen ω, bei denen $Ring^{(l)}$ vollbesetzt ist.

Definition 3.5 ($\overline{R^{(l)}}$) $\overline{R^{(l)}} = \Omega \setminus R^{(l)}$
$\overline{R^{(l)}}$ ist das Gegenereignis von $R^{(l)}$, die Menge aller n-Auswahlen, bei denen in $Ring^{(l)}$ eine Ecke nicht besetzt ist.

$\mathcal{W}_l$ läßt sich dann auch anders darstellen:

$$\mathcal{W}_l \quad = \quad R^{(l)} \cap \overline{R^{(l+1)}} \cap \overline{R^{(l+2)}} \cap \overline{R^{(l+3)}} \cap \ldots,$$

woraus

$$\mathcal{W}_l \subseteq R^{(l)} \qquad \text{und} \qquad \mathcal{W}_l \subseteq \overline{R^{(l+1)}}$$

folgt. Im Gegensatz zu $P(\mathcal{W}_l)$ ist $P(R^{(l)})$, die Wahrscheinlichkeit, daß alle vier Eckfelder von $Ring^{(l)}$ besetzt sind, leicht anzugeben. Wir werden nun $P(R^{(l)})$ und $P(\overline{R^{(l)}})$ abschätzen. In Abschnitt 2 ist bereits darauf hingewiesen worden, daß $Ring^{(l)}$ aus $8 \cdot l$ Feldern zusammengesetzt ist. Da die Wahrscheinlichkeitsdichte Eigenschaft **E2** erfüllt, gilt für jeden Rechteck-Ring $Ring^{(l)}$, $l = 1, 2, \ldots$

$$q(l) \geq \frac{p \cdot (1-p)^l}{8l},$$

[4] siehe Abschnitt 5
[5] siehe Abschnitt 3.5

wobei $q(l)$ die Wahrscheinlichkeit jedes Eckfeldes von $Ring^{(l)}$ ist.

Es sei $\omega_1, \ldots, \omega_n$ eine n-Auswahl und h_i bezeichne wie oft das i-te Eckfeld von $Ring^{(l)}$ gewählt wurde. Da $\omega_1, \ldots, \omega_n$ unabhängig voneinander gewählt werden, können wir die Polynomialverteilung für die Berechnung von $P(R^{(l)})$ verwenden[6].

$$\begin{aligned}
P(R^{(l)}) &= \sum_{\substack{1 \leq h_1, \ldots, h_4 \leq n \\ h_1 + \cdots + h_4 \leq n}} \frac{n!}{h_1!\, h_2!\, h_3!\, h_4!}\, q(l)^{h_1 + \cdots + h_4}\, (1 - q(l))^{n-(h_1+\cdots+h_4)} \\
&= 1 - 4(1 - q(l))^n + 6(1 - 2q(l))^n - 4(1 - 3q(l))^n + (1 - 4q(l))^n.
\end{aligned}$$

Es ist immer $6(1 - 2q(l))^n \geq 4(1 - 3q(l))^n$ und daher

$$P(R^{(l)}) \geq 1 - 4(1 - q(l))^n.$$

Soll diese Wahrscheinlichkeit größer als $\frac{1}{2}$ sein, muß zwischen n und $q(l)$ folgende Beziehung bestehen:

Lemma 3.1 *Für $n \geq \frac{\ln 8}{q(l)}$ ist $P(R^{(l)}) \geq \frac{1}{2}$.*

Beweis

Aus $n \geq \frac{\ln 8}{q(l)}$ folgt unter Verwendung von $\ln x \leq x - 1$ zuerst

$$n \cdot \ln(1 - q(l)) \leq n \cdot -q(l) \leq -\ln 8$$

und dann

$$n \cdot \ln(1 - q(l)) \leq -\ln 8.$$

Zwei einfache Äquivalenzumformungen ergeben

$$\begin{aligned}
(1 - q(l))^n &\leq \frac{1}{8} \\
P(R^{(l)}) \geq 1 - 4(1 - q(l))^n &\geq \frac{1}{2}. \square
\end{aligned}$$

Da in unserem Falle $q(l) \geq \frac{p(1-p)^l}{8l}$ gewählt wurde, erhält man somit

$$n \geq \frac{\ln 8}{p(1-p)^l} \cdot 8l.$$

Für welche Schachtelungstiefen l Lemma 3.1 gilt, sagt Lemma 3.2.

Lemma 3.2 *Falls n groß und $0 < p < 1$ ist, ist für Schachtelungstiefen $l \leq \lfloor \alpha \log n \rfloor$, wobei $0 < \alpha < -\frac{1}{\log(1-p)}$ sein muß, die Voraussetzung von Lemma 3.1 erfüllt.*

[6] vgl. [23]

Beweis

Für $l' \leq l$ ist $\frac{\ln 8}{p(1-p)^l} \cdot 8l \geq \frac{\ln 8}{p(1-p)^{l'}} \cdot 8l'$. Es genügt also zu zeigen, daß für $l_n = \lfloor \alpha \log n \rfloor$ die Behauptung gilt. Die Ungleichung

$$n \geq \frac{8 \ln 8}{p} \cdot \alpha \log n \cdot n^{-\alpha \log (1-p)}$$

ist bei hinreichend großem n erfüllbar. Denn durch die spezielle Wahl von α läßt sich die Ungleichung umformen in

$$\begin{array}{rcll} n^{c'} & \geq & c \cdot \log n & , \text{mit } n,\ c,\ c' > 0 \\ n^{c'} & \geq & \log n^c & \\ \sqrt[c']{n} & \leq & b^{(n^c)} & , \text{Umkehrfunktionen, } b > 1, \end{array}$$

was für $n \to \infty$ gültig ist[7]. □

Es sei n groß genug gewählt, so daß für Schachtelungstiefe l_n die Wahrscheinlichkeit $P(R^{(l_n)}) \geq 1 - 4(1-q(l_n))^n \geq \frac{1}{2}$ ist.Dann muß $P(\overline{R^{(l_n)}}) \leq \frac{1}{2}$ sein. Denn

$$\begin{array}{rcl} P(\overline{R^{(l_n)}}) & = & 1 - P(R^{(l_n)}) \\ & \leq & 1 - (1 - 4(1-q(l_n))^n) \\ & = & 4(1-q(l_n))^n \quad \leq \quad \frac{1}{2}. \end{array}$$

Konstruktionsbedingt wird $Ring^{(l_n)}$ im Mittel $\frac{1}{1-p}$-mal so viele Punkte wie $Ring^{(l_n+1)}$ enthalten. Setzen wir diese Idee fort: $Ring^{(l_n-r)}$ wird dann im Mittel $(1-p)^{-r}$-mal mehr Punkte als $Ring^{(l_n)}$ enthalten und somit das Ereignis „$Ring^{(l_n-r)}$ `hat ein leeres Eckfeld`“ entsprechend unwahrscheinlicher sein. Folglich wird der Anschauung nach die Wahrscheinlichkeit, daß $Ring^{(l_n-r)}$ eine leere Ecke besitzt, für $r = 0, \ldots, l_n - 1$ kleiner als $\left(\frac{1}{2}\right)^{(1-p)^{-r}}$ sein. Aber man kann dies auch formal nachweisen.

Lemma 3.3 *Ist n groß und $0 < p < 1$, dann gilt für Schachtelungstiefe $l_n = \lfloor \alpha \log n \rfloor$, $0 < \alpha < -\frac{1}{\log(1-p)}$, und $r = 0, \ldots, l_n - 1$*

$$P(\overline{R^{(l_n-r)}}) \leq \left(\frac{1}{2}\right)^{(1-p)^{-r}}.$$

Beweis

Zunächst zeigen wir, daß für alle $(1-p)^{-r} \in \mathbb{R}$

$$1 - q(l_n)(1-p)^{-r} \leq (1 - q(l_n))^{(1-p)^{-r}}$$

[7] vgl. auch Seite 79

gilt. Sei c eine Konstante, $0 < c < 1$, und $y \geq 1$. Für $y = 1$ ist

$$1 - c \cdot y = (1-c)^y.$$

Für $y \geq 1$ gilt aufgrund des Monotonieverhaltens $1 - c \cdot y \leq (1-c)^y$, wenn $\frac{d}{dy}(1 - c \cdot y) \leq \frac{d}{dy}(1-c)^y$ ist. Wegen $\ln x \geq 1 - \frac{1}{x}$ ist

$$-\frac{c}{1-c} \leq \ln(1-c).$$

Woraus

$$\begin{array}{rrcl} & -c & \leq & (1-c) \cdot \ln(1-c) \\ \Rightarrow & -c & \leq & (1-c)^y \cdot \ln(1-c) \end{array}$$

und schließlich

$$\frac{d}{dy}(1 - c \cdot y) \leq \frac{d}{dy}(1-c)^y$$

folgt. Für $r = 0, \ldots, l_n - 1$ ist $(1-p)^{-r} \geq 1$ und somit

$$1 - q(l_n)(1-p)^{-r} \leq (1 - q(l_n))^{(1-p)^{-r}}.$$

Aus

$$q(l_n - r) = \frac{p(1-p)^{l_n - r}}{8(l_n - r)} \geq \frac{p(1-p)^{l_n - r}}{8(l_n)} = q(l_n)(1-p)^{-r}$$

folgt

$$\begin{array}{rrcl} & 1 - q(l_n - r) & \leq & 1 - q(l_n)(1-p)^{-r} \;\leq\; (1 - q(l_n))^{(1-p)^{-r}} \\ \Rightarrow & 1 - q(l_n - r) & \leq & (1 - q(l_n))^{(1-p)^{-r}} \\ & (1 - q(l_n - r))^n & \leq & ((1 - q(l_n))^{(1-p)^{-r}})^n \\ & 4(1 - q(l_n - r))^n & \leq & 4(1 - q(l_n))^{n \cdot (1-p)^{-r}} \\ \Rightarrow & 4(1 - q(l_n - r))^n & \leq & (4(1 - q(l_n))^n)^{(1-p)^{-r}}. \end{array}$$

Wegen Lemma 3.2 ist $4(1 - q(l_n))^n \leq \frac{1}{2}$, woraus die Behauptung folgt. □

Wenden wir uns mit diesem Wissen wieder der erwarteten Anzahl der Punkte zu, die zur Berechnung der konvexen Hülle benötigt werden.

$$\begin{array}{rcl} E[Y] & = & \sum_{l=0}^{\infty} P(\mathcal{W}_l) \cdot E[Y_l | \mathcal{W}_l] \\ & \leq^8 & E[Y_{l'} | \mathcal{W}_{l'}] \cdot \sum_{l=l'}^{\infty} P(\mathcal{W}_l) + \sum_{l=0}^{l'-1} P(\mathcal{W}_l) \cdot E[Y_l | \mathcal{W}_l]. \end{array}$$

Sei $l' \leq l_0$ aus Vermutung 3.1, dann gilt die Ungleichung

$$\begin{aligned} E[Y] &\leq E[Y_{l'}] + \sum_{l=0}^{l'-1} P(\overline{R^{(l+1)}}) \cdot E[Y_l] \\ &= E[Y_{l'}] + \sum_{r=0}^{l'-1} P(\overline{R^{(l'-r)}}) \cdot E[Y_{l'-(r+1)}]. \end{aligned}$$

An dieser Stelle ist es hilfreich, mehr über die Form von $E[Y_l]$ zu erfahren. Laut Definition ist Y_l die Anzahl der Punkte jenseits von $Ring^{(l)}$. Dies entspricht genau der Anzahl der Punkte in dem Bereich $\bigcup_{j=l}^{\infty} Ring^{(j)}$. Die Wahrscheinlichkeit, daß dieses Gebiet i Punkte enthält, beträgt $\binom{n}{i}(1-p)^{l \cdot i}\left(1-(1-p)^l\right)^{n-i}$. Die erwartete Anzahl der Punkte jenseits von $Ring^{(l)}$ beträgt somit

$$\begin{aligned} E[Y_l] &= \sum_{i=0}^{n} \binom{n}{i} (1-p)^{l \cdot i}(1-(1-p)^l)^{n-i} \cdot i \\ &= n \cdot (1-p)^l. \end{aligned}$$

Wir sind aber nur sekundär an der erwarteten Anzahl der Punkte jenseits von $Ring^{(l)}$ interessiert. Was wir eigentlich erfahren wollen, ist, wie groß die erwarteten *Kosten* $E[Y_C]$ zur Berechnung der konvexen Hülle sind. Wir werden daher in der bisherigen Rechnung statt mit der Anzahl i von Punkten mit der Laufzeit $T(i)$ eines Algorithmus zur Berechnung der konvexen Hülle von i Punkten rechnen. Die worst-case Laufzeit zur Berechnung der konvexen Hülle von i Punkten ist bei Verwendung eines Standardalgorithmus im Zweidimensionalen, z.B. *Graham's Scan*,

$$T(i) = c \cdot i \log i \qquad , c \text{ konstant.}$$

Setzen wir dies in unsere bisherige Rechnung ein, so erhalten wir

Lemma 3.4 *Strebt $n \to \infty$, dann sind für $0 < p < 1$ und eine Konstante $c' > 0$ die erwarteten Kosten von Schritt 2*

$$E[Y_C] \leq c' \cdot n \log n \cdot (1-p)^{l_n}.$$

Beweis

[8] Aufgrund der Definition der Wahrscheinlichkeitsverteilung P ist es anschaulich, daß $E[Y_i|\mathcal{W}_i] \leq E[Y_j|\mathcal{W}_j]$ ist, falls $i > j$ ist. Eine empirische Bestätigung liefern die Simulationsergebnisse aus Abschnitt 5.

$$
\begin{aligned}
E[Y_C] &= \sum_{i=0}^{n} P(Y=i)\cdot T(i) \\
&= \sum_{i=0}^{n} P(Y=i)\cdot c\cdot i\log i \\
&\leq c\cdot\log n\cdot\sum_{l=0}^{\infty} P(\mathcal{W}_l)\cdot\sum_{i=0}^{n}\frac{P(\{Y_l=i\}\cap\mathcal{W}_l)}{P(\mathcal{W}_l)}\cdot i \\
&= c\cdot\log n\cdot\sum_{l=0}^{\infty} P(\mathcal{W}_l)\cdot\sum_{i=0}^{n} P(Y_l=i|\mathcal{W}_l)\cdot i \\
&= c\cdot\log n\cdot\sum_{l=0}^{\infty} P(\mathcal{W}_l)\cdot E[Y_l|\mathcal{W}_l].
\end{aligned}
$$

Wählen wir wieder $l'\leq l_0$ aus Vermutung 3.1, dann ist

$$
\begin{aligned}
E[Y_C] &\leq c\cdot\log n\cdot\left(E[Y_{l'}]+\sum_{r=0}^{l'-1}P(\overline{R^{(l'-r)}})\cdot E[Y_{l'-(r+1)}]\right) \\
&= c\cdot\log n\cdot\left(n\cdot(1-p)^{l'}+\sum_{r=0}^{l'-1}P(\overline{R^{(l'-r)}})\cdot\frac{n\cdot(1-p)^{l'}}{(1-p)^{r+1}}\right) \\
&= c\cdot n\log n\cdot(1-p)^{l'}\cdot\left(1+\sum_{r=0}^{l'-1}\frac{P(\overline{R^{(l'-r)}})}{(1-p)^{r+1}}\right).
\end{aligned}
$$

Mit Vermutung 3.2 können wir für l' die Schachtelungstiefe l_n verwenden und mit Lemma 3.3 gilt

$$
E[Y_C]\leq c\cdot n\log n\cdot(1-p)^{l_n}\cdot\left(1+\sum_{r=0}^{l_n-1}\frac{1}{2^{\left(\frac{1}{1-p}\right)^r}}\cdot\frac{1}{(1-p)^{r+1}}\right).
$$

Die Summe $\sum_{r=0}^{l_n-1}\frac{1}{2^{\left(\frac{1}{1-p}\right)^r}}\cdot\frac{1}{(1-p)^{r+1}}$ ist auch für große n beschränkt, da die Reihe $\frac{1}{1-p}\cdot\sum_{r=0}^{\infty}\frac{1}{2^{\left(\frac{1}{1-p}\right)^r}}\cdot\frac{1}{(1-p)^r}$ für $0<p<1$ absolut konvergiert und damit unterhalb einer Konstanten liegt. Dies zeigen wir mit dem Wurzelkriterium.

$$
\begin{aligned}
\varrho &:= \overline{\lim_{n\to\infty}}\sqrt[n]{|a_n|} \\
&= \lim_{n\to\infty}\frac{1}{\sqrt[n]{2^{\left(\frac{1}{1-p}\right)^n}\cdot(1-p)^n}}<1 \\
\Leftrightarrow\ & \lim_{n\to\infty}\frac{1}{2^{\left[\left(\frac{1}{1-p}\right)^n\cdot\frac{1}{n}\right]}}<(1-p)
\end{aligned}
$$

$$\begin{aligned}
&\Leftrightarrow \quad \lim_{n\to\infty} -\left(\frac{1}{1-p}\right)^n \cdot \frac{1}{n} \cdot \log 2 < \log(1-p) \\
&\Leftrightarrow \quad \lim_{n\to\infty} n \log|(1-p)| - \log n + \log\log 2 > \log|\log(1-p)|.
\end{aligned}$$

Für $p \to 1$ divergiert die Summe wegen des Faktors $\frac{1}{1-p}$ und bei $p = 0$ strebt sie wegen $\sum_{r=0}^{l_n-1} \frac{1}{2\left(\frac{1}{1-p}\right)^r} \cdot \frac{1}{(1-p)^r} = \sum_{r=0}^{l_n-1} \frac{1}{2}$ für große n gegen Unendlich. □

Es ist nun leicht, Lemma 3.5 zu zeigen.

Lemma 3.5 *Für $n \to \infty$, $0 < p < 1$ und eine Konstante $c' > 1$ ist $E[Y_C] \leq n^{\left(\frac{1}{c'}\right)}$.*

Beweis

Da $c' > 1$ ist, existiert ein α, so daß

$$0 < \left(\frac{1-c'}{c'}\right) \cdot \frac{1}{\log(1-p)} < \alpha < \frac{1}{|\log(1-p)|}$$

ist. Daraus folgt

$$\begin{aligned}
\left(\frac{1-c'}{c'}\right) \cdot \frac{1}{\log(1-p)} &< \alpha \\
\frac{1-c'}{c'} - \alpha \log(1-p) &> 0.
\end{aligned}$$

Strebt $n \to \infty$, dann ist daher für jede Konstante $c > 0$

$$\begin{aligned}
n^{\frac{1}{c'}-1-\alpha\log(1-p)} &> \log(n^c) \\
n^{\left(\frac{1}{c'}\right)} &> n^{1+\alpha\log(1-p)} \cdot \log(n^c).
\end{aligned}$$

Da $\lfloor \alpha \log n \rfloor = l_n$ ist, ist dies gleichbedeutend mit

$$n^{\left(\frac{1}{c'}\right)} > \frac{c}{1-p} \cdot n \log n \cdot (1-p)^{l_n}.$$

Mit Lemma 3.4 erhält man die Behauptung

$$n^{\left(\frac{1}{c'}\right)} > E[Y_C].$$

□

Zusammenfassend ergibt sich die Gültigkeit von

Satz 3.1 *Es seien die n Punkte $a_1, \ldots, a_n$ aus $\mathbb{R}^2$ unabhängig voneinander gewählt und f sei eine Wahrscheinlichkeitsdichte auf $\mathbb{R}^2$, für die eine Rechteck-Ring-Schachtelung existiert, so daß die Eigenschaften* **E1** *und* **E2** *erfüllt sind. Dann berechnet Algorithmus* `A` *für $0 < p < 1$ und große Werte n die konvexe Hülle in erwarteter Laufzeit*

$$E[\texttt{konvexe Hülle}] \leq (2 + \frac{2}{p} + c_{oh}) \cdot n + n.$$

Beweis

Da die Zufallsvariablen X und Y_C über demselben Ergebnisraum Ω definiert sind, belaufen sich die erwarteten Kosten asymptotisch auf

$$\begin{aligned} E[\texttt{konvexe Hülle}] &= E[X + Y_C] \\ &= E[X] + E[Y_C] \\ &\leq (2 + \frac{2}{p} + c_{oh}) \cdot n + n. \end{aligned}$$

■

3.4. Laufzeit des Verfahrens bei verschiedenen Wahrscheinlichkeiten p

Die Untersuchungen des letzten Abschnitts zeigen, daß sowohl die erwarteten Kosten des ersten Schrittes als auch die des zweiten Schrittes von Algorithmus `A` von der Wahrscheinlichkeit p des zentralen Rechtecks abhängen. Die Frage ist nun, welchen Einfluß p auf die erwartete Laufzeit des Verfahrens hat.

Hält man n fest und variiert p, so erkennt man leicht, daß sich die Kosten für Schritt 1, $E[X] = (2 + \frac{2}{p} + c_{oh}) \cdot n$, hyperbolisch verhalten; sie werden minimal für $p \to 1$ und wachsen ins Unendliche für $p \to 0$.

Für Schritt 2 ist die Angabe nicht so einfach. Es ist wegen Lemma 3.5

$$\begin{aligned} E[Y_C] &\leq c \cdot n \log n \cdot (1-p)^{l_n} \cdot \left(1 + \sum_{r=0}^{l_n - 1} \frac{1}{2^{\left(\frac{1}{1-p}\right)^r}} \cdot \frac{1}{(1-p)^{r+1}}\right) \\ &\leq n^{\left(\frac{1}{c'}\right)} \cdot \left(1 + \sum_{r=0}^{l_n - 1} \frac{1}{2^{\left(\frac{1}{1-p}\right)^r}} \cdot \frac{1}{(1-p)^{r+1}}\right), \end{aligned}$$

wobei $c, c' > 0$ beliebige Konstanten sind. Man sieht, daß $E[Y_C]$ von dem Wert der Summe abhängt.

Untersuchen wir nun die Abhängigkeit der Summe von p. Zunächst stellen wir fest

Lemma 3.6

$$a_r = \left(\frac{1}{2}\right)^{\left(\frac{1}{1-p}\right)^r} \left(\frac{1}{1-p}\right)^{r+1}$$

ist eine Nullfolge.

Beweis

Da alle Summenglieder $a_r > 0$ sind und die erste Ableitung der Summenglieder

$$a_r{}' = \left(\frac{1}{2}\right)^{\left(\frac{1}{1-p}\right)^r} \left(\frac{1}{1-p}\right)^{r+1} \ln(1-p) \left(\left(\frac{1}{1-p}\right)^r \ln 2 - 1\right)$$

für große r kleiner als 0 ist, gibt es ein r, ab dem a_r monoton fällt. Benutzt man die Regel von de L'Hospital, dann sieht man, daß wegen

$$\begin{aligned} \lim_{r\to\infty} a_r &= \lim_{r\to\infty} \frac{\left(\frac{1}{1-p}\right)^{r+1}}{2^{\left(\frac{1}{1-p}\right)^r}} \\ &= \lim_{r\to\infty} \frac{\left(\frac{1}{1-p}\right)}{2^{\left(\frac{1}{1-p}\right)^r} \ln 2} \\ &\to 0 \end{aligned}$$

a_r eine Nullfolge ist. □

Da aufgrund

$$\begin{aligned} \left(\frac{1}{1-p}\right)^r \ln 2 &\leq 1 \\ \left(\frac{1}{1-p}\right)^r &\leq \frac{1}{\ln 2} \\ -r \cdot \ln(1-p) &\leq -\ln\ln 2 \end{aligned}$$

a_r bis $r = \left\lfloor \frac{\ln\ln 2}{\ln(1-p)} \right\rfloor$ steigt, ist

$$\sum_{r=0}^{l_n-1} \left(\frac{1}{2}\right)^{\left(\frac{1}{1-p}\right)^r} \left(\frac{1}{1-p}\right)^{r+1} \leq \frac{1}{2}\left(\frac{1}{1-p}\right)^{\frac{\ln\ln 2}{\ln(1-p)}} + \int_{r=0}^{\infty} \left(\frac{1}{2}\right)^{\left(\frac{1}{1-p}\right)^r} \left(\frac{1}{1-p}\right)^{r+1} dr.$$

Daraus ergibt sich die Abschätzung

Lemma 3.7

$$\begin{aligned} \sum_{r=0}^{l_n-1} \left(\frac{1}{2}\right)^{\left(\frac{1}{1-p}\right)^r} \left(\frac{1}{1-p}\right)^{r+1} &\leq \frac{1}{2}\left(\frac{1}{1-p}\right)^{\frac{\ln\ln 2}{\ln(1-p)}} + \frac{1}{2\ln(1-p)\ln(2)(p-1)} \\ &= S(p). \end{aligned}$$

Beweis

$$\begin{aligned} \frac{1}{1-p} \cdot \int_{r=0}^{\infty} \left(\frac{1}{2}\right)^{\left(\frac{1}{1-p}\right)^r} \left(\frac{1}{1-p}\right)^r dr &= \frac{1}{1-p} \cdot \frac{1}{\ln(2)\ln(1-p)} \cdot \left(\frac{1}{2}\right)^{\left(\frac{1}{1-p}\right)^r} \Bigg|_0^\infty \\ &= \frac{1}{2\ln(1-p)\ln(2)(p-1)}. \end{aligned}$$

□

Der Graph von $S(p)$ für $0 < p < 1$ ist in Abbildung 5 dargestellt.

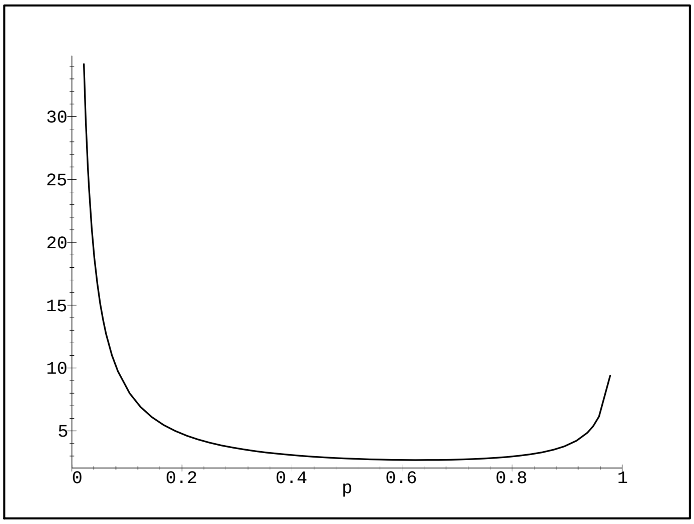

Abbildung 5: Graph von $S(p)$

Der Graph bestätigt das Verhalten der Reihe für $0 < p < 1$, wie wir es bereits im Beweis von Lemma 3.4 auf Seite 33 diskutiert haben. Für $p \to 0$, bzw. $p \to 1$, divergiert die Reihe.

Qualitativ wird $E[Y_C]$ dem Verlauf von $S(p)$ folgen. Man darf jedoch nicht aus den Augen verlieren, daß der wahre Verlauf von $E[Y_C]$ nicht so extrem sein wird. Denn für $p = 1$ liegen alle n Punkte im innersten Rechteck, $D^{(0)}$, so daß uns die Ringkonstruktion überhaupt keinen Rechenzeitgewinn bringt. $E[Y_C]$ beträgt dann $c \cdot n \log n$. Gleiches gilt für $p \to 0$. So groß die Wahrscheinlichkeit, daß sich die Punkte in sehr weit außenliegenden Ringen befinden, auch sein mag; es werden nie mehr als n Punkte jenseits des äußersten vollbesetzten Ringes liegen. Somit gilt auch für $p \to 0$: $\quad E[Y_C] \leq c \cdot n \log n$.

Was aber der Verlauf des Graphen nahelegt, ist, daß das Minimum von $E[Y_C]$ mehr zur 1 hin liegt. D.h., daß $E[Y_C]$ für „größere" p klein wird. Diese Vermutung werden wir in Abschnitt 5 durch numerische Simulationen bestätigt sehen.

Insgesamt stellen wir fest, daß für die meisten p die Kosten von Schritt 1 die Laufzeit des Verfahrens dominieren werden. Für p nahe bei 1 und kleine Eingabegrößen n können die Kosten von Schritt 2 dominieren. Für große n gilt jedoch als direkte Folge von Lemma 3.5

Satz 3.2 *Strebt $n \to \infty$, dann dominieren – unabhängig von p – die erwarteten Kosten von Schritt 1 die erwartete Laufzeit von Algorithmus* `A`.

3.5. Ein einfacheres Experiment zur Abschätzung von $E[Y_l|\mathcal{W}_l]$

Im Beweis von Lemma 3.4 haben wir Vermutung 3.1 benutzt. In diesem Abschnitt wollen wir untersuchen, wie sich die involvierten Wahrscheinlichkeiten in Abhängigkeit von n verhalten, so daß für kleine Werte n diese Ungleichung nicht gilt, während sie – abhängig von der Schachtelungstiefe l – für große n gilt. Dazu betrachten wir ein erheblich simpleres Wahrscheinlichkeitsexperiment.

Die Rechteck-Ring-Schachtelung bestehe diesmal lediglich aus vier Ringen $Ring^{(0)}$, $Ring^{(1)}$, $Ring^{(2)}$ und $Ring^{(3)}$. Ein Ring sei *vollbesetzt* genau dann, wenn ein ausgezeichnetes Eckfeld – o.B.d.A. das linke obere Eckfeld jedes Ringes, Eckfeld 1, – einen Punkt enthält. Die Schachtelung besteht aus vier Ringen, damit zum einen die Bestimmung der Wahrscheinlichkeiten nicht zu schwer wird. Zum anderen haben wir bei einer Schachtelung mit mehreren Schachtelungstiefen die Möglichkeit, zu sehen, wie bei einem vorgegebenen Wert von n die Schachtelungstiefe auf die Gültigkeit der Ungleichung

$$E[Y_l|\mathcal{W}_l] \leq E[Y_l]$$

Einfluß nimmt.

Der Übersichtlichkeit halber werden wir die Nomenklatur vereinfachen. $Ring^{(0)}$ sei $D^{(0)}$. Eckfeld 1 von $Ring^{(1)}$ sei $D^{(1)}$; alle restlichen Felder von $Ring^{(1)}$ fassen wir unter $D^{(2)}$ zusammen. Ähnlich verfahren wir mit $Ring^{(2)}$ und $Ring^{(3)}$:

$$\begin{aligned} D^{(2i-1)} &\hat{=} \text{Eckfeld 1 von } Ring^{(i)} \\ D^{(2i)} &\hat{=} \text{Vereinigung aller restlichen Felder von } Ring^{(i)} \\ & \quad i = 1, \ldots, 3. \end{aligned}$$

Die Definition des elementaren Wahrscheinlichkeitsraumes $(\Omega_e, \mathcal{A}_e, P_e)$ ist analog zu Abschnitt 3.2:

$$\begin{aligned} \Omega_e &= \{D^{(0)}, \ldots, D^{(6)}\}, \\ \mathcal{A}_e &= Pot(\Omega_e) \end{aligned}$$

und für $0 < p < 1$

$$\begin{aligned} P_e(D^{(0)}) &= P_e(Ring^{(0)}), \\ P_e(D^{(2i-1)}) + P_e(D^{(2i)}) &= P_e(Ring^{(i)}), \\ P_e(Ring^{(0)}) &= p, \\ P_e(Ring^{(i+1)}) &= (1-p) \cdot P_e(Ring^{(i)}), \\ P_e(Ring^{(3)}) &= 1 - 3p + 3p^2 - p^3. \end{aligned}$$

Im Gegensatz zu Abschnitt 3.2 bevorteiligen wir diesmal Eckfeld 1 gegenüber den restlichen Feldern eines Ringes. Die Wahrscheinlichkeit sei viermal so groß wie im vorherigen Modell:

$$\begin{aligned}
P_e(D^{(1)}) &= 4 \cdot \frac{p(1-p)}{8} \\
P_e(D^{(3)}) &= 4 \cdot \frac{p(1-p)^2}{8 \cdot 2} \\
P_e(D^{(5)}) &= 4 \cdot \frac{1-3p+3p^2-p^3}{8 \cdot 3}.
\end{aligned}$$

Davon erhoffen wir uns, daß die Wahrscheinlichkeit, daß ein Ring vollbesetzt ist, nicht zu sehr von der im ursprünglichen Modell abweicht.

Zur Modellierung der n-Auswahlen benutzen wir wieder den Produktraum $(\Omega, \mathcal{A}, P)$ und übernehmen die Notationen und Definitionen aus Abschnitt 3.3.2. Wir werden nun für das vorliegende Wahrscheinlichkeitsexperiment $E[Y_l|\mathcal{W}_l]$ exakt ausrechnen und mit $E[Y_l]$ vergleichen. Dazu benötigt man die Wahrscheinlichkeiten $P(\{Y_l = i\} \cap \mathcal{W}_l)$ und $P(\mathcal{W}_l)$.

Zunächst bestimmen wir $P(\mathcal{W}_l)$. $\mathcal{W}_l$ ist definitionsgemäß die Menge aller n-Auswahlen, die $Ring^{(l)}$ „erzeugen". Wir werden die Elemente von $\mathcal{W}_1$ in aller Ausführlichkeit angeben und für $\mathcal{W}_2$ und $\mathcal{W}_3$ nur deren Wahrscheinlichkeiten nennen; denn die Konstruktion dieser Mengen verläuft analog. Der Einfachheit halber soll $D^{(j)}$ als j und $P_e(D^{(j)})$ als p_j bezeichnet werden.

Jedes Element $\omega = (\omega_1, \ldots, \omega_n)$ von $\mathcal{W}_1$ muß mindestens eine 1 enthalten; die übrigen Einträge dürfen beliebig aus $\{0, 2, 4, 6\}$ gewählt werden. So wird garantiert, daß $Ring^{(1)}$ der äußerste vollbesetzte Ring ist. Sei k, $1 \leq k \leq n$, die Anzahl der Einsen von ω. Man hat dann $\binom{n}{k}$ Möglichkeiten, die k Einsen auf n Plätze zu verteilen. Auf die restlichen $n-k$ Plätze werden (mit Wiederholung) Elemente aus $\{0, 2, 4, 6\}$ gesetzt, das sind 4^{n-k} Möglichkeiten. Die Mächtigkeit von $\mathcal{W}_1$ ist daher

$$|\mathcal{W}_1| = \sum_{k=1}^{n} \binom{n}{k} \cdot 4^{n-k} = 5^n - 4^n.$$

$P(\mathcal{W}_1)$ beträgt aufgrund der Gestalt der Elemente von $\mathcal{W}_1$

$$\begin{aligned}
P(\mathcal{W}_1) &= \sum_{k=1}^{n} \binom{n}{k} \cdot {p_1}^k \cdot (p_0 + p_2 + p_4 + p_6)^{n-k} \\
&= (1 - p_3 - p_5)^n - (1 - p_1 - p_3 - p_5)^n.
\end{aligned}$$

Analog erhält man

$$\begin{aligned}
P(\mathcal{W}_2) &= \sum_{k=1}^{n} \binom{n}{k} \cdot {p_3}^k \cdot (p_0 + p_1 + p_2 + p_4 + p_6)^{n-k} \\
&= (1 - p_5)^n - (1 - p_3 - p_5)^n
\end{aligned}$$

und

$$
\begin{aligned}
P\left(\mathcal{W}_3\right) &= \sum_{k=1}^{n} \binom{n}{k} \cdot {p_5}^k \cdot (p_0+p_1+p_2+p_3+p_4+p_6)^{n-k} \\
&= 1-(1-p_5)^n .
\end{aligned}
$$

Nun zu $P\left(\{Y_l=i\}\cap\mathcal{W}_l\right)$; $\{Y_l=i\}\cap\mathcal{W}_l$ ist die Menge aller n-Auswahlen, die $Ring^{(l)}$ „erzeugen" und bei denen i Punkte jenseits von $Ring^{(l)}$ liegen. Wir behandeln wieder $\{Y_1=i\}\cap\mathcal{W}_1$ ausführlich und skizzieren die anderen Fälle.

Elemente $\omega=(\omega_1,\ldots,\omega_n)$ von $\{Y_1=i\}\cap\mathcal{W}_1$ besitzen i Einträge ungleich 0, von denen mindestens einer gleich 1 sein muß. Damit $Ring^{(1)}$ der äußerste vollbesetzte Ring ist, dürfen die restlichen Einträge ungleich 0 nur aus $\{2,4,6\}$ stammen.

Sei k die Anzahl der Einsen. Es bestehen $\binom{n}{k}$ Möglichkeiten, die Einsen zu plazieren. Anschließend verteilen wir die übrigen $i-k$ Einträge ungleich 0, d.h. aus $\{2,4,6\}$, auf noch $n-k$ freie Plätze. Das sind $\binom{n-k}{i-k}\cdot 3^{i-k}$ Möglichkeiten. Die restlichen Plätze werden mit Nullen aufgefüllt. $\{Y_1=i\}\cap\mathcal{W}_1$ hat daher die Mächtigkeit

$$
\left|\{Y_1=i\}\cap\mathcal{W}_1\right| = \sum_{k=1}^{i}\binom{n}{k}\cdot\binom{n-k}{i-k}\cdot 3^{i-k}.
$$

Daraus läßt sich $P\left(\{Y_1=i\}\cap\mathcal{W}_1\right)$ ableiten:

$$
\begin{aligned}
P\left(\{Y_1=i\}\cap\mathcal{W}_1\right) &= \sum_{k=1}^{i}\binom{n}{k}{p_1}^k\cdot\binom{n-k}{i-k}(p_2+p_4+p_6)^{i-k}\cdot{p_0}^{n-i} \\
&= \binom{n}{i}(p_2+p_4+p_6)^i{p_0}^{n-i}\left(\left(1+\frac{p_1}{p_2+p_4+p_6}\right)^i-1\right).
\end{aligned}
$$

Die Konstruktion für $\{Y_2=i\}\cap\mathcal{W}_2$ und $\{Y_3=i\}\cap\mathcal{W}_3$ ist ähnlich.

$$
\begin{aligned}
P\left(\{Y_2=i\}\cap\mathcal{W}_2\right) &= \sum_{k=1}^{i}\binom{n}{k}{p_3}^k\cdot\binom{n-k}{i-k}(p_4+p_6)^{i-k}\cdot(p_0+p_1+p_2)^{n-i} \\
&= \binom{n}{i}(p_4+p_6)^i(p_0+p_1+p_2)^{n-i}\left(\left(1+\frac{p_3}{p_4+p_6}\right)^i-1\right)
\end{aligned}
$$

und

$$
\begin{aligned}
P\left(\{Y_3=i\}\cap\mathcal{W}_3\right) &= \sum_{k=1}^{i}\binom{n}{k}{p_5}^k\cdot\binom{n-k}{i-k}{p_6}^{i-k}\cdot(p_0+p_1+p_2+p_3+p_4)^{n-i} \\
&= \binom{n}{i}{p_6}^i(1-p_5-p_6)^{n-i}\left(\left(1+\frac{p_5}{p_6}\right)^i-1\right).
\end{aligned}
$$

Wir können nun

$$E[Y_l|\mathcal{W}_l] = \frac{1}{P(\mathcal{W}_l)} \cdot \sum_{i=0}^{n} P(\{Y_l = i\} \cap \mathcal{W}_l) \cdot i$$

ausrechnen, bzw. uns die Graphen anzeigen lassen, und mit $E[Y_l]$ vergleichen. Die Gleichungen der Erwartungswerte von Y_1, Y_2 und Y_3 sind, leicht aufzustellen.

$$\begin{aligned} E[Y_1] &= (1-p) \cdot n \\ E[Y_2] &= (1-2p+p^2) \cdot n \\ E[Y_3] &= (1-3p+3p^2-p^3) \cdot n. \end{aligned}$$

In den folgenden Abbildungen 6, 7 und 8 ist der Verlauf der Differenz $E[Y_l] - E[Y_l|\mathcal{W}_l]$, $l = 1, 2, 3$, für $n = 60$ und $p \in (0, 1)$ wiedergegeben.

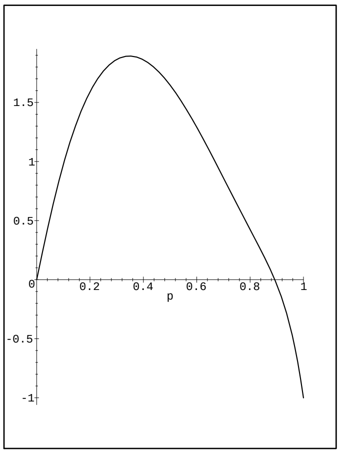

Abbildung 6:
$E[Y_1] - E[Y_1|\mathcal{W}_1]$
$n = 60$

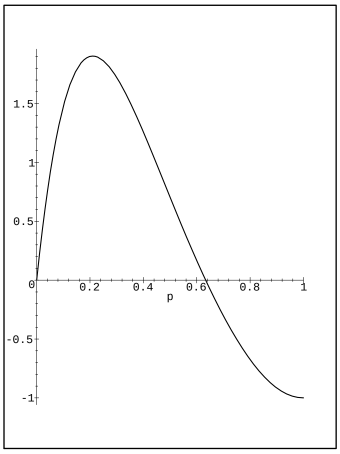

Abbildung 7:
$E[Y_2] - E[Y_2|\mathcal{W}_2]$
$n = 60$

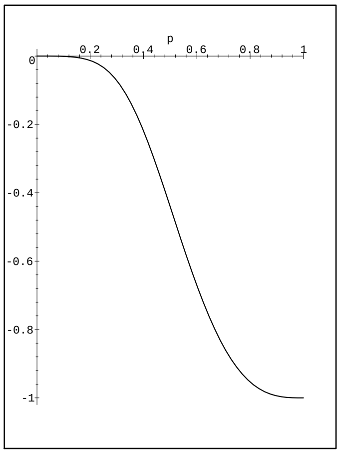

Abbildung 8:
$E[Y_3] - E[Y_3|\mathcal{W}_3]$
$n = 60$

Man erkennt, daß die Differenz von der Schachtelungstiefe l, der Wahrscheinlichkeit p und der Eingabegröße n abhängt. Wächst l und läßt man n und p konstant, so verringert sich der Wert der Differenz und wird sogar negativ. Läßt man jedoch l und p konstant, dann ist bei sehr großen Eingabewerten n auch für p nahe bei 1 die Ungleichung $E[Y_l|\mathcal{W}_l] \geq E[Y_l]$ gültig. Für $p = 0.99$ ist dies in den Abbildungen 9, 10 und 11 illustriert.

Die Auswertung dieses einfacheren Wahrscheinlichkeitsexperiments zeigt also, daß, falls n groß und p fest ist, ein l_0 existiert, so daß für $l \leq l_0$

$$E[Y_l|\mathcal{W}_l] \leq E[Y_l]$$

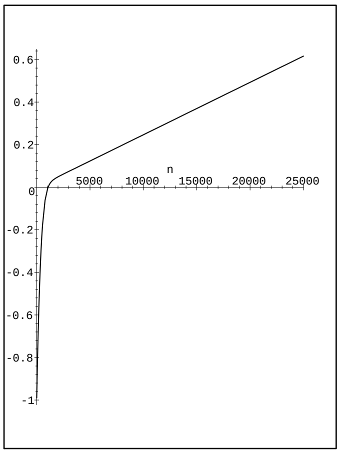

Abbildung 9:
$E[Y_1] - E[Y_1|\mathcal{W}_1]$
$p = 0.99$
$n = 2.5 \cdot 10^4$

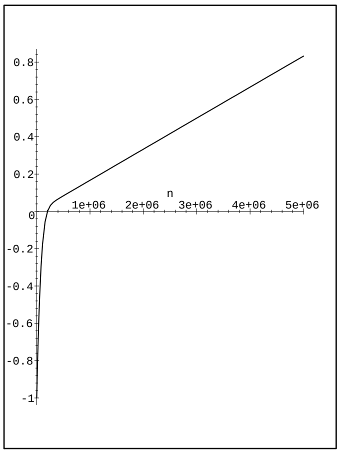

Abbildung 10:
$E[Y_2] - E[Y_2|\mathcal{W}_2]$
$p = 0.99$
$n = 5 \cdot 10^6$

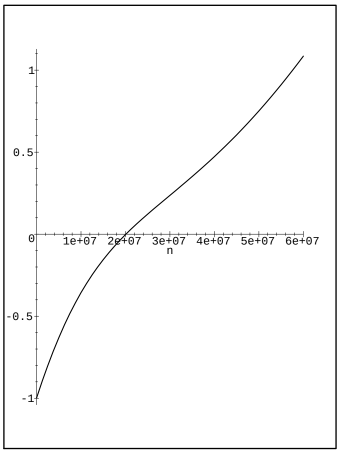

Abbildung 11:
$E[Y_3] - E[Y_3|\mathcal{W}_3]$
$p = 0.99$
$n = 6 \cdot 10^7$

ist.

Aber warum ist das so? Vergleichen wir dazu $P(Y_l = i|\mathcal{W}_l)$ und $P(Y_l = i)$ für $l = 1, 2, 3$ in den Abbildungen 12-20. Die durchgezogene Linie gibt den Verlauf von $P(Y_l = i|\mathcal{W}_l)$ und die gestrichelte Linie den von $P(Y_l = i)$ wieder.

Wir vergleichen nun exemplarisch den Verlauf von $P(Y_l = i|\mathcal{W}_l)$ und $P(Y_l = i)$ für $l = 2$. Ist n klein, dann existiert ein $i_0 \in [1, n]$, so daß für $i > i_0$ die Wahrscheinlichkeit $P(Y_2 = i|\mathcal{W}_2) \geq P(Y_2 = i)$ ist (vgl. Abbildung 15). Für kleine Werte n und einige Werte von p ist deshalb $E[Y_2|\mathcal{W}_2] \geq E[Y_2]$. Wird n größer, dann kehrt sich das Verhältnis der Wahrscheinlichkeiten um. Es gibt dann ein i_0, so daß für $i > i_0$ die Wahrscheinlichkeit $P(Y_2 = i) \geq P(Y_2 = i|\mathcal{W}_2)$ ist (vgl. Abbildungen 16, 17). Das erklärt, weshalb für beliebige, aber feste p und große n, abhängig von der Schachtelungstiefe l, die Ungleichung $E[Y_l|\mathcal{W}_l] \leq E[Y_l]$ gilt.

In Abschnitt 5 werden wir sehen, daß beim Wahrscheinlichkeitsexperiment „**Ziehung einer n-Auswahl aus einer Rechteck-Ring-Schachtelung**“ für große n ein Wert $b > 1$ existiert, so daß die Schachtelungstiefe $l_n = \lfloor \alpha \log_b n \rfloor$ ausreicht, damit für $l \leq l_n$ die Ungleichung $E[Y_l|\mathcal{W}_l] \leq E[Y_l]$ erfüllt ist[9].

Die Auswertung dieses einfacheren Experiments gibt aber auch Anlaß zu einer anderen Vermutung, nämlich daß stets $E[Y_l|\mathcal{W}_l] \leq k + E[Y_l]$ ist, wobei k die Anzahl der Eckfelder sei, die belegt sein müssen (vgl. Abbildung 7 und 8; $E[Y_l] - E[Y_l|\mathcal{W}_l]$ ist immer größer als -1. Bei dem vereinfachten Wahrscheinlichkeitsexperiment war

[9] Es geht also um eine empirische Überprüfung von Vermutung 3.2.

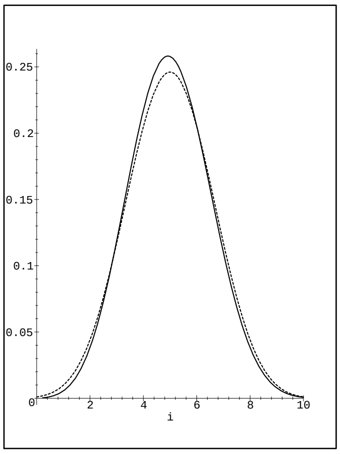

Abbildung 12:
$P(Y_1 = i)$, $P(Y_1 = i|\mathcal{W}_1)$
$p = 0.5$
$n = 10$

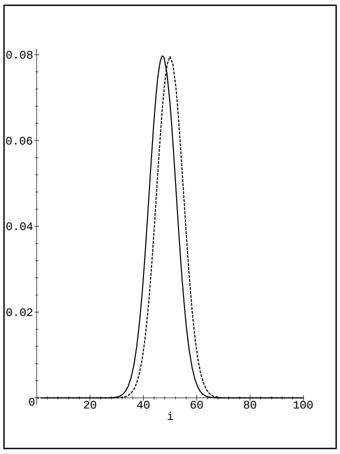

Abbildung 13:
$P(Y_1 = i)$, $P(Y_1 = i|\mathcal{W}_1)$
$p = 0.5$
$n = 100$

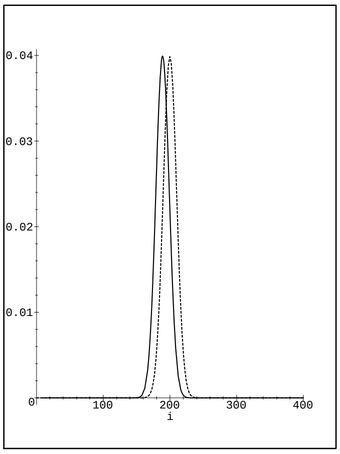

Abbildung 14:
$P(Y_1 = i)$, $P(Y_1 = i|\mathcal{W}_1)$
$p = 0.5$
$n = 400$

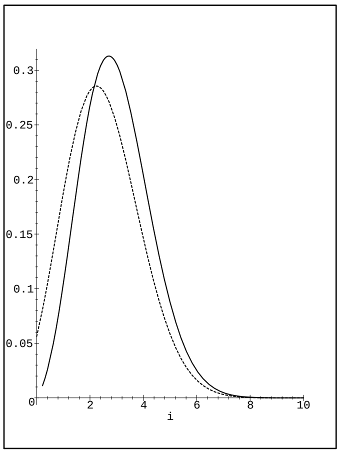

Abbildung 15:
$P(Y_2 = i)$, $P(Y_2 = i|\mathcal{W}_2)$
$p = 0.5$
$n = 10$

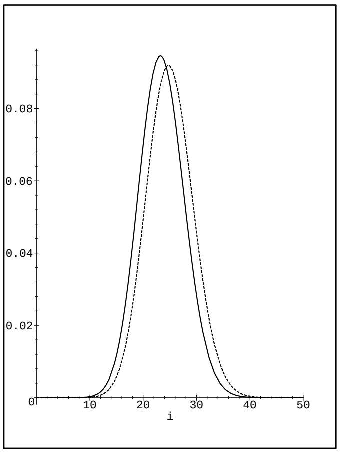

Abbildung 16:
$P(Y_2 = i)$, $P(Y_2 = i|\mathcal{W}_2)$
$p = 0.5$
$n = 100$

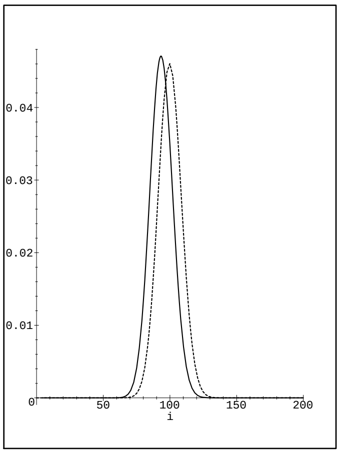

Abbildung 17:
$P(Y_2 = i)$, $P(Y_2 = i|\mathcal{W}_2)$
$p = 0.5$
$n = 400$

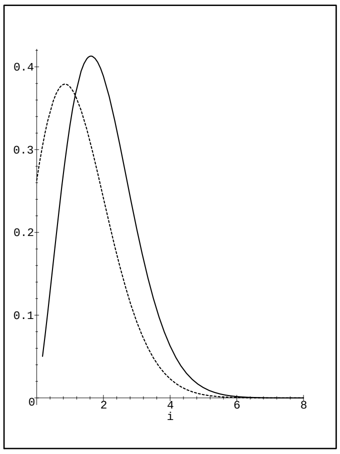

Abbildung 18:
$P(Y_3 = i)$, $P(Y_3 = i|\mathcal{W}_3)$
$p = 0.5$
$n = 10$

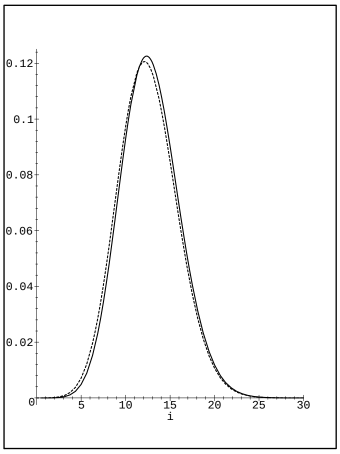

Abbildung 19:
$P(Y_3 = i)$, $P(Y_3 = i|\mathcal{W}_3)$
$p = 0.5$
$n = 100$

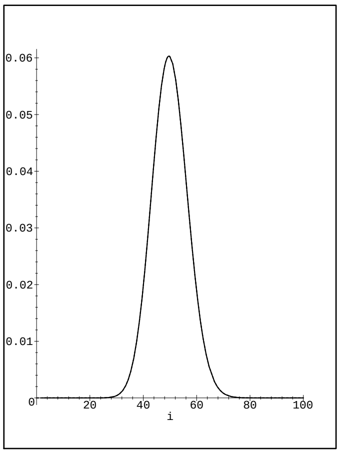

Abbildung 20:
$P(Y_3 = i)$, $P(Y_3 = i|\mathcal{W}_3)$
$p = 0.5$
$n = 400$

$k = 1$. Ob dies für größere k immer noch der Fall ist, werden wir ebenfalls in Abschnitt 5 sehen.) Daß diese Vermutung sinnvoll ist, liegt auf der Hand. $E[Y_l|\mathcal{W}_l]$ kann man sich nämlich wie folgt zusammengesetzt denken: besetze zunächst die k Eckfelder, dann verteile die restlichen Punkte gemäß der gegebenen Wahrscheinlichkeitsverteilung, was im Mittel auf jedenfall den Erwartungswert $E[Y_l]$ als obere Schranke haben wird. Insgesamt erhält man daher $k + E[Y_l]$ als obere Abschätzung.

4. Konvexe Hüllen im $\mathbb{R}^2$ bei einer Polygon-Ring-Schachtelung

Wir haben gesehen, daß man mit einer Rechteck-Ring-Schachtelung die konvexe Hülle in linearer, erwarteter Laufzeit bestimmen kann. Nun stellt sich die Frage, ob dieses Resultat auch mit einem anderen Aussiebverfahren möglich ist. Welche Laufzeit ergibt sich z.B., wenn man ein Aussiebverfahren benutzt, das nicht auf Rechtecken sondern auf geschachtelten Polygonen beruht?

Anstatt den $\mathbb{R}^2$ mit Rechteck-Ringen zu überdecken wie in Abschnitt 3, verwenden wir diesmal seitenparallele, konzentrisch geschachtelte reguläre Polygone. Die Konstruktion verläuft wie folgt:

Das innerste k-Eck, $Ring^{(0)}$, sei durch k Geraden $l_1, \ldots, l_k$ gegeben. Um $Ring^{(0)}$ legen wir nun ein weiteres reguläres k-Eck, P_1, das durch k Geraden $l_{k+1}, \ldots, l_{2k}$ begrenzt sein soll, wobei jeweils zwei Geraden l_i und l_{k+i} parallel zueinander seien, d.h. $l_1 \parallel l_{k+1}$, $l_2 \parallel l_{k+2}$, …, $l_k \parallel l_{2k}$. Die Fläche $P_1 - Ring^{(0)}$ bezeichnen wir nun als $Ring^{(1)}$. Um $Ring^{(1)}$ legen wir ein weiteres Polygon, P_2, welches durch die Linien $l_{2k+1}, \ldots, l_{3k}$ begrenzt ist und für die ebenfalls gilt: $l_i \parallel l_{k+i}$. $Ring^{(2)}$ ist dann $P_2 - P_1$. Analog werden nun die Polygon-Ringe $Ring^{(l)}$, $l = 3, 4, \ldots$, konstruiert, d.h. $Ring^{(l)} = P_l - P_{l-1}$. Im Prinzip handelt es sich also um die gleiche Konstruktion wie bei den Rechteck-Ringen aus Abschnitt 3. Abbildung 21 auf Seite 46 zeigt eine Sechseck-Ring-Konstruktion.

Definition 4.1 (Polygon-Ring-Schachtelung) *Die oben beschriebene Konstruktion heiße* Polygon-Ring-Schachtelung*.*
Der Ring mit Schachtelungstiefe l sei $Ring^{(l)}$.

Betrachten wir nun $Ring^{(1)}$ etwas genauer. Offensichtlich wird $Ring^{(1)}$ von den Begrenzungsgeraden von $Ring^{(0)}$, $l_1, \ldots, l_k$, in Felder unterteilt. Ebenso werden die äußeren Kanten von $Ring^{(1)}$ in *Kantenabschnitte* eingeteilt. Die Frage ist nun, in wieviele Felder ein Polygon-Ring unterteilt wird. Betrachtet man den Übergang von $Ring^{(1)}$ zu $Ring^{(2)}$ (i.a. den Übergang von $Ring^{(i)}$ zu $Ring^{(i+1)}$), dann fällt auf, daß jedes Feld von $Ring^{(1)}$ – außer den Eckfeldern – ein Feld von $Ring^{(2)}$ erzeugt[10]. Jedes Eckfeld erzeugt statt nur einem Feld drei Felder. Dies ist leicht verständlich, wenn man berücksichtigt, daß jedes Eckfeld nach außenhin von zwei Kantenabschnitten

[10]Dies gilt in uneingeschränkter Form nur für $k = 3, 4$. Für k-Ecke mit $k \geq 5$ sieht man, daß sich die Begrenzungslinien der Polygone gegenseitig schneiden. Es kann dann passieren, daß bei ungeschickter Plazierung der Polygon-Ringe solch ein Schnittpunkt gerade innerhalb eines Polygon-Ringes liegt. Dieses Problem kann man aber umgehen, indem man die Polygone so legt, daß Schnittpunkte von Begrenzungslinien genau *auf* den Polygon*kanten* liegen.
Die Folge ist unweigerlich, daß man bei der Zerlegung des $\mathbb{R}^2$ keine freie Wahl mehr hat, wenn man mit dem oben beschriebenen Ansatz die Anzahl der Felder eines Polygon-Ringes bestimmen will.

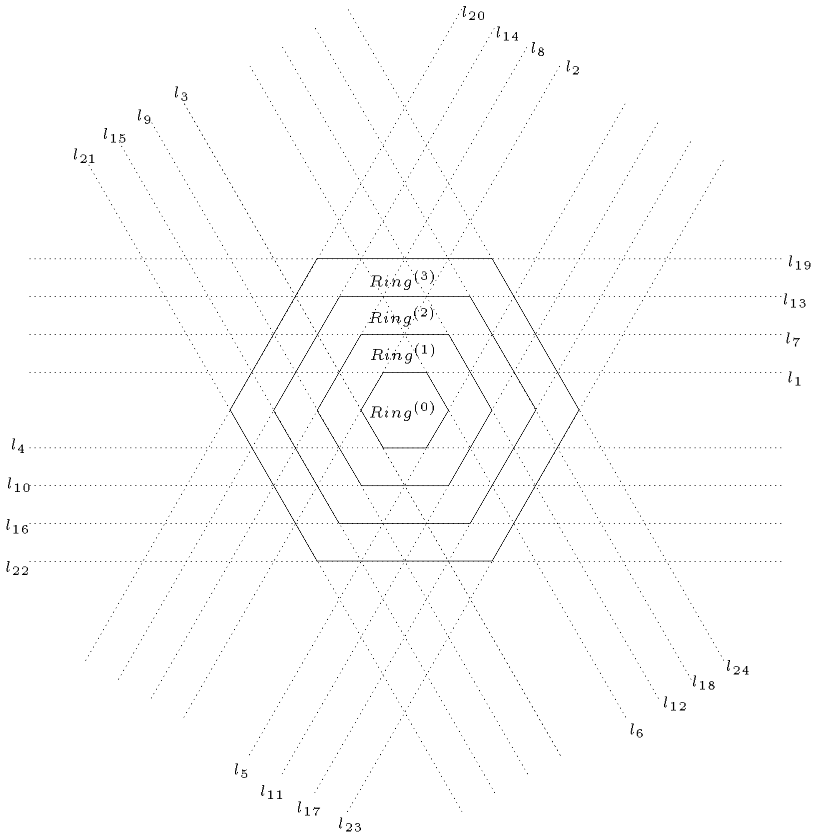

Abbildung 21: Sechseck-Ring-Konstruktion

abgeschlossen wird, und somit zwei „zusätzliche“ Felder erzeugt. Die Frage ist dann: In wieviele Felder wird $Ring^{(1)}$ unterteilt? Es läßt sich jedoch leicht überprüfen, daß $Ring^{(1)}$ aus $2k$ Feldern besteht[11]. Zusammenfassend erhält man folgende Rekursionsgleichung:

Sei $F(i)$ die Anzahl der Felder von $Ring^{(i)}$, dann gilt:

$$F(1) = 2k,$$
$$F(i) = F(i-1) + 2k, \qquad i \geq 2$$

Löst man diese Rekursionsgleichung, so erhält man $F(i) = i \cdot 2k$.

Seien $a_1, \ldots, a_n$ Punkte des $\mathbb{R}^2$. Wie bei der Rechteck-Ring-Schachtelung erreicht man mit der Polygon-Ring-Schachtelung eine Reduktion der Ausgangspunktmenge: Sind in einem Polygon-Ring $Ring^{(l)}$ alle Eckfelder besetzt, so werden alle Punkte in den Polygon-Ringen $Ring^{(0)}, \ldots, Ring^{(l-1)}$ bei der Berechnung der konvexen Hülle

[11] Offensichtlich besitzt $Ring^{(0)}$ k Ecken und k Kanten. Da jede Kante von $Ring^{(0)}$ ein Feld von $Ring^{(1)}$ erzeugt und ebenso jede Ecke von $Ring^{(0)}$, besteht $Ring^{(1)}$ aus $2k$ Feldern.

nicht mehr benötigt. Wir untersuchen nun die erwarteten Kosten von Algorithmus `A` bei einer Polygon-Ring-Schachtelung.

4.1. erwartete Kosten bei Polygon-Ring-Schachtelung

Diesmal gehen wir davon aus, daß für die gegebene Wahrscheinlichkeitsdichte eine Polygon-Ring-Schachtelung existiert, so daß die Eigenschaften **E1** und **E2** erfüllt sind. Das bedeutet: erstens, die Wahrscheinlichkeit, daß ein Punkt $a \in \mathbb{R}^2$ in $Ring^{(l)}$ liegt, beträgt für $0 < p < 1$

$$P(Ring^{(0)}) = p$$

und

$$P(Ring^{(l+1)}) = (1-p) \cdot P(Ring^{(l)})$$

für $l = 0, 1, 2, \ldots$. Zweitens, alle Eckfelder haben dieselbe Wahrscheinlichkeit $q(l)$, welche gegenüber den Wahrscheinlichkeiten der anderen Felder von $Ring^{(l)}$ stochastisch nicht benachteiligt ist: $q(l) \geq \frac{p(1-p)^l}{2kl}$.

Betrachten wir wieder als Wahrscheinlichkeitsexperiment n unabhängig voneinander gewählte Punkte des $\mathbb{R}^2$. Genauso wie bei der Rechteck-Ring-Schachtelung erhält man durch die Polygon-Ring-Schachtelung eine Verringerung der Beobachtungstiefe[12]. Wir werden daher ähnlich zu Abschnitt 3 als Wahrscheinlichkeitsexperiment die „`Ziehung einer` n`-Auswahl aus einer Polygon-Ring-Schachtelung`“ betrachten. Der dazugehörige Wahrscheinlichkeitsraum definiert sich genauso wie in Abschnitt 3.2. Ebenso übernehmen wir die Definitionen und Notationen aus Abschnitt 3.

4.1.1. erwartete Kosten von Schritt 1

Zuerst müssen wir wissen, wieviele Schritte man benötigt, Punkte des $\mathbb{R}^2$ Polygon-Ringen zuzuordnen. Trivialerweise kann man alle Punkte, die in $Ring^{(0)}$ liegen, mit k Abfragen lokalisieren (nämlich durch k Vergleiche mit den Linien $l_1, \ldots, l_k$). Ein Punkt in einem Eckfeld von $Ring^{(1)}$ wird höchstens zwei dieser Abfragen verletzen und ein Punkt in $Ring^{(1)}$, der nicht in einem Eckfeld liegt, höchstens eine. Also lassen sich alle Punkte von $Ring^{(1)}$ in $k+2$ Schritten lokalisieren. Mit demselben Argument stellt man fest, daß sich die Punkte von $Ring^{(2)}$ mit höchstens $k+2+2$ Abfragen lokalisieren lassen müssen usw. Es läßt sich daher wie in Abschnitt 2 ein Suchbaum aufbauen, der zur Lokalisation eines Punktes in $Ring^{(l)}$ höchstens $k+2l$ Abfragen benötigt.

[12] Im Gegensatz zu Abschnitt 3.1 erzeugt eine Polygon-Ring-Schachtelung für $k \neq 4$ keine Felder, die Rechtecke $\subseteq \mathbb{R}^2$ sind. Der Urbildraum ist daher diesmal nicht $(\mathbb{R}^2, \mathcal{B}^2)$ sondern im allgemeinen $(\mathbb{R}^2, Pot(D))$, wobei $D = \{D^{(0)}, D^{(1)}, \ldots\}$ die Menge der Felder der Polygon-Ring-Schachtelung ist.

Sei der Mehraufwand pro Punktlokalisation, um den äußersten vollbesetzten Ring zu bestimmen, wieder durch eine Konstante c_{oh} abgeschätzt[13]. Die Kosten zur Lokalisierung des äußersten vollbesetzten Ringes betragen dann

$$k + 2l + c_{oh} \quad \text{Abfragen pro Punkt.}$$

Seien die Zufallsvariablen $X_1, \ldots, X_n, X$ so definiert wie in Abschnitt 3.3.1. Die erwarteten Kosten zur Lokalisierung eines Punktes des $\mathbb{R}^2$ in einem Polygon-Ring $Ring^{(l)}$ mit gleichzeitiger Abfrage, ob $Ring^{(l)}$ vollbesetzt ist, sind dann:

$$\begin{aligned} E[X_i] &= \sum_{j=0}^{\infty} P(X_i = j) \cdot j \\ &\leq \sum_{l=0}^{\infty} P(Ring^{(l)}) \cdot (k + 2l + c_{oh}) \\ &= \sum_{l=0}^{\infty} p \cdot (1-p)^l \cdot (k + 2l + c_{oh}) \\ &= (k-2) + \frac{2}{p} + c_{oh} \; . \end{aligned}$$

Die erwarteten Kosten von Schritt 1 für n Punkte liegen somit unterhalb von

$$E[X] \quad \leq \quad \left((k-2) + \frac{2}{p} + c_{oh} \right) \cdot n$$

Abfragen.

4.1.2. erwartete Kosten von Schritt 2

Die Laufzeitanalyse von Schritt 2 des Verfahrens läuft formal analog zu Abschnitt 3.3.2 ab. Entscheidend ist daher wieder die Abschätzung von $P(\mathcal{W}_l)$ durch $P(\overline{R^{(l+1)}})$ und $E[Y_l|\mathcal{W}_l]$ durch $E[Y_l]$.

Vermutung 4.1 *Gegeben eine Polygon-Ring-Schachtelung. Für große Werte n und festes $p \in (0,1)$ gibt es eine maximale Schachtelungstiefe l_0, so daß für alle Schachtelungstiefen $l \leq l_0$ die Ungleichung*

$$E[Y_l|\mathcal{W}_l] \leq E[Y_l]$$

gilt.

[13] Die Idee ist dieselbe, die wir schon in Fußnote 1 auf Seite 19 beschrieben haben. Nehmen wir also wieder $c_{oh} = 2$ an. Der einzige Unterschied besteht diesmal darin, daß bei Polygon-Ringen jedem Eckfeld $k-1$ Zeiger auf die Speicherzellen der anderen Eckfelder zur Verfügung gestellt werden müssen.

Vermutung 4.2 *Sei f eine Wahrscheinlichkeitsdichte, für die die Eigenschaften* **E1** *und* **E2** *erfüllt sind. Dann existiert ein Wert $b > 1$, so daß für große Werte von n, $0 < p < 1$ und alle Schachtelungstiefen $l \leq l_n = \lfloor \alpha \log_b n \rfloor$, $0 < \alpha < -\frac{1}{\log_b(1-p)}$, die Abschätzung*

$$E[Y_l|\mathcal{W}_l] \leq E[Y_l]$$

erfüllt ist.

Die Gültigkeit von Vermutung 4.1 ergibt sich aus der Plausibilitätsbetrachtung für Vermutung 3.1. Die Gültigkeit von Vermutung 4.2 stützen wir wieder auf empirische Ergebnisse[14].Bestimmen wir zunächst $P(R^{(l)})$, die Wahrscheinlichkeit, daß alle Eckfelder eines Polygon-Ringes $Ring^{(l)}$ belegt sind. Wir werden zeigen, daß man diese mit der Siebformel von Poincaré-Sylvester bestimmen kann. Die Siebformel von Poincaré-Sylvester erlaubt es, für k gegebene (nicht notwendigerweise paarweise disjunkte) Wahrscheinlichkeits-Ereignisse $A_1, \ldots, A_k \in Pot(\Omega)$ die Wahrscheinlichkeit auszurechnen, daß keines dieser $A_1, \ldots, A_k$ eintreten wird.

$$P\,(\texttt{keines der } A_i \texttt{ tritt ein}) = \sum_{m=0}^{k} (-1)^m \cdot \mathcal{S}_m$$

wobei

$$\begin{aligned} \mathcal{S}_0 &:= 1 = P(\Omega) \\ \mathcal{S}_1 &:= \sum_{i=1}^{k} P(A_i) \\ \mathcal{S}_2 &:= \sum_{1 \leq i < j \leq k} P(A_i \cap A_j) \\ \mathcal{S}_m &:= \sum_{\{i_1,\ldots,i_m\} \subseteq \{1,\ldots,k\}} P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_m}) \end{aligned}$$

Bemerkung 4.1 *$\mathcal{S}_m$ besteht aus $\binom{k}{m}$ Summanden.*

Wie sehen die Ereignisse A_i in unserem Falle aus? Nun, es wird

A_i : „`Eckfeld` i `in` $Ring^{(l)}$ `ist nicht besetzt`“,

$A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_m}$: „`Eckfelder` i_1 `und` i_2 `...` `und` i_m `in` $Ring^{(l)}$ `sind nicht besetzt`“

bedeuten.

Die Wahrscheinlichkeiten dieser Ereignisse lassen sich leicht berechnen. Betrachten wir dazu wieder eine n-Auswahl. Der Wert $q(l)$ sei die Wahrscheinlichkeit eines

[14] siehe Abschnitt 5

Eckfeldes von $Ring^{(l)}$ (und alle Eckfelder seien gleichwahrscheinlich). Für die Wahrscheinlichkeiten ergeben sich dann folgende Werte:

$$\begin{aligned} P(A_i) &= (1-q(l))^n \\ P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_m}) &= (1 - m \cdot q(l))^n \end{aligned}$$

$$\begin{aligned} \Rightarrow P\left(\texttt{kein Eckfeld in } Ring^{(l)} \texttt{ leer}\right) &= P(\texttt{keines der } A_i \texttt{ tritt ein}) \\ = \sum_{m=0}^{k} (-1)^m \cdot \mathcal{S}_m &= \sum_{m=0}^{k} (-1)^m \binom{k}{m} (1 - m \cdot q(l))^n \,. \end{aligned}$$

Aufgrund des Beweises der Siebformel von Poincaré-Sylvester gilt:

$$P\left(\texttt{kein Eckfeld in } Ring^{(l)} \texttt{ leer}\right) \geq 1 - \sum_{i=1}^{k} P(A_i)\,.$$

Die Wahrscheinlichkeit $P(R^{(l)})$ läßt sich abschätzen.

Lemma 4.1 *Für* $n \geq \frac{\ln(2k)}{q(l)}$ *ist* $P(R^{(l)}) \geq \frac{1}{2}$.

Beweis

Aus

$$n \geq \frac{\ln(2k)}{q(l)}$$

folgt wegen $\ln x \leq x - 1$

$$\begin{aligned} n \cdot \ln(1 - q(l)) &\leq n \cdot -q(l) \leq -\ln(2k) \\ \Rightarrow \quad -\ln(2k) &\geq n \cdot \ln(1-q(l)) \\ \frac{1}{2k} &\geq (1-q(l))^n \\ \Rightarrow \quad P(R^{(l)}) &\geq 1 - k\,(1-q(l))^n \geq \frac{1}{2}. \end{aligned}$$

□

Da die Wahrscheinlichkeitsdichte Eigenschaft **E2** erfüllt, ist $q(l) \geq \frac{p(1-p)^l}{2kl}$. Es stellt sich nun wieder die Frage, für welche Schachtelungstiefen l Lemma 4.1 gilt. Der Beweis von Lemma 3.2 zeigt jedoch die Gültigkeit von

Lemma 4.2 *Für große n und $0 < p < 1$ ist bei Schachtelungstiefe $l_n = \lfloor \alpha \log n \rfloor$, $0 < \alpha < -\frac{1}{\log(1-p)}$, die Voraussetzung von Lemma 4.1 erfüllt.*

Damit gilt auch bei Polygon-Ring-Schachtelungen

Lemma 4.3 *Ist n groß und $0 < p < 1$, dann gilt für Schachtelungstiefe $l_n = \lfloor \alpha \log n \rfloor$, $0 < \alpha < -\frac{1}{\log(1-p)}$, und $r = 0, \ldots, l_n - 1$*

$$P(\overline{R^{(l_n - r)}}) \leq 2^{-(1-p)^{-r}}.$$

Beweis

Der Nachweis verläuft quasi identisch zu dem von Lemma 3.3. Wegen

$$P(\overline{R^{(l_n)}}) \leq k\,(1 - q(l_n))^n \leq \frac{1}{2}$$

und Lemma 4.1 ergibt sich die Behauptung, falls

$$k(1 - q(l_n - r))^n \leq (k(1 - q(l_n))^n)^{(1-p)^{-r}}$$

gilt.

Da $q(l_n - r) = \frac{p(1-p)^{l_n - r}}{2k(l_n - r)} \geq \frac{p(1-p)^{l_n - r}}{2k(l_n)} = q(l_n)(1-p)^{-r}$ ist, folgt

$$\begin{aligned} 1 - q(l_n - r) &\leq 1 - q(l_n)(1-p)^{-r} \leq (1 - q(l_n))^{(1-p)^{-r}} \\ \Rightarrow \quad k(1 - q(l_n - r))^n &\leq k(1 - q(l_n))^{n \cdot (1-p)^{-r}} \\ \Rightarrow \quad k(1 - q(l_n - r))^n &\leq (k(1 - q(l_n))^n)^{(1-p)^{-r}} \end{aligned}$$

und damit die Behauptung. □

Wie wir gesehen haben gilt auch bei einer Polygon-Ring-Schachtelung, daß für große n und bei geeigneter Wahl von l_n

$$E[Y_l | \mathcal{W}_l] \leq E[Y_l]$$

und

$$P(\overline{R^{(l_n - r)}}) \leq \left(\frac{1}{2}\right)^{\left(\frac{1}{1-p}\right)^r}$$

ist. Daher gelten auch bei Polygon-Ring-Schachtelungen die Aussagen von Lemma 3.4 und Lemma 3.5.

Folglich gilt bei Polygon-Ring-Schachtelungen

Satz 4.1 *Seien $a_1, \ldots, a_n$ unabhängig voneinander gewählte Punkte des $\mathbb{R}^2$ und f eine Wahrscheinlichkeitsdichte auf $\mathbb{R}^2$, für die eine Polygon-Ring-Schachtelung existiert, so daß f die Eigenschaften* **E1** *und* **E2** *erfüllt. Es sei k die Anzahl der Ecken der Polygone. Für $0 < p < 1$ und genügend große n berechnet Algorithmus* `A` *dann die konvexe Hülle von $a_1, \ldots, a_n$ in erwarteter Laufzeit*

$$\begin{aligned} E[\texttt{konvexe Hülle}] &= E[X] + E[Y_C] \\ &\leq ((k-2) + \frac{2}{p} + c_{oh}) \cdot n + n. \end{aligned}$$

Wie man sieht erhält man das Ergebnis der Rechteck-Ring-Schachtelung, wenn man $k = 4$ wählt.

4.2. Optimale Polygon-Ring Form bei einer rotationssymmetrischen Verteilung

Die Idee, den $\mathbb{R}^2$ mit regulären Polygon-Ringen anstatt Rechteck-Ringen zu überdecken, entsprang der Hoffnung, daß bei speziellen Wahrscheinlichkeitsverteilungen durch eine andere Ringform ein Laufzeitgewinn erzielbar sei. Was heißt das konkret? Gehen wir davon aus, daß wir eine konzentrische, kreisförmige, exponentiell abfallende Wahrscheinlichkeitsverteilung über dem $\mathbb{R}^2$ haben. Es liegt also eine rotationssymmetrische Stufenverteilung vor.

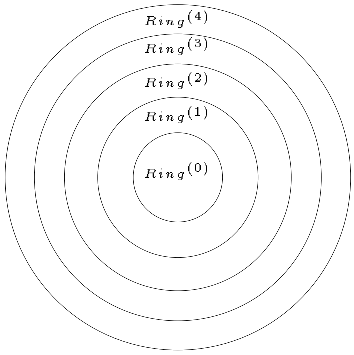

Abbildung 22: Stufenverteilung

Unter bestimmten Umständen, auf die wir gleich näher eingehen werden, sind bei einer solchen Verteilung die erwarteten Laufzeiten von Algorithmus A für verschiedenförmige Polygon-Ring-Schachtelungen miteinander vergleichbar, z.B. die erwartete Laufzeit bei einer 3-Eck-Ring-Schachtelung mit der bei einer 6-Eck-Ring-Schachtelung.

Da ein 6-Eck einen Kreis besser approximiert, schmiegt sich eine 6-Eck-Ring-Schachtelung besser an eine solche Verteilung an als eine 3-Eck-Ring-Schachtelung. Dies hat zur Folge, daß bei einem vollbesetzten 6-Eck-Ring mehr Punkte bei der Berechnung der konvexen Hülle *ausgesiebt* werden können als bei einem vollbesetzten 3-Eck-Ring, wie in Abbildung 23 verdeutlicht werden soll.

Gehen wir nun näher auf die Rahmenbedingungen ein. Gegeben sei eine rotationssymmetrische Stufenverteilung mit exponentiellem Wahrscheinlichkeitsabfall, d.h. der Wahrscheinlichkeitsverlust in den *Kreis-Ringen* verhalte sich genauso wie vorher in den Polygon-Ringen. Für $0 < p < 1$ ist

$$P(K\text{-}Ring^{(0)}) = p$$

und

$$P(K\text{-}Ring^{(l+1)}) = (1-p) \cdot P(K\text{-}Ring^{(l)}).$$

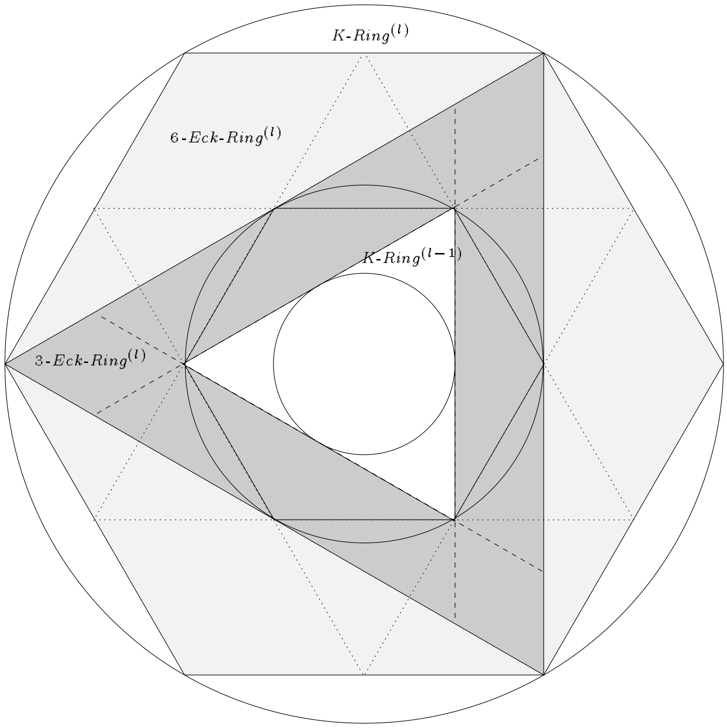

Abbildung 23:

Bemerkung 4.2 *Zur besseren Unterscheidbarkeit wollen wir einen Polygon-Ring als* P-Ring *und einen Kreis-Ring als* K-Ring *bezeichnen.*

Weiterhin fordern wir, daß die Punkte in *K-Ring*$^{(l)}$ gleichverteilt sind und die Radien der Kreise mit einem konstanten Faktor μ wachsen. Der Radius r_{i+1} von Kreis K_{i+1} sei daher $\mu \cdot r_i$.

Diese Stufenverteilung können wir anschließend derart mit einer Polygon-Ring-Schachtelung überdecken, daß jedes Polygon P_l vollständig von Kreis K_l und Kreis K_{l-1} vollständig von Polygon P_l umfaßt wird[15]. Damit die oben beschriebene Schachtelung auch für 3-Eck-Ringe möglich ist, muß $\mu \geq 2$ sein. Ansonsten ist es nicht möglich, daß das gleichseitige Dreieck D_l vollständig von Kreis K_l und Kreis K_{l-1} vollständig von D_l umfaßt wird.

Wegen der geänderten Ausgangsvoraussetzungen müssen wir die erwarteten Kosten von Schritt 1 und 2 des Verfahrens neu berechnen. Wir betrachten wieder n Punkte des $\mathbb{R}^2$.

[15] Wie in Abbildung 23, Seite 53 dargestellt. Die Punkte von Polygon P_l liegen *co-circular* auf dem Radius von Kreis K_l.

erwartete Kosten von Schritt 1

Formal berechnen sich die Kosten von Schritt 1 genauso wie in Abschnitt 4.1.1. Das Problem ist nur, daß wir zwar die Wahrscheinlichkeit von Kreis-Ring $K\text{-}Ring^{(l)}$ kennen, aber nicht die von Polygon-Ring $P\text{-}Ring^{(l)}$. $P(P\text{-}Ring^{(l)})$ läßt sich jedoch anhand von $P(K\text{-}Ring^{(l)})$ wie folgt berechnen. Bei genauerer Betrachtung der Konstruktion sieht man, daß $P\text{-}Ring^{(l)}$ zu einem Teil, den wir A_l nennen, in $K\text{-}Ring^{(l)}$ und zu einem anderen Teil, den wir B_l nennen wollen, in $K\text{-}Ring^{(l-1)}$ liegt. $\mathbb{A}(F)$ sei die Fläche von F. Es gilt

$$\mathbb{A}(P\text{-}Ring^{(l)}) = A_l + B_l$$

und

$$\begin{aligned} A_l &= \mathbb{A}(P_l) - \mathbb{A}(K_{l-1}) \\ B_l &= \mathbb{A}(K_{l-1}) - \mathbb{A}(P_{l-1}). \end{aligned}$$

Aus der Gleichverteilung innerhalb von $K\text{-}Ring^{(l)}$ ergibt sich, daß

$$\begin{aligned} P(P\text{-}Ring^{(l)}) &= \frac{A_l}{\mathbb{A}(K\text{-}Ring^{(l)})} \cdot p(1-p)^l \\ &\quad + \frac{B_l}{\mathbb{A}(K\text{-}Ring^{(l-1)})} \cdot p(1-p)^{l-1}, \quad \text{für } l \geq 1 \\ P(P\text{-}Ring^{(0)}) &= \frac{\mathbb{A}(P_0)}{\mathbb{A}(K_0)} \cdot p. \end{aligned}$$

Es ist bekanntermaßen

$$\begin{aligned} \mathbb{A}(K\text{-}Ring^{(l)}) &= \mathbb{A}(K_l) - \mathbb{A}(K_{l-1}) \\ &= \pi \cdot {r_l}^2 - \pi \cdot {r_{l-1}}^2 \\ &= \pi \cdot \mu^2 \cdot {r_{l-1}}^2 - \pi \cdot {r_{l-1}}^2, \quad r_l = \mu \cdot r_{l-1} \\ &= (\mu^2 - 1) \cdot \mathbb{A}(K_{l-1}). \end{aligned}$$

Die Fläche eines regulären k-Ecks mit Umkreisradius r läßt sich leicht als $k \cdot \sin\left(\frac{\pi}{k}\right) \cos\left(\frac{\pi}{k}\right) \cdot r^2$ herleiten. Somit sind nun

$$\begin{aligned} P(P\text{-}Ring^{(1)}) &= \frac{\mathbb{A}(P_1) - \mathbb{A}(K_0)}{(\mu^2 - 1) \cdot \mathbb{A}(K_0)} \cdot p(1-p) + \frac{\mathbb{A}(K_0) - \mathbb{A}(P_0)}{\mathbb{A}(K_0)} \cdot p \\ &= \frac{\mu^2 \cdot \mathbb{A}(P_0) - \mathbb{A}(K_0)}{(\mu^2 - 1) \cdot \mathbb{A}(K_0)} \cdot p(1-p) + \left(1 - \frac{\mathbb{A}(P_0)}{\mathbb{A}(K_0)}\right) \cdot p \\ &= \frac{1}{\mu^2 - 1} \left(\mu^2 \frac{\mathbb{A}(P_0)}{\mathbb{A}(K_0)} - 1\right) \cdot p(1-p) + \left(1 - \frac{\mathbb{A}(P_0)}{\mathbb{A}(K_0)}\right) \cdot p \end{aligned}$$

und für $l = 2, 3, \ldots$

$$\begin{aligned}
P(P\text{-}Ring^{(l)}) &= \frac{\mathbb{A}(P_l) - \mathbb{A}(K_{l-1})}{(\mu^2-1)\cdot\mathbb{A}(K_{l-1})}\cdot p(1-p)^l \\
&\quad + \frac{\mathbb{A}(K_{l-1}) - \mathbb{A}(P_{l-1})}{(\mu^2-1)\cdot\mathbb{A}(K_{l-2})}\cdot p(1-p)^{l-1} \\
&= \frac{\mu^2\cdot\mathbb{A}(P_{l-1}) - \mathbb{A}(K_{l-1})}{(\mu^2-1)\cdot\mathbb{A}(K_{l-1})}\cdot p(1-p)^l \\
&\quad + \frac{\mu^2\,\mathbb{A}\,(K_{l-2}) - \mu^2\,\mathbb{A}\,(P_{l-2})}{(\mu^2-1)\cdot\mathbb{A}\,(K_{l-2})}\cdot p(1-p)^{l-1} \\
&= \frac{1}{\mu^2-1}\left(\mu^2\,\frac{\mathbb{A}\,(P_{l-1})}{\mathbb{A}\,(K_{l-1})} - 1\right)\cdot p(1-p)^l \\
&\quad + \frac{\mu^2}{\mu^2-1}\left(1 - \frac{\mathbb{A}\,(P_{l-2})}{\mathbb{A}\,(K_{l-2})}\right)\cdot p(1-p)^{l-1}.
\end{aligned}$$

Da das Verhältnis

$$v = \frac{\mathbb{A}\,(P_l)}{\mathbb{A}\,(K_l)} = \frac{k\cdot\sin\left(\frac{\pi}{k}\right)\cos\left(\frac{\pi}{k}\right)}{\pi}$$

für alle l gleichbleibt, erhält man für die erwarteten Kosten von Schritt 1 bei n Eingabepunkten

$$\begin{aligned}
E[X] &\leq n\cdot\sum_{l=0}^{\infty} P(P\text{-}Ring^{(l)})\cdot(k+2l+c_{oh}) \\
&= \left[v\,p\cdot(k+c_{oh}) + \left(\frac{\mu^2 v - 1}{\mu^2-1}\,p(1-p) + (1-v)\,p\right)\cdot(k+2+c_{oh})\right. \\
&\quad + \frac{\mu^2 v - 1}{\mu^2-1}\,p\cdot\sum_{l=2}^{\infty}(1-p)^l\cdot(k+2l+c_{oh}) \\
&\quad \left. + \frac{\mu^2(1-v)}{\mu^2-1}\,p\cdot\sum_{l=2}^{\infty}(1-p)^{l-1}\cdot(k+2l+c_{oh})\right]\cdot n \\
&= \left[k + \frac{2}{p} + c_{oh} + \frac{2(1-p)}{\mu^2-1} - \frac{2\left(\mu^2-p\right)\,k\,\sin\left(\frac{\pi}{k}\right)\cos\left(\frac{\pi}{k}\right)}{(\mu^2-1)\,\pi}\right]\cdot n.
\end{aligned}$$

Läßt man k fest, so wird dieser Ausdruck monoton fallend für $p \to 1$ minimal und geht für $p \to 0$ gegen Unendlich. Ebenso wächst $E[X]$ bei konstantem p für $k \to \infty$ ins Unendliche, d.h. je mehr Ecken die Polygone haben, desto größer werden die erwarteten Kosten von Schritt 1.

Seien $E[X']$ die erwarteten Kosten für Schritt 1 des Verfahrens bei einer k'-Eck-Ring-Schachtelung und $E[X]$ die Kosten bei einer k-Eck-Ring-Schachtelung.

Lemma 4.4 *Für $0 < p < 1$ ist $E[X]$ bei $k = 3$ minimal.*

Beweis

Wäre dies nicht so, dann müßte ein $k' > 3$ existieren, so daß

$$E[X'] - E[X] \quad = \quad k' - 3 + \frac{2(\mu^2 - p)}{\mu^2 - 1}\left(\frac{3/4\sqrt{3} - k' \sin\left(\frac{\pi}{k'}\right)\cos\left(\frac{\pi}{k'}\right)}{\pi}\right) < 0$$

ist. Nun ist für alle $\mu \geq 2$, $0 < p < 1$ und $k' > 3$

$$\frac{2(\mu^2 - p)}{\mu^2 - 1}\left(\frac{3/4\sqrt{3} - k' \sin\left(\frac{\pi}{k'}\right)\cos\left(\frac{\pi}{k'}\right)}{\pi}\right) \geq 2\left(\frac{3\sqrt{3}}{4\pi} - 1\right) \approx -1.17.$$

Folglich kann für $k' \geq 5$ die Differenz $E[X'] - E[X]$ nicht kleiner als 0 sein. Da für $k' = 4$ der Wert von $\frac{2(\mu^2-p)}{\mu^2-1}\left(\frac{3/4\sqrt{3}-k'\sin\left(\frac{\pi}{k'}\right)\cos\left(\frac{\pi}{k'}\right)}{\pi}\right) \geq -0.59$ ist, muß $E[X]$ für $k = 3$ tatsächlich minimal sein. □

erwartete Kosten von Schritt 2

Vorab ein paar Bemerkungen. Abbildung 23 zeigt,

1. daß alle Eckfelder von *P-Ring*$^{(l)}$ in *K-Ring*$^{(l)}$ liegen und
2. daß für $k_1 > k_2$ die Eckfelder von k_1-*Eck-Ring*$^{(l)}$ größer sind als die von k_2-*Eck-Ring*$^{(l)}$.

Dies müssen wir bei unserer Berechnung berücksichtigen. Wir suchen zunächst einen Ring *Ring*$^{(l)}$ für den $P(R^{(l)}) \geq \frac{1}{2}$ ist. Zur Berechnung von $P(R^{(l)})$ benötigen wir die Wahrscheinlichkeit $q(l)$ der Eckfelder von *Ring*$^{(l)}$. Da wir von einer Gleichverteilung innerhalb eines Kreis-Ringes ausgegangen sind, ist die Wahrscheinlichkeit der Eckfelder von *Ring*$^{(l)}$ durch

$$q(l) = \frac{\mathbb{A}(P\text{-}Ring^{(l)}\,Eckfeld)}{\mathbb{A}(K\text{-}Ring^{(l)})}\,P(K\text{-}Ring^{(l)}) = \frac{(\mu - 1)\,\tan\left(\frac{\pi(k-2)}{2k}\right)}{\pi}\,p(1-p)^l$$

gegeben. Es ist[16] $\mathbb{A}(P\text{-}Ring^{(l)}Eckfeld) = r_{l-1}^2 \cdot (\mu - 1)^2 \cdot \tan\left(\frac{\pi(k-2)}{2k}\right)$. Im Prinzip schließt sich nun dieselbe Betrachtung an, wie in den Abschnitten 3.3.2, bzw. 4.1.2. Zuerst überprüft man, für welche Schachtelungstiefe $P(R^{(l)})$, die Wahrscheinlichkeit, daß kein Eckfeld in *P-Ring*$^{(l)}$ leer ist, größer als $\frac{1}{2}$ ist.

[16] Mit diesem Ausdruck wird $P(R^{(l)})$ für $k \geq 6$ durch $1 - k(1 - q(l))^n$ überschätzt, da sich dann die „Eckfelder" von *Ring*$^{(l)}$ überlappen. Diesen Fehler nehmen wir jedoch in Kauf, um die Wahrscheinlichkeiten leicht ausrechnen zu können.

Lemma 4.5 *Für große n und $0 < p < 1$ ist für Schachtelungstiefen $l \leq l_n$ und $l_n = \lfloor \alpha \log n \rfloor$, $0 < \alpha < -\frac{1}{\log(1-p)}$, die Wahrscheinlichkeit $P(R^{(l)}) \geq \frac{1}{2}$.*

Beweis

Es ist wegen Lemma 4.1 die Wahrscheinlichkeit $P(R^{(l)}) \geq \frac{1}{2}$, falls

$$n \geq \frac{\ln(2k) \cdot \pi}{p(1-p)^l (\mu - 1) \tan\left(\frac{\pi(k-2)}{2k}\right)}$$

ist. Aufgrund der Wahl von $l_n = \lfloor \alpha \log n \rfloor$ folgt

$$n^{1+\alpha \log(1-p)} \geq \frac{\ln(2k) \cdot \pi}{p \cdot (\mu - 1) \cdot \tan\left(\frac{\pi(k-2)}{2k}\right)}.$$

Diese Ungleichung ist jedoch für $k \geq 3$, $0 < p < 1$, $\mu \geq 2$ und n genügend groß immer erfüllbar. □

Ebenfalls läßt sich leicht überprüfen, daß immer noch Lemma 4.3

$$P(\overline{R^{(l_n - r)}}) \leq \left(\frac{1}{2}\right)^{\left(\frac{1}{1-p}\right)^r}$$

gilt.

Es gibt aber einen Unterschied zur vorherigen Berechnung, was $E[Y_l]$ anbelangt. Sei Polygon-Ring *P-Ring*$^{(l)}$ als äußerster vollbesetzt. *P-Ring*$^{(l)}$ liegt nicht vollständig in Kreis-Ring *K-Ring*$^{(l)}$, sondern schneidet Kreis-Ring *K-Ring*$^{(l-1)}$. Die Punkte jenseits von *P-Ring*$^{(l)}$ sind daher die Punkte jenseits von *K-Ring*$^{(l)}$ zuzüglich der Punkte, die im Schnitt von *K-Ring*$^{(l-1)}$ und *P-Ring*$^{(l)}$ liegen:

$$\begin{aligned} E[Y_0] &= n \\ E[Y_1] &= (1-p)n + (1-v)p\,n \\ E[Y_l] &= (1-p)^l\, n + \frac{\mu^2}{\mu^2 - 1}(1-v)\,p\,(1-p)^{l-1}\, n \\ &\quad \text{für } l = 2, \ldots, \infty. \end{aligned}$$

Wir erhalten daher folgende Abschätzung[17] für die erwarteten Kosten $E[Y_C]$ von Schritt 2:

$$\begin{aligned} E[Y_C] &\leq c \cdot \log n \left(E[Y_{l'}] + \sum_{r=0}^{l'-1} P(\overline{R^{(l'-r)}}) \cdot E[Y_{l'-(r+1)}] \right) \\ &\leq c \cdot \log n \cdot \left(1 + \frac{\mu^2}{\mu^2 - 1}(1-v)\frac{p}{1-p}\right)(1-p)^{l'} n \cdot \left(1 + \sum_{r=0}^{l'-1} \frac{P(\overline{R^{(l'-r)}})}{(1-p)^{r+1}}\right) \end{aligned}$$

[17] vgl. Lemma 3.4

und daher mit Lemma 3.4

$$E[Y_C] \leq \tilde{c} \cdot \log n \left(1 + \frac{\mu^2}{\mu^2 - 1}\left(1 - \frac{k \sin(\frac{\pi}{k})\cos(\frac{\pi}{k})}{\pi}\right)\frac{p}{1-p}\right) \cdot (1-p)^{l'}\, n.$$

Für $\mu \geq 2$ und festes k gilt wieder Lemma 3.5, d.h. die erwarteten Kosten von Schritt 2 sind sublinear.

Die Frage ist nun, wann $E[Y_C]$ minimal wird.

Lemma 4.6 *Die erwarteten Kosten $E[Y_C]$ von Schritt 2 sind monoton fallend in der Anzahl der Ecken k, d.h. je mehr Ecken die Polygone haben, desto besser ist die erwartete Laufzeit.*

Beweis

Betrachten wir dazu zunächst den ersten Faktor,

$$\tilde{c} \cdot \log n \left(1 + \frac{\mu^2}{\mu^2 - 1}\left(1 - \frac{k \sin(\frac{\pi}{k})\cos(\frac{\pi}{k})}{\pi}\right)\frac{p}{1-p}\right),$$

von $E[Y_C]$. Dieser Faktor ist monoton fallend in k.

Der zweite Faktor, $(1-p)^{l'} \cdot n$, hängt von l' ab. Es ist also notwendig, das Verhalten von l' in Abhängigkeit von k zu untersuchen. Bisher wurde l' so gewählt, daß $P(R^{(l')}) \geq \frac{1}{2}$ war. Dazu mußte $n \geq \frac{\ln(2)\cdot\pi}{p\cdot(1-p)^{l'}\cdot(\mu-1)\cdot\tan(\frac{\pi(k-2)}{2k})}$ erfüllt sein. Verwenden wir anstelle dieser Ungleichung die Gleichung

$$n = \frac{\ln(2)\cdot\pi}{p\cdot(1-p)^{l'}\cdot(\mu-1)\cdot\tan(\frac{\pi(k-2)}{2k})},$$

mit Lösung

$$l' = \frac{\ln(\pi) + \ln\left(\frac{\ln(2)+\ln(k)}{n\,p\,\tan(\frac{\pi(k-2)}{2k})\,(\mu-1)}\right)}{\ln(1-p)} = \log_{(1-p)}\left(\frac{\pi\,(\ln(2)+\ln(k))}{n\,p\,\tan(\frac{\pi(k-2)}{2k})\,(\mu-1)}\right),$$

dann folgt daraus

$$E[Y_C] \leq \tilde{c} \cdot \log n \cdot \left(1 + \frac{\mu^2}{\mu^2 - 1}\left(1 - \frac{k \sin(\frac{\pi}{k})\cos(\frac{\pi}{k})}{\pi}\right)\frac{p}{1-p}\right) \cdot \frac{\pi\,(\ln(2)+\ln(k))}{p\,\tan(\frac{\pi(k-2)}{2k})\,(\mu-1)}.$$

Interessant für unsere Betrachtung ist nur $\frac{\ln(k)}{\tan(\frac{\pi(k-2)}{2k})}$. Da $\tan(\frac{\pi(k-2)}{2k})$ für wachsende k erheblich schneller zunimmt als $\ln(k)$, wird l' für größere k monoton steigen. Damit wird auch der zweite Faktor, $(1-p)^{l'} \cdot n$, für große n kleine Werte annehmen. □

Da jedoch asymptotisch in n, der Eingabegröße, Schritt 1 die Laufzeit dominiert, verschlechtert sich die Laufzeit des Verfahrens für Polygone mit vielen Ecken.

Satz 4.2 *Bei einer rotationssymmetrischen Stufenverteilung mit $\mu \geq 2$ und für große Werte n sind die erwarteten Kosten von Algorithmus* `A` *bei einer 3-Eck-Ring-Schachtelung am geringsten.*

5. Numerische Simulationen und Varianten des Verfahrens

Wir haben zur Abschätzung der erwarteten Kosten des zweiten Schrittes von Algorithmus `A` bei einer Rechteck-Ring-Schachtelung Vermutung 3.1 und 3.2 und bei einer Polygon-Ring-Schachtelung Vermutung 4.1 und 4.2 benutzt. Die Aussagen der Vermutungen stimmen für beide Schachtelungstypen überein:

A1 Für große Werte n gibt es eine maximale Schachtelungstiefe l_0, so daß für alle Schachtelungstiefen $l \leq l_0$

$$E[Y_l|\mathcal{W}_l] \leq E[Y_l]$$

gilt.

A2 Es existiert ein Wert $b > 1$, so daß für große n die Schachtelungstiefe $l_n = \lfloor \alpha \log_b n \rfloor$, $0 < \alpha < -\frac{1}{\log_b(1-p)}$, hinreichend für die Gültigkeit der Ungleichung von **A1** ist.

Bemerkung 5.1 *Da eine Rechteck-Ring-Schachtelung eine spezielle Polygon-Ring-Schachtelung ist, werden wir im weiteren nur noch den allgemeineren Fall der Polygon-Ring-Schachtelung betrachten.*

In Abschnitt 3.5 haben wir gesehen, daß bei einem vereinfachten Wahrscheinlichkeitsexperiment für große n und beliebige, aber feste Werte $p \in (0, 1)$ die Aussage **A1** gilt und haben dies darauf zurückgeführt, daß für große n ein $i_0 \in [1, n]$ existiert, so daß für alle $i \geq i_0$ die Wahrscheinlichkeit $P(Y_l = i|\mathcal{W}_l) \leq P(Y_l = i)$ ist. Die Frage ist nun, ob dies auch bei dem nicht-vereinfachten Wahrscheinlichkeitsexperiment aus Abschnitt 4.1 bzw. 3.2 zutrifft.

Es wurden numerische Simulationen durchgeführt, um experimentell zu überprüfen, daß die Aussage **A1** beim nicht-vereinfachten Wahrscheinlichkeitsexperiment immer noch gilt. Dazu wurde für unterschiedliche Werte von p jeweils 100-mal die Ziehung einer n-Auswahl aus einer k-Eck-Ring-Schachtelung simuliert[18]. Bei der Simulation auf dem Rechner wurde $q(l)$, die Wahrscheinlichkeit jedes Eckfeldes von $Ring^{(l)}$, gleich $\frac{p(1-p)^l}{2kl}$ gesetzt. Die Ergebnisse sind in den Tabellen 1-3 zusammengefaßt. Die erste Spalte gibt an, bei welchen Schachtelungstiefen ein Ring vollbesetzt war. Die Wahrscheinlichkeit $P(\mathcal{W}_l)$ wurde durch die relative Häufigkeit des Ereignisses $\mathcal{W}_l$ angenähert. Für $E[Y_l|\mathcal{W}_l]$ wurde das arithmetische Mittel der 100 Simulationen genommen. Werten wir als Beispiel Tabelle 1 aus. Für $p = 0.1$ wurden 100-mal 10000-Auswahlen aus einer 3-, 4-, 10- und 20-Eck-Ring-Schachtelung simuliert. Man kann ablesen, daß bei 3-Eck-Ringen Vermutung 4.1 für Schachtelungstiefen $l \leq 20$, bei 4-Eck-Ringen für $l \leq 16$ und bei 10-Eck-Ringen für $l \leq 7$ gilt. Bei 20-Eck-Ring-Schachtelungen ist die Gültigkeit aus diesen Werten nicht ersichtlich.

[18] Es wurden auch Simulationen mit mehr als 100 Wiederholungen durchgeführt, diese lieferten jedoch das gleiche Ergebnis.

äußerster *voll-besetzter* Ring				$P(\mathcal{W}_l)$				$E[Y_l \mid \mathcal{W}_l]$				$E[Y_l]$
$k=3$	4	10	20	$k=3$	4	10	20	$k=3$	4	10	20	
			2				0.03				8115	8100
			3				0.08				7289	7290
			4				0.48				6566	6561
		5	5			0.01	0.27			5804	5912	5905
		6	6			0.08	0.10			5344	5305	5314
		7	7			0.18	0.03			4770	4750	4783
		8	8			0.25	0.01			4316	4314	4305
		9				0.23				3884		3874
		10				0.13				3492		3487
		11				0.06				3108		3138
		12				0.05				2790		2824
	13				0.02				2507			2542
	14				0.10				2254			2288
15	15	15		0.03	0.16	0.01		2031	2055	1974		2059
16	16			0.04	0.19			1855	1844			1853
17	17			0.04	0.17			1679	1679			1668
18	18			0.08	0.08			1491	1505			1501
19	19			0.14	0.07			1346	1353			1351
20	20			0.15	0.05			1205	1212			1216
21	21			0.06	0.05			1103	1092			1094
22	22			0.11	0.06			996	971			985
23	23			0.07	0.02			885	875			886
24	24			0.07	0.02			788	803			798
25	25			0.03	0.01			695	752			718
26				0.06				646				646
27				0.03				576				581
28				0.04				538				523
29				0.04				476				471
30				0.04				426				424
34				0.01				273				278

Tabelle 1: $p = 0.1$, $n = 10^4$

äußerster *vollbesetzter* Ring				$P(\mathcal{W}_l)$				$E[Y_l \mid \mathcal{W}_l]$				$E[Y_l]$
$k=3$	4	10	20	$k=3$	4	10	20	$k=3$	4	10	20	
			2				0.12				2492	2500
		3	3			0.15	0.84			1233	1250	1250
	4	4	4		0.05	0.75	0.04		621	622	625	625
5	5	5		0.24	0.45	0.10		310	316	319		313
6	6			0.59	0.45			153	157			156
7	7			0.14	0.05			78	83			78
8				0.03				43				39

Tabelle 2: $p = 0.5$, $n = 10^4$

äußerster *vollbesetzter* Ring				$P(\mathcal{W}_l)$				$E[Y_l \mid \mathcal{W}_l]$				$E[Y_l]$
$k=3$	4	10	20	$k=3$	4	10	20	$k=3$	4	10	20	
			1				0.01				3041	3000
		2	2			0.34	0.98			892	901	900
3	3	3	3	0.26	0.51	0.66	0.01	268	267	268	294	270
4	4			0.61	0.47			80	82			81
5	5			0.13	0.02			25	26			24

Tabelle 3: $p = 0.7$, $n = 10^4$

Ring	$P(\mathcal{W}_l)$	$E[Y_l \vert \mathcal{W}_l]$	$E[Y_l]$
$n = 10^5$			
7	0.08	789	781
8	0.63	392	391
9	0.22	188	196
10	0.06	99	97
11	0.01	56	49
$n = 10^6$			
10	0.09	974	976
11	0.55	487	488
12	0.31	244	244
13	0.05	131	122
$n = 10^8$			
16	0.08	1507	1525
17	0.55	755	762
18	0.27	377	381
19	0.10	193	190
$n = 10^9$			
19	0.12	1860	1907
20	0.42	929	953
21	0.41	446	476
22	0.05	207	238

Tabelle 4: $p = 0.5$, $k = 4$

Ring	$P(\mathcal{W}_l)$	$E[Y_l \vert \mathcal{W}_l]$	$E[Y_l]$
27	0.03	57935	58150
28	0.07	52336	52335
29	0.14	47048	47101
30	0.21	42338	42391
31	0.17	38117	38152
32	0.16	34320	34336
33	0.15	30973	30903
34	0.05	27861	27812
35	0.02	25080	25031

Tabelle 5: $p = 0.1$, $k = 20$, $n = 10^6$

Wir haben daher den Versuch mit $n = 10^6$ wiederholt. Die Ergebnisse befinden sich in Tabelle 5. Diesmal war n groß genug, so daß Vermutung 4.1 für Schachtelungstiefen $l \leq 32$ gilt.

A2, die Aussage von Vermutung 3.2 bzw. 4.2, war, daß ein Wert $b > 1$ existiert, so daß für große Werte n die Wahl $l_n = \lfloor \alpha \log_b n \rfloor$, $0 < \alpha < -\frac{1}{\log_b(1-p)}$, genügt, damit die Ungleichung von **A1** gilt. Die Gültigkeit dieser Aussage wurde ebenfalls experimentell bestätigt. Exemplarisch sind mehrere Meßreihen für $k = 4$ und $p = 0.5$ in Tabelle 4 abgebildet. Wie man sieht, ist die Wahl $l_n = \ln(n)$ hinreichend für die Gültigkeit von **A1**. Man sieht anhand der Tabellen 1-3 jedoch auch, daß man für große Werte von k und je nach Wahl von p einen Logarithmus mit großer Basis wählen muß, damit **A2** gilt.

Es soll noch erwähnt werden, daß die Vermutung aus Abschnitt 3.5: $E[Y_l|\mathcal{W}_l] \leq k + E[Y_l]$ – bis auf ein paar Ausreißer – ebenfalls durch Tabellen 1-5 experimentell bestätigt wird.

Als nächstes haben wir das Laufzeitverhalten von Schritt 2 für die Rechteck-Ring-Schachtelung in Abhängigkeit von p, wie es in Abschnitt 3.4 bereits angerissen wurde, verifiziert. Wie man sieht, liegt das Minimum bedeutend näher bei 1 als die

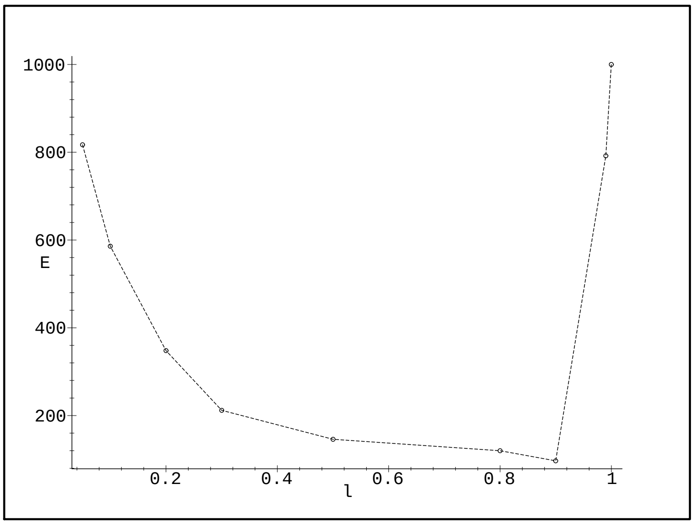

Abbildung 24: Erwartungswert in Abhängigkeit von p

Abschätzung in Abschnitt 3.4 vermuten ließ. Die erwartete Laufzeit des Algorithmus verringert sich also bis zu einem Wert p' nahe bei 1 und steigt für $p \geq p'$ steil an.

5.1. Varianten

Wir führen nun zwei Varianten des Verfahrens ein, die eine wollen wir *Segment-Methode* und die andere *variable Segment-Methode* nennen. Sie unterscheiden sich vom ursprünglichen Verfahren, das wir im folgenden als *Eckfeld-Methode* bezeichnen werden, dadurch, daß bei ihnen der Begriff *vollbesetzter* Ring anders definiert ist.

5.1.1. Die Segment-Methode

Bei der Segment-Methode heißt ein Ring *vollbesetzt*, wenn statt jedem *Eckfeld* jedes *Ecksegment* belegt ist. Was das bedeutet, soll in Abbildung 26 verdeutlicht werden.

Wir werden die Segment-Methode nur für Rechteck-Ring-Schachtelungen betrachten. Bei Polygon-Ring-Schachtelungen mit mehr Ecken werden sich die Ecksegmente überlappen und dadurch die Bestimmung der Wahrscheinlichkeiten eines vollbesetzten Ringes erschweren.

Laut Definition[19] ist $Ring^{(i)} = P_i - P_{i-1}$. Die Ecksegmente von $Ring^{(i)}$ – die schraffierten Flächen in Abbildung 26 – werden daher durch die Begrenzungslinien von

[19] vgl. Seite 45.

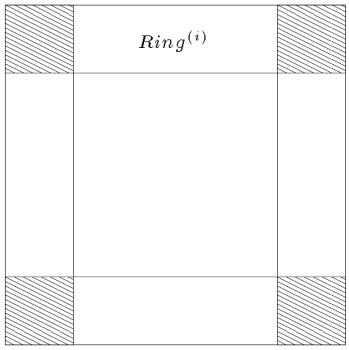

Abbildung 25: Eckfeld-Methode

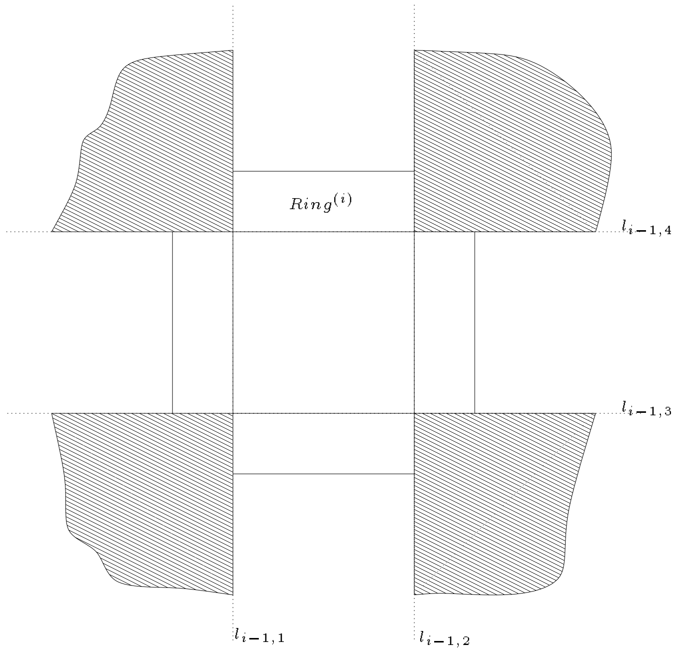

Abbildung 26: Segment-Methode

Polygon P_{i-1} gebildet. Z.B. ist das obere rechte Ecksegment von $Ring^{(i)}$ der Bereich „$> l_{i-1,2}$, $> l_{i-1,4}$“. Ein Ecksegment ist *belegt*, sobald ein Punkt in einen solchen Bereich fällt.

Ecksegmente decken einen größeren Bereich ab als Eckfelder. Die Hoffnung ist deshalb, daß bei der Ecksegment-Methode der äußerste vollbesetzte Ring weiter außen liegt als bei der Eckfeld-Methode.

5.1.2. Die variable Segment-Methode

Bei der *variablen Segment-Methode* handelt es sich um eine Erweiterung der *Segment-Methode*. Man schiebt dabei die Ecksegmente *unabhängig* voneinander solange nach außen bis jede weitere Verschiebung dazu führen würde, daß das Ecksegment nicht mehr belegt ist. Wie man in Abbildung 27 sehen kann, führt das dazu, daß nicht mehr alle Segmente demselben Ring zuzuordnen sind.

Wir müssen uns daher auch vom Konzept des vollbesetzten Ringes lösen. Die Punkte, die nun zur Berechnung der konvexen Hülle herangezogen werden, sind die, die außerhalb des Vierecks V liegen, das durch die vier äußersten belegten Ecksegmente definiert wird.

Dadurch, daß die Ecksegmente unabhängig voneinander verschoben werden dürfen, ist zu erwarten, daß weniger Punkte als bei der Segment-Methode zur Berechnung

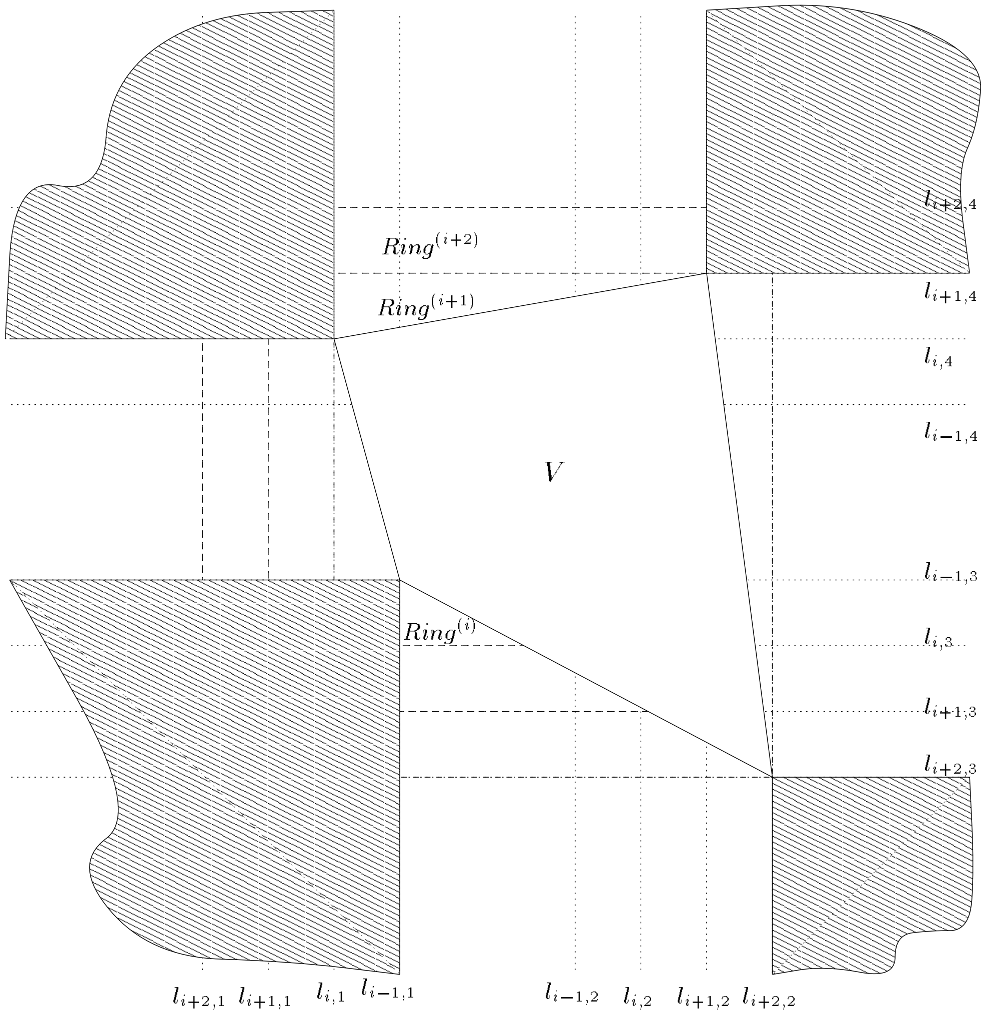

Abbildung 27: variable Segment-Methode

der konvexen Hülle benötigt werden. Der Grund ist, daß bei der Segment-Methode der irrelevante Bereich durch das Ecksegment, das am weitesten innen liegt, diktiert wird. Bei der variablen Segment-Methode können die anderen Ecksegmente davon unbeirrt weiter nach außen verschoben werden. Im allgemeinen wird daher der irrelevante Bereich bei der variablen Segment-Methode immer größer sein als bei der Segment-Methode.

5.1.3. Laufzeitvergleich der Varianten

Die Varianten werden ihre Stärken im zweiten Schritt des Algorithmus zeigen, da erst dort die konvexe Hülle berechnet wird. In Abschnitt 3 haben wir gesehen, daß

$E[Y]$, die erwartete Anzahl der zur Berechnung notwendigen Punkte, ein Maß für $E[Y_C]$, die erwarteten Kosten zur Berechnung der konvexen Hülle, ist. Im folgenden werden daher die $E[Y]$'s der Varianten für verschiedene Werte von p, der Wahrscheinlichkeit des innersten Rechtecks, verglichen. Dazu wurde für 100 n-Auswahlen die Anzahl der *aktiven* Punkte experimentell bestimmt. $E[Y]$ wurde dann durch den Mittelwert dieser 100 Auswahlen approximiert. Zur Vereinfachung wurde die Verteilung so gewählt, daß alle Felder $D^{(i)}$ eines Ringes $Ring^{(l)}$ dieselbe Wahrscheinlichkeit besitzen; d.h. $P_e(D^{((2l-1)^2)}) = \ldots = P_e(D^{((2l+1)^2-1)})$.

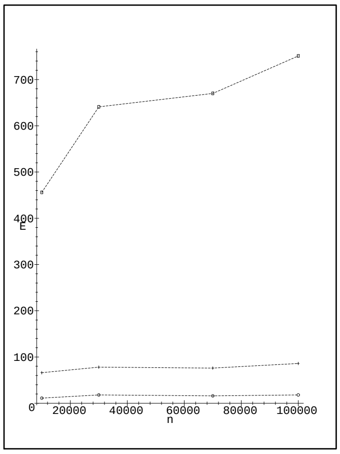

Abbildung 28:
$E[Y]$
$p = 0.3$

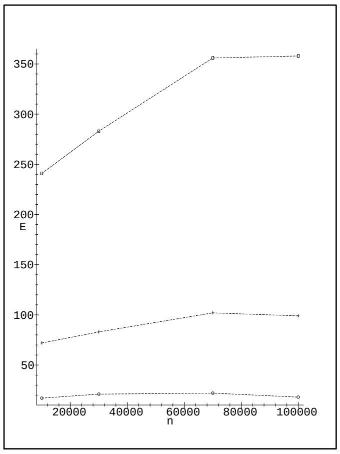

Abbildung 29:
$E[Y]$
$p = 0.5$

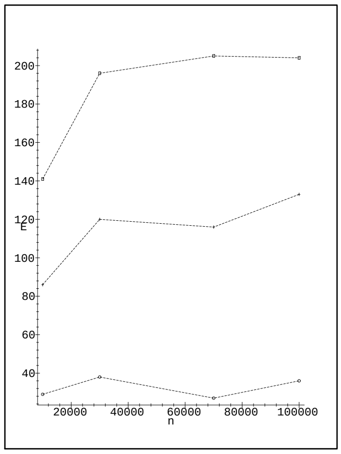

Abbildung 30:
$E[Y]$
$p = 0.8$

□: Eckfeld-Methode +: Segment-Methode ∘: variable Segment-Methode

Wie man sieht liefert die Eckfeld-Methode immer die schlechtesten und die variable Segment-Methode immer die besten Resultate. Für kleine p ist die Diskrepanz zwischen den beiden Segment-Methoden und der Eckfeld-Methode am größten, denn für solche Werte ist der Verlust an Wahrscheinlichkeitsmasse in den äußeren Ringen gering. Dadurch liegen noch verhältnismäßig viele Punkte weiter außen. Für p nahe bei 1 ist die Wahrscheinlichkeit, daß Punkte außen liegen nur noch sehr klein. Dadurch verlieren die Segment-Methoden ihren Vorteil gegenüber der Eckfeld-Methode und die Laufzeiten rücken näher zusammen.

5.2. Eine Näherung für $P(\mathcal{W}_l)$

Bisher hatten wir der Einfachheit halber $P(\mathcal{W}_l)$ nach oben durch $P(\overline{R^{(l+1)}})$ abgeschätzt. Das werden wir auch weiterhin tun. Hier sollen dennoch zwei Ansätze demonstriert werden, wie man $P(\mathcal{W}_l)$ feiner abschätzen kann.

Zunächst leiten wir uns aus der Definition von $\mathcal{W}_l$ eine *naive* Formel für $P(\mathcal{W}_l)$ her:

$$
\begin{aligned}
P(\mathcal{W}_l) \quad &= \quad P(R^{(l)}) \cdot \prod_{i=l+1}^{\infty} P(\overline{R^{(i+1)}}) \\
&\approx^{20} \quad \sum_{m=0}^{k} (-1)^m \binom{k}{m} (1 - m\,q(l))^n \cdot \prod_{i=l+1}^{\infty} \sum_{m=1}^{k} (-1)^{m+1} \binom{k}{m} (1 - m\,q(i))^n \\
&\qquad \text{wobei, } q(j) = \frac{p(1-p)^j}{2\,k\,j}
\end{aligned}
$$

Es handelt sich dabei um einen naiven Ansatz, da die Ereignisse $R^{(l)}$, $\overline{R^{(i)}}$ i.a. nicht stochastisch unabhängig sein werden. Vergleiche mit experimentellen Bestimmungen von $P(\mathcal{W}_l)$ zeigen jedoch , daß diese naive Näherung trotzdem recht gut ist. Dies kann man den Abbildungen 31-33 entnehmen. Zum Vergleich wurden hier die Werte aus Tabelle 1-3 auf Seite 60f für $k = 4$ gewählt.

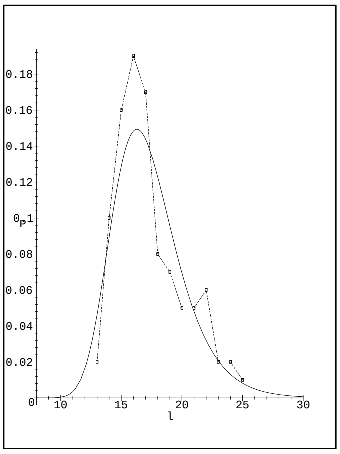

Abbildung 31:
$P(\mathcal{W}_l)$
$p = 0.1$

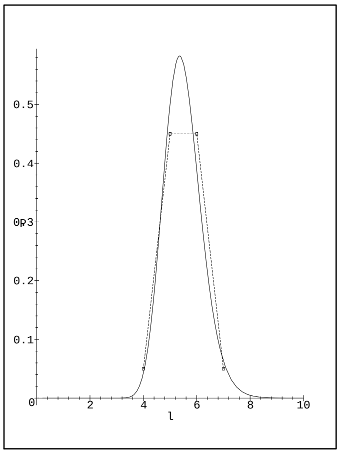

Abbildung 32:
$P(\mathcal{W}_l)$
$p = 0.5$

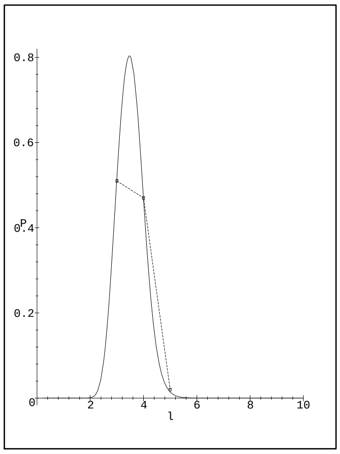

Abbildung 33:
$P(\mathcal{W}_l)$
$p = 0.7$

Eine andere Möglichkeit zur Abschätzung von $P(\mathcal{W}_l)$ ist folgende: Ebenfalls in Abschnitt 3 – auf Seite 28 – sah man bereits, daß $P(\mathcal{W}_l) \leq P(\overline{R^{(l+1)}})$ und $P(\mathcal{W}_l) \leq P(R^{(l)})$ ist. Darüber hinaus gilt asymptotisch

$$\lim_{l \to 0} P(\overline{R^{(l+1)}}) = 0 \quad \text{und} \quad \lim_{l \to \infty} P(\overline{R^{(l+1)}}) = 1$$

und

$$\lim_{l \to 0} P(R^{(l)}) = 1 \quad \text{und} \quad \lim_{l \to \infty} P(R^{(l)}) = 0.$$

[20] vgl. Abschnitt 4.1.2, Seite 50f; k ist die Anzahl der Ecken des Polygons.

$P(\mathcal{W}_l)$ läßt sich daher von $P(R^{(l)})$ und $P(\overline{R^{(l+1)}})$ einfassen. Dies ist in Abbildung 34 dargestellt.

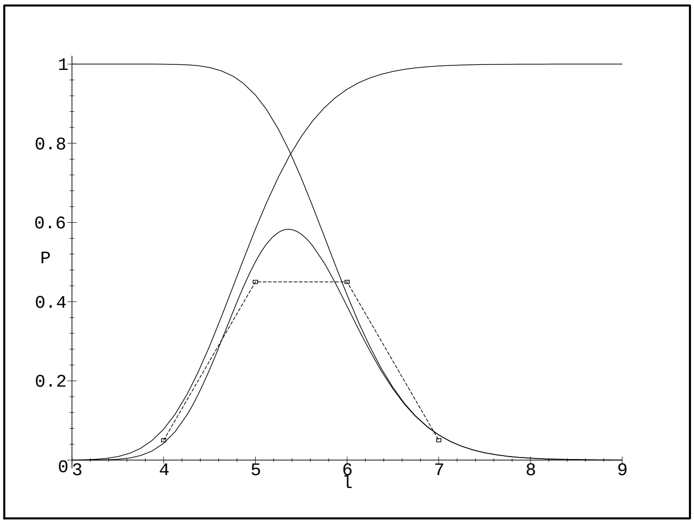

Abbildung 34: $p = 0.5$, $k = 4$, $n = 10^4$

Sei l_s der Wert bei dem sich $P(R^{(l)})$ und $P(\overline{R^{(l+1)}})$ schneiden. Dann läßt sich $P(\mathcal{W}_l)$ mit unseren bisherigen Ansätzen am genauesten durch

$$P(l) = \begin{cases} P(\overline{R^{(l+1)}}) : l \leq l_s \\ P(R^{(l)}) : l > l_s \end{cases}$$

abschätzen.

Zum Abschluß seien hier noch zwei Bemerkungen zu $P(l)$. Die Schwierigkeit, $P(l)$ anstelle von $P(\overline{R^{(l+1)}})$ in den bisherigen Abschätzung von $E[Y_C]$ zu verwenden, liegt darin, daß bei unserer Wahl von $q(l) = \frac{p(1-p)^l}{2\,k\,l}$ die Gleichung

$$P(\overline{R^{(l+1)}}) = P(R^{(l)})$$

schwer zu lösen und damit die Bestimmung von l_s nicht trivial ist.

Außerdem sei darauf hingewiesen, daß die Abschätzung $E[Y_l|\mathcal{W}_l] \leq E[Y_l]$ nicht uneingeschränkt gilt; für festes n und $l \to \infty$ gilt sie, wie wir durch die Simulationen gesehen haben, nicht. Schätzt man daher $P(\mathcal{W}_l)$ durch $P(l)$ ab, sollte man für große l in der Berechnung von $E[Y_C]$ statt $E[Y_l]$ den Term $k + E[Y_l]$ einsetzen.

6. Laufzeitanalyse der Segment-Methode

In Abschnitt 5 haben wir zwei Varianten des Verfahrens vorgestellt: die *Segment-Methode* und die *variable Segment-Methode*. Leider würde eine Analyse der variablen Segment-Methode Ansätze erfordern, die völlig von den bisherigen verschieden sind; da es bei dieser Methode keinen *vollbesetzten* Ring gibt, können wir nicht mehr ausrechnen, wie groß die Wahrscheinlichkeit ist, daß ein Ring *vollbesetzt* ist, oder wieviele Punkte *jenseits* dieses Ringes liegen. Stattdessen müßte man betrachten, welche Position die Ecksegmente einnehmen können und wie groß die Wahrscheinlichkeit dieser Position ist. Zu jeder möglichen Positionierung der Ecksegmente müßte man anschließend angeben, wieviele Punkte man im Mittel zur Berechnung der konvexen Hülle braucht. Dies ist offensichtlich ein höchst kompliziertes Unterfangen und daher werden wir uns auf eine Analyse der Segment-Methode beschränken.

Zunächst werden wir untersuchen, in welchen Schritten sich die Segment-Methode von der Eckfeld-Methode unterscheidet. Dabei setzen wir die Notationen und Definitionen aus Abschnitt 3 als bekannt voraus. Um zu konkreten Aussagen zu gelangen, werden wir hier sowie in Abschnitt 6.1 davon ausgehen, daß eine Wahrscheinlichkeitsverteilung vorliegt, bei der alle Felder $D^{(i)}$ von $Ring^{(l)}$ die gleiche Wahrscheinlichkeit haben; also $P_e(D^{((2l-1)^2)}) = \ldots = P_e(D^{((2l+1)^2-1)})$. Gehen wir nun die Schritte von Algorithmus `A` im einzelnen durch.

Schritt 1 Der Unterschied zur Eckfeld-Methode besteht in der Bestimmung des äußersten vollbesetzten Ringes. Wir brauchten bisher, lediglich zu überprüfen, ob ein Punkt, der in ein Eckfeld fällt, einen Ring *vollbesetzt*. Bei der Segment-Methode ist der Mehraufwand auch nicht erheblich größer. Der Mehraufwand (*overhead*) liegt daher unterhalb einer Konstanten c_{oh}[21]. Die erwarteten Kosten für Schritt 1 belaufen sich somit ebenfalls auf

$$E[X] \quad \leq \quad n \cdot \sum_{l=0}^{\infty} p \cdot (1-p)^l \cdot (4 + 2l + c_{oh}) \quad = \quad (2 + \frac{2}{p} + c_{oh}) \cdot n.$$

Schritt 2 Beginnen wir mit dem offensichtlichsten Unterschied zwischen den beiden Methoden, nämlich der Wahrscheinlichkeit $q(l)$ eines Eckfeldes und der Wahrscheinlichkeit $s(l)$ eines Ecksegmentes von $Ring^{(l)}$. Unter obigen Annahmen ist $q(l) = \frac{p(1-p)^l}{8l}$. Die Wahrscheinlichkeit eines Ecksegmentes genügt dann

[21] Sei wieder der Suchbaum $\mathcal{B}$ zur Punktlokalisation gegeben, z.B. so wie auf Seite 18. Man sieht, daß jedes Ecksegment von $Ring^{(l)}$ einem Ast des Suchbaumes $\mathcal{B}$ entspricht. Offensichtlich sind diese Äste disjunkt, d.h. kein Ast enthält einen anderen. Man kann daher jedem Ecksegment von $Ring^{(l)}$ einen Teilbaum $T_i^{(l)}$ mit Wurzel $W_i^{(l)}$, $i = 1, \ldots, 4$, zuordnen. Jeder Knoten von $T_i^{(l)}$ besitze nun drei Zeiger auf die Wurzeln $W_j^{(l)}$ der anderen Teilbäume $T_j^{(l)}$, $j \neq i$. Liegt nun ein Punkt in einem Ecksegment, d.h. einem solchen Teilbaum, dann läßt sich wieder in konstanter Zeit prüfen, ob alle Ecksegmente belegt sind. Im Falle derselben Idealisierungen wie in Fußnote 1 auf Seite 19 ist daher auch bei der Segment-Methode $c_{oh} = 2$.

folgender Abschätzung

$$
\begin{aligned}
s(l) &= \sum_{i=0}^{\infty}(1+2i)\cdot q(l+i) \;=\; \sum_{i=0}^{\infty}(1+2i)\cdot\frac{p(1-p)^{l+i}}{8\,(l+i)} \\
&\geq \frac{p(1-p)^l}{8\,l}+\frac{p(1-p)^l}{4}\left(\sum_{i=1}^{\infty}\frac{i\,(1-p)^i}{l+i}\right) \\
&\geq \frac{p(1-p)^l}{8\,l}+\frac{p(1-p)^l}{8\,l}\left(\sum_{i=1}^{l-1} i\,(1-p)^i\right)+\frac{p(1-p)^l}{8}\left(\sum_{i=l}^{\infty}(1-p)^i\right) \\
&= \frac{p^2(1-p)^l-(1-p)^{2l}+p(1-p)^{2l}-(1-p)^l p+(1-p)^l}{8\,l\,p} \\
&= \frac{(1-p)^l\left(1-p+p^2-(1-p)^{l+1}\right)}{8\,l\,p}.
\end{aligned}
$$

Im vorherigen Abschnitt haben wir – anhand von numerischen Simulationen – bereits gesehen, daß die Segment-Methode deutlich schneller ist als die Eckfeld-Methode. Wir werden daher von einer weiteren Analyse der Laufzeit des Verfahrens absehen und lediglich feststellen daß auch für die Segment-Methode die erwarteten Kosten von Schritt 2 asymptotisch sublinear sind.

Lemma 6.1 *Für genügend große n, $0 < p < 1$ und eine Konstante $c' > 1$ sind die erwarteten Kosten $E[Y_C]$ von Schritt 2 der Segment-Methode*

$$E[Y_C] \leq n^{\left(\frac{1}{c'}\right)}.$$

Die viel interessantere Frage ist, um wieviel die Segment-Methode im zweiten Schritt schneller ist als die Eckfeld-Methode. Dieser Frage werden wir im nächsten Abschnitt nachgehen.

6.1. Laufzeitvergleich der Eckfeld- und Segment-Methode

Der Laufzeitvergleich zwischen beiden Methoden soll an folgender Überlegung angesetzt werden: offensichtlich ist die Wahrscheinlichkeit eines Ecksegmentes größer als die eines Eckfeldes und somit auch die Wahrscheinlichkeit, daß alle Ecksegmente besetzt sind, größer als die Wahrscheinlichkeit, daß alle Eckfelder besetzt sind. Es sollte daher der Fall sein, daß der äußerste vollbesetzte Ring bei der Segment-Methode weiter außen liegt als bei der Eckfeld-Methode.

Zunächst definieren wir die Menge $S^{(l)}$ als das Gegenstück der Segment-Methode zu $R^{(l)}$ bei der Eckfeld-Methode.

Definition 6.1 ($S^{(l)}$)

$$S^{(l)} = \left\{ \omega \in \Omega \mid \text{ bei } \omega \text{ ist } Ring^{(l)} vollbesetzt \right\}$$

$S^{(l)}$ sei die Menge aller n-Auswahlen ω, bei denen bei der Segment-Methode $Ring^{(l)}$ vollbesetzt ist.

Bei der Eckfeld-Methode haben wir statt dem äußersten vollbesetzten Ring einen Ring $Ring^{(l')}$ gewählt für den gilt, daß $P(R^{(l')}) \geq \frac{1}{2}$ ist[22]. Bei der Segment-Methode können wir wieder denselben Ansatz wählen und schauen, wie sich l' dabei verändert.

Lemma 6.2 *Für*

$$n \geq \frac{\ln 8}{(1-p)^l} \cdot 8l$$

ist $P(S^{(l)}) \geq \frac{1}{2}$.

Beweis

$$\begin{aligned} n &\geq \frac{\ln 8}{(1-p)^l} \cdot 8l \\ n &\geq \frac{\ln 8}{(1-p)^l \, (1-p+p^2-(1-p)^2)} \cdot 8lp \\ \Rightarrow \quad n &\geq \frac{\ln 8}{(1-p)^l \, (1-p+p^2-(1-p)^{l+1})} \cdot 8lp. \end{aligned}$$

Da $\frac{(1-p)^l \left(1-p+p^2-(1-p)^{l+1}\right)}{8lp}$ eine untere Schranke für $s^{(l)}$ ist, folgt

$$\begin{aligned} n &\geq \frac{\ln 8}{s(l)} \\ \Rightarrow \quad P(S^{(l)}) &\geq 1 - 4(1-s(l))^n \geq \frac{1}{2}. \end{aligned}$$

□

Wie wir sehen, liefert die Segment-Methode eine andere hinreichende Ungleichung für l' als die Eckfeld-Methode. Unsere Vermutung war, daß l' bei der Segment-Methode größer ist als bei der Eckfeld-Methode. Um diese Vermutung zu verifizieren, sind Ungleichungen nicht exakt genug. Im weiteren Verlauf werden wir das l' der Eckfeld-Methode l_e und das l' der Segment-Methode l_s nennen. Wandeln wir die Ungleichungen für l_e bzw. l_s in Gleichungen um, dann erhalten wir eindeutige Werte.

[22] vgl. Lemma 3.1

Damit wären wir dann in der Lage, die Laufzeiten der beiden Verfahren exakt zu vergleichen. Betrachten wir also

$$\begin{aligned} n &= \frac{8\ln 8 \cdot l_e}{p(1-p)^{l_e}} \\ \text{und} \quad n &= \frac{8\ln 8 \cdot l_s}{(1-p)^{l_s}}. \end{aligned}$$

Satz 6.1 *Für Gleichungen der Form*

$$x\,e^x = t \qquad , t > 0,\ x > 0$$

existiert eine eindeutige Lösung. Asymptotisch, d.h. für $t \to \infty$, läßt sich eine Näherungslösung angeben

$$x \approx \ln t - \ln\ln t.$$

Ein Beweis von Satz 6.1 findet sich in [5] auf den Seiten 25–28. Mit Satz 6.1 lassen sich Lösungen für l_e und l_s bestimmen.

Lemma 6.3 *Es ist für genügend große n und $0 < p < 1$*

$$l_e = -\log_{(1-p)}\left(\frac{-np\,\ln(1-p)}{\ln\left(-\frac{np\,\ln(1-p)}{8\ln 8}\right)\ 8\ln 8}\right)$$

und

$$l_s = -\log_{(1-p)}\left(\frac{-n\,\ln(1-p)}{\ln\left(-\frac{n\,\ln(1-p)}{8\ln 8}\right)\ 8\ln 8}\right).$$

Beweis

$$\begin{aligned} n &= \frac{8\ln 8 \cdot l_e}{p(1-p)^{l_e}} \\ &= \frac{8\ln 8}{p} \cdot l_e \left(\frac{1}{1-p}\right)^{l_e} \\ \Leftrightarrow \quad \frac{n\,p}{8\ln 8} &= l_e\left(\frac{1}{1-p}\right)^{l_e} \\ &= l_e \cdot e^{\ln\left(\frac{1}{1-p}\right) l_e} \\ \Leftrightarrow \quad \frac{n\,p}{8\ln 8}\cdot\ln\left(\frac{1}{1-p}\right) &= \ln\left(\frac{1}{1-p}\right) l_e \cdot e^{\ln\left(\frac{1}{1-p}\right) l_e}. \end{aligned}$$

Eine Variablentransformation, $x = \ln\left(\frac{1}{1-p}\right) l_e$, liefert uns die gewünschte Form

$$x\, e^x = \frac{n\, p}{8 \ln 8} \cdot \ln\left(\frac{1}{1-p}\right) \qquad , x > 0.$$

Daraus erhalten wir unmittelbar die Lösung

$$l_e \;=\; -\frac{x}{\ln(1-p)} \;=\; -\frac{\ln\left(-\frac{n\,p\,\ln(1-p)}{8\ln 8}\right) - \ln\ln\left(-\frac{n\,p\,\ln(1-p)}{8\ln 8}\right)}{\ln(1-p)}.$$

Mit dem gleichen Ansatz löst man nach l_s auf:

$$l_s \;=\; -\frac{\ln\left(-\frac{n\,\ln(1-p)}{8\ln 8}\right) - \ln\ln\left(-\frac{n\,\ln(1-p)}{8\ln 8}\right)}{\ln(1-p)} \;=\; -\log_{(1-p)}\left(\frac{-n\,\ln(1-p)}{\ln\left(-\frac{n\,\ln(1-p)}{8\ln 8}\right)\, 8\ln 8}\right).$$

$\square$

Daraus folgt

Satz 6.2 *Unter der Annahme, daß alle Felder eines Ringes dieselbe Wahrscheinlichkeit besitzen, verhalten sich für* $0 < p < 1$ *die erwarteten Kosten* $E_e[Y_C]$ *von Schritt 2 der Eckfeld-Methode zu den erwarteten Kosten* $E_s[Y_C]$ *von Schritt 2 der Segment-Methode wie*

$$\frac{E_e[Y_C]}{E_s[Y_C]} \approx \frac{1}{p} \cdot \frac{\ln\left(-\frac{n\,p\,\ln(1-p)}{8\ln 8}\right)}{\ln\left(-\frac{n\,\ln(1-p)}{8\ln 8}\right)}.$$

Beweis

Mit Lemma 6.3 berechnen sich die erwarteten Kosten zu

$$\begin{aligned}
E_e[Y_C] \;&\leq\; \tilde{c} \cdot n \log n \cdot (1-p)^{-\log_{(1-p)}\left(\frac{-n\,p\,\ln(1-p)}{\ln\left(-\frac{n\,p\,\ln(1-p)}{8\ln 8}\right) 8\ln 8}\right)} \\
&=\; \tilde{c} \cdot n \log n \cdot \frac{\ln\left(-\frac{n\,p\,\ln(1-p)}{8\ln 8}\right) 8\ln 8}{-n\,p\,\ln(1-p)}
\end{aligned}$$

und

$$E_s[Y_C] \;\leq\; \tilde{c} \cdot n \log n \cdot \frac{\ln\left(-\frac{n\,\ln(1-p)}{8\ln 8}\right) 8\ln 8}{-n\,\ln(1-p)}.$$

Das Verhältnis der Laufzeiten ist daher der Quotient:

$$\frac{E_e[Y_C]}{E_s[Y_C]} \;\approx\; \frac{1}{p} \cdot \frac{\ln\left(-\frac{n\,p\,\ln(1-p)}{8\ln 8}\right)}{\ln\left(-\frac{n\,\ln(1-p)}{8\ln 8}\right)}.$$

■

$\frac{E_e[Y_C]}{E_s[Y_C]}$ ist zwar abhängig von der Eingabegröße n, konvergiert aber für $n \to \infty$ gegen einen Grenzwert. $N_0 = 10^{20}$ ist groß genug, so daß sich für $n \geq N_0$ der Quotient $\frac{E_e[Y_C]}{E_s[Y_C]}$ nicht mehr nennenswert ändert. Der Verlauf von $\frac{E_e[Y_C]}{E_s[Y_C]}$ für $n = 10^{20}$ und $0 < p < 1$ ist in Abbildung 35 wiedergegeben.

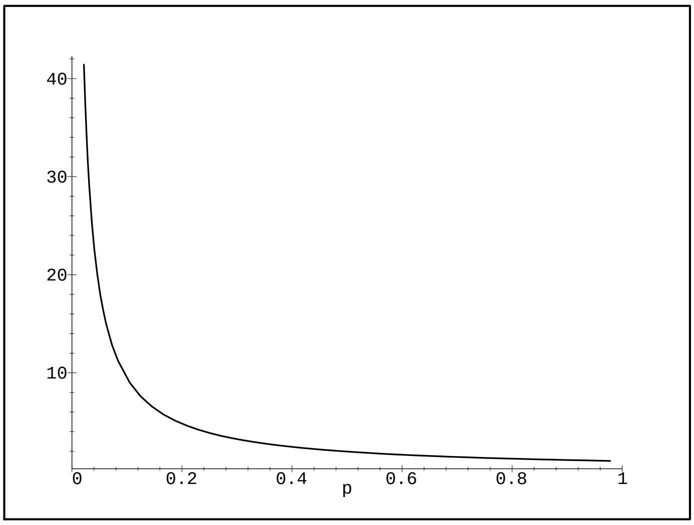

Abbildung 35: Verlauf von $\frac{E_e[Y_C]}{E_s[Y_C]}$ für $n = 10^{20}$

Wir haben gesehen, daß auch für die Schachtelungstiefe $l_n \approx \alpha \log n$ die Wahrscheinlichkeit $P(R^{(l_n)}) \geq \frac{1}{2}$ ist. Eine interessante Frage ist nun, ob es günstiger ist, l_n oder $l_e \approx -\frac{\log n - \log\log n}{\ln(1-p)}$ für die Abschätzung von $E[Y_C]$ zu benutzen. Es läßt sich leicht zeigen, daß asymptotisch in n die Schachtelungstiefe $l_n \leq l_e$ ist. Unter Vernachlässigung aller konstanten Faktoren erhält man

$$\begin{aligned}
\alpha \log n &= \frac{c}{|\ln(1-p)|} \cdot \log n \leq \frac{\log\left(\frac{n}{\log n}\right)}{|\ln(1-p)|} \quad , 0 < c < 1 \\
c \cdot \log n &\leq \log\left(\frac{n}{\log n}\right) \\
n^c &\leq \frac{n}{\log n} \\
\log n &\leq n^{1-c}.
\end{aligned}$$

Die Schachtelungstiefe l_e ist daher günstiger als l_n. Da aber $l_n = \lfloor \alpha \log n \rfloor$ mathematisch einfacher handhabbar ist und darüber hinaus ebenfalls die gewünschten Resultate liefert, werden wir weiterhin an dieser Abschätzung festhalten.

7. Konvexe Hüllen im $\mathbb{R}^d$ bei einer Hyperquader-Schalen-Schachtelung

Im folgenden wollen wir das in Abschnitt 2 vorgestellte Verfahren, Algorithmus `A`, im d-Dimensionalen anwenden. Anstatt aber Schachtelungen mit beliebigen Körpern zuzulassen, werden wir uns auf Hyper-Quader beschränken[23].

Der $\mathbb{R}^d$ wird also mit einem Gitter aus seitenparallelen, ineinander geschachtelten Hyperquadern überdeckt – ganz analog zu der Vorgehensweise in den Abschnitten 3 und 4; man hat also einen innersten Hyperquader Q_0, legt darum seitenparallel einen weiteren Q_1 etc. Die Gebiete, die wir vorher *Ringe* genannt haben, sind im d-Dimensionalen *Schalen* und offenbar werden diese Schalen nicht mehr von Linien, sondern von Hyperebenen begrenzt. Konstruktionsbedingt zerschneiden die Begrenzungs-Hyperebenen die Schalen wieder in viele, kleine Raumelemente, die wir *Quader* nennen wollen.

Definition 7.1 (Hyperquader-Schalen-Schachtelung) *Die oben beschriebene Konstruktion heiße* Hyperquader-Schalen-Schachtelung.
Die Schale der Schachtelungstiefe l sei $Schale^{(l)}$.

Die Anzahl der Hyperebenenvergleiche, um die Punkte des innersten Hyperquaders, $Schale^{(0)}$, zu lokalisieren, ist $2d$. Denn $Schale^{(0)}$ grenzt jede Dimension eines kartesischen Koordinatensystems des $\mathbb{R}^d$ in beide Richtungen ab. Ein Punkt, der in $Schale^{(1)}$ liegt, kann in jeder Dimension höchstens eine der beiden Abgrenzungen verletzen; er liegt daher weiterhin innerhalb der anderen Abgrenzung. Liegt ein Punkt in einem Eckquader von $Schale^{(1)}$, dann sind gegenüber $Schale^{(0)}$ genau d Abgrenzungen nicht mehr gültig, und, wie wir bereits erwähnten, kann ein Punkt in $Schale^{(1)}$ auch nicht mehr als d Abgrenzungen verletzen. Folglich kann man mit maximal $2d+d$ Hyperebenen-Abfragen alle Punkt in $Schale^{(1)}$ bestimmen. Mit derselben Betrachtung sieht man, daß alle Punkte in $Schale^{(2)}$ höchstens d Abfragen gegenüber $Schale^{(1)}$ verletzen können und somit mit höchstens $2d + d + d$ Abfragen lokalisierbar sind. Allgemein gilt,daß sich mit $2d + l \cdot d$ Hyperebenen-Abfragen feststellen läßt, ob ein Punkt in $Schale^{(l)}$ liegt. Zur Lokalisation der Punkte in den Schalen bauen wir wieder einen Suchbaum auf, der obige Abfragenanzahl einhält. Den Mehraufwand pro Punktlokalisation zur Bestimmung der äußersten *vollbesetzten* Schale schätzen wir wieder durch eine Konstante c_{oh} ab[24].

[23] In Abschnitt 4 haben wir beliebige reguläre Polygone zur Schachtelung zugelassen. Es liegt daher nahe, dies auch im $\mathbb{R}^d$ zu gestatten. Das Problem ist jedoch, daß es in höheren Dimensionen nicht zu jeder beliebigen Anzahl von Ecken einen regulären Körper gibt (Im $\mathbb{R}^3$ gibt es z.B. nur noch fünf reguläre Körper). Damit wir also nicht für jede Dimension angeben müssen, welche regulären Körper existieren, werden wir uns aus Gründen der Einfachheit in höheren Dimensionen auf eine Schachtelung mit Hyperquadern beschränken.

[24] Im Idealfall beträgt auch diesmal $c_{oh} = 2$. Die Idee ist die gleiche wie in Fußnote 1, Seite 19. Im $\mathbb{R}^d$ braucht man jedoch $2^d - 1$ Zeiger auf die anderen Speicherzellen.

7.1. erwartete Laufzeit bei einer Hyperquader-Schalen-Schachtelung

Wir betrachten Wahrscheinlichkeitsdichten $f : \mathbb{R}^d \to \mathbb{R}$, für die es Hyperquader-Schalen-Schachtelungen gibt, so daß die Entsprechungen im $\mathbb{R}^d$ zu den Eigenschaften **E1** und **E2** erfüllt sind.

D1 Die Wahrscheinlichkeit, daß ein Punkt $a \in \mathbb{R}^d$ in $Schale^{(l)}$ liegt, beträgt für $0 < p < 1$

$$\begin{aligned} P(Schale^{(0)}) &= p, \\ P(Schale^{(l+1)}) &= (1-p) \cdot P(Schale^{(l)}) \end{aligned}$$

für $l = 0, 1, 2, \ldots$.

D2 Die Eckquader von $Schale^{(l)}$ sind gegenüber den anderen Quadern von $Schale^{(l)}$ stochastisch nicht benachteiligt und haben alle dieselbe Wahrscheinlichkeit.

Ebenso wie bei der Rechteck-Ring-Schachtelung ergibt sich durch die Hyperquader-Schalen-Schachtelung eine Verringerung der Beobachtungstiefe[25]. Es genügt daher, als Wahrscheinlichkeitsexperiment die Ziehung einer n-Auswahl aus einer Hyperquader-Schalen-Schachtelung zu betrachten. Der dazugehörige Wahrscheinlichkeitsraum ist formal identisch mit dem Produktraum aus Abschnitt 3.2. Wir werden die Notationen und Definitionen aus Abschnitt 3 wiederverwenden; es sollte jedoch bedacht werden, daß wir nun von Schalen statt Ringen ausgehen.

7.1.1. erwartete Kosten von Schritt 1

Die erwartete Anzahl von Hyperebenen-Abfragen zur Lokalisierung von n Punkten in der Hyperquader-Schalen-Schachtelung und der gleichzeitigen Bestimmung der äußersten vollbesetzten Schale beträgt unter Verwendung obiger Betrachtungen:

$$E[X] \quad \leq \quad n \cdot \sum_{l=0}^{\infty} p \cdot (1-p)^l \cdot (2d + ld + c_{oh}) \quad = \quad (d + \frac{d}{p} + c_{oh}) \cdot n,$$

wobei die Zufallsvariable X so definiert sei wie auf Seite 24.

7.1.2. erwartete Kosten von Schritt 2

Zur Ermittlung der Laufzeit von Schritt 2 werden wir die bereits bekannte Abschätzung[26] $P(\mathcal{W}_l) \leq P(\overline{R^{(l+1)}})$ benutzen.

[25] Die Hyperquader-Schalen-Schachtelung definiert als Gitterfelder Quader $\subset \mathbb{R}^d$. Bei der Verringerung handelt es sich daher um eine Zufallsvariable von $((\mathbb{R}^d)^n, (\mathcal{B}^d)^n)$ nach $(D^n, Pot(D^n))$, wobei $\mathcal{B}^d$ die d-dimensionale Borel-σ-Algebra ist.

[26] siehe Seite 28

Bei der Berechnung von $P(\overline{R^{(l+1)}})$, bzw. $P(R^{(l)})$, der Wahrscheinlichkeit, daß alle Eck*quader* besetzt sind, haben wir bisher immer angenommen, daß die Eckquader einer Schale stochastisch nicht benachteiligt seien. Diese Forderung wollen wir beibehalten. Dazu müssen wir aber wissen, aus wievielen Quadern eine Schale besteht.

Betrachtet man die Konstruktionsweise des Hyperquader-Gitters, sieht man, daß sich die Anzahl, K_l, der Quader in $Schale^{(l)}$ wie folgt berechnet: Es sei $(2l+1)^d$ die Anzahl der Quader in Hyperquader Q_l. Die Anzahl der Quader in $Schale^{(l)}$ ist die Differenz der Anzahl der Quader von Hyperquader Q_l und Hyperquader Q_{l-1}. Somit lautet die Gleichung für K_l:

$$\begin{aligned} K_l &= (2l+1)^d - (2(l-1)+1)^d \\ &= (2l+1)^d - (2l-1)^d \qquad \text{für } l = 1, 2, \ldots. \end{aligned}$$

Verwendet man nun, daß

$$\begin{aligned} (2l+1)^d &= \sum_{k=0}^{d} \binom{d}{k} (2l)^{d-k} \qquad \text{und} \\ (2l-1)^d &= \sum_{k=0}^{d} (-1)^k \binom{d}{k} (2l)^{d-k} \end{aligned}$$

ist, erhält man

$$K_l = \sum_{k=0}^{d} \binom{d}{k} (2l)^{d-k} (1 - (-1)^k).$$

Dadurch haben wir zwar keine geschlossene Formel für K_l, aber wir wissen nun, daß

Lemma 7.1 *Im $\mathbb{R}^d$ gilt $K_l \in O(l^{d-1})$.*

Beweis

Es ist

$$\begin{aligned} K_l &= \sum_{k=0}^{d} \binom{d}{k} (2l)^{d-k} (1 - (-1)^k) \\ &= \sum_{k=1}^{d} \binom{d}{k} (2l)^{d-k} (1 - (-1)^k) \\ &\leq (2l)^{d-1} \cdot \sum_{k=1}^{d} \binom{d}{k} (1 - (-1)^k) \\ &= (2l)^{d-1} \cdot 2 \cdot 2^{d-1} \\ &= 2^{2d-1} \cdot l^{d-1}. \end{aligned}$$

□

Wie groß ist nun die Wahrscheinlichkeit $P(R^{(l)})$, daß alle Eckquader von $Schale^{(l)}$ besetzt sind? Oder anders ausgedrückt: gewährleistet die Wahl $l_n = \lfloor \alpha \log n \rfloor$, $0 < \alpha < -\frac{1}{\log(1-p)}$, wieder, daß $P(R^{(l)}) \geq \frac{1}{2}$ ist?

$P(R^{(l)})$ läßt sich wieder mit der Siebformel von Poincaré-Sylvester berechnen. Folglich darf man die Wahrscheinlichkeit also genauso abschätzen, wie wir es in Abschnitt 4.1.2 auf Seite 50 getan haben. Das einzige, was wir dazu noch benötigen, ist die Anzahl der Eckfelder einer Schale.

Dieses Problem läßt sich jedoch leicht lösen, wenn man die Repräsentation eines Einheits-Hyperquaders durch seine Eckpunkte in einem d-dimensionalen, kartesischen Koordinatensystem betrachtet.

Hier sei jeweils ein Beispiel im Zwei- und Dreidimensionalen:

- im Zweidimensionalen lautet die Koordinatenrepräsentation der Ecken

$$(0,0);\ (0,1);\ (1,0);\ (1,1)$$

- im Dreidimensionalen

$$(0,0,0);\ (0,0,1);\ (0,1,0);\ (0,1,1);\ (1,0,0);\ (1,0,1);\ (1,1,0);\ (1,1,1).$$

Es gibt also eine Bijektion zwischen den Ecken eines Hyperquaders und dem $\mathbb{B}^d$. Somit besitzt ein d-dimensionaler Hyperquader $|\mathbb{B}^d|$ Ecken.

Wir sind nun in Kenntnis aller Angaben, um die Wahrscheinlichkeit $P(R^{(l)})$ abzuschätzen. Es seien $a_1, \ldots, a_n$ Punkte des $\mathbb{R}^d$. Da die Wahrscheinlichkeitsdichte auf $\mathbb{R}^d$ die Eigenschaft **D2** erfüllt, besitzen alle Eckquader dieselbe Wahrscheinlichkeit $q(l)$. Es gilt

Lemma 7.2 *Für $n \geq \frac{(d+1)\ln 2}{q(l)}$ ist $P(R^{(l)}) \geq \frac{1}{2}$.*

Beweis

Aus $n \geq \frac{(d+1)\ln 2}{q(l)}$ folgt unter Verwendung von $\ln x \leq x - 1$

$$\begin{aligned}
n \cdot \ln(1-q(l)) &\leq n \cdot -q(l) \leq -\ln 2^{(d+1)} \\
\Rightarrow \quad -\ln 2^{(d+1)} &\geq n \cdot \ln(1-q(l)) \\
\frac{1}{2^{d+1}} &\geq (1-q(l))^n \\
\Rightarrow \quad P(R^{(l)}) &\geq 1 - 2^d (1-q(l))^n \geq \frac{1}{2}.
\end{aligned}$$

□

In den Abschnitten 3 und 4 konnten wir bei der Abschätzung $q(l) \geq \frac{p(1-p)^l}{K_l}$ die Anzahl K_l der Felder von $Ring^{(l)}$ genau angeben. Da wir diesmal K_l nicht exakt bestimmen konnten, wollen wir K_l durch $c \cdot l^{d-1}$ abschätzen, wobei c eine hinreichend große Konstante sei. Damit erhalten wir:

$$n \geq \frac{(d+1)\ln 2}{p(1-p)^l} \cdot c \cdot l^{d-1}.$$

Lemma 7.3 *Ist n groß und $0 < p < 1$, dann ist für Schachtelungstiefen $l \leq l_n = \lfloor \alpha \log n \rfloor$, $0 < \alpha < -\frac{1}{\log(1-p)}$, die Voraussetzung von Lemma 7.2 erfüllt.*

Beweis

Setzen wir wieder $l_n = \lfloor \alpha \log n \rfloor$, $0 < \alpha < -\frac{1}{\log(1-p)}$, dann muß n die Bedingung

$$n \geq \frac{(d+1)\ln 2 \cdot c}{p} \cdot \frac{(\alpha \log n)^{d-1}}{n^{\alpha \log(1-p)}}$$

erfüllen. Die Frage ist nun, ob obige Ungleichung für beliebige Dimensionen $d \geq 3$ erfüllbar ist. Formen wir die Ungleichung dazu etwas um:

$$\begin{array}{rcll} n & \geq & c' \cdot \frac{(\alpha \log n)^{d-1}}{n^{\alpha \log(1-p)}} & , c' = \frac{(d+1)\ln 2 \cdot c}{p} \\ n^{1+\alpha \log(1-p)} & \geq & c' \cdot (\alpha \log n)^{d-1} & \\ n^{\frac{1+\alpha \log(1-p)}{d-1}} & \geq & c'^{\frac{1}{d-1}} \cdot \alpha \log n. & \end{array}$$

Wir können also die Frage umformulieren; wir suchen demnach nach einem $n > 1$, so daß für Konstanten $\tilde{c}, c^* > 0$

$$n^{\tilde{c}} \geq c^* \cdot \log n$$

gilt. Aber das ist selbstverständlich, denn $x^\alpha = o(e^{\beta x})$, d.h. $e^{\beta x}$ wächst schneller als jede Potenz x^α für $\alpha, \beta > 0$ und $x \to \infty$. Woraus wir entnehmen können:

$$\begin{array}{crcl} & n^{\tilde{c}} & \geq & c^* \cdot \log n \\ & n^{\tilde{c}} & \geq & \log n^{c^*} \\ \Leftrightarrow & \sqrt[\tilde{c}]{n} & \leq & b^{\left(n^{c^*}\right)} = e^{\left(n^{c^*} \cdot \ln b\right)} \\ & & & , \text{Umkehrfunktionen}, b > 1, \end{array}$$

falls n groß genug ist. □

Daraus folgt sofort

Lemma 7.4 *Für große n und $0 < p < 1$ ist für Schachtelungstiefe $l_n = \lfloor \alpha \log n \rfloor$, $0 < \alpha < \frac{1}{\log(1-p)}$, und $r = 0, \ldots, l_n - 1$*

$$P(\overline{R^{(l_n - r)}}) \leq \left(\frac{1}{2}\right)^{\left(\frac{1}{1-p}\right)^r}.$$

Beweis

Der Beweis verläuft identisch zu dem Beweis von Lemma 4.3. Es ist

$$q(l_n - r) = \frac{p(1-p)^{l_n - r}}{c \cdot (l_n - r)^{d-1}} \geq \frac{p(1-p)^{l_n - r}}{c \cdot {l_n}^{d-1}} = q(l_n)(1-p)^{-r}.$$

Als Folge gilt die Ungleichungskette

$$1 - q(l_n - r) \leq 1 - q(l_n)(1-p)^{-r} \leq (1 - q(ln))^{(1-p)^{-r}}.$$

Setzt man nun im Beweis von Lemma 4.3 den Wert $k = 2^d$, dann erhält man mit Lemma 7.3 sofort die Behauptung. □

Wir können damit die erwarteten Kosten $E[Y_C]$ von Schritt 2 bei einer Hyperquader-Schalen-Schachtelung abschätzen.

Lemma 7.5 *Für große n, $0 < p < 1$ und eine hinreichend große Konstante $c > 1$ ist*

$$E[Y_C] \leq c \cdot \left(n(1-p)^{l_n}\right)^{\left(\lfloor \frac{d}{2} \rfloor + 1\right)}.$$

Beweis

$$\begin{aligned}
E[Y_C] &= \sum_{i=0}^{n} P\left(Y = i\right) \cdot T(i) \\
&= \sum_{l=0}^{\infty} P(\mathcal{W}_l) \cdot \sum_{i=0}^{n} P\left(Y_l = i | \mathcal{W}_l\right) \cdot c \cdot i^{\left(\lfloor \frac{d}{2} \rfloor + 1\right)} \quad , T(i) = c \cdot i^{\left(\lfloor \frac{d}{2} \rfloor + 1\right)} \\
&\leq^{27} c \cdot \sum_{i=0}^{n} P\left(Y_{l'} = i | \mathcal{W}_{l'}\right) \, i^{\left(\lfloor \frac{d}{2} \rfloor + 1\right)} \cdot \sum_{l=l'}^{\infty} P\left(\mathcal{W}_l\right) \\
&\quad + c \cdot \sum_{l=0}^{l'-1} P\left(\mathcal{W}_l\right) \cdot \sum_{i=0}^{n} P\left(Y_l = i | \mathcal{W}_l\right) \, i^{\left(\lfloor \frac{d}{2} \rfloor + 1\right)}
\end{aligned}$$

Hier stoßen wir nun auf ein Problem. Nehmen wir einen Standardalgorithmus zur Bestimmung der konvexen Hülle von i Punkten im $\mathbb{R}^d$, z.B. das *Beneath-Beyond-*

[27] Da die Wahrscheinlichkeitsverteilung P exponentiell in den Schalen abnimmt, haben wir bereits gesehen, daß für $j > k$ $\sum_{i=0}^{n} P(Y_j = i|\mathcal{W}_j) \cdot i = E[Y_j|\mathcal{W}_j] < E[Y_k|\mathcal{W}_k] = \sum_{i=0}^{n} P(Y_k = i|\mathcal{W}_k) \cdot i$ ist (vgl. Abschnitt 3, Seite 31, Fußnote 8). Das liegt daran, daß es ein i_0 gibt, so daß für $i > i_0$ $P(Y_j = i|\mathcal{W}_j) < P(Y_k = i|\mathcal{W}_k)$ ist. Dann muß aber auch $\sum_{i=0}^{n} P(Y_j = i|\mathcal{W}_j) \cdot i^{(\lfloor \frac{d}{2} + 1 \rfloor)} \leq \sum_{i=0}^{n} P(Y_k = i|\mathcal{W}_k) \cdot i^{(\lfloor \frac{d}{2} + 1 \rfloor)}$ sein.

Verfahren, so beträgt die worst-case Laufzeit $O(i^{\lfloor\frac{d}{2}\rfloor+1})$. Die Abschätzung von Vermutung 3.1, $E[Y_l|\mathcal{W}_l] \leq E[Y_l]$, hilft uns daher nicht weiter. In Abschnitt 7.2 werden wir jedoch sehen, daß wir ,falls l' geeignet gewählt ist, für große n und $l \leq l'$ anscheinend die Abschätzung $\sum_{i=0}^{n} P\left(Y_l = i|\mathcal{W}_l\right) \cdot i^{\left(\lfloor\frac{d}{2}\rfloor+1\right)} \leq \sum_{i=0}^{n} P\left(Y_l = i\right) \cdot i^{\left(\lfloor\frac{d}{2}\rfloor+1\right)}$ verwenden dürfen. Daraus folgt

$$\begin{aligned} E[Y_C] &\leq c \cdot \sum_{i=0}^{n} P\left(Y_{l'} = i\right)\, i^{\left(\lfloor\frac{d}{2}\rfloor+1\right)} \cdot \sum_{l=l'}^{\infty} P\left(\mathcal{W}_l\right) \\ &\quad + c \cdot \sum_{l=0}^{l'-1} P\left(\mathcal{W}_l\right) \cdot \sum_{i=0}^{n} P\left(Y_l = i\right)\, i^{\left(\lfloor\frac{d}{2}\rfloor+1\right)}. \end{aligned}$$

Mit Lemma 7.7 aus Abschnitt 7.2 läßt sich $E[Y_C]$ weiter abschätzen.

$$E[Y_C] \leq c \cdot c' \cdot \left(\left(n(1-p)^{l'}\right)^{\left(\lfloor\frac{d}{2}\rfloor+1\right)} + \sum_{l=0}^{l'-1} P\left(\mathcal{W}_l\right) \cdot \left(n(1-p)^{l}\right)^{\left(\lfloor\frac{d}{2}\rfloor+1\right)} \right)$$

und numerische Simulationen[28] legen nahe, daß wieder ein Wert $b > 1$ existiert, so daß für große n die Wahl $l' = l_n$ hinreichend ist.

$$\begin{aligned} E[Y_C] &\leq c^* \cdot \left(\left(n(1-p)^{l_n}\right)^{\left(\lfloor\frac{d}{2}\rfloor+1\right)} + n^{\left(\lfloor\frac{d}{2}\rfloor+1\right)} \cdot \sum_{r=0}^{l_n-1} P\left(\overline{R^{(l_n-r)}}\right) \cdot (1-p)^{(l_n-1-r)\cdot\left(\lfloor\frac{d}{2}\rfloor+1\right)} \right) \\ &\leq c^* \cdot \left(n(1-p)^{l_n}\right)^{\left(\lfloor\frac{d}{2}\rfloor+1\right)} \cdot \left(1 + \sum_{r=0}^{l_n-1} \frac{1}{2^{(1-p)^{-r}}} \cdot \frac{1}{(1-p)^{(r+1)\cdot\left(\frac{d}{2}+1\right)}}\right). \end{aligned}$$

Da wir Aussagen für große Werte von n machen, sind wir an dem Konvergenzverhalten der Reihe

$$\frac{1}{(1-p)^{\left(\frac{d}{2}+1\right)}} \cdot \sum_{r=0}^{\infty} \frac{1}{2^{(1-p)^{-r}}} \cdot \frac{1}{(1-p)^{r\cdot\left(\frac{d}{2}+1\right)}}$$

interessiert. Die Konvergenz dieser Reihe für $0 < p < 1$ erhält man mit dem Wurzelkriterium:

$$\begin{aligned} \varrho &:= \overline{\lim_{n\to\infty}} \sqrt[n]{|a_n|} \\ &= \lim_{n\to\infty} \frac{1}{\sqrt[n]{2^{\left(\frac{1}{1-p}\right)^n} \cdot (1-p)^{n\cdot\left(\frac{d}{2}+1\right)}}} < 1 \\ \Leftrightarrow &\quad \lim_{n\to\infty} \frac{1}{2^{\left[\left(\frac{1}{1-p}\right)^n \cdot \frac{1}{n}\right]}} < (1-p)^{\left(\frac{d}{2}+1\right)} \\ \Leftrightarrow &\quad \lim_{n\to\infty} -\left(\frac{1}{1-p}\right)^n \cdot \frac{1}{n} \cdot \log 2 < \left(\frac{d}{2}+1\right) \cdot \log(1-p) \\ \Leftrightarrow &\quad \lim_{n\to\infty} n \log|(1-p)| - \log n + \log\log 2 > \log\left(\frac{d}{2}+1\right) + \log|\log(1-p)|. \end{aligned}$$

[28] vgl. Abschnitt 7.2

$\left(1 + \sum_{r=0}^{l_n-1} \frac{1}{2^{(1-p)^{-r}}} \cdot \frac{1}{(1-p)^{(r+1)\cdot(\frac{d}{2}+1)}}\right)$ läßt sich also für $0 < p < 1$ durch einen konstanten Faktor abschätzen. Aus dem Grund, den wir in Abschnitt 3.3.2 auf Seite 34 bereits genannt haben, divergiert die Reihe für $p \to 1$ und $p \to 0$. □

Wir werden nun zeigen, daß für $d \geq 3$ die erwarteten Kosten von Schritt 2 asymptotisch sublinear sind.

Lemma 7.6 *Für $d \geq 3$, große n, eine Konstante $c' > 1$ und $0 < p < 1$ ist*

$$E[Y_C] \leq n^{\frac{1}{c'}}.$$

Beweis

Es existiert ein $\alpha > 0$ für das

$$\frac{\lfloor\frac{d}{2}\rfloor + 1 - \frac{1}{c'}}{\lfloor\frac{d}{2}\rfloor + 1} \cdot \frac{1}{|\log(1-p)|} < \alpha < \frac{1}{|\log(1-p)|}$$

ist. Somit erhält man

$$\begin{aligned} \left\lfloor\frac{d}{2}\right\rfloor + 1 - \frac{1}{c'} &< \alpha \cdot |\log(1-p)| \cdot \left(\left\lfloor\frac{d}{2}\right\rfloor + 1\right) \\ \alpha \cdot \log(1-p) \cdot \left(\left\lfloor\frac{d}{2}\right\rfloor + 1\right) &< \frac{1}{c'} - \left\lfloor\frac{d}{2}\right\rfloor - 1. \end{aligned}$$

Mit der Wahl von $l_n = \lfloor \alpha \log n \rfloor$ gilt daher für $n \to \infty$ und eine Konstante $c > 1$

$$\begin{aligned} \frac{c}{1-p} \cdot n^{\alpha \cdot \log(1-p) \cdot (\lfloor\frac{d}{2}\rfloor + 1)} &\leq n^{(\frac{1}{c'} - \lfloor\frac{d}{2}\rfloor - 1)} \\ c \cdot (1-p)^{l_n \cdot (\lfloor\frac{d}{2}\rfloor + 1)} &\leq n^{(\frac{1}{c'} - \lfloor\frac{d}{2}\rfloor - 1)} \\ c \cdot \left(n(1-p)^{l_n}\right)^{(\lfloor\frac{d}{2}\rfloor + 1)} &\leq n^{\frac{1}{c'}}. \end{aligned}$$

Woraus mit Lemma 7.5 die Behauptung folgt. □

Zerlegt man also den $\mathbb{R}^d$ mittels eines Hyperquader-Gitters, erhalten wir zusammenfassend das Ergebnis:

Satz 7.1 *Seien $a_1, \ldots, a_n$ unabhängig voneinander gewählte Punkte des $\mathbb{R}^d$ und f eine Wahrscheinlichkeitsdichte auf $\mathbb{R}^d$, für die eine Hyperquader-Schalen-Schachtelung existiert, so daß die Eigenschaften* **D1** *und* **D2** *erfüllt sind. Für $0 < p < 1$ und hinreichend große Werte n berechnet Algorithmus* `A` *dann die konvexe Hülle der Punkte $a_1, \ldots, a_n$ mit erwarteten Kosten*

$$\begin{aligned} E[\texttt{konvexe Hülle}] &= E[X] + E[Y_C] \\ &\leq \left(d + \frac{d}{p} + c_{oh}\right) \cdot n + o(n). \end{aligned}$$

7.2. Ein einfacheres Experiment zur Abschätzung von $\sum_{i=0}^{n} P\left(Y_l = i|\mathcal{W}_l\right) \cdot i^k$

In Abschnitt 7.1.2 sahen wir uns gezwungen, eine Abschätzung von $\sum_{i=0}^{n} P\left(Y_l = i|\mathcal{W}_l\right) \cdot i^k$ anzugeben. Wir bedienten uns dazu einer Abschätzung durch $\sum_{i=0}^{n} P\left(Y_l = i\right) \cdot i^k$. Zur Rechtfertigung dieser Abschätzung kehren wir zum vereinfachten Experiment aus Abschnitt 3.5 zurück und modifizieren die Wahrscheinlichkeiten, so daß sie passend für das Wahrscheinlichkeitsexperiment „`Ziehung einer` n`-Auswahl aus einer vereinfachten Hyperquader-Schalen-Schachtelung`“ sind. Wie bei der vereinfachten Polygon-Ring-Schachtelung sei auch bei der vereinfachten Hyperquader-Schalen-Schachtelung eine Schale vollbesetzt, wenn ein ausgezeichneter Eckquader besetzt ist.

Im Gegensatz zur Polygon-Ring-Schachtelung ist bei der Hyperquader-Schalen-Schachtelung der Nenner der Wahrscheinlichkeit der Eckquader polynomiell in der Schachtelungstiefe: $q(l) \geq \frac{p(1-p)^l}{c \cdot l^{d-1}}$. Wir müssen daher überprüfen, ob bei der Hyperquader-Schalen-Schachtelung immer noch dieselben Beobachtungen möglich sind wie bei der Polygon-Ring-Schachtelung in Abschnitt 3.5. Generell geht es jedoch um die Gültigkeit von

Vermutung 7.1 *Gegeben sei eine Hyperquader-Schalen-Schachtelung. Für große Werte n und festes $p \in (0,1)$ gibt es eine maximale Schachtelungstiefe l_0, so daß für alle Schachtelungstiefen $l \leq l_0$ ein $i_0 \leq n$ existiert, ab dem für alle $i \geq i_0$ die Wahrscheinlichkeit $P(Y_l = i|\mathcal{W}_l) \leq P(Y_l = i)$ ist.*

Vermutung 7.2 *Sei f eine Wahrscheinlichkeitsdichte, für die die Eigenschaften* **D1** *und* **D2** *erfüllt sind. Dann existiert ein Wert $b > 1$, so daß für große Werte n, $0 < p < 1$ die Schachtelungstiefe $l_n = \lfloor \alpha \log_b n \rfloor$, $0 < \alpha < \frac{1}{|\log_b(1-p)|}$, kleiner als l_0 aus Vermutung 7.1 ist.*

Aufgrund der Plausibilitätsbetrachtung für Vermutung 3.1 können wir davon ausgehen, daß auch Vermutung 7.1 gilt. Zur Veranschaulichung werden wir das vereinfachte Wahrscheinlichkeitsexperiment auswerten. Wir betrachten dazu die Ziehung einer n-Auswahl aus einer vereinfachten Hyperquader-Schalen-Schachtelung für $d = 3$. Die einzigen Änderungen zu dem Wahrscheinlichkeitsexperiment aus Abschnitt 3.5 sind die Wahrscheinlichkeiten der Felder bzw. Quader $D^{(1)}$, $D^{(3)}$ und $D^{(5)}$. Es ist mit der Formel

$$K_l = \sum_{k=0}^{d} \binom{d}{k} (2l)^{d-k} \left(1 - (-1)^k\right)$$

für die Anzahl der Quader von $Schale^{(l)}$

$$P_e(D^{(1)}) \quad = \quad p(1-p)\frac{4}{26}$$

$$
\begin{aligned}
P_e(D^{(3)}) &= p(1-p)^2 \frac{4}{98} \\
P_e(D^{(5)}) &= (1 - 3p + 3p^2 - p^3) \frac{4}{218}.
\end{aligned}
$$

Für $p = 0.5$ und verschiedene Werte von n sind die Verläufe von $P(Y_l = i)$ und $P(Y_l = i|\mathcal{W}_l)$ für $i = 0, ..., n$ in den Abbildungen 36-44 dargestellt. Die durchgezogene Linie gibt den Verlauf von $P(Y_l = i|\mathcal{W}_l)$ und die gestrichelte Linie den von $P(Y_l = i)$ wieder.

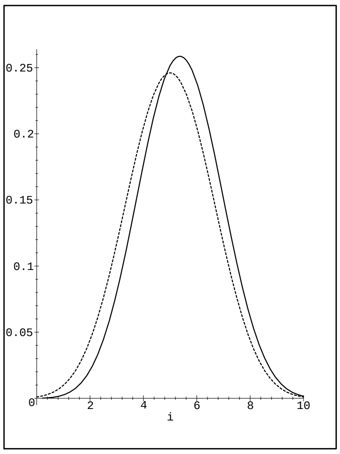

Abbildung 36:
$P(Y_1 = i)$, $P(Y_1 = i|\mathcal{W}_1)$
$p = 0.5$
$n = 10$

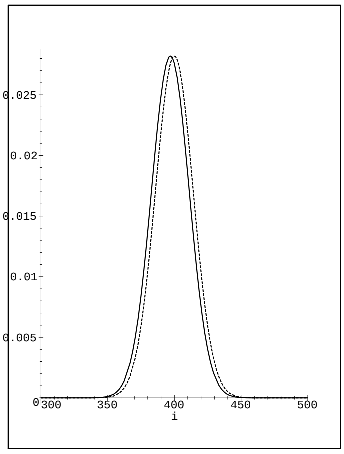

Abbildung 37:
$P(Y_1 = i)$, $P(Y_1 = i|\mathcal{W}_1)$
$p = 0.5$
$n = 800$

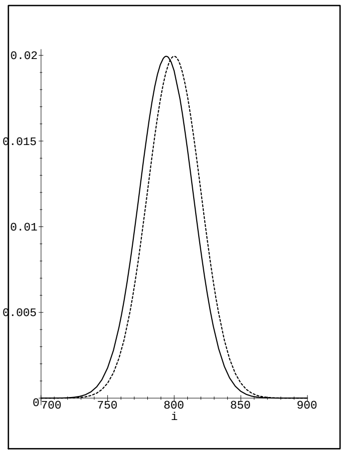

Abbildung 38:
$P(Y_1 = i)$, $P(Y_1 = i|\mathcal{W}_1)$
$p = 0.5$
$n = 1600$

Vergleicht man die Kurven mit denen aus Abschnitt 3.5, so sieht man, daß der Peak von $P(Y_l = i)$ den Peak von $P(Y_l = i|\mathcal{W}_l)$ erst bei noch größeren Werten von n „überholt". Die Frage ist daher, ob Vermutung 7.2 gilt. Dazu haben wir wie in Abschnitt 5 numerische Simulationen durchgeführt. Für $d = 3, 4, 5, 6$ wurden jeweils für wachsende Werte von n 100 n-Auswahlen simuliert. Der Einfachheit halber wurde jedem Eckquader von $Schale^{(l)}$ dieselbe Wahrscheinlichkeit $q(l) = \frac{p(1-p)^l}{K_l}$ zugewiesen, so daß Eigenschaft **D2** erfüllt war. Die Ergebnisse sind in den Tabellen 6-9 abzulesen. Wählt man für $d = 3, 4$ die Schachtelungstiefe $l_n = \log_{10} n$ und für $d = 5, 6$ Schachtelungstiefe $l_n = \log_{100} n$, so bestätigen die Simulationsergebnisse die Aussage von Vermutung 7.2. Denn im Fall $d = 3, n = 10^7$ gilt für Schachtelungstiefen $l \leq 10$ und für $d = 4, n = 10^7$ für Schachtelungstiefen $l \leq 7$ die Ungleichung $E[Y_l|\mathcal{W}_l] \leq E[Y_l]$ (der Wert von $E[Y_6|\mathcal{W}_6]$ für $n = 10^7$ ist ein Ausreißer, was durchaus vorkommen kann, da hier statistische Experimente gemacht wurden). Für $d = 5$ und $n = 10^7$ ist die

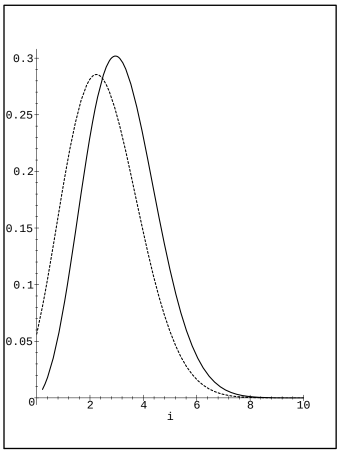

Abbildung 39:
$P(Y_2 = i)$, $P(Y_2 = i|\mathcal{W}_2)$
$p = 0.5$
$n = 10$

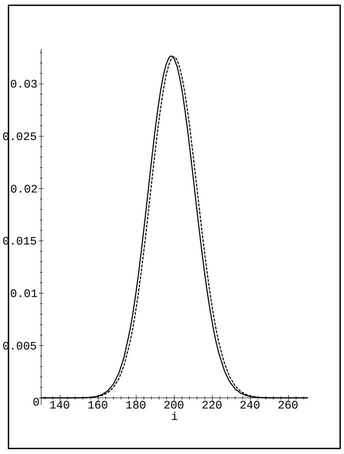

Abbildung 40:
$P(Y_2 = i)$, $P(Y_2 = i|\mathcal{W}_2)$
$p = 0.5$
$n = 800$

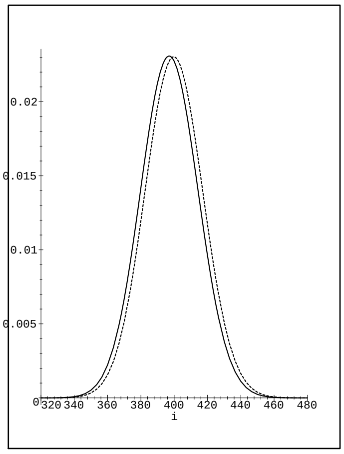

Abbildung 41:
$P(Y_2 = i)$, $P(Y_2 = i|\mathcal{W}_2)$
$p = 0.5$
$n = 1600$

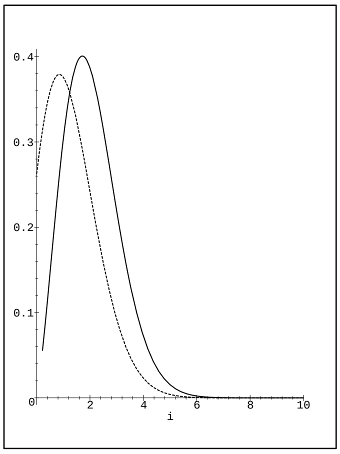

Abbildung 42:
$P(Y_3 = i)$, $P(Y_3 = i|\mathcal{W}_3)$
$p = 0.5$
$n = 10$

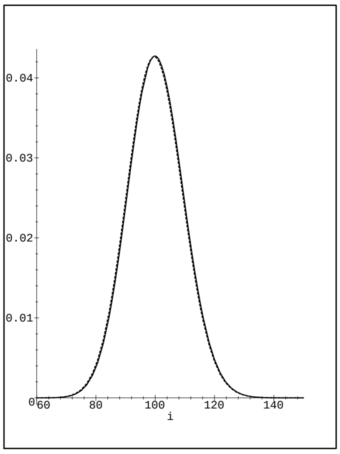

Abbildung 43:
$P(Y_3 = i)$, $P(Y_3 = i|\mathcal{W}_3)$
$p = 0.5$
$n = 800$

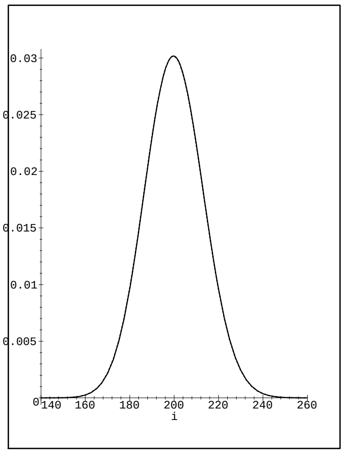

Abbildung 44:
$P(Y_3 = i)$, $P(Y_3 = i|\mathcal{W}_3)$
$p = 0.5$
$n = 1600$

Ungleichung für $l = 4$ zwar erfüllt, für $l = 3$ aber nicht mehr. Da jedoch die relative Häufigkeit bei $l = 3$ sehr gering ist, sollte dieser Wert als Ausreißer angesehen werden. Gleiches gilt für das Ergebnis von $d = 6$ und $n = 10^7$ bei Schachtelungstiefe 2. Simulationen mit größeren Werten für n würden eindeutigere Ergebnisse liefern; aus Zeitgründen konnten jedoch keine weiteren Simulationen durchgeführt werden.

Ring	$P(\mathcal{W}_l)$	$E[Y_l \mid \mathcal{W}_l]$	$E[Y_l]$
$n = 10^3$			
1	0.93	501	500
2	0.07	253	250
$n = 10^4$			
2	0.32	2502	2500
3	0.65	1252	1250
4	0.03	656	625
$n = 10^5$			
4	0.39	6261	6250
5	0.58	3115	3125
6	0.03	1581	1563
$n = 10^6$			
6	0.27	15633	15625
7	0.67	7813	7812
8	0.06	3889	3906
$n = 10^7$			
9	0.65	19528	19531
10	0.35	9745	9765

Tabelle 6: $p = 0.5$, $d = 3$

Ring	$P(\mathcal{W}_l)$	$E[Y_l \mid \mathcal{W}_l]$	$E[Y_l]$
$n = 10^3$			
0	0.51	1000	1000
1	0.49	503	500
$n = 10^4$			
1	0.81	4999	5000
2	0.19	2499	2500
$n = 10^5$			
2	0.40	24973	25000
3	0.60	12519	12500
$n = 10^6$			
3	0.03	124980	125000
4	0.89	62479	62500
5	0.08	31098	31250
$n = 10^7$			
5	0.06	312471	312500
6	0.86	156285	156250
7	0.08	78090	78125

Tabelle 7: $p = 0.5$, $d = 4$

Es gibt aber auch eine andere Möglichkeit, sich von der Gültigkeit von Vermutung 7.2 zu überzeugen. Beim Vergleich der Auswertungen der Wahrscheinlichkeitsexperimente „`Ziehung einer` n`-Auswahl aus einer vereinfachten Hyperquader-SchalenSchachtelung`“ und „`Ziehung einer` n`-Auswahl aus einer vereinfachten Polygon-Ring-Schachtelung`“ sieht man deutlich, daß die Rate, mit der $q(l)$ abnimmt, ausschlaggebend dafür ist, für welche Schachtelungstiefe $l(n)$ bei vorgegebenem Wert n die Ungleichung $E[Y_{l(n)} \mid \mathcal{W}_{l(n)}] \leq E[Y_{l(n)}]$ gilt. Damit Vermutung 7.2 nicht gilt, muß es also ein N_0 geben, so daß für $l_{N_0} = \lfloor \alpha \log N_0 \rfloor$ die Ungleichung $E[Y_{l_{N_0}} \mid \mathcal{W}_{l_{N_0}}] \leq E[Y_{l_{N_0}}]$ gilt, aber für alle $n \geq N_0$ mit $l_n = \lfloor \alpha \log n \rfloor$ die Ungleichung $E[Y_{l_n} \mid \mathcal{W}_{l_n}] \leq E[Y_{l_n}]$ nicht erfüllbar ist. Insbesondere ist dann die Ungleichung für alle $n \geq N_0$ mit $l_n = l_{N_0} + 1$ nicht erfüllt. In anderen Worten, beim Übergang von l_{N_0} nach $l_{N_0} + 1$ nimmt $q(l)$ so stark ab, daß die Ungleichung nicht mehr gilt.

Offenbar nimmt bei einer Polygon-Ring-Schachtelung mit den Eigenschaften **E1** und **E2** für große Schachtelungstiefen l die Wahrscheinlichkeit $q(l)$ beim Übergang von l

Ring	$P(\mathcal{W}_l)$	$E[Y_l\|\mathcal{W}_l]$	$E[Y_l]$
$n = 10^3$			
0	1.00	1000	1000
$n = 10^4$			
1	1.00	5002	5000
$n = 10^5$			
1	0.39	50011	50000
2	0.61	25054	25000
$n = 10^6$			
2	0.37	250098	250000
3	0.63	125034	125000
$n = 10^7$			
3	0.03	1250048	1250000
4	0.97	624977	625000

Tabelle 8: $p = 0.5$, $d = 5$

Ring	$P(\mathcal{W}_l)$	$E[Y_l\|\mathcal{W}_l]$	$E[Y_l]$
$n = 10^3$			
0	1.00	1000	1000
$n = 10^4$			
0	0.90	10000	10000
1	0.10	5016	5000
$n = 10^5$			
1	1.00	50017	50000
$n = 10^6$			
1	0.02	499603	500000
2	0.98	250018	250000
$n = 10^7$			
2	0.16	2500028	2500000
3	0.84	1249957	1250000

Tabelle 9: $p = 0.5$, $d = 6$

nach $l+1$ mit einem Verhältnis ab, das die Gültigkeit von Vermutung 4.2 bzw. für $l_n = \lfloor \alpha \log n \rfloor$ die Gültigkeit der Ungleichung $E[Y_{l_n}|\mathcal{W}_{l_n}] \leq E[Y_{l_n}]$ gewährleistet. Im ungünstigsten Fall, weil $q(l)$ dann am kleinsten ist, beträgt $q(l) = \frac{p(1-p)^l}{2kl}$. Die Rate, mit der $q(l)$ beim Übergang von Schachtelungstiefe l nach $l+1$ abnimmt, beträgt somit bei der Polygon-Ring-Schachtelung

$$\frac{p(1-p)^l}{2kl} \cdot \frac{2k(l+1)}{p(1-p)^{l+1}} = \frac{l+1}{l}\left(\frac{1}{1-p}\right) \xrightarrow{l\to\infty} \frac{1}{1-p}.$$

Der Quotient konvergiert gegen eine Konstante. Es ist daher anzunehmen, daß Vermutung 4.2 gilt, weil das Verhältnis $\frac{q(l)}{q(l+1)}$ für große Schachtelungstiefen konstant wird. Der Wert $b > 1$, der als Basis des Logarithmus verwendet wird, hängt offenbar von dieser Konstanten ab. Daher sieht man in Abschnitt 5 auch, daß für große Werte von p eine große Basis gewählt werden muß.

Vergleichen wir nun für große l dieses Verhältnis bei der Hyperquader-Schalen-Schachtelung. Es ist

$$\begin{aligned} K_l &= \sum_{k=1}^{d} \binom{d}{k} (2l)^{d-k}(1-(-1)^k) \\ &= 2^{d-1} \cdot l^{d-1} \sum_{k=1}^{d} \binom{d}{k} \frac{1}{(2l)^{k-1}}(1-(-1)^k) \\ &\xrightarrow{l\to\infty} 2d \cdot 2^{d-1} \cdot l^{d-1}. \end{aligned}$$

Das Verhältnis bei der Hyperquader-Schalen-Schachtelung beträgt folglich im ungünstigsten Falle, wenn $q(l) = \frac{p(1-p)^l}{K_l}$ ist,

$$\frac{p(1-p)^l}{K_l} \cdot \frac{K_{l+1}}{p(1-p)^{l+1}} \overset{l\to\infty}{\longrightarrow} \frac{p(1-p)^l}{2d\,2^{d-1}l^{d-1}} \cdot \frac{2d\,2^{d-1}(l+1)^{d-1}}{p(1-p)^{l+1}}$$
$$\overset{l\to\infty}{\longrightarrow} \frac{(l+1)^{d-1}}{l^{d-1}} \left(\frac{1}{1-p}\right) \overset{l\to\infty}{\longrightarrow} \frac{1}{1-p}.$$

Da auch bei der Hyperquader-Schalen-Schachtelung das Verhältnis konvergiert, wird wieder eine Basis $b > 1$ existieren, so daß für große Werte n die Wahl $l_n = \lfloor \alpha \log_b n \rfloor$ ausreicht. Daher gilt für große n und alle Schachtelungstiefen $l \leq l_n$ die Ungleichung

$$\sum_{i=0}^{n} P\left(Y_l = i | \mathcal{W}_l\right) \cdot i^k \leq \sum_{i=0}^{n} P\left(Y_l = i\right) \cdot i^k.$$

Nun müssen wir nur noch $\sum\limits_{i=0}^{n} P\left(Y_l = i\right) \cdot i^k$ abschätzen. Es ist

Lemma 7.7

$$\sum_{i=0}^{n} P\left(Y_l = i\right) \cdot i^k = O\ (\ \left(n\,(1-p)^l\right)^k\ )\ .$$

Beweis[29]

Die Wahrscheinlichkeit, daß i Punkte jenseits von $Schale^{(l)}$ liegen beträgt

$$P\left(Y_l = i\right) \ = \ \binom{n}{i} (1-p)^{l\,i} \left(1 - (1-p)^l\right)^{n-i}.$$

Zur Vereinfachung setzen wir $p = (1-p)^l$. Daraus folgt

$$\begin{aligned}
\sum_{i=0}^{n} P\left(Y_l = i\right) \cdot i^k \ &= \ \sum_{i=1}^{n} \frac{n!}{i!\,(n-i)!}\, p^i\, (1-p)^{n-i}\, i^k \\
&= \ n\,p \sum_{i=1}^{n} \frac{(n-1)!}{(i-1)!\,(n-i)!}\, p^{i-1}\, (1-p)^{n-i}\, i^{k-1}
\end{aligned}$$

ersetze $i-1$ durch j:

$$= \ n\,p \sum_{j=0}^{n-1} \frac{(n-1)!}{j!\,(n-j)!}\, p^j\, (1-p)^{n-1-j}\, (j+1)^{k-1}$$

[29] Ein Ähnliches Ergebnis findet man auch in [16].

Ersetzen wir j durch i und benutzen für $(j+1)^{k-1}$ die binomische Formel, dann erhält man

$$\begin{aligned}\sum_{i=0}^{n} P(Y_l = i) \cdot i^k &= n\,p \sum_{i=0}^{n-1} \binom{n-1}{i} p^i (1-p)^{n-1-i} \left(\sum_{l=0}^{k-1} \binom{k-1}{l} i^{k-1-l} \right) \\ &= n\,p \cdot \sum_{l=0}^{k-1} \binom{k-1}{l} \cdot \sum_{i=0}^{n-1} \binom{n-1}{i} p^i (1-p)^{n-1-i}\, i^{k-1-l}.\end{aligned}$$

Schaut man sich das Ergebnis der obigen Umformungen an, so sieht man, daß sich eine Rekursionsgleichung ergibt. Sei

$$f(n,p,k) \;=\; \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i}\, i^k,$$

dann lautet die Rekursionsgleichung

$$f(n,p,k) \;=\; n\,p \cdot \left(\sum_{l=0}^{k-1} \binom{k-1}{l} \cdot f(n-1, p, k-1-l) \right)$$

und

$$f(\eta, p, 1) \;=\; \eta\, p.$$

Wir wollen nun $f(n,p,k)$ weiter abschätzen. Dabei betrachten wir k als eine Konstante

$$\begin{aligned}f(n,p,k) &\leq n\,p \cdot \left(\sum_{l=0}^{k-1} \binom{k-1}{l} \cdot f(n-1,p,k-1) \right) \\ &\leq n\,p \cdot c \cdot f(n-1,p,k-1).\end{aligned}$$

Durch rekursives Einsetzen gelangt man schließlich zu

$$f(n,p,k) \quad \leq \quad c^* \cdot (n\,p)^k.$$

Es ist daher

$$f(n,p,k) \;=\; O\left((n\,p)^k \right).$$

Setzt man für p wieder $(1-p)^l$ ein, erhält man

$$\sum_{i=0}^{n} P(Y_l = i) \cdot i^k \;=\; O\left(\left(n\,(1-p)^l \right)^k \right).$$

$\square$

7.3. Laufzeitverhalten für verschiedene Wahrscheinlichkeiten p

In Abschnitt 3.4 haben wir untersucht, wie sich die Laufzeit im Zweidimensionalen bei verschiedenen Werten von p, der Wahrscheinlichkeit des zentralen Rechtecks, verhält. Nun wollen wir eine entsprechende Untersuchung in höheren Dimensionen durchführen. Es geht also darum, zu betrachten, welchen Einfluß p auf die erwartete Laufzeit der Schritte 1 und 2 hat, wenn wir das Verfahren im $\mathbb{R}^d$ verwenden. Da die erwartet Laufzeit von Schritt 1

$$E[X] = (d + \frac{d}{p} + c_{oh}) \cdot n$$

beträgt, verhält sie sich hyperbolisch für $0 < p < 1$. Die erwartete Laufzeit von Schritt 2, $E[Y_C]$, läßt sich laut Lemma 7.6 durch

$$E[Y_C] \leq \tilde{c} \cdot n \cdot \left(1 + \sum_{r=0}^{l_n - 1} \frac{1}{2^{(1-p)^{-r}}} \cdot \frac{1}{(1-p)^{(r+1)\cdot(\lfloor \frac{d}{2} \rfloor + 1)}}\right)$$

angeben. Es gilt das Analogon zu Lemma 3.6

Lemma 7.8

$$a_r = \left(\frac{1}{2}\right)^{\left(\frac{1}{1-p}\right)^r} \left(\frac{1}{1-p}\right)^{r\left(\lfloor \frac{d}{2} \rfloor + 1\right)}$$

ist eine Nullfolge.

Der Beweis verläuft genauso wie der von Lemma 3.6.

Bildet man die erste Ableitung $a_r{}'$, dann sieht man, daß a_r für $r \leq \left\lfloor \frac{\ln\ln 2 - \ln\left(\lfloor \frac{d}{2} \rfloor + 1\right)}{\ln(1-p)} \right\rfloor$ monoton steigt und für größere r monoton fällt. Somit ist

$$\frac{1}{2}\left(\frac{1}{1-p}\right)^{\left(\frac{\ln\ln 2 - \ln\left(\lfloor \frac{d}{2} \rfloor + 1\right)}{\ln(1-p)} + 1\right)\cdot\left(\lfloor \frac{d}{2} \rfloor + 1\right)} + \int_{r=0}^{\infty} \left(\frac{1}{2}\right)^{\left(\frac{1}{1-p}\right)^r} \left(\frac{1}{1-p}\right)^{(r+1)\left(\lfloor \frac{d}{2} \rfloor + 1\right)} dr$$

eine obere Schranke für

$$\sum_{r=0}^{l_n - 1} \left(\frac{1}{2}\right)^{\left(\frac{1}{1-p}\right)^r} \left(\frac{1}{1-p}\right)^{(r+1)\left(\lfloor \frac{d}{2} \rfloor + 1\right)}.$$

Dafür gibt es eine kompaktere Schreibweise

Lemma 7.9

$$\begin{aligned}
&\sum_{r=0}^{l_n-1}\left(\frac{1}{2}\right)^{\left(\frac{1}{1-p}\right)^r}\left(\frac{1}{1-p}\right)^{(r+1)\left(\lfloor\frac{d}{2}\rfloor+1\right)} \\
\leq\ & \frac{1}{2}\left(\frac{1}{1-p}\right)^{\left(\frac{\ln\ln 2-\ln\left(\lfloor\frac{d}{2}\rfloor+1\right)}{\ln(1-p)}+1\right)\cdot\left(\lfloor\frac{d}{2}\rfloor+1\right)} \\
& -\left(\frac{1}{1-p}\right)^{\left(\lfloor\frac{d}{2}\rfloor+1\right)}\frac{1}{2\ln(1-p)}\cdot\sum_{i=0}^{\lfloor\frac{d}{2}\rfloor}\frac{\left(\lfloor\frac{d}{2}\rfloor\right)_{\left(\lfloor\frac{d}{2}\rfloor-i\right)}}{(\ln 2)^{i+1}} \\
=\ & S(p,d).
\end{aligned}$$

Beweis

Die Substitution $y=\left(\frac{1}{1-p}\right)^r$ liefert für das Integral

$$\int\left(\frac{1}{2}\right)^{\left(\frac{1}{1-p}\right)^r}\left(\frac{1}{1-p}\right)^{r\cdot\left(\lfloor\frac{d}{2}\rfloor+1\right)}\,dr$$

den Ausdruck

$$-\frac{1}{\ln(1-p)}\cdot\int\left(\frac{1}{2}\right)^y\cdot y^{\lfloor\frac{d}{2}\rfloor}\,dy.$$

Dieser Ausdruck läßt sich partiell integrieren

$$\int\left(\frac{1}{2}\right)^y\cdot y^{\lfloor\frac{d}{2}\rfloor}\,dy \quad=\quad -\frac{1}{\ln 2}\left(\frac{1}{2}\right)^y y^{\lfloor\frac{d}{2}\rfloor}+\frac{\lfloor\frac{d}{2}\rfloor}{\ln 2}\int\left(\frac{1}{2}\right)^y y^{\left(\lfloor\frac{d}{2}\rfloor-1\right)}\,dy.$$

Mit vollständiger Induktion zeigt man leicht, daß

$$\int\left(\frac{1}{2}\right)^y\cdot y^{\lfloor\frac{d}{2}\rfloor}\,dy \quad=\quad -\sum_{i=0}^{\lfloor\frac{d}{2}\rfloor}\frac{\left(\lfloor\frac{d}{2}\rfloor\right)_{\left(\lfloor\frac{d}{2}\rfloor-i\right)}}{(\ln 2)^{i+1}}\cdot\left(\frac{1}{2}\right)^y y^{\left(\lfloor\frac{d}{2}\rfloor-i\right)},$$

wobei für $m\leq n$

$$\begin{aligned}
(n)_m &= \frac{n!}{m!} \\
\text{und}\quad (n)_0 &= n!
\end{aligned}$$

ist. Durch Rücksubstitution und Einsetzen der Grenzen erhalten wir

$$\begin{aligned}
&\left(\frac{1}{1-p}\right)^{\left(\lfloor\frac{d}{2}\rfloor+1\right)}\int_{r=0}^{\infty}\left(\frac{1}{2}\right)^{\left(\frac{1}{1-p}\right)^r}\left(\frac{1}{1-p}\right)^{r\cdot\left(\lfloor\frac{d}{2}\rfloor+1\right)}\,dr \\
=\ & -\left(\frac{1}{1-p}\right)^{\left(\lfloor\frac{d}{2}\rfloor+1\right)}\frac{1}{2\ln(1-p)}\cdot\sum_{i=0}^{\lfloor\frac{d}{2}\rfloor}\frac{\left(\lfloor\frac{d}{2}\rfloor\right)_{\left(\lfloor\frac{d}{2}\rfloor-i\right)}}{(\ln 2)^{i+1}}.
\end{aligned}$$

□

Für $d = 3$, 10, 20 und $0 < p < 1$ ist der Graph von $S(p,d)$ in den Abbildungen 45 bis 47 dargestellt[30].

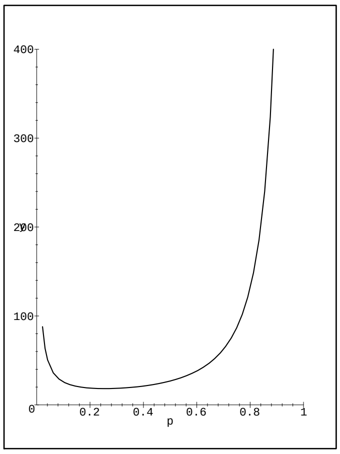

Abbildung 45:
$d = 3$

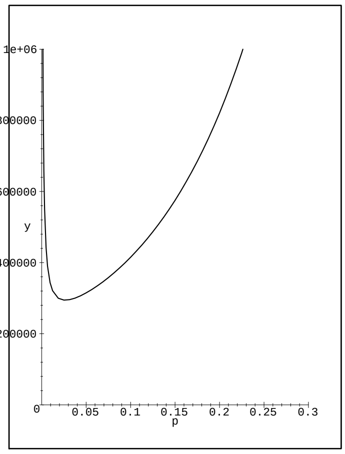

Abbildung 46:
$d = 10$

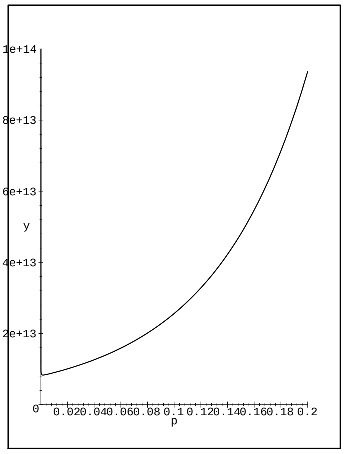

Abbildung 47:
$d = 20$

Wie man sieht, verlagert sich das Minimum der Kurve für größere d mehr zur 0 hin. D.h. in höheren Dimensionen verbessert sich die erwartete Laufzeit von Schritt 3, wenn wir p klein wählen.

Satz 7.2 *Asymptotisch, für $n \to \infty$, bestimmt wegen Lemma 7.6 die Laufzeit von Schritt 2 die erwartete Laufzeit von Algorithmus* `A`.

[30] In Abbildung 47 überlagert sich der steile Anstieg von $S(p,d)$ für kleine Werte p mit der Ordinate.

8. Erwartete Laufzeit bei *zulässigen* Riemann-Dichten

In den bisherigen Abschnitten wurde gezeigt, daß man für bestimmte Wahrscheinlichkeitsdichten im Mittel die konvexe Hülle von n Punkten asymptotisch in Linearzeit berechnen kann. Die Voraussetzung ist, daß die Wahrscheinlichkeitsdichte so beschaffen ist, daß man durch eine Polygon-Ring-Schachtelung auf dem $\mathbb{R}^2$, bzw. eine Hyperquader-Schalen-Schachtelung auf dem $\mathbb{R}^d$ erzeugt, so daß Eigenschaften **D1** und **D2** erfüllt sind.

Definition 8.1 (gerecht) *Wahrscheinlichkeitsdichten, die* **D1** *und* **D2** *erfüllen, wollen wir künftig als* gerecht *bezeichnen.*

In den folgenden Abschnitten werden wir erörtern, ob man entscheiden kann, bei welchen Dichten die beiden Bedingungen erfüllbar sind, ob man eventuell eine der beiden Bedingungen **D1** oder **D2** abschwächen kann und ob man Aussagen über die erwarteten Kosten des Verfahrens bei allgemeineren Wahrscheinlichkeitsdichten machen kann. Da die meisten reell-wertigen Wahrscheinlichkeitsexperimente mit Riemann-Dichten modelliert werden, werden wir in der nun folgenden Betrachtung nur auf Riemann-Dichten oder kurz *Dichten* eingehen.

In Abschnitt 4.2 haben wir gesehen, daß man im $\mathbb{R}^2$ bei bestimmten rotationssymmetrischen Stufenverteilungen erwartete Linearzeit erreichen kann. Benutzt man eine Hyperquader-Schalen-Schachtelung, so ist auch im $\mathbb{R}^d$ bei solchen rotationssymmetrischen Stufenverteilungen lineare erwartete Laufzeit möglich[31].

Betrachten wir nun allgemeine Wahrscheinlichkeitsdichten auf dem $\mathbb{R}^d$. Da Integrale stetig sind, ist Eigenschaft **D1** als Folge des Zwischenwertsatzes bei jeder Wahrscheinlichkeitsdichte erfüllbar. Ob Eigenschaft **D2** im allgemeinen erfüllbar ist, läßt sich offensichtlich schwer entscheiden. Wofür benötigten wir **D2**? War **D2** erfüllt, konnten wir erstens mit der Wahl $l_n = \lfloor \alpha \log n \rfloor$ sicherstellen, daß für große n die Wahrscheinlichkeit $P(R^{(l_n)}) \geq \frac{1}{2}$ ist. Damit konnten wir für $l_n \to \infty$ die Konvergenz der Reihe

$$\sum_{r=0}^{l_n-1} \frac{P(\overline{R^{(l_n-r)}})}{(1-p)^{(r+1)\cdot(\frac{d}{2}+1)}}$$

zeigen. Ebenfalls war dann für alle Schachtelungstiefen $l \leq l_n$ die Ungleichung

$$\sum_{i=0}^{n} P(Y_l = i | \mathcal{W}_l) i^k \leq \sum_{i=0}^{n} P(Y_l = i) i^k$$

erfüllt, wodurch die Kosten $E[Y_C]$ linear wurden. Es stellt sich also die Frage, was passiert, wenn die zweite Bedingung – $q_i(l) \geq \frac{p(1-p)^l}{K_l}$ für alle Eckquader $i = 1, \ldots, 2^d$ von $Schale^{(l)}$ – aufgeweicht wird.

[31] Im $\mathbb{R}^d$ gibt es $d-1$ Rotationsachsen, um die die Verteilung symmetrisch sein muß.

Sei $m(l) = \min_{1 \leq i \leq 2^d} q_i(l)$ die kleinste Wahrscheinlichkeit aller Eckquader von $Schale^{(l)}$. Dann gilt

$$\begin{aligned} P(R^{(l)}) &\geq 1 - \sum_{i=1}^{2^d} (1 - q_i(l))^n \\ &\geq 1 - 2^d \, (1 - m(l))^n. \end{aligned}$$

Woraus folgt, daß

$$P(\overline{R^{(l)}}) \leq 2^d \, (1 - m(l))^n$$

ist. Lösen wir uns also von der zweiten Bedingung, dann sollten wir uns fragen, wie sich die $m(l)$ verhalten müssen, damit

$$\sum_{r=0}^{l'-1} \frac{P(\overline{R^{(l'-r)}})}{(1-p)^{(r+1)\cdot(\frac{d}{2}+1)}}$$

, wobei l' beliebig sei, immer noch konvergiert. Da wir das Konvergenzverhalten für $l' \to \infty$ ermitteln wollen, untersuchen wir die Konvergenz einer Reihe. Benutzen wir das Quotientenkriterium, dann konvergiert die Reihe falls

$$\begin{aligned} R &:= \overline{\lim_{r \to \infty}} \left| \frac{a_{r+1}}{a_r} \right| \\ &= \lim_{r \to \infty} \frac{2^d (1 - m(l' - (r+1)))^n}{(1-p)^{(r+2)\cdot(\frac{d}{2}+1)}} \cdot \frac{(1-p)^{(r+1)\cdot(\frac{d}{2}+1)}}{2^d \, (1 - m(l' - r))^n} < 1 \\ \Leftrightarrow \quad & \frac{1}{1-p} \cdot \lim_{r \to \infty} \left(\frac{1 - m(l' - r - 1)}{1 - m(l' - r)} \right)^n < 1. \end{aligned}$$

Dies ist gewährleistet, wenn für alle r

$$\frac{1 - m(l' - r - 1)}{1 - m(l' - r)} < \sqrt[n]{1-p}$$

gilt.

Das veranlaßt uns zu einer neuen Definition

Definition 8.2 (zulässig) *Eine Wahrscheinlichkeitsdichte heißt* zulässig, *falls es eine Hyperquader-Schalen-Schachtelung mit folgenden beiden Eigenschaften gibt:*

Z1 $P(Schale^{(l)}) = p(1-p)^l$

Z2 $\frac{1-m(l-1)}{1-m(l)} < \sqrt[n]{1-p}$

für alle Schachtelungstiefen l.

Bemerkung 8.1 *Wegen des Zwischenwertsatzes ist die erste Bedingung bei jeder Riemann-Dichte erfüllt.*

Satz 8.1 *Seien $a_1, \ldots, a_n$ unabhängig voneinander gewählte Punkte des $\mathbb{R}^d$. Ist $f : \mathbb{R}^d \to \mathbb{R}$ eine zulässige Wahrscheinlichkeitsdichte, so betragen die erwarteten Kosten von Algorithmus* `A` *zur Berechnung der konvexen Hülle von $a_1, \ldots, a_n$*

$$\textit{für } d \leq 3\text{: } (d + \frac{d}{p} + c_{oh})n + o(n \log n)$$

und

$$\textit{für } d > 3\text{: } (d + \frac{d}{p} + c_{oh})n + o(n^{\lfloor \frac{d}{2} \rfloor}).$$

Beweis

Da die Wahrscheinlichkeitsdichte **Z1** erfüllt, bleiben die Kosten von Schritt 1 gleich. Bestimmen wir daher die erwarteten Kosten von Schritt 2. Im Gegensatz zur bisherigen Betrachtung werden wir statt einem Standardverfahren nun ein Verfahren von Chazelle [6] mit worst-case Laufzeit $O(n \log n + n^{\lfloor \frac{d}{2} \rfloor})$ benutzen.

$d \leq 3$:

Es sei l_0 so gewählt wie in Vermutung 7.1. Laut Voraussetzung ist die Wahrscheinlichkeitsdichte zulässig. Daher gilt für die erwarteten Kosten[32]

$$E[Y_C] \leq c \cdot n \log n \cdot (1-p)^{l_0}.$$

Da Eigenschaft **D2** nicht erfüllt ist, können wir im allgemeinen nicht davon ausgehen, daß Vermutung 7.2 gilt. In Bemerkung 3.4 haben wir bereits darauf hingewiesen, daß l_0 monoton steigend in n ist. Folglich erhält man

$$\frac{E[Y_C]}{n \log n} \stackrel{n\to\infty}{\longrightarrow} c \cdot (1-p)^{l_0} \stackrel{n\to\infty}{\longrightarrow} 0.$$

$d > 3$:

Wir wählen wieder l_0 so wie in Vermutung 7.1. Aufgrund der Zulässigkeit der Wahrscheinlichkeitsdichte ist[33]

$$E[Y_C] \leq c \cdot \left(n(1-p)^{l_0}\right)^{\left(\frac{d}{2}\right)}.$$

Analog zum Fall $d \leq 3$ gilt

$$\frac{E[Y_C]}{n^{\left(\frac{d}{2}\right)}} \stackrel{n\to\infty}{\longrightarrow} c \cdot (1-p)^{l_0} \stackrel{n\to\infty}{\longrightarrow} 0.$$

■

[32] vgl. Beweis von Lemma 3.4.
[33] vgl. Beweis von Lemma 7.5.

8.1. *gutartige* Riemann-Dichten

Es soll nun gezeigt werden, daß es sehr viele zulässige Wahrscheinlichkeitsdichten gibt. Es ist beispielsweise jede *gutartige* Wahrscheinlichkeitsdichte zulässig.

Definition 8.3 (gutartig) *Eine Wahrscheinlichkeitsdichte $f : \mathbb{R}^d \to \mathbb{R}$ heißt* gutartig *genau dann, wenn*

1. *ein achsenorientierter Hyperquader $H \subseteq \mathbb{R}^d$ existiert, wobei auch $H = \mathbb{R}^d$ zugelassen sei, so daß*

$$\int_{\mathbb{R}^d} f(x)\, dx \;=\; \int_H f(x)\, dx \;=\; 1$$

   und

2. *für jeden achsenorientierten Hyperquader $H' \subseteq H$*

$$\int_{H'} f(x)\, dx \;>\; 0$$

   gilt.

Bevor wir zeigen, daß jede gutartige Riemann-Dichte $f : \mathbb{R}^d \to \mathbb{R}$ zulässig ist, führen wir ein paar wichtige Definitionen und Notationen ein.

Definition 8.4 ($x \leq y$) *Seien $x, y \in \mathbb{R}^d$, dann sei $x \leq y$ genau dann, wenn x in jeder Komponente von y dominiert wird: $x_i \leq y_i$, $\forall\, i$, $1 \leq i \leq d$.*

Definition 8.5 (definierende Punkte) *Sei $H_i \subseteq \mathbb{R}^d$ ein d-dimensionaler, achsenorientierter Hyperquader. Es seien $h_i^a = \left(h_{i,1}^a, \ldots, h_{i,d}^a\right)^T < h_i^b = \left(h_{i,1}^b, \ldots, h_{i,d}^b\right)^T$ die beiden zueinander raumdiagonal liegenden Eckpunkte, die H_i eindeutig bestimmen. Die Punkte h_i^a und h_i^b heißen dann* definierende Punkte *von H_i.*

Definition 8.6 (Punktquadrant) *Sei $v \in \mathbb{R}^d$. Verlegt man das kartesische Koordinatensystem so, daß v nun das Zentrum bildet, dann gehen 2^d Quadranten von v aus. Diese heißen* Punktquadranten *von v.*
Ist in einem Punktquadranten von v ein Hyper-Quader H ganz enthalten, dann heiße dieser Punktquadrant Q_v^{+H}. Der Q_v^{+H} raumdiagonal gegenüberliegende Quadrant heiße Q_v^{-H}.

Wir beweisen nun den

Satz 8.2 *Ist f eine gutartige Riemann-Dichte und $p < \frac{1}{2^d}$, dann ist sie auch zulässig.*

Beweis

Wir führen einen Induktionsbeweis über den schrittweisen Aufbau der Gitterkonstruktion. Zur besseren Verständlichkeit wird der Beweis mit Abbildungen der Konstruktion im $\mathbb{R}^2$ illustriert.

Induktionsverankerung:

Gegeben eine gutartige Riemann-integrierbare Funktion $f : \mathbb{R}^d \to \mathbb{R}$. Es sei

$$\int_{\mathbb{R}^d} f(x)\,dx = \int_H f(x)\,dx = 1$$

und $h^a < h^b$ seien die definierenden Punkte von H. Mit dem Zwischenwertsatz existiert dann ein Hyperquader $H_0 \subset H$ (im $\mathbb{R}^2$ ein Rechteck) definiert durch die beiden Punkte $h_0^a > h^a$ und $h_0^b < h^b$ (siehe Abbildung 48), für den gilt

$$\int_{h_0^a}^{h_0^b} f(x)\,dx = p.$$

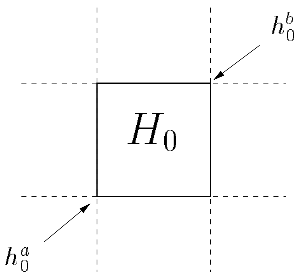

Abbildung 48:

Als nächstes legen wir um H_0 einen weiteren Hyperquader H_1. Da

$$\int_H f(x)\,dx - \int_{H_0} f(x)\,dx \;=\; 1 - p$$

ist, muß es laut Zwischenwertsatz einen Hyperquader H_1 mit den definierenden Punkten h_1^a, $h^a < h_1^a < h_0^a$, und h_1^b, $h^b > h_1^b > h_0^b$, geben, so daß

$$\int_{H_1} f(x)\,dx - \int_{H_0} f(x)\,dx \;=\; p(1 - p)$$

ist (wie in Abbildung 49 dargestellt). Auf diese Weise haben wir nun $Schale^{(0)}$ und

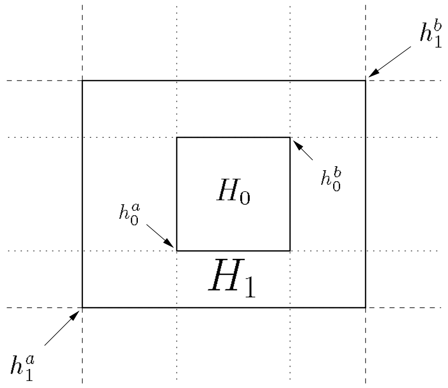

Abbildung 49:

$Schale^{(1)}$ des Gitters konstruiert.

***Induktionsschluß* $(l \to l+1)$:**

Es seien die Schalen $Schale^{(0)}$ bis $Schale^{(l)}$ bereits konstruiert. Wir versuchen nun, $Schale^{(l+1)}$ so zu konstruieren, daß die beiden Bedingungen

Z1 $P\left(Schale^{(l+1)}\right) \;=\; p(1-p)^{l+1}$ und

Z2 $\frac{1-m(l)}{1-m(l+1)} \;<\; \sqrt[n]{1-p}$

erfüllt sind. Dazu suchen wir einen Punkt $\tilde{h}^a_{l+1}$, $h^a < \tilde{h}^a_{l+1} < h^a_l$, so daß die Bedingungen

$\left[\tilde{h}^a_{l+1}.1\right]$:

$$\int_{Q^{+H_l}_{\tilde{h}^a_{l+1}}} f(x)\,dx - \int_{H_l} f(x)\,dx \;\geq\; p(1-p)^{l+1}$$

und

$\left[\tilde{h}^a_{l+1}.2\right]$:

$$\frac{1-m(l)}{1-\int_{\tilde{h}^a_{l+1}}^{h^a_l} f(x)\,dx} \;<\; \sqrt[n]{1-p}$$

erfüllt sind (vgl. Abbildung 50). Existiert so ein Punkt $\tilde{h}^a_{l+1}$, dann garantiert – weil $\left[\tilde{h}^a_{l+1}.1\right]$ erfüllt ist – der Zwischenwertsatz die Existenz eines Hyperquaders H_{l+1}, $H_l \subset H_{l+1} \subseteq Q^{+H_l}_{\tilde{h}^a_{l+1}}$, mit definierenden Punkten h^a_{l+1}, $h^a_{l+1} \geq \tilde{h}^a_{l+1}$, und h^b_{l+1} mit der Eigenschaft

$$P(Schale^{(l+1)}) \;=\; \int_{H_{l+1}} f(x)\,dx - \int_{H_l} f(x)\,dx \;=\; p(1-p)^{l+1}.$$

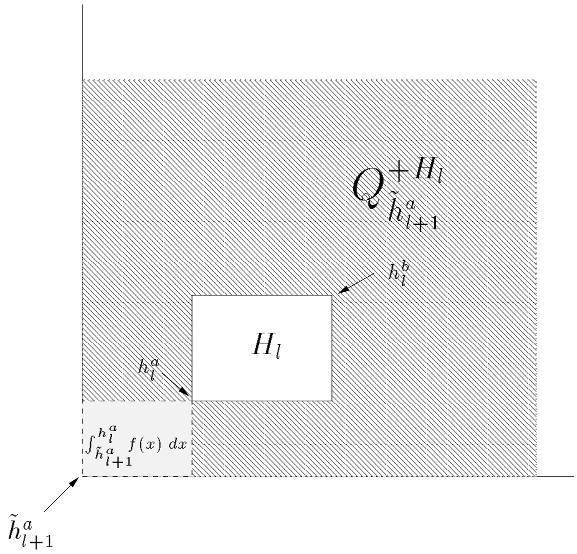

Abbildung 50:

Wegen $m(l+1) \leq \int_{\tilde{h}^a_{l+1}}^{h^a_l} f(x)\,dx$ gilt auch **Z2** (vgl. Abbildung 51). Einen Punkt, der die Bedingung $\left[\tilde{h}^a_{l+1}.1\right]$ erfüllt, gibt es immer; z.B. erfüllt jeder Eckpunkt von Hyperquader H diese Bedingung. Falls es also keinen Punkt $\tilde{h}^a_{l+1}$ gibt, der $\left[\tilde{h}^a_{l+1}.1\right]$ und $\left[\tilde{h}^a_{l+1}.2\right]$ gleichzeitig erfüllt, dann kann dies nur den Grund haben, daß jeder Punkt, der $\left[\tilde{h}^a_{l+1}.1\right]$ erfüllt, die Bedingung $\left[\tilde{h}^a_{l+1}.2\right]$ nicht erfüllt.

Bemerkung 8.2 *Trifft dies zu, dann gilt für alle Punkte v, die im durch $\tilde{h}^a_{l+1}$ und h^a_l definierten Hyperquader liegen, daß $\int_{Q_v^{+H_l}} f(x)\,dx - \int_{H_l} f(x)\,dx \;<\; p(1-p)^{l+1}$ ist.*

In diesem Falle wählen wir $\tilde{h}^a_{l+1} = h^a_l$, so daß $\left[\tilde{h}^a_{l+1}.1\right]$ nicht erfüllt ist, d.h.

$$\int_{Q^{+H_l}_{\tilde{h}^a_{l+1}}} f(x)\,dx - \int_{H_l} f(x)\,dx \;<\; p(1-p)^{l+1}$$

ist. Anschließend suchen wir einen Punkt $\tilde{h}^b_{l+1}$, $h^b_l < \tilde{h}^b_{l+1} < h^b$, der

$\left[\tilde{h}^b_{l+1}.1\right]$:

$$\int_{Q^{+H_l}_{\tilde{h}^b_{l+1}}} f(x)\,dx - \int_{H_l} f(x)\,dx \;\geq\; p(1-p)^{l+1}$$

und

$\left[\tilde{h}^b_{l+1}.2\right]$:

$$\frac{1 - m(l)}{1 - \int_{h^b_l}^{\tilde{h}^b_{l+1}} f(x)\, dx} \;<\; \sqrt[n]{1-p}$$

gewährleistet (siehe Abbildung 52).

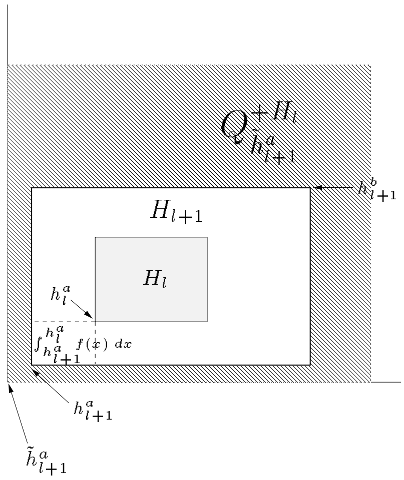

Abbildung 51:

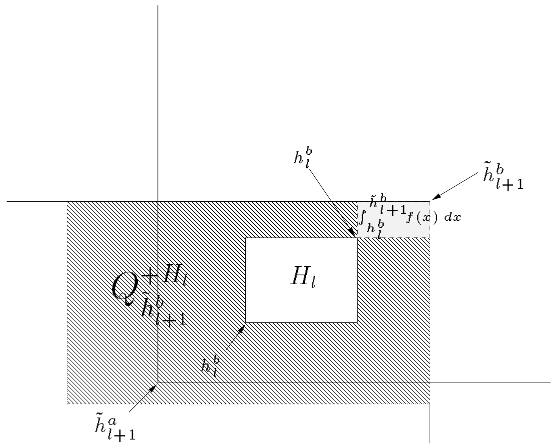

Abbildung 52:

Existiert ein solcher Punkt $\tilde{h}^b_{l+1}$, dann sind wir – genauso wie bei Punkt $\tilde{h}^a_{l+1}$ – in der Lage, einen Hyperquader H_{l+1} mit den definierenden Punkten h^a_{l+1}, $h^a < h^a_{l+1} < h^a_l$, und h^b_{l+1}, $h^b_l < h^b_{l+1} \leq \tilde{h}^b_{l+1}$, zu konstruieren, der die Bedingungen

1. $P\left(Schale^{(l+1)}\right) \;=\; p(1-p)^{l+1}$ und
2. $\frac{1-m(l)}{1-m(l+1)} \;<\; \sqrt[n]{1-p}$

für $Schale^{(l+1)}$ erfüllt. Gibt es einen solchen Punkt $\tilde{h}^b_{l+1}$ nicht, dann wähle auch $\tilde{h}^b_{l+1} = h^b_l$ so, daß $\left[\tilde{h}^b_{l+1}.1\right]$ nicht erfüllt ist. Es ist daher

$$\int_{Q^{+H_l}_{\tilde{h}^b_{l+1}}} f(x)\, dx - \int_{H_l} f(x)\, dx \;<\; p(1-p)^{l+1}.$$

Die Frage ist nun, ob wir es überhaupt noch schaffen können, einen Hyperquader H_{l+1} zu finden, mit dem sich die Gitterkonstruktion fortführen läßt. Die Punkte $\tilde{h}^a_{l+1}$ und $\tilde{h}^b_{l+1}$ definieren einen Hyperquader $\tilde{H}_{l+1}$, der neben $\tilde{h}^a_{l+1}$ und $\tilde{h}^b_{l+1}$ die $2^d - 2$ weiteren Eckpunkte $\tilde{v}^{l+1}_1, \ldots, \tilde{v}^{l+1}_{2^d-2}$ (im $\mathbb{R}^2$ $\tilde{v}^{l+1}_1$ und $\tilde{v}^{l+1}_2$) besitzt. Betrachten wir zunächst den Bereich des $\mathbb{R}^d$, der nicht in den Punktquadranten $Q^{+H_l}_{\tilde{h}^a_{l+1}}$ und $Q^{+H_l}_{\tilde{h}^b_{l+1}}$ liegt. Es ist

$$\mathbb{R}^d - Q^{+H_l}_{\tilde{h}^a_{l+1}} - Q^{+H_l}_{\tilde{h}^b_{l+1}} \;=\; \sum_{i=1}^{2^d-2} Q^{-H_l}_{\tilde{v}^{l+1}_i}$$

(zur Verdeutlichung siehe Abbildung 53).

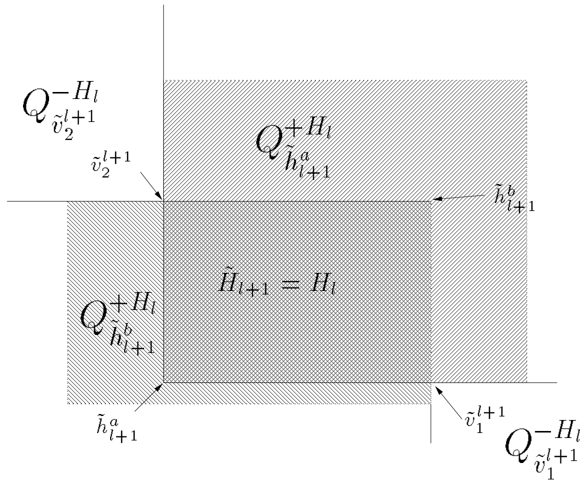

Abbildung 53:

Wenn es einen Eckpunkt von $\tilde{H}_{l+1}$ mit

$$\int_{Q^{-H_l}_{\tilde{v}^{l+1}_i}} f(x)\,dx \;>\; p(1-p)^{l+1}$$

gibt, dann liegt in dem Gebiet $Q^{-H_l}_{\tilde{v}^{l+1}_i}$ ein Punkt v, so daß

$$\int_{Q^{+H_l}_v} f(x)\,dx > p(1-p)^{l+1} + \int_{H_l} f(x)\,dx$$

ist. Daher kann man immer einen Hyperquader H_{l+1} finden, mit dem sich die Gitterkonstruktion fortführen läßt. $\tilde{H}_{l+1}$ besitzt auf jedenfall einen solchen Eckpunkt, wenn die Ungleichung

$$\sum_{i=1}^{2^d-2} \int_{Q^{-H_l}_{\tilde{v}^{l+1}_i}} f(x)\,dx \;>\; (2^d-2)\,p(1-p)^{l+1}$$

gilt.

Lemma 8.1 *Für $p < \frac{1}{2^d}$ ist die Ungleichung*

$$\sum_{i=1}^{2^d-2} \int_{Q^{-H_l}_{\tilde{v}_i^{l+1}}} f(x)\,dx \;>\; (2^d-2)\,p(1-p)^{l+1}$$

erfüllt.

Beweis

Es ist aufgrund der Wahl von $\tilde{h}^a_{l+1}$ und $\tilde{h}^b_{l+1}$

$$\begin{aligned}\sum_{i=1}^{2^d-2} \int_{Q^{-H_l}_{\tilde{v}_i^{l+1}}} f(x)\,dx \;&=\; 1 - \left(\int_{Q^{+H_l}_{\tilde{h}^a_{l+1}}} f(x)\,dx - \int_{\tilde{H}_{l+1}} f(x)\,dx\right)\\ &\quad - \left(\int_{Q^{+H_l}_{\tilde{h}^b_{l+1}}} f(x)\,dx - \int_{\tilde{H}_{l+1}} f(x)\,dx\right)\\ &\quad - \int_{\tilde{H}_{l+1}} f(x)\,dx.\end{aligned}$$

Da weder $\left[\tilde{h}^a_{l+1}.1\right]$ noch $\left[\tilde{h}^b_{l+1}.1\right]$ erfüllt sind, muß

$$\begin{aligned}\sum_{i=1}^{2^d-2} \int_{Q^{-H_l}_{\tilde{v}_i^{l+1}}} f(x)\,dx \;&>\; 1 - 2\,p(1-p)^{l+1} - \int_{\tilde{H}_{l+1}} f(x)\,dx\\ &=\; 1 - 2\,p(1-p)^{l+1} - (1-(1-p)^{l+1})\\ &=\; (1-p)^{l+1} - 2\,p(1-p)^{l+1}\end{aligned}$$

sein. Somit gilt die Behauptung, wenn

$$\begin{aligned}(1-p)^{l+1} - 2\,p(1-p)^{l+1} \;&>\; (2^d-2)\,p(1-p)^{l+1}\\ \Leftrightarrow \quad p \;&<\; \frac{1}{2^d}\end{aligned}$$

ist. $\square$

Sei o.B.d.A. für Punktquadrant $Q^{-H_l}_{\tilde{v}_\alpha^{l+1}}$

$$\int_{Q^{-H_l}_{\tilde{v}_\alpha^{l+1}}} f(x)\,dx \;>\; p(1-p)^{l+1}.$$

$\tilde{v}_\beta^{l+1}$ sei der Eckpunkt von $\tilde{H}_{l+1}$, der $\tilde{v}_\alpha^{l+1}$ raumdiagonal gegenüberliegt. Da $Q^{-H_l}_{\tilde{v}_\alpha^{l+1}} \subset Q^{+H_l}_{\tilde{v}_\beta^{l+1}}$ ist, muß

$$\int_{Q^{+H_l}_{\tilde{v}_\beta^{l+1}}} f(x)\,dx - \int_{H_l} f(x)\,dx \;>\; p(1-p)^{l+1}$$

sein. Wegen des Zwischenwertsatzes kann man daher einen Hyperquader H_{l+1} mit $H_l \subset H_{l+1} \subset Q^{+H_l}_{\tilde{v}^{l+1}_\beta}$ wählen, so daß

$$\int_{H_{l+1}} f(x)\,dx - \int_{H_l} f(x)\,dx \;=\; p(1-p)^{l+1}$$

beträgt und durch geschickte Plazierung von H_{l+1} erreicht man auch, daß $\frac{1-m(l)}{1-m(l+1)} < \sqrt[n]{1-p}$ ist (vgl. Abbildung 54).

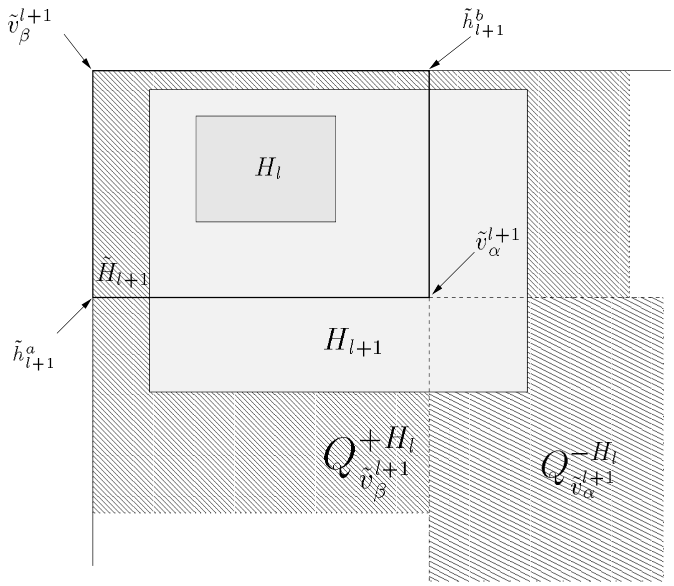

Abbildung 54:

Die Gutartigkeit von $f : \mathbb{R}^d \to \mathbb{R}$ und $p < \frac{1}{2^d}$ sind deshalb ein hinreichendes Kriterium für die Existenz einer Hyperquader-Schalen-Konstruktion, die die Bedingungen

Z1 $P\left(Schale^{(l)}\right) \;=\; p(1-p)^l$ und

Z2 $\frac{1-m(l-1)}{1-m(l)} \;<\; \sqrt[n]{1-p}$

erfüllt. ∎

8.2. Zulässigkeit *nicht-gutartiger* Riemann-Dichten

Wir haben gesehen, daß gutartige Riemann-Dichten immer zulässig sind. Die Frage ist nun, ob es *nicht-gutartige* Dichten gibt, die trotzdem zulässig sind. Diese Frage läßt sich nicht so einfach mit ja oder nein beantworten. Um diese Frage für bestimmte Klassen von nicht-gutartigen Dichten beantworten zu können, führen wir den Begriff des *Signifikanzbereichs* einer Dichte ein.

Definition 8.7 (Signifikanzbereich) *Sei $f : \mathbb{R}^d \to \mathbb{R}$ eine Riemann-Dichte. Der* Signifikanz-Bereich *S der Dichte sei die Menge aller $x \in \mathbb{R}^d$ mit $f(x) > 0$.*

Bei gutartigen Wahrscheinlichkeitsdichten ist der Signifikanzbereich ein zusammenhängendes, hyperquaderförmiges Gebiet. Bei nicht-gutartigen Dichten ist der Signifikanzbereich entweder nicht hyperquaderförmig oder nicht zusammenhängend oder mehrfach-zusammenhängend, d.h. er enthält Löcher. Je nachdem, wie der Signifikanzbereich aussieht, kann eine nicht-gutartige Wahrscheinlichkeitsdichte zulässig sein. Betrachten wir z.B. eine nicht-gutartige Wahrscheinlichkeitsdichte f_L mit mehrfach-zusammenhängendem, unbeschränktem Signifikanzbereich, d.h.

$$S = \mathbb{R}^d \setminus L, \; L \subset \mathbb{R}^d.$$

Die Wahrscheinlichkeitsdichte f_L kann zulässig sein, wenn kein Eckquader einer Schale ganz in einem Loch des Signifikanzbereiches liegt. D.h. es gibt keinen Eckquader q mit $P(q) = 0$. Die Schalen selbst hingegen dürfen Löcher enthalten, solange die Bedingung $P(Schale^{(l)}) = p(1-p)^l$ unverletzt bleibt.

Charakterisieren wir nun eine Klasse von Wahrscheinlichkeitsdichten $f : \mathbb{R}^d \to \mathbb{R}$, die nicht zulässig ist. Solche Dichten sind Dichten, bei denen die Dichte $f(x)$ nur auf dem *Rand* von Gebieten oder auf Körpern der Dimension kleiner oder gleich $d-1$ größer als 0 ist. Ein konkretes Beispiel für eine solche Dichte ist die Dichte der Gleichverteilung *auf* dem Einheits*kreis*, d.h.

$$f(x) > 0 \iff |x| = 1.$$

8.3. Eine Relaxierung von Eigenschaft D2

Die Wahrscheinlichkeit $P(\mathcal{W}_l)$, daß $Schale^{(l)}$ als äußerste vollbesetzt ist, ist genauso wie ihre obere Abschätzung $P(R^{(l)})$ abhängig von den Wahrscheinlichkeiten $q_i(l), i = 1, \ldots, 2^d$, der Eckquader von $Schale^{(l)}$. Besonders groß ist die Abhängigkeit von $m(l) = \min_{1 \leq i \leq 2^d} q_i(l)$, der kleinsten Wahrscheinlichkeit eines Eckquaders. Dies ist intuitiv; denn wenn o.B.d.A. die Wahrscheinlichkeit von Eckquader 1, $q_1(l)$, sehr klein ist, wird – so groß die Wahrscheinlichkeit eines anderen Eckquaders auch sein mag – auch $P(\mathcal{W}_l)$ bzw. $P(R^{(l)})$ sehr klein sein[34].

In Abschnitt 7.2 wurde herausgearbeitet, daß Vermutung 7.2 gilt, falls für große Schachtelungstiefen l das Verhältnis der Eckquaderwahrscheinlichkeiten $\frac{q(l)}{q(l+1)}$ gegen eine Konstante konvergiert, wobei wegen der Gültigkeit von **D2** alle Eckquader dieselbe Wahrscheinlichkeit besaßen. Da im allgemeinen **D2** nicht gilt und $m(l)$ für $P(\mathcal{W}_l)$ die einflußreichste Wahrscheinlichkeit ist, ist es ein hinreichendes Kriterium für die Gültigkeit von Vermutung 7.2, wenn für große l der Quotient $\frac{m(l)}{m(l+1)}$ konstant wird.

[34] Formal ergibt sich dies, wenn man die Abschätzung $P(R^{(l)}) \geq 1 - \sum_{i=1}^{2^d} (1 - q_i(l))^n$ betrachtet. Wahrscheinlichkeiten $q_i(l)$ mit kleinem Wert sind ausschlaggebender als $q_i(l)$ mit großen Werten.

Definition 8.8 (relaxiert) *Eine zulässige Wahrscheinlichkeitsdichte, für die eine Hyperquader-Schalen-Schachtelung existiert, so daß eine Konstante $c > 0$ mit*

R $\lim_{l \to \infty} \frac{m(l)}{m(l+1)} = c$

existiert, heißt relaxiert.

Es gilt dann wegen der Gültigkeit von Vermutung 7.2

Satz 8.3 *Ist $f : \mathbb{R}^d \to \mathbb{R}$ eine relaxierte Dichte, dann berechnet für $0 < p < 1$ Algorithmus* `A` *die konvexe Hülle von n unabhängig voneinander gewählten Punkten $a_1, \ldots, a_n \in \mathbb{R}^d$ mit erwarteten Kosten*

$$E[\texttt{konvexe Hülle}] = (d + \frac{d}{p} + c_{oh})n + o(n).$$

8.3.1. Beispiele *relaxierter* Dichten

8.3.1.1. Gleichverteilung Wir werden nun zeigen, daß die Dichte der Gleichverteilung auf dem d-dimensionalen Einheitshyperwürfel E relaxiert ist. Da bei einer Gleichverteilung die Wahrscheinlichkeit eines Gebietes bis auf einen konstanten Faktor seinem Volumen entspricht, sind die Eigenschaften **Z1** und **Z2** erfüllt, wenn man die Hyperwürfel $W_0, W_1, \ldots$ konzentrisch um den Ursprung des Koordinatensystems schachtelt. Die Frage ist nun, ob Eigenschaft **R** gilt. Dazu betrachten wir zunächst die Seitenlängen der Hyperwürfel.

Lemma 8.2 *Sei s_0 die Seitenlänge von Hyperwürfel W_0. Für $l \geq 1$ beträgt die Differenz δ_l der Seitenlänge von Hyperwürfel W_{l+1} und Hyperwürfel W_l*

$$\delta_l = s_0 \left(\sqrt[d]{\frac{1-(1-p)^{l+1}}{p}} - \sqrt[d]{\frac{1-(1-p)^l}{p}} \right).$$

Beweis

Der Beweis erfolgt per Induktion.

Induktionsverankerung ***(**$l = 1$**)**:

Es ist wegen **Z1**

$$\frac{{s_0}^d}{(s_0 + \delta_1)^d - {s_0}^d} = \frac{1}{1-p}.$$

Löst man dies nach δ_1 auf, erhält man

$$\begin{aligned} {s_0}^d(1-p) &= (s_0+\delta_1)^d - {s_0}^d \\ \Leftrightarrow \quad \delta_1 &= s_0 \sqrt[d]{2-p} - s_0. \end{aligned}$$

Induktionsschluß *($l \to l+1$)*:

Die Gleichung für δ_{l+1} lautet wegen **Z1**

$$\frac{\left(s_0 + \sum_{j=1}^{l} \delta_j\right)^d - \left(s_0 + \sum_{j=1}^{l-1} \delta_j\right)^d}{\left(s_0 + \sum_{j=1}^{l} \delta_j + \delta_{l+1}\right)^d - \left(s_0 + \sum_{j=1}^{l} \delta_j\right)^d} = \frac{1}{1-p}.$$

Für $k \leq l$ ist die Summe

$$s_0 + \sum_{j=1}^{k} \delta_j = s_0 \sqrt[d]{\frac{1-(1-p)^{k+1}}{p}}$$

eine Teleskopsumme. Die Gleichung vereinfacht sich demnach zu

$$\begin{aligned}
\frac{{s_0}^d \left(\frac{1-(1-p)^{l+1}}{p}\right) - {s_0}^d \left(\frac{1-(1-p)^{l}}{p}\right)}{\left(s_0 \sqrt[d]{\frac{1-(1-p)^{l+1}}{p}} + \delta_{l+1}\right)^d - {s_0}^d \left(\frac{1-(1-p)^{l+1}}{p}\right)} &= \frac{1}{1-p} \\
\Leftrightarrow \quad \frac{{s_0}^d (1-p)^{l+1}}{\left(s_0 \sqrt[d]{\frac{1-(1-p)^{l+1}}{p}} + \delta_{l+1}\right)^d - {s_0}^d \left(\frac{1-(1-p)^{l+1}}{p}\right)} &= 1
\end{aligned}$$

mit der Lösung

$$\delta_{l+1} = s_0 \sqrt[d]{\frac{1-(1-p)^{l+2}}{p}} - s_0 \sqrt[d]{\frac{1-(1-p)^{l+1}}{p}}.$$

$\square$

Die Dichte f ist die Dichte einer Gleichverteilung. Daher ist der Quotient der Eckquaderwahrscheinlichkeiten von $Schale^{(l)}$ und $Schale^{(l+1)}$

$$\frac{m(l)}{m(l+1)} = \frac{q(l)}{q(l+1)} = \left(\frac{\delta_l}{\delta_{l+1}}\right)^d.$$

Lemma 8.3 *Der Quotient* $\frac{\delta_l}{\delta_{l+1}}$ *konvergiert. Es ist*

$$\lim_{l\to\infty} = \frac{\delta_l}{\delta_{l+1}} = \frac{1}{1-p}.$$

Beweis

Ist eine Funktion $g : D \to B$ konkav in D (d.h. $g''(x) < 0$ für alle $x \in D$), dann ist für $x_0 \in D$ und alle $x \neq x_0$ in D

$$g(x_0) + g'(x_0)\,(x - x_0) > g(x).$$

Daraus folgt

$$g'(x)\,(x - x_0) < g(x) - g(x_0) < g'(x_0)\,(x - x_0).$$

Die Funktion

$$g := \left(x \mapsto \frac{s_0}{\sqrt[d]{p}}\, x^{\frac{1}{d}}\right) : \mathbb{R} \to \mathbb{R}$$

ist konkav in $\mathbb{R}^+$, den positiven reellen Zahlen. Setzen wir

$$x_l = 1 - (1-p)^{l+1}$$

dann konvergiert für $l \to \infty$ die Folge x_l gegen 1 und die Ableitung $g'(x_l)$ gegen $\frac{s_0}{d\sqrt[d]{p}}$. Somit konvergiert für $l \to \infty$ die Differenz $g(x_l) - g(x_{l-1})$ gegen $\frac{s_0}{d\sqrt[d]{p}}\,(x_l - x_{l-1})$. Es ist

$$g(x_l) - g(x_{l-1}) = \frac{s_0}{\sqrt[d]{p}}\left(1 - (1-p)^{l+1}\right)^{\frac{1}{d}} - \frac{s_0}{\sqrt[d]{p}}\left(1 - (1-p)^{l}\right)^{\frac{1}{d}} = \delta_l$$

und deshalb

$$\lim_{l\to\infty} \delta_l = \frac{s_0}{d\sqrt[d]{p}}\left((1-p)^l - (1-p)^{l+1}\right).$$

Daraus folgt

$$\lim_{l\to\infty} \frac{\delta_l}{\delta_{l+1}} = \frac{(1-p)^l\,(1-(1-p))}{(1-p)^{l+1}\,(1-(1-p))} = \frac{1}{1-p}.$$

□

Damit ist **R** erfüllt, weil $\lim_{l\to\infty} \frac{m(l)}{m(l+1)} = \left(\frac{1}{1-p}\right)^d$ ist, und es gilt

Satz 8.4 *Die Dichte der Gleichverteilung auf dem d-dimensionalen Einheitshyperwürfel ist relaxiert.*

Da unter affinen Abbildungen eine Hyperquader-Schalen-Schachtelung in eine andere Hyperquader-Schalen-Schachtelung – jedoch mit denselben relativen Abständen zwischen den Begrenzungsgeraden – überführt wird und folglich für alle $l \geq 0$ die Wahrscheinlichkeiten $P(Schale^{(l)})$ und $m(l)$ gleich bleiben, folgt

Korollar 8.1 *Jede Dichte, die eine affine Abbildung der Dichte der Gleichverteilung auf dem d-dimensionalen Einheitshyperwürfel ist, ist relaxiert.*

8.3.1.2. Standard-Normalverteilung Betrachten wir die Dichte der d-dimensionalen Standard-Normalverteilung:

$$f := \left(x \mapsto \left(\frac{1}{2\pi}\right)^{\frac{d}{2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{d} {x_i}^2\right)\right) : \mathbb{R}^d \to \mathbb{R}.$$

Die Ableitung von f ist der Gradient

$$\operatorname{grad} f(x) = \Big(- x_1 f(x), -x_2 f(x), \ldots, -x_d f(x)\Big).$$

Sei $|x|$ der Euklidische Abstand des Vektors $x \in \mathbb{R}^d$ vom Zentrum des Koordinatensystems. Man sieht sofort, daß für große Werte $|x|$ der Gradient von f gegen Null strebt. Da der Gradient die Richtung der größten Steigung angibt und für große Werte $|x|$ annähernd Null ist, ist die d-dimensionale Standard-Normalverteilung weit außerhalb vom Zentrum annähernd gleichverteilt. Bei Gleichverteilung haben wir bereits gesehen, daß man mit einer konzentrischen Schachtelung Bedingung **R** erfüllen kann. Es genügt also, wenn man in diesem Bereich der Standard-Normalverteilung konzentrische Hyperwürfel zur Schachtelung verwendet. Die Dichte der d-dimensionalen Standard-Normalverteilung ist daher relaxiert.

Nun ist es jedoch bei jeder Dichte $f : \mathbb{R}^d \to \mathbb{R}$, deren Signifikanzbereich $S = \mathbb{R}^d$ ist, der Fall, daß sie weit außerhalb vom Zentrum annähernd eine Gleichverteilung definiert. Ansonsten wäre die Normierungsbedingung

$$\int_{\mathbb{R}^d} f(x)\, dx = 1$$

nicht erfüllbar.

Satz 8.5 *Jede Dichte, deren Signifikanzbereich $S = \mathbb{R}^d$ ist, ist relaxiert.*

Wir sehen also, daß die Forderung **R** von einer Vielzahl von Dichten erfüllt werden kann. Als Beispiel ist für eine einfache Dichte über $\mathbb{R}^2$ mit Signifikanzbereich $S = \mathbb{R}^2$ im Anhang eine Schachtelungskonstruktion angegeben.

8.4. Erwartete Laufzeit bei *komponenten-unabhängigen* Dichten

Eine Dichte $f : \mathbb{R}^d \to \mathbb{R}$ bei der alle d Komponenten stochastisch unabhängig voneinander sind, d.h.

$$f(x_1, \ldots, x_n) = \prod_{i=1}^{n} f^i(x_i),$$

wobei $f^i : \mathbb{R} \to \mathbb{R}$ Dichten sind, heißt komponenten-unabhängig. Bei einer zulässigen Dichte beträgt die Anzahl der erwarteten Punkte nach dem Aussiebschritt[35]

$$\begin{aligned} E[Y] &\leq E[Y_{l'}] + \sum_{r=0}^{l'-1} P(\overline{R^{(l'-r)}}) \cdot E[Y_{l'-(r+1)}] \\ &\leq n(1-p)^{l'} \cdot \left(1 + \sum_{r=0}^{l'-1} \frac{P(\overline{R^{(l'-r)}})}{(1-p)^{r+1}}\right) \\ &\leq c \cdot n(1-p)^{l'}. \end{aligned}$$

Da für große n die Schachtelungstiefe l' größer wird, ist asymptotisch die erwartete Anzahl $E[Y] \leq n$.

Bentley und Shamos [3] haben einen Algorithmus angegeben, der die Maxima[36] einer Punktmenge aus einer komponenten-unabhängigen Dichte in erwarteter Zeit $O(n)$ ermittelt. Für $d \leq 3$ läßt sich mit dem Algorithmus sogar die konvexe Hülle berechnen. Da jeder Punkt der konvexen Hülle unter jeder der 2^d Vorzeichenzuweisungen an das Koordinatensystem ein Maximum ist[37], kann man mit erwarteten Kosten $O(n)$ eine Obermenge der konvexen Hülle von $a_1, \ldots, a_n$ berechnen.

Wir modifizieren nun Algorithmus `A`. Statt in Schritt 2 ein Standardverfahren zur Berechnung der konvexen Hülle zu benutzen, rufen wir 2^d-mal – jedesmal mit einer anderen Vorzeichenzuweisung an das Koordinatensystem – das Verfahren von Bentley und Shamos zur Berechnung der Maxima auf. Den modifizierten Algorithmus nennen wir `A'`.

Satz 8.6 [38] *Sei f eine komponenten-unabhängige, zulässige Dichte. Algorithmus* `A'` *berechnet dann für $d \leq 3$ die konvexe Hülle und für $d > 3$ eine Obermenge der konvexen Hülle von n unabhängig voneinander gewählten Punkten $a_1, \ldots, a_n$ mit erwarteten Kosten $O(n)$.*

[35] vgl. Abschnitt 3.3.2, S. 31

[36] Gegeben eine Punktmenge $M \subseteq \mathbb{R}^d$. Wir betrachten die maximale Teilmenge $T \subseteq M$, deren Elemente untereinander alle bezüglich $\leq$ aus Definition 8.4 vergleichbar sind. Das größte Element von T ist dann ein Maximum von M. Beachte: M kann mehrere Maxima haben, die untereinander jedoch bezüglich $\leq$ nicht vergleichbar sind.

[37] vgl. [21], S. 152

[38] vgl. [2]

9. Schluß

Wir haben ein Verfahren angegeben, mit dem sich für bestimmte Wahrscheinlichkeitsdichten die konvexe Hülle einer Punktmenge des $\mathbb{R}^d$ im Mittel in asymptotisch linearer Laufzeit berechnen läßt. Die Schlüsselannahme war, daß eine unendliche Schachtelung vorliegt, die den $\mathbb{R}^d$ in geeigneter Weise unterteilt. Es wurde gezeigt, daß für eine Vielzahl von Dichten eine solche Schachtelung existiert. Damit haben wir aber noch nicht gesagt, wie man diese Schachtelung nun konkret konstruiert. Abhängig von der gegebenen Wahrscheinlichkeitsdichte kann es sein, daß die Konstruktion einer solchen Schachtelung entweder nicht möglich oder nicht-trivial ist. Besonders störend ist die Forderung, daß die Schachtelung unendlich sei. Dadurch ist es effektiv unmöglich, daß die Schachtelung von Anfang an in aller Vollständigkeit vorliegt. In der vorliegenden Form ist es daher durchaus berechtigt, zu sagen, die Analyse zeige lediglich, daß man bei bestimmten Wahrscheinlichkeitsdichten die konvexe Hülle in erwarteter linearer Zeit berechnen könnte.

Nun ist es aber nicht der Fall, daß die vollständige Schachtelung bereits am Anfang vorliegen muß. Bei funktionalen Programmiersprachen gibt es das Konzept der *lazy evaluation* (vgl. [13], S. 183ff), mit dem es möglich ist, auch potentiell unendliche Datenstrukturen zu verarbeiten; die Datenstruktur wird dann nur bis zu dem Wert ausgewertet, der auch tatsächlich benötigt wird. In unserem Falle kann man daher die Konstruktion der Schachtelung als *online*-Schritt betrachten.Man beginnt zuerst nur mit $Schale^{(0)}$. Sobald ein Punkt außerhalb von $Schale^{(0)}$ liegt, werden so viele Schalen berechnet, bis der Punkt lokalisiert werden kann.

Eine andere Möglichkeit, die Schachtelung effektiv auszurechnen, besteht darin, die Schachtelung nur bis zu einer festen Tiefe zu ermitteln. Bei der Bestimmung der erwarteten Kosten von Schritt 2 von Algorithmus A haben wir ohnehin die Schachtelung nur bis Schachtelungstiefe $l_n = \lfloor \alpha \log_b n \rfloor$ betrachtet, wobei n die Eingabegröße, $0 < \alpha < -\frac{1}{log_b(1-p)}$ und b eine geeignete Basis größer als 1 war (vgl. Abschnitt 7.1.2, S. 81). Bei einer endlichen Schachtelung bis Schachtelungstiefe $l_n = \lfloor \alpha \log_b n \rfloor$ würde daher die Abschätzung der erwarteten Kosten von Schritt 2 unverändert bleiben und die erwarteten Kosten von Schritt 1 würden sogar verringert werden.

Vergleichen wir nun den hier vorgestellten Ansatz mit anderen Verfahren, so ist der Vergleich mit den Verfahren von Golin & Sedgewick; Bentley, Clarkson & Levine und Borgwardt interessant, da für diese Verfahren auch erwartete Laufzeiten im $\mathbb{R}^d$ angegeben wurden. Das Verfahren von Golin & Sedgewick berechnet bei Gleichverteilung auf einem d-dimensionalen Hyperquader eine Obermenge der konvexen Hülle in erwarteter Linearzeit. Bentley, Clarkson & Levine's Verfahren berechnet für $d \leq 3$ die konvexe Hülle und für $d > 3$ eine Obermenge der konvexen Hülle in erwarteter Linearzeit, falls die Verteilung komponenten-unabhängig ist (dazu gehört z.B. auch jede Gleichverteilung auf einem d-dimensionalen Hyperquader). Im Vergleich dazu stellt das hier vorgestellte Verfahren eine Verbesserung dar, da bei Gleichverteilung auf einem d-dimensionalen Hyperquader keine Obermenge, sondern die konvexe

Hülle selbst in erwarteter Linearzeit berechnet wird. Bei zulässigen, komponenten-unabhängigen Dichten wird wie beim Verfahren von Bentley, Clarkson & Levine in erwarteter Zeit $O(n)$ eine Obermenge der konvexen Hülle berechnet. Die Bedingung der Zulässigkeit der Dichte ist keine große Einschränkung, da die meisten interessanten Dichten, z.B. die Dichte der d-dimensionalen Standard-Normalverteilung, zulässig und komponenten-unabhängig sind. Im Falle der Dichte der d-dimensionalen Standard-Normalverteilung wird von unserem Verfahren sogar die konvexe Hülle selbst in erwarteter Linearzeit berechnet.

Ein Vergleich mit dem Verfahren von Borgwardt ist schwieriger, da Borgwardts Analyse auf rotationssymmetrischen Verteilungen beruht. Unser Ansatz garantiert bei bestimmten rotationssymmetrischen Verteilungen zwar lineare erwartete Laufzeit, eine Analyse bei parametrisierten Verteilungen wie bei Borgwardt ist jedoch kompliziert. Generell kann man bei unserem Ansatz die erwartete Laufzeit bei Gleichverteilungen auf nicht-hyperquaderförmigen Gebieten schwer angeben. Das Problem besteht darin, daß die erste Bedingung der Zulässigkeit

$$\textbf{Z1}\ P(Schale^{(l)}) = p(1-p)^l \text{ für } 0 < p < 1,\ l \geq 0$$

immer erfüllbar ist. Es kann jedoch sein, daß ab einer bestimmten Schachtelungstiefe die Eckquader der Schalen aus dem Bereich der Gleichverteilung herausragen und somit ihre Wahrscheinlichkeit gleich Null wird. D.h. es ist dann nicht mehr möglich, daß die Schalen vollbesetzt werden.

Ist die Wahrscheinlichkeitsdichte $f : \mathbb{R}^d \to \mathbb{R}$ jedoch zulässig – und das sind alle Wahrscheinlichkeitsdichten, für die ein Hyperquader H existiert, der die gesamte Wahrscheinlichkeitsmasse enthält, so daß

$$\int_{\mathbb{R}^d} f(x)\,dx = \int_H f(x)\,dx = 1$$

und $f(x) > 0$ für alle $x \in H$ ist –, dann ist die erwartete Laufzeit des Verfahrens mit $O(n) + o(n \log n)$ für $d \leq 3$ und $O(n) + o(n^{\lfloor \frac{d}{2} \rfloor})$ für $d > 3$ immer noch erheblich geringer als die worst-case Laufzeit der bekannten Algorithmen zur Berechnung der konvexen Hülle.

A. Anhang

Wir geben nun für die Dichte

$$f := \left((x,y) \mapsto \frac{1}{4}\, e^{-(|x|+|y|)}\right) : \mathbb{R}^2 \to \mathbb{R}$$

eine Schachtelungskonstruktion an, die zeigt, daß f relaxiert ist. Da f rotationssymmetrisch ist, genügt es, wenn man eine Schachtelung aus um den Nullpunkt zentrierten, achsenorientierten Quadraten aufbaut. Dies hat darüber hinaus den Vorteil, daß alle Eckfelder eines Ringes dieselbe Wahrscheinlichkeit besitzen werden, so daß man stellvertretend nur die Wahrscheinlichkeit eines Eckfeldes auszurechnen braucht. Der Einfachheit halber berechnen wir daher nur die Wahrscheinlichkeit der Eckfelder im Quadranten $x \geq 0, y \geq 0$. Seien (x_{l-1}, x_{l-1}) und (x_l, x_l) die definierenden Punkte des entsprechenden Eckfeldes in $Ring^{(l)}$. Die Wahrscheinlichkeit $q(l)$ dieses Eckfeldes ist

$$\begin{aligned}
\int_{x_{l-1}}^{x_l} \int_{x_{l-1}}^{x_l} \frac{1}{4} e^{-(x+y)}\, dx\, dy &= \frac{1}{4} \int_{x_{l-1}}^{x_l} -e^{-(x+y)}\Big|_{x_{l-1}}^{x_l}\, dy \\
&= \frac{1}{4}\left(e^{-2x_l} - 2\, e^{-(x_l+x_{l-1})} + e^{-2x_{l-1}}\right).
\end{aligned}$$

Das Quadrat Q_l sei durch den Punkt (x_l, x_l) gegeben. Da es die Wahrscheinlichkeit

$$\begin{aligned}
P(Q_l) &= \int_0^{x_l} \int_0^{x_l} e^{-(x+y)}\, dx\, dy = e^{-2x_l} - 2\, e^{-x_l} + 1 \\
&= \sum_{i=0}^{l} p(1-p)^i = 1 - (1-p)^{l+1}
\end{aligned}$$

besitzt, läßt sich x_l leicht ausrechnen.

$$x_l = -\ln\left(1 - \sqrt{1 - (1-p)^{l+1}}\right).$$

Für $p = 0.1$ und $l = 1, \ldots, 24$ ist der Quotient $\frac{m(l)}{m(l+1)} = \frac{q(l)}{q(l+1)}$ in Abbildung 55 dargestellt. Man sieht deutlich, daß die Folge der Quotienten gegen $\left(\frac{1}{1-p}\right)^2 \approx 1.235$ konvergiert (vgl. Lemma 8.3). Die Dichte ist daher relaxiert. Der Grund dafür ist, daß der Signifikanzbereich S von f gleich $\mathbb{R}^2$ ist und f somit weit außerhalb vom Zentrum annähernd eine Gleichverteilung definiert.

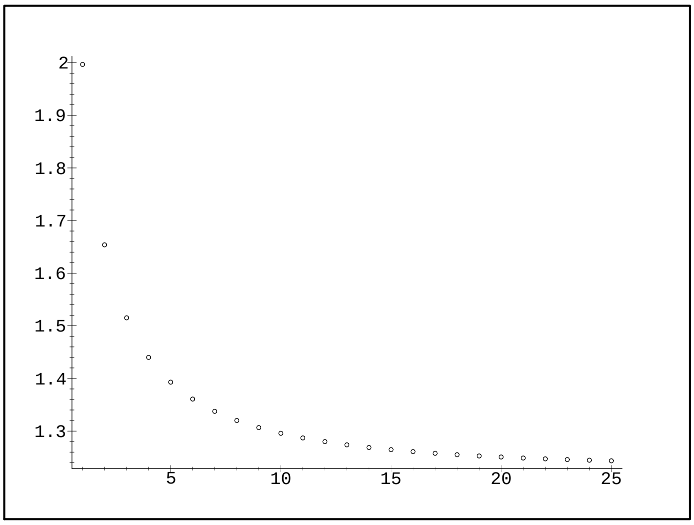

Abbildung 55: $\frac{m(l)}{m(l+1)}$

Literatur

[1] **Akl, G. und Toussaint, G.T.:** `A Fast Convex Hull Algorithm`, Information Processing Letters, Bd.7 - 1978

[2] **Bentley, J.L.; Clarkson, K.L. und Levine, D.B.:** `Fast Linear Expected-Time Algorithms for Computing Maxima and Convex Hulls`, Algorithmica, Bd.9 - 1993

[3] **Bentley, J.L. und Shamos, M.I.:** `Divide and Conquer for Linear Expected Time`, Information Processing Letters, Bd.7 - 1978

[4] **Borgwardt, K.H.:** `Average Complexity of a Gift-Wrapping Algorithm for Determining the Convex Hull of Randomly Given Points`, Discrete & Computational Geometry, Bd.17 - 1997

[5] **Bruijn, Nicolaas G. De:** `Asymptotic Methods In Analysis`, North-Holland Publishing Co., Amsterdam 1958

[6] **Chazelle, B.:** `An optimal Convex Hull Algorithm in Any Fixed Dimension`, Discrete & Computational Geometry, Bd.10 - 1993

[7] **Clarkson, K.L.; Mehlhorn, K. und Seidel, R.:** `Four Results on Randomized Incremental Constructions`, Computational Geometry: Theory and Applications, Bd.3 - 1993

[8] **Devroye, L. und Toussaint, G.T.:** `A Note on Finding Convex Hulls via Maximal Vectors`, Information Processing Letters, Bd.11 - 1980

[9] **Dwyer, Rex:** `Kinder, Gentler Average-Case Analysis for Convex Hulls and Maximal Vectors`, SIGACT News, Bd.21 - 1990

[10] **Graham, Ronald L.:** `An Efficient Algorithm for Determining the Convex Hull of a Finite Planar Set`, Information Processing Letters, Bd.1 - 1972

[11] **Graham, Ronald L. und Yao, Frances:** `Finding the Convex Hull of a Simple Polygon`, Technical Report CS-TR-81-887, Stanford University - Dept. of Comp. Science, 1995

[12] **Golin, M. und Sedgewick, R.:** `Analysis of a Simple yet Efficient Convex Hull Algorithm`, Proceedings of the 4th Annual Symposium on Computational Geometry, 1988

[13] **Hinze, Ralf:** `Einführung in die funktionale Programmierung mit Miranda`, B.G. Teubner, Stuttgart 1992

[14] **Hotz, Günter:** `Algorithmische Informationstheorie`, B.G. TEUBNER Verlagsgesellschaft, Stuttgart/Leipzig 1997

[15] **Jarvis, R.A.:** `On the Identification of the Convex Hull of a Finite Set of Points in the Plane`, Information Processing Letters, Bd.2 - 1973

[16] **Jünger, M.; Borgwardt, K.H.; Gaffke, N. und Reinelt, G.:** `Computing the Convex Hull in the Euclidean Plane in Linear Expected Time`, in: Applied Geometry and Discrete Mathematics (Gritzmann, P. und Sturmfels, B. Hrsg.), Bd.4 - 1991

[17] **Kirkpatrick, David G. und Seidel, Raimund:** `The Ultimate Planar Convex Hull Algorithm ?`, Technical Report TR83-577, Cornell University - Comp. Science Dept., 1983

[18] **Maus, Arne:** Delauney Triangulation and the Convex Hull of n Points in Expected Linear Time, BIT, Bd.24 - 1984

[19] **Mehlhorn, Kurt:** `Data Structures and Algorithms 3: Multi-dimensional Searching and Computational Geometry`, S. 93-102, Springer-Verlag, Berlin 1984

[20] **Min, Hla; Zheng, Si-Qing:** `Time-Space Optimal Convex Hull Algorithms`, Proceedings of the Symposium on Applied Computing (Deaton, Ed; George, K.M.; Berghel, Hal und Hedrick, George Hrsg.), S. 687-693, ACM Press, Indianapolis 1993

[21] **Preparata, Franco P.; Shamos, Michael Ian :** `Computational Geometry, An Introduction`, Springer-Verlag, New York 1985

[22] **Seidel, Raimund:** `A Convex Hull Algorithm Optimal for Point Sets in Even Dimensions`, Technical Report TR-81-14, University of British Columbia - Dept. of Comp. Science, 1981

[23] **Schulz, Frank:** `Sortieren und Suchen unter dem Gesichtspunkt der statistischen Informationstheorie`, 31. Workshop Komplexitätstheorie, Datenstrukturen und Effiziente Algorithmen (Lautemann, Clemens und Schwentick, Thomas Hrsg.), 1997

Index