

FernUniversität in Hagen

Bachelorarbeit Sommersemester 2022

---

**Stochastische Tverberg-Sätze  
mit Anwendungen in der Data Science**

---

**Michael Koller**

Matrikelnummer: 9198300

Betreuung: Prof. Dr. Hochstättler

Datum: 6. August 2022



# Inhaltsverzeichnis

<b>1</b>	<b>Einführung</b>	<b>5</b>
1.1	Hauptquelle . . . . .	5
1.2	Motivation . . . . .	5
1.3	Gliederung . . . . .	6
<b>2</b>	<b>Notation, Definitionen, Hintergrund</b>	<b>7</b>
2.1	Notation . . . . .	7
2.2	Farbige Punktmenge, Tverberg-Zerlegung, Tverberg-Punkt . . . . .	7
2.3	Deterministische Tverberg-Sätze . . . . .	8
2.4	Zwei Modelle zufälliger Punktmenge . . . . .	14
2.5	Für Beweise wichtige Aussagen . . . . .	15
<b>3</b>	<b>Stochastische Radon- und Tverberg-Sätze</b>	<b>21</b>
3.1	In der Literatur bekannte Aussagen . . . . .	21
3.2	Zufällige Gleichverteilung . . . . .	22
3.3	Zufällige Zuordnung . . . . .	29
<b>4</b>	<b>Anwendungsbeispiele</b>	<b>33</b>
4.1	Zentralwert . . . . .	33
4.2	Separierbarkeit . . . . .	35
4.3	Maximum-Likelihood-Schätzung . . . . .	38
4.3.1	Farberkennung bei mehreren Farben . . . . .	39
4.3.2	Binäre logistische Regression . . . . .	40
4.3.3	Existenz der Maximum-Likelihood-Schätzung . . . . .	40
<b>5</b>	<b>Schluss</b>	<b>45</b>



# 1 Einführung

Wir beschäftigen uns in dieser Arbeit mit Tverberg-Zerlegungen. Wir betrachten dabei eine Menge  $\mathcal{S} \subset \mathbb{R}^d$  und wollen sie derart in  $m \in \mathbb{N}$  paarweise disjunkte und nichtleere Teilmengen  $\mathcal{S}_1, \dots, \mathcal{S}_m$  zerlegen, dass der Schnitt aller konvexen Hüllen nichtleer ist. Es soll also  $\bigcap_{i=1}^m \text{conv}(\mathcal{S}_i) \neq \emptyset$  gelten. Wir sprechen dann von einer  $m$ -Tverberg-Zerlegung.

## 1.1 Hauptquelle

Es handelt sich bei dieser Arbeit um eine Darstellung der Ergebnisse aus [10]. Im Folgenden stammen sämtliche Aussagen und Beweise, die nicht mit einer eigenen Quellenangabe versehen sind, aus [10]. Um den Zusammenhang zu [10] noch expliziter zu machen, ist bei der Formulierung von Sätzen und Definitionen, die auch in [10] aufgeführt werden, in Klammern angegeben, welche Bezeichnung dafür in [10] verwendet wird.

## 1.2 Motivation

Es ist aus [22] bekannt, dass für jede Menge  $\mathcal{S} \subset \mathbb{R}^d$  mit mindestens  $(d+1)(m-1)+1$  Punkten eine  $m$ -Tverberg-Zerlegung existiert. Es sind jedoch keine effizienten Methoden bekannt, um zu einer gegebenen Menge eine  $m$ -Tverberg-Zerlegung zu bestimmen. Diese Tatsache motiviert die Betrachtung von stochastischen Methoden, die Zufälligkeit ausnutzen, um die gewünschte Eigenschaft mit hoher Wahrscheinlichkeit zu erreichen. Zu diesem Zweck stellen die Autoren von [10] zunächst zwei Modelle zur zufälligen Erzeugung von  $n$ -elementigen Mengen  $\mathcal{S} \subset \mathbb{R}^d$  vor. Zusätzlich wird jeder der  $n$  Punkte zufällig in eine von  $m$  Farben eingefärbt. Auf diese Weise wird  $\mathcal{S}$  in  $m$  den Farben entsprechende Teilmengen  $\mathcal{S}_1, \dots, \mathcal{S}_m$  zerlegt. Die Autoren beantworten dann beispielsweise die Frage, in welchem Verhältnis  $m$  und  $n$  stehen müssen, um mit hoher Wahrscheinlichkeit eine  $m$ -Tverberg-Zerlegung zu erhalten.

Eine weitere Motivation von stochastischen Tverberg-Aussagen stellt das Lösen eines Klassifizierungsproblems dar. Wir beschreiben die Details in Abschnitt 4.3. Prinzipiell liegt ein Datenpunkt im  $\mathbb{R}^d$  vor und es ist bekannt (oder wird angenommen), dass dieser genau einer von  $m$  Klassen zugeordnet werden kann. Das Klassifizierungsproblem besteht darin, dem Punkt die richtige Klasse zuzuordnen. Ein Ansatz, die Klassifizierung durchzuführen, stellt die logistische Regression dar. Hierfür sind zunächst  $n$  Datenpunkte im  $\mathbb{R}^d$  gegeben, wobei jeder Punkt bereits genau einer von  $m$  Klassen zugeordnet ist. Diese Datenpunkte werden zusammen mit ihren Klassen benutzt, um die Koeffizienten der logistischen Regression zu bestimmen. Im Anschluss kann für neue („klassenlose“) Datenpunkte eine Klasse bestimmt werden. Der Zusammenhang zu stochastischen Tverberg-Aussagen ergibt sich wie folgt. Wir teilen die  $n$  vorliegenden Datenpunkte auf  $m$  ihren Klassen entsprechende Teilmengen  $\mathcal{S}_1, \dots, \mathcal{S}_m$  auf. Falls es sich bei  $\mathcal{S}_1, \dots, \mathcal{S}_m$  um eine  $m$ -Tverberg-Zerlegung von  $\mathcal{S}$  handelt, finden wir optimale Koeffizienten der logistischen Regression. Details und die Antwort auf die Frage, was „optimale“ Koeffizienten sind, erklären wir in Abschnitt 4.3. Stochastische Tverberg-Aussagen können nun helfen, um

beispielsweise zu klären, wie groß  $n$  (die Anzahl der vorliegenden Datenpunkte) sein muss, um das Klassifizierungsproblem mithilfe der logistischen Regression lösen zu können.

### 1.3 Gliederung

Wir definieren die für diese Arbeit wichtigen Begriffe in Abschnitt 2. Dieser Abschnitt führt auch die zwei Modelle zur zufälligen Erzeugung von  $n$ -elementigen Mengen im  $\mathbb{R}^d$  ein, enthält deterministische Tverberg-Aussagen und erklärt für die Beweise der stochastischen Aussagen wichtige Sätze. Abschnitt 3 enthält alle stochastischen Tverberg-Sätze. Wir konzentrieren uns in Abschnitt 3.2 auf das erste Modell zur zufälligen Erzeugung von  $n$ -elementigen Mengen und wir konzentrieren uns in Abschnitt 3.3 auf das zweite Modell. Abschnitt 3.1 führt stochastische Tverberg-Aussagen auf, welche in der Literatur bekannt sind und nicht in [10] bewiesen werden. Abschnitt 4 zeigt drei Anwendungsbeispiele der stochastischen Tverberg-Sätze. Unter anderem kommen wir in Abschnitt 4.3 auf die bereits erwähnte logistische Regression zurück.

## 2 Notation, Definitionen, Hintergrund

Wir führen in diesem Abschnitt die im Rest der Arbeit verwendete Notation ein, definieren Begriffe wie die Tverberg-Zerlegung und diskutieren die deterministischen Tverberg-Aussagen, welche die in Abschnitt 3 vorgestellten stochastischen Tverberg-Aussagen motivieren. Um stochastische Tverberg-Aussagen zu treffen, greifen wir auf zwei Modelle zur Erzeugung zufälliger Punktmenge zurück, welche das Thema von Abschnitt 2.4 sind. Wir schließen Abschnitt 2 mit einer Auflistung von Aussagen, die für Beweise im Rest der Arbeit eine Rolle spielen.

### 2.1 Notation

Wir schreiben abkürzend  $[m] = \{1, \dots, m\}$  für  $m \in \mathbb{N}$ . Wir sagen bei einer Folge  $(X_n)_{n \in \mathbb{N}}$  von Ereignissen, dass diese mit hoher Wahrscheinlichkeit eintreten, falls  $\lim_{n \rightarrow \infty} \mathbb{P}(X_n) = 1$  erfüllt ist. Im Speziellen finden wir dann zu jedem  $\varepsilon > 0$  ein  $N \in \mathbb{N}$ , sodass  $\mathbb{P}(X_n) \geq 1 - \varepsilon$  für alle  $n \geq N$  erfüllt ist. Außerdem schreiben wir  $f(x) \gg g(x)$  für zwei Funktionen  $f$  und  $g$ , falls  $\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} \geq 1$  erfüllt ist.

### 2.2 Farbige Punktmenge, Tverberg-Zerlegung, Tverberg-Punkt

Zu einer gegebenen nichtleere Menge  $\mathcal{S} \subset \mathbb{R}^d$  bezeichnet

$$\text{aff}(\mathcal{S}) = \left\{ \sum_{i=1}^l \lambda_i \mathbf{x}_i \mid \mathbf{x}_i \in \mathcal{S}, l \in \mathbb{N}, \sum_{i=1}^l \lambda_i = 1 \right\} \quad (2.1)$$

deren *affine Hülle* [16]. Weiter bezeichnet

$$\text{conv}(\mathcal{S}) = \left\{ \sum_{i=1}^l \lambda_i \mathbf{x}_i \mid \mathbf{x}_i \in \mathcal{S}, l \in \mathbb{N}, \sum_{i=1}^l \lambda_i = 1, \lambda_i \geq 0 \right\} \quad (2.2)$$

die *konvexe Hülle* [1]. Die konvexe Hülle einer endlichen Menge ist kompakt [16] und auch als *beschränktes Polyeder* bekannt [5]. Wir kommen auf diese Tatsachen in zwei Beweisen zurück. Zuletzt definieren wir noch die Dimension einer konvexen Menge  $\mathcal{C} \subset \mathbb{R}^d$  als  $\dim(\mathcal{C}) = \dim(\text{aff}(\mathcal{C}))$ .

Mit diesen Definitionen können wir die Tverberg-Zerlegung (siehe Definition 2.2) definieren. Zuvor führen wir noch den Begriff *farbige Punktmenge* ein. Da es anschaulich ist, sprechen wir im weiteren Verlauf von *Farben*. Eine Alternative wäre, nicht von *Farben*, sondern von *Klassen* oder von *Kategorien* zu sprechen.

**Definition 2.1.** *Wir bezeichnen eine Teilmenge  $\mathcal{X} \times \mathcal{Y}$  von  $\mathbb{R}^d \times [m]$  als  $m$ -farbige Punktmenge in  $\mathbb{R}^d$ . Jedes Element  $(\mathbf{x}, y) \in \mathcal{X} \times \mathcal{Y}$  ist ein Punkt  $\mathbf{x} \in \mathbb{R}^d$  zusammen mit einer Farbe  $y \in [m]$ . Wir sprechen auch von einem  $y$ -eingefärbten Punkt  $\mathbf{x}$ . Je nach Kontext wird abkürzend schon  $\mathcal{X}$  als  $m$ -farbige Punktmenge in  $\mathbb{R}^d$  bezeichnet.*

Wir sind in dieser Arbeit daran interessiert, Aussagen darüber zu treffen, ob eine gegebene farbige Punktmenge eine Tverberg-Zerlegung (gegebenenfalls mit Toleranz) darstellt. Wir nennen  $\mathcal{S}_1, \dots, \mathcal{S}_m$  eine *Zerlegung der Menge*  $\mathcal{S} \subset \mathbb{R}^d$  *in  $m$  Teilmengen*, falls alle Teilmengen nichtleer und paarweise disjunkt sind und  $\mathcal{S} = \bigcup_{i \in [m]} \mathcal{S}_i$  gilt.

**Definition 2.2** (Tverberg-Zerlegung mit Toleranz, Definition 2.2). *Es sei eine Menge  $\mathcal{S} \subset \mathbb{R}^d$  gegeben. Eine  $m$ -Tverberg-Zerlegung von  $\mathcal{S}$  mit Toleranz  $t$  ist eine Zerlegung von  $\mathcal{S}$  in  $m$  Teilmengen  $\mathcal{S}_1, \dots, \mathcal{S}_m$ , sodass für alle  $\mathcal{C} \subset \mathcal{S}$  mit  $|\mathcal{C}| \leq t$  gilt:*

$$\bigcap_{i \in [m]} \text{conv}(\mathcal{S}_i \setminus \mathcal{C}) \neq \emptyset. \quad (2.3)$$

*Wir sagen in diesem Fall auch „ $\mathcal{S}$  ist Tverberg mit Toleranz  $t$ “. Wir nennen jedes  $\mathbf{x} \in \bigcap_{i \in [m]} \text{conv}(\mathcal{S}_i)$  einen Tverberg-Punkt.*

Eine 2-Tverberg-Zerlegung ( $m = 2$ ) ist auch unter dem Namen *Radon-Zerlegung* bekannt. Wir können uns vorstellen, dass in Definition 2.2 den Punkten der Menge  $\mathcal{S}_i$  die Farbe  $i$  zugeordnet wird,  $i \in [m]$ . Es geht also anschaulich bei einer  $m$ -Tverberg-Zerlegung (mit Toleranz  $t = 0$ ) darum, eine gegebene Punktmenge derart in eine  $m$ -farbige Punktmenge einzufärben, dass sich die konvexen Hüllen aller Farben schneiden. Im Fall einer  $t$ -toleranten  $m$ -Tverberg-Zerlegung können  $t$  beliebige Punkte entfernt werden, ohne dass die Tverberg-Eigenschaft verloren geht.

### 2.3 Deterministische Tverberg-Sätze

Die hier vorgestellten deterministischen Tverberg-Sätze bestimmen, wie viele Elemente eine gegebene Menge  $\mathcal{S} \subset \mathbb{R}^d$  enthalten muss, damit eine Tverberg-Zerlegung (gegebenenfalls mit Toleranz) existiert. Der zweifarbige ( $m = 2$ ) Fall ist in einem Lemma von Radon behandelt worden [14]. Der allgemeine Fall ( $m \in \mathbb{N}$ ) mit Toleranz  $t = 0$  ist als Satz von Tverberg bekannt.

**Satz 2.3** (Tverberg). *Jede Menge  $\mathcal{S}$  mit mindestens  $(d + 1)(m - 1) + 1$  Punkten in  $\mathbb{R}^d$  hat mindestens eine  $m$ -Tverberg-Zerlegung.*

Es gibt mittlerweile mehrere Beweise dieses Satzes. Zwei davon sind vom Namensgeber selbst: einer wurde 1966 veröffentlicht [22] und ein weiterer 1981 [23]. Zudem gibt es einen Beweis, der 1993 aus der Zusammenarbeit von Tverberg und Vrećica hervorgegangen ist [24]. Des Weiteren gibt es Beweise von Sarkaria [18] und Zvageľ'skii [26]. Wir präsentieren einen Beweis von Roudneff [17], der auch in [3] gezeigt wird. In [3] gibt es außerdem diesen historischen Überblick und weiteren Kontext zu deterministischen Tverberg-Aussagen.

**Bemerkung 2.4.** *Im Beweis von Satz 2.3 wird das relative Innere einer Menge benutzt. Für eine konvexen Menge  $\mathcal{C} \in \mathbb{R}^d$  ist dieses definiert als [5]*

$$\text{relint}(\mathcal{C}) = \{\mathbf{x} \in \mathcal{C} \mid \text{es gibt ein } r > 0 \text{ mit } B_r(\mathbf{x}) \cap \text{aff}(\mathcal{C}) \subset \mathcal{C}\}, \quad (2.4)$$

wobei  $B_r(\mathbf{x})$  die Kugel um  $\mathbf{x}$  mit Radius  $r$  bezeichnet. Außerdem wird  $\dim(\emptyset) = -1$  gesetzt und die für  $m$  Unterräume  $U_1, U_2, \dots, U_m$  des  $\mathbb{R}^d$  geltende Ungleichung

$$\text{codim} \left( \bigcap_{i \in [m]} U_i \right) \leq \sum_{i \in [m]} \text{codim}(U_i) = md - \sum_{i \in [m]} \dim(U_i) \quad (2.5)$$

(siehe beispielsweise [16]) verwendet. Hier ist  $\text{codim}(U_i) = d - \dim(U_i)$  die Kodimension von  $U_i$ . Es wird außerdem im Beweis angenommen, dass sich die Punkte in  $\mathcal{S}$  in einer ausreichend allgemeinen Lage befinden, womit beispielsweise keine  $k \leq d + 1$  Punkte in einem  $(k - 2)$ -dimensionalen affinen Unterraum liegen [11].

Zuletzt benötigen wir für den Beweis von Satz 2.3 (und auch für den von Lemma 2.9) noch folgende Überlegung.

**Bemerkung 2.5.** Es sei  $\tilde{\mathcal{S}}$  eine endliche Menge in  $\mathbb{R}^d$  und  $\mathbf{z} \in \mathbb{R}^d$ . Wir definieren den Abstand von  $\mathbf{z}$  zu  $\text{conv}(\tilde{\mathcal{S}})$  als

$$\text{dist}(\mathbf{z}, \text{conv}(\tilde{\mathcal{S}})) = \min_{\mathbf{y} \in \text{conv}(\tilde{\mathcal{S}})} \|\mathbf{z} - \mathbf{y}\|. \quad (2.6)$$

Da  $\tilde{\mathcal{S}}$  endlich ist, gibt es ein eindeutiges  $\mathbf{y} \in \text{conv}(\tilde{\mathcal{S}})$  mit  $\text{dist}(\mathbf{z}, \text{conv}(\tilde{\mathcal{S}})) = \|\mathbf{z} - \mathbf{y}\|$ , ein  $\mathbf{y}$  also, das  $\mathbf{z}$  am nächsten ist [5]. Für  $\mathbf{z} \notin \text{conv}(\tilde{\mathcal{S}})$  liegt dieses  $\mathbf{y}$  auf dem Rand von  $\text{conv}(\tilde{\mathcal{S}})$  und damit in einer echten Seitenfläche des Polyeders, weshalb es mit dem Satz von Carathéodory (siehe beispielsweise [1]) eine Teilmenge  $\tilde{\mathcal{S}}' \subset \tilde{\mathcal{S}}$  mit  $|\tilde{\mathcal{S}}'| \leq d$  gibt, sodass  $\mathbf{y} \in \text{relint}(\text{conv}(\tilde{\mathcal{S}}'))$  gilt.

*Beweis von Satz 2.3.* Für eine beliebige Zerlegung  $\mathcal{P} = \{\mathcal{S}_i\}_{i \in [m]}$  von  $\mathcal{S}$  mit  $|\mathcal{S}_i| \leq d + 1$  definieren wir eine Abstandsfunktion  $f$  über

$$f(\mathbf{x}, \mathcal{P}) = \sum_{i \in [m]} \text{dist}^2(\mathbf{x}, \text{conv}(\mathcal{S}_i)), \quad (2.7)$$

wobei  $\text{dist}$  wie in Bemerkung 2.5 definiert ist. Falls es eine Zerlegung  $\mathcal{P}$  und ein  $\mathbf{x} \in \mathbb{R}^d$  mit  $f(\mathbf{x}, \mathcal{P}) = 0$  gibt, gilt  $\mathbf{x} \in \text{conv}(\mathcal{S}_i)$  für jedes  $i \in [m]$ . Demnach ist  $\mathbf{x}$  ein Tverberg-Punkt und  $\mathcal{P}$  also eine Tverberg-Zerlegung. Wir wollen deshalb im Folgenden die Existenz einer Zerlegung  $\mathcal{P}$  mit  $\min_{\mathbf{x} \in \mathbb{R}^d} f(\mathbf{x}, \mathcal{P}) = 0$  zeigen.

Die Distanz  $\text{dist}$  ist bezüglich des ersten Arguments konvex (siehe beispielsweise [5]) und es gilt  $\lim_{\|\mathbf{x}\| \rightarrow \infty} \text{dist}(\mathbf{x}, \text{conv}(\mathcal{S}_i)) = \infty$  für jedes  $i \in [m]$ . Aus diesem Grund ist auch  $f$  bezüglich des ersten Arguments konvex und nimmt bei festem  $\mathcal{P}$  ein Minimum an, vergleiche beispielsweise [1]. Es bezeichne ab jetzt  $\mathcal{P}$  die Partition, die zum kleinsten Minimum  $\mu = \min_{\mathbf{x} \in \mathbb{R}^d} f(\mathbf{x}, \mathcal{P})$  von  $f$  führt.

Nach der Überlegung unter (2.7) ist  $\mu = 0$  zu zeigen. Der Rest des Beweises beschäftigt sich deswegen damit, die Annahme  $\mu > 0$  zu einem Widerspruch zu führen. Es sei  $\mathbf{z} \in \mathbb{R}^d$  so, dass  $f(\mathbf{z}, \mathcal{P}) = \mu > 0$  gilt. Für jedes  $i \in [m]$  gibt es genau ein  $\mathbf{y}_i \in \text{conv}(\mathcal{S}_i)$  mit  $\text{dist}(\mathbf{z}, \text{conv}(\mathcal{S}_i)) = \|\mathbf{z} - \mathbf{y}_i\|$  (siehe Bemerkung 2.5). Es nimmt nun  $g : \mathbf{x} \mapsto \sum_{i \in [m]} \|\mathbf{x} -$

$\mathbf{y}_i\|^2$  bei  $\mathbf{z}$  ein Minimum an. Da  $g$  bei  $\mathbf{z}$  minimal ist, verschwindet dort der Gradient:  $2\sum_{i\in[m]}(\mathbf{z}-\mathbf{y}_i)=\mathbf{0}$ . Beachte, die Summe über alle  $\mathbf{z}-\mathbf{y}_i$  verschwindet zwar, aber es gibt mindestens ein  $j\in[m]$  mit  $\mathbf{z}\neq\mathbf{y}_j$ , da wir  $\mu>0$  annehmen.

Für jedes  $i\in[m]$  definieren wir  $\mathcal{Y}_i$  als die kleinste Teilmenge von  $\mathcal{S}_i$ , sodass noch  $\mathbf{y}_i\in\text{relint}(\text{conv}(\mathcal{Y}_i))$  erfüllt ist. Dann gilt  $\bigcap_{i\in[m]}\text{aff}(\mathcal{Y}_i)=\emptyset$ . Denn, falls es ein  $\mathbf{v}\in\bigcap_{i\in[m]}\text{aff}(\mathcal{Y}_i)$  gibt, finden wir wie folgt einen Widerspruch. Für jedes  $\mathbf{y}_i$  mit  $\mathbf{z}=\mathbf{y}_i$  gilt offenbar  $(\mathbf{z}-\mathbf{v})^\text{T}(\mathbf{z}-\mathbf{y}_i)=0$ . Wir erklären unten noch, weshalb für jedes  $\mathbf{y}_i$  mit  $\mathbf{z}\neq\mathbf{y}_i$  die Ungleichung  $(\mathbf{z}-\mathbf{v})^\text{T}(\mathbf{z}-\mathbf{y}_i)>0$  gilt. Es folgt dann insgesamt  $\sum_{i\in[m]}(\mathbf{z}-\mathbf{v})^\text{T}(\mathbf{z}-\mathbf{y}_i)=(\mathbf{z}-\mathbf{v})^\text{T}(\sum_{i\in[m]}(\mathbf{z}-\mathbf{y}_i))>0$ , was aber wegen des verschwindenden Gradienten nicht sein kann. Zur Erklärung der Ungleichung: Es bezeichne  $\tilde{\mathbf{y}}_i$  den Punkt in  $\text{aff}(\mathcal{Y}_i)$ , welcher  $\mathbf{z}$  am nächsten ist. Dann ist  $\tilde{\mathbf{y}}_i-\mathbf{z}$  orthogonal zu  $\text{aff}(\mathcal{Y}_i)$ . Wegen  $\text{conv}(\mathcal{Y}_i)\subset\text{aff}(\mathcal{Y}_i)$  gilt  $\tilde{\mathbf{y}}_i=\mathbf{y}_i$ , falls  $\tilde{\mathbf{y}}_i\in\text{conv}(\mathcal{Y}_i)$ . Falls jedoch  $\tilde{\mathbf{y}}_i\notin\text{conv}(\mathcal{Y}_i)$  gilt, muss  $\mathbf{y}_i$  der Punkt in  $\text{conv}(\mathcal{Y}_i)$  sein, der  $\tilde{\mathbf{y}}_i$  am nächsten ist. Damit wäre  $\mathbf{y}_i$  aber nicht mehr in  $\text{relint}(\text{conv}(\mathcal{Y}_i))$ , sondern auf dem Rand von  $\text{conv}(\mathcal{Y}_i)$ . Es ist also  $\mathbf{y}_i-\mathbf{z}$  orthogonal zu  $\text{aff}(\mathcal{Y}_i)$ . Weil  $\mathbf{v}\in\text{aff}(\mathcal{Y}_i)$  und  $\mathbf{y}_i\in\text{aff}(\mathcal{Y}_i)$  gelten, folgt nun  $(\mathbf{z}-\mathbf{v})^\text{T}(\mathbf{z}-\mathbf{y}_i)>0$ .

Aus (2.5) folgt dank der leeren Schnittmenge  $\bigcap_{i\in[m]}\text{aff}(\mathcal{Y}_i)$  die Ungleichung  $d+1\leq md-\sum_{i\in[m]}\dim(\text{aff}(\mathcal{Y}_i))$ . Aus der Minimalität von  $\mathcal{Y}_i$  und der allgemeinen Lage der Punkte in  $\mathcal{S}$  folgt  $\dim(\text{aff}(\mathcal{Y}_i))=|\mathcal{Y}_i|-1$  und damit  $d+1\leq md-\sum_{i\in[m]}(|\mathcal{Y}_i|-1)$ , weshalb schließlich

$$\sum_{i\in[m]}|\mathcal{Y}_i|\leq md+m-(d+1)=(d+1)(m-1) \quad (2.8)$$

erfüllt ist. Wegen  $|\mathcal{S}|>(d+1)(m-1)$  gibt es ein  $\mathbf{x}\in\mathcal{S}$  mit  $\mathbf{x}\notin\mathcal{Y}_i$  für jedes  $i\in[m]$ . Es muss

$$(\mathbf{x}-\mathbf{y}_i)^\text{T}(\mathbf{z}-\mathbf{y}_i)\leq 0 \quad (2.9)$$

für jedes  $i\in[m]$  gelten. Denn, falls für ein  $j\in[m]$  mit  $\mathbf{z}\neq\mathbf{y}_j$  die Ungleichung  $(\mathbf{x}-\mathbf{y}_j)^\text{T}(\mathbf{z}-\mathbf{y}_j)>0$  gilt, können wir  $\mu$  wie folgt verkleinern (und das ist dann ein Widerspruch zur Definition von  $\mu$  als kleinstes Minimum). Wir betrachten  $\lambda\mathbf{x}+(1-\lambda)\mathbf{y}_j\in\text{conv}(\mathcal{Y}_j\cup\{\mathbf{x}\})$  für  $\lambda\in]0,1[$  und berechnen

$$\|\mathbf{z}-\lambda\mathbf{x}-(1-\lambda)\mathbf{y}_j\|^2=\|\mathbf{z}-\mathbf{y}_j-\lambda(\mathbf{x}-\mathbf{y}_j)\|^2 \quad (2.10)$$

$$=\|\mathbf{z}-\mathbf{y}_j\|^2-\lambda(2(\mathbf{z}-\mathbf{y}_j)^\text{T}(\mathbf{x}-\mathbf{y}_j)-\lambda\|\mathbf{x}-\mathbf{y}_j\|^2). \quad (2.11)$$

Für  $\lambda$  klein genug wird der Subtrahend der letzten Gleichung positiv, sodass  $\lambda\mathbf{x}+(1-\lambda)\mathbf{y}_j$  näher an  $\mathbf{z}$  ist als  $\mathbf{y}_j$ . Das widerspricht aber der Optimalität von  $\mu$ . Wir haben also (2.9) und beobachten deswegen

$$0\geq\sum_{i\in[m]}(\mathbf{x}-\mathbf{y}_i)^\text{T}(\mathbf{z}-\mathbf{y}_i)=\sum_{i\in[m]}(\mathbf{x}-\mathbf{z}+\mathbf{z}-\mathbf{y}_i)^\text{T}(\mathbf{z}-\mathbf{y}_i) \quad (2.12)$$

$$=(\mathbf{x}-\mathbf{z})^\text{T}\left(\sum_{i\in[m]}(\mathbf{z}-\mathbf{y}_i)\right)+\sum_{i\in[m]}(\mathbf{z}-\mathbf{y}_i)^\text{T}(\mathbf{z}-\mathbf{y}_i) \quad (2.13)$$

$$=0+\sum_{i\in[m]}\|\mathbf{z}-\mathbf{y}_i\|^2=\mu>0. \quad (2.14)$$

Wegen dieses Widerspruchs muss  $\mu = 0$  gelten. Von (2.13) auf (2.14) haben wir wieder ausgenutzt, dass der Gradient von  $g$  bei  $\mathbf{z}$  verschwindet.  $\square$

Es gibt Mengen mit  $(d+1)(m-1)$  Punkten, für die keine  $m$ -Tverberg-Zerlegung existiert [22]. In diesem Sinne ist die Anzahl der Punkte in Satz 2.3 optimal. Eine Verallgemeinerung auf Tverberg-Zerlegungen mit Toleranz  $t \geq 0$  lautet wie folgt [21]:

**Satz 2.6** (Soberón, Strausz). *Jede Menge  $\mathcal{S} = \{\mathbf{x}_j\}_{j \in [n]}$  mit mindestens  $n = (t+1)(m-1)(d+1) + 1$  Punkten in  $\mathbb{R}^d$  hat mindestens eine  $m$ -Tverberg-Zerlegung mit Toleranz  $t$ .*

**Bemerkung 2.7.** *Wir verwenden im Beweis das für zwei Vektoren  $\mathbf{b} \in \mathbb{R}^{d+1}$  und  $\mathbf{u} = (u_1, u_2, \dots, u_{m-1})^\top \in \mathbb{R}^{m-1}$  als*

$$\mathbf{u} \otimes \mathbf{b} = (u_1 \mathbf{b}^\top, u_2 \mathbf{b}^\top, \dots, u_{m-1} \mathbf{b}^\top)^\top \in \mathbb{R}^{(m-1)(d+1)} \quad (2.15)$$

definierte Kronecker-Produkt der beiden Vektoren, siehe beispielsweise [16].

Weiter machen wir für den Beweis noch folgende Beobachtung: Falls für Vektoren  $\mathbf{u}^{(0)}, \mathbf{u}^{(1)}, \dots, \mathbf{u}^{(m-1)} \in \mathbb{R}^{m-1}$  Koeffizienten  $\alpha_0, \alpha_1, \dots, \alpha_{m-1} \in \mathbb{R}$  mit  $\sum_{i=0}^{m-1} \alpha_i \mathbf{u}^{(i)} = \mathbf{0} \in \mathbb{R}^{m-1}$  existieren, gilt auch

$$\sum_{i=0}^{m-1} \alpha_i (\mathbf{u}^{(i)} \otimes \mathbf{b}) = \left( \sum_{i=0}^{m-1} \alpha_i u_0^{(i)} \mathbf{b}^\top, \sum_{i=0}^{m-1} \alpha_i u_1^{(i)} \mathbf{b}^\top, \dots, \sum_{i=0}^{m-1} \alpha_i u_{m-1}^{(i)} \mathbf{b}^\top \right)^\top \quad (2.16)$$

$$= \left( \left( \sum_{i=0}^{m-1} \alpha_i u_0^{(i)} \right) \mathbf{b}^\top, \left( \sum_{i=0}^{m-1} \alpha_i u_1^{(i)} \right) \mathbf{b}^\top, \dots, \left( \sum_{i=0}^{m-1} \alpha_i u_{m-1}^{(i)} \right) \mathbf{b}^\top \right)^\top \quad (2.17)$$

$$= \mathbf{0} \in \mathbb{R}^{(m-1)(d+1)}. \quad (2.18)$$

Wir beweisen Satz 2.6 mit Hilfe des nachfolgenden Lemma 2.9. Dieses wiederum baut auf [2, Theorem 2.1] auf:

**Satz 2.8.** *Es seien  $\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_{p+1}$  Teilmengen des  $\mathbb{R}^p$  und  $\mathbf{a} \in \bigcap_{i \in [p+1]} \text{conv}(\mathcal{V}_i)$ . Dann existieren Vektoren  $\mathbf{v}_i \in \mathcal{V}_i, i \in [p+1]$ , sodass  $\mathbf{a} \in \text{conv}(\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{p+1}\})$  erfüllt ist.*

Wir führen noch eine Sprechweise ein: Eine Menge  $\mathcal{C} \subset \mathbb{R}^d$  erfasst den Ursprung  $\mathbf{0} \in \mathbb{R}^d$  mit Toleranz  $t$ , falls  $\mathbf{0} \in \text{conv}(\mathcal{C} \setminus \mathcal{C}')$  für jedes  $\mathcal{C}' \subset \mathcal{C}$  mit  $|\mathcal{C}'| \leq t$  erfüllt ist. Damit formulieren wir das nachfolgende Lemma, das wir im Beweis von Satz 2.6 anwenden. Es handelt sich dabei um [21, Lemma 1].

**Lemma 2.9.** *Es seien  $p \geq 1$  und  $t \geq 0$  ganze Zahlen und es sei  $\mathcal{S}_0 = \{\mathbf{a}_i\}_{i \in [n]} \subset \mathbb{R}^p$  eine Menge mit  $n = p(t+1) + 1$  Punkten. Es sei weiter  $G$  eine Gruppe mit  $|G| \leq p$ . Falls es eine Wirkung von  $G$  auf einer Menge  $\mathcal{S}' \subset \mathbb{R}^p$  gibt, welche mit  $\mathcal{S}_0 \subset \mathcal{S}'$  kompatibel ist, gibt es für jedes  $\mathbf{a}_i \in \mathcal{S}_0$  ein  $g_i \in G$ , sodass  $\{g_i \mathbf{a}_i\}_{i \in [n]}$  den Ursprung mit Toleranz  $t$  erfasst. Dabei ist  $G$  mit  $\mathcal{S}_0$  kompatibel, falls die folgenden zwei Punkte erfüllt sind:*

1. Für jedes  $\mathbf{a} \in \mathcal{S}_0$  erfasst  $G\mathbf{a}$  den Ursprung.

2. Falls  $\mathcal{A} \subset \mathcal{S}'$  den Ursprung erfasst, erfasst  $g\mathcal{A}$  den Ursprung für jedes  $g \in G$ .

*Beweis nach [21].* Es handelt sich um einen Induktionsbeweis. Die Induktionsvoraussetzung für  $t = 0$  folgt aus Satz 2.8: Für  $t = 0$  haben wir mit  $G\mathbf{a}_i$ ,  $i \in [n]$ , genau  $n = p + 1$  Teilmengen des  $\mathbb{R}^p$  sowie  $\mathbf{0} \in \bigcap_{i \in [p+1]} \text{conv}(G\mathbf{a}_i)$ , womit Satz 2.8 die gewünschte Aussage impliziert.

Es sei jetzt  $t \in \mathbb{N}$  und das Lemma sei zwar für  $t - 1$  erfüllt, für  $t$  aber nicht mehr. Wir wollen in dieser Situation einen Widerspruch herleiten, sodass das Lemma also für  $t$  erfüllt sein muss, wenn es für  $t - 1$  schon erfüllt ist.

Es seien  $g_1, g_2, \dots, g_r$  die  $r \leq p$  Elemente der Gruppe  $G$ . Für einen Vektor  $\boldsymbol{\alpha} = (g_1, g_2, \dots, g_n)$  mit Elementen in  $G$  schreiben wir  $\boldsymbol{\alpha} \cdot \mathcal{S}_0 = \{g_i \mathbf{a}_i\}_{i \in [n]}$ . Nach Annahme gibt es für beliebiges  $\boldsymbol{\alpha}$  eine  $t$ -elementige Teilmenge  $\mathcal{C} \subset (\boldsymbol{\alpha} \cdot \mathcal{S}_0)$  mit der Eigenschaft, dass  $(\boldsymbol{\alpha} \cdot \mathcal{S}_0) \setminus \mathcal{C}$  den Ursprung  $\mathbf{0} \in \mathbb{R}^p$  nicht erfasst:  $\mathbf{0} \notin \text{conv}((\boldsymbol{\alpha} \cdot \mathcal{S}_0) \setminus \mathcal{C})$ . Für jedes  $\boldsymbol{\alpha}$  definieren wir nun (siehe auch Bemerkung 2.5):

$$P(\boldsymbol{\alpha}) = \max_{|\mathcal{C}|=t} \text{dist}(\mathbf{0}, \text{conv}((\boldsymbol{\alpha} \cdot \mathcal{S}_0) \setminus \mathcal{C})). \quad (2.19)$$

Nach Annahme ist  $P(\boldsymbol{\alpha}) > 0$  für jedes  $\boldsymbol{\alpha}$  erfüllt. Da es nur endlich viele verschiedene  $\boldsymbol{\alpha}$  gibt, gilt  $P(\boldsymbol{\alpha}_0) > 0$  auch für einen Minimierer  $\boldsymbol{\alpha}_0$  von  $P$ .

Es sei nun  $\boldsymbol{\alpha}_0$  ein Minimierer von  $P$  und  $\mathcal{C}_0 \subset (\boldsymbol{\alpha}_0 \cdot \mathcal{S}_0)$  eine zugehörige  $t$ -elementige Menge mit  $P(\boldsymbol{\alpha}_0) = \text{dist}(\mathbf{0}, \text{conv}((\boldsymbol{\alpha}_0 \cdot \mathcal{S}_0) \setminus \mathcal{C}_0))$ . Es sei  $\mathbf{p}_0 \in \text{conv}((\boldsymbol{\alpha}_0 \cdot \mathcal{S}_0) \setminus \mathcal{C}_0)$  ein Punkt, der  $\mathbf{0}$  am nächsten ist. Mit Bemerkung 2.5 gibt es eine Teilmenge  $\mathcal{A} \subset \text{conv}((\boldsymbol{\alpha}_0 \cdot \mathcal{S}_0) \setminus \mathcal{C}_0)$  aus höchstens  $p$  Punkten, sodass  $\mathbf{p}_0 \in \text{relint}(\mathcal{A})$  erfüllt ist. Wir schreiben  $\mathcal{B} = (\boldsymbol{\alpha}_0 \cdot \mathcal{S}_0) \setminus \mathcal{A}$  und wählen eine Hyperebene, die  $\mathcal{A}$  auf einer Seite enthält, aber  $\mathbf{0}$  im offenen Halbraum  $h^-$  auf der anderen Seite enthält.

Die Gruppe  $G$  ist mit  $\mathcal{B}$  kompatibel: Punkt 2 ist unverändert erfüllt und Punkt 1 gilt, weil jedes  $\mathbf{b} \in \mathcal{B}$  ein Element der Form  $\mathbf{b} = g\mathbf{a}$  mit  $g \in G$  und  $\mathbf{a} \in \mathcal{S}_0$  ist, sodass wegen  $Gg = G$  weiterhin  $G\mathbf{b}$  den Ursprung erfasst. Die Menge  $\mathcal{B}$  hat außerdem mindestens  $n - p = pt + 1$  Elemente. Wir können also das Lemma anwenden (wir sind hier im Fall  $t - 1$  und nutzen die Induktionsvoraussetzung) und finden einen Vektor  $\boldsymbol{\beta}$  mit  $|\mathcal{B}|$  Elementen in  $G$ , sodass  $\boldsymbol{\beta} \cdot \mathcal{B}$  den Ursprung mit Toleranz  $t - 1$  erfasst. Da  $G\mathbf{b}$  den Ursprung für jedes  $\mathbf{b} \in \mathcal{B}$  erfasst, gibt es zu jedem  $\mathbf{b} \in \mathcal{B}$  ein  $g_i$  mit  $g_i \mathbf{b} \in h^-$ . Wir ordnen nun jedes  $\mathbf{b}$  einem  $g_i$  so zu, dass  $g_i \mathbf{b} \in h^-$  gilt. Da es  $r \leq p$  Gruppenelemente  $g_i$  gibt, teilen wir hier die mindestens  $pt + 1$  Elemente aus  $\mathcal{B}$  auf die  $r \leq p$  Gruppenelemente auf. Es folgt nun, dass es ein Gruppenelement gibt, dem mindestens  $t + 1$  Elemente aus  $\mathcal{B}$  zugeordnet werden. Deshalb finden wir ein  $g \in G$ , sodass mindestens  $t + 1$  Punkte aus  $(g\boldsymbol{\beta}) \cdot \mathcal{B}$  in  $h^-$  liegen.

Es sei nun  $\boldsymbol{\alpha}_1 \in G^n$  der Vektor, der entsteht, indem jene Elemente in  $\boldsymbol{\alpha}_0 \in G^n$ , die zur Menge  $\mathcal{B}$  gehören, durch die Elemente  $g\boldsymbol{\beta}$  ersetzt werden. Das führt nun wie folgt zum Widerspruch  $P(\boldsymbol{\alpha}_1) < P(\boldsymbol{\alpha}_0)$ . Es sei  $\mathcal{C}' \subset (\boldsymbol{\alpha}_1 \cdot \mathcal{S}_0)$  eine  $t$ -elementige Menge. Falls darin höchstens  $t - 1$  Elemente aus  $(g\boldsymbol{\beta}) \cdot \mathcal{B}$  enthalten sind, erfasst  $(\boldsymbol{\alpha}_1 \cdot \mathcal{S}_0) \setminus \mathcal{C}'$  den Ursprung (weil  $(g\boldsymbol{\beta}) \cdot \mathcal{B}$  schon den Ursprung erfasst und zwar mit Toleranz  $t - 1$ ) und wir haben  $P(\boldsymbol{\alpha}_1) = 0 < P(\boldsymbol{\alpha}_0)$ . Andernfalls gilt  $\mathcal{C}' \subset ((g\boldsymbol{\beta}) \cdot \mathcal{B})$ . Es gelten  $|\mathcal{C}'| \leq t$  und  $|h^- \cap ((g\boldsymbol{\beta}) \cdot \mathcal{B})| \geq t + 1$ , weshalb es einen Punkt  $\mathbf{x} \in h^- \cap ((g\boldsymbol{\beta}) \cdot \mathcal{B})$  gibt, der nicht

in  $\mathcal{C}'$  enthalten ist. Dann ist aber  $\text{conv}(\mathcal{A} \cup \{\mathbf{x}\})$  näher am Ursprung als  $\text{conv}(\mathcal{A})$ , also  $P(\boldsymbol{\alpha}_1) < P(\boldsymbol{\alpha}_0)$ .  $\square$

*Beweis von Satz 2.6 nach [21].* Es seien  $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{m-1}$  die Ecken des um  $\mathbf{0} \in \mathbb{R}^{m-1}$  zentrierten Einheitssimplex in  $\mathbb{R}^{m-1}$ . Da alle Ecken gleich weit voneinander entfernt sind, gilt  $\sum_{i=0}^{m-1} \alpha_i \mathbf{u}_i = \mathbf{0}$  genau dann, wenn  $\alpha_0 = \alpha_1 = \dots = \alpha_{m-1}$ . Wir erweitern jeden Punkt  $\mathbf{x}_j \in \mathcal{S}$  über  $\mathbf{b}_j = (\mathbf{x}_j^T, 1)^T \in \mathbb{R}^{d+1}$  zu einem  $(d+1)$ -dimensionalen Vektor,  $j \in [n]$ .

Das prinzipielle Vorgehen im Beweis ist nun das folgende. Wir definieren zwei Mengen  $\mathcal{S}_0, \mathcal{S}'$  und eine Gruppenwirkung  $G$  und zeigen, dass diese die Bedingungen von Lemma 2.9 erfüllen. Durch Anwendung von Lemma 2.9 finden wir dann eine Zerlegung von  $\mathcal{S}$  in  $m$  Teilmengen, für welche wir schließlich zeigen, dass es sich um die gesuchte  $t$ -tolerante  $m$ -Tverberg-Zerlegung handelt.

Die zwei Mengen sind:

$$\mathcal{S}' = \{\mathbf{u}_i \otimes \mathbf{b}_j \mid i = 0, 1, \dots, m-1, j \in [n]\} \subset \mathbb{R}^p \quad (2.20)$$

$$\mathcal{S}_0 = \{\mathbf{u}_0 \otimes \mathbf{b}_j\}_{j \in [n]} \subset \mathbb{R}^p \quad (2.21)$$

mit  $p = (m-1)(d+1)$ . Es gilt dann  $|\mathcal{S}_0| = n = p(t+1) + 1$  wie es in Lemma 2.9 vorausgesetzt wird. Wir schreiben  $\mathbb{Z}_m$  für den Restklassenring modulo  $m$  und definieren eine Gruppenwirkung über

$$\mathbb{Z}_m \times \mathcal{S}' \rightarrow \mathcal{S}', (g, \mathbf{u}_i \otimes \mathbf{b}_j) \mapsto g(\mathbf{u}_i \otimes \mathbf{b}_j) = \mathbf{u}_{(i+g) \bmod m} \otimes \mathbf{b}_j. \quad (2.22)$$

Die Vektoren  $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{m-1}$  erfassen den Ursprung  $\mathbf{0} \in \mathbb{R}^{m-1}$ . Deshalb erfassen auch die Vektoren  $\mathbb{Z}_m \mathbf{a}$  den Ursprung  $\mathbf{0} \in \mathbb{R}^p$  für beliebiges  $\mathbf{a} \in \mathcal{S}_0$  (siehe auch Bemerkung 2.7). Damit ist Punkt 1 von Lemma 2.9 erfüllt.

Es sei  $\mathcal{A} \subset \mathcal{S}'$  eine Teilmenge, die  $\mathbf{0} \in \mathbb{R}^p$  erfasst. Da alle Vektoren  $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{m-1}$  gleich weit voneinander entfernt und um  $\mathbf{0}$  zentriert sind, führen jene Koeffizienten, die zu einer Darstellung von  $\mathbf{0} \in \mathbb{R}^p$  als Konvexkombination der Elemente in  $\mathcal{A}$  führen, zu einer Darstellung von  $\mathbf{0} \in \mathbb{R}^p$  als Konvexkombination der Elemente in  $g\mathcal{A}$  für ein beliebiges  $g \in \mathbb{Z}_m$ . Damit ist auch Punkt 2 von Lemma 2.9 erfüllt.

Wir können also Lemma 2.9 anwenden und erhalten  $g_1, g_2, \dots, g_n \in \mathbb{Z}_m$ , sodass  $\{g_j(\mathbf{u}_0 \otimes \mathbf{b}_j)\}_{j \in [n]}$  den Ursprung mit Toleranz  $t$  erfasst. Wir zeigen im Rest des Beweises, dass die Mengen  $\mathcal{S}_i = \{\mathbf{x}_j \in \mathcal{S} \mid g_j = i\}$ ,  $i \in \{0, 1, \dots, m-1\}$ , die gesuchte  $t$ -tolerante  $m$ -Tverberg-Zerlegung bilden. Dafür seien  $\mathcal{H}_i = \{j \mid g_j = i\}$  die Indizes der Mengen  $\mathcal{S}_i$  und es sei  $\mathcal{C} \subset [n]$  eine beliebige  $t$ -elementige Menge. Wir betrachten die Mengen  $\mathcal{S}'_i = \{\mathbf{x}_j \mid j \in \mathcal{H}_i \setminus \mathcal{C}\}$  und wollen  $\bigcap_{i=0}^{m-1} \text{conv}(\mathcal{S}'_i) \neq \emptyset$  zeigen.

Wegen der  $t$ -Toleranz in Lemma 2.9 erfasst die Menge  $\{\mathbf{u}_{g_j} \otimes \mathbf{b}_j \mid j \in [n] \setminus \mathcal{C}\}$  den Ursprung und es gibt also eine Konvexkombination der Elemente, sodass

$$\sum_{j \in [n] \setminus \mathcal{C}} \alpha_j (\mathbf{u}_{g_j} \otimes \mathbf{b}_j) = \mathbf{0} \quad (2.23)$$

gilt. Wegen der Definition  $\mathbf{b}_j = (\mathbf{x}_j^T, 1)^T$  lesen wir aus (2.23)

$$\sum_{j \in [n] \setminus \mathcal{C}} \alpha_j \mathbf{u}_{g_j} = \mathbf{0} \quad (2.24)$$

ab. Wir haben also eine Konvexkombination des Ursprungs in  $\mathbb{R}^{m-1}$  vorliegen. Mit dem Kommentar im ersten Absatz dieses Beweises schließen wir auf

$$\sum_{j \in \mathcal{H}_0 \setminus \mathcal{C}} \alpha_j = \sum_{j \in \mathcal{H}_1 \setminus \mathcal{C}} \alpha_j = \dots = \sum_{j \in \mathcal{H}_{m-1} \setminus \mathcal{C}} \alpha_j = T > 0. \quad (2.25)$$

Wegen der Definition  $\mathbf{b}_j = (\mathbf{x}_j^T, 1)^T$  lesen wir aus (2.23) außerdem für jedes  $1 \leq r \leq d$

$$\sum_{j \in [n] \setminus \mathcal{C}} \alpha_j (\mathbf{x}_j)_r \mathbf{u}_{g_j} = \mathbf{0} \quad (2.26)$$

ab, wobei mit dem Index  $r$  das  $r$ te Element eines Vektors ausgewählt wird. Analog folgt jetzt

$$\sum_{j \in \mathcal{H}_0 \setminus \mathcal{C}} \alpha_j (\mathbf{x}_j)_r = \sum_{j \in \mathcal{H}_1 \setminus \mathcal{C}} \alpha_j (\mathbf{x}_j)_r = \dots = \sum_{j \in \mathcal{H}_{m-1} \setminus \mathcal{C}} \alpha_j (\mathbf{x}_j)_r. \quad (2.27)$$

Wir können also für jedes  $j \in [n]$  die Koeffizienten  $\beta_j = \frac{\alpha_j}{T}$  definieren, für welche einerseits  $\sum_{j \in \mathcal{H}_i \setminus \mathcal{C}} \beta_j = 1$  für jedes  $i \in \{0, 1, \dots, m-1\}$  und andererseits

$$\sum_{j \in \mathcal{H}_0 \setminus \mathcal{C}} \beta_j \mathbf{x}_j = \sum_{j \in \mathcal{H}_1 \setminus \mathcal{C}} \beta_j \mathbf{x}_j = \dots = \sum_{j \in \mathcal{H}_{m-1} \setminus \mathcal{C}} \beta_j \mathbf{x}_j = \hat{\mathbf{x}} \quad (2.28)$$

erfüllt ist. Damit ist  $\hat{\mathbf{x}}$  ein Tverberg-Punkt:  $\hat{\mathbf{x}} \in \bigcap_{i=0}^{m-1} \text{conv}(\mathcal{S}'_i)$ . Der Schnitt der konvexen Hüllen ist also nichtleer.  $\square$

Es handelt sich hier um deterministische Aussagen, die jedoch nicht erklären, wie eine gegebene Menge  $\mathcal{S}$  entsprechend eingefärbt werden kann, um eine  $m$ -Tverberg-Zerlegung zu erhalten. In dieser Arbeit betrachten wir zwei Modelle zur zufälligen Erzeugung von  $m$ -farbigen Punktmengen und erhalten Aussagen über die Wahrscheinlichkeit, dass es sich bei diesen zufällig generierten Mengen um  $m$ -Tverberg-Zerlegungen handelt.

## 2.4 Zwei Modelle zufälliger Punktmengen

Wir führen zwei Modelle zur Erzeugung zufälliger Punktmengen, die in  $m$  Farben eingefärbt werden, ein. In beiden Modellen werden alle Punkte gemäß einer stetigen Wahrscheinlichkeitsverteilung erzeugt. Im ersten Modell wird garantiert, dass es gleich viele Punkte jeder Farbe gibt:

**Definition 2.10** (Modell der zufälligen Gleichverteilung, random equipartition model). *Es seien  $m, n \in \mathbb{N}$  und  $D$  eine stetige Wahrscheinlichkeitsverteilung auf  $\mathbb{R}^d$ . Dann bezeichnet  $\mathcal{E}_{m,n,D}$  eine Menge bestehend aus  $mn$  Punkten, die in  $m$  Farben mit jeweils genau  $n$  Punkten eingefärbt sind und die unabhängig voneinander gemäß  $D$  realisiert wurden.*

Im zweiten Modell wird jeder Punkt mit der gleichen Wahrscheinlichkeit  $1/m$  einer der  $m$  Farben zugeordnet:

**Definition 2.11** (Modell der zufälligen Zuordnung, random allocation model). *Es seien  $m, k \in \mathbb{N}$  und  $D$  eine stetige Wahrscheinlichkeitsverteilung auf  $\mathbb{R}^d$ . Dann bezeichnet  $\mathcal{R}_{m,k,D}$  eine Menge bestehend aus  $k$  Punkten mit der Eigenschaft, dass jeder Punkt unabhängig gemäß  $D$  realisiert und mit Wahrscheinlichkeit  $1/m$  in eine von  $m$  Farben eingefärbt wurde.*

Für beide Modelle stellen wir in Abschnitt 3 Aussagen über die Wahrscheinlichkeit, dass es sich bei einer gemäß einem der beiden Modelle zufällig realisierten Mengen um eine  $m$ -Tverberg-Zerlegung (gegebenenfalls mit Toleranz) handelt, vor. Genauer stellen wir untere und obere Schranken an Wahrscheinlichkeiten der Form  $\mathbb{P}(\mathcal{E}_{m,n,D}$  ist Tverberg) oder  $\mathbb{P}(\mathcal{R}_{m,k,D}$  ist Tverberg) vor.

Wir können das Modell  $\mathcal{R}_{m,k,D}$  mit dem Modell  $\mathcal{E}_{m,n,D}$  in Verbindung bringen, indem wir im Modell  $\mathcal{R}_{m,k,D}$  so viele Punkte erzeugen, dass es mindestens  $n$  Punkte jeder Farbe gibt. Wenn wir die Wahrscheinlichkeit für dieses Ereignis kennen und die Wahrscheinlichkeit, dass  $\mathcal{E}_{m,n,D}$  Tverberg ist, erhalten wir damit eine untere Schranke an die Wahrscheinlichkeit, dass  $\mathcal{R}_{m,k,D}$  Tverberg ist:

**Beobachtung 2.12** (Observation 2.3). *Es seien  $k, m, n \in \mathbb{N}$  und  $D$  eine stetige Wahrscheinlichkeitsverteilung auf  $\mathbb{R}^d$ . Es sei  $\mathcal{R}_{m,k,D}$  eine zufällige Zuordnung (Definition 2.11) von  $k$  Punkten in  $m$  Farben. Weiter sei  $\mathcal{E}_{m,n,D}$  eine zufällige Gleichverteilung (Definition 2.10) von  $nm$  Punkten in  $m$  Farben. Wir schreiben  $P$  für die Wahrscheinlichkeit, dass  $\mathcal{R}_{m,k,D}$  mindestens  $n$  Punkte jeder Farbe hat. Dann gilt*

$$\mathbb{P}(\mathcal{R}_{m,k,D} \text{ ist Tverberg mit Toleranz } t) \geq P \cdot \mathbb{P}(\mathcal{E}_{m,n,D} \text{ ist Tverberg mit Toleranz } t).$$

In manchen Beweisen schätzen wir zunächst die Wahrscheinlichkeit, dass  $\mathcal{E}_{m,n,D}$  Tverberg ist, ab und nutzen dann Beobachtung 2.12, um eine Aussagen über die Wahrscheinlichkeit, dass  $\mathcal{R}_{m,k,D}$  Tverberg ist, zu treffen.

## 2.5 Für Beweise wichtige Aussagen

Wir führen zum Abschluss von Abschnitt 2 Aussagen auf, die jeweils in mindestens einem der nachfolgend dargestellten Beweisen eine Rolle spielen.

In einem Beweis stellt sich die Frage, wie wahrscheinlich es ist, dass der Ursprung in der konvexen Hülle von  $n$  zufällig auf einer Kugeloberfläche verteilten Punkten liegt. Satz 2.13 gibt eine Wahrscheinlichkeit an, dass das nicht passiert.

**Satz 2.13** (Wendel). *Es werden  $n$  Punkte im  $\mathbb{R}^d$  gemäß Gleichverteilung auf der Oberfläche einer  $d$ -dimensionalen Kugel um den Ursprung verteilt. Die Wahrscheinlichkeit, dass alle  $n$  Punkte auf einer Halbkugel liegen, ist*

$$p_{d,n} = 2^{-n+1} \sum_{k=0}^{d-1} \binom{n-1}{k}. \quad (2.29)$$

*Beweis nach [25].* Im Fall  $d = 1$  gilt

$$p_{1,n} = 2^{-n+1}, \quad (2.30)$$

was einerseits der Wahrscheinlichkeit entspricht, bei  $n$  Würfeln einer fairen Münze stets Kopf oder stets Zahl zu beobachten, und andererseits wegen  $\binom{n-1}{0} = 1$  die rechte Seite von (2.29) für  $d = 1$  ist. Damit ist (2.29) für  $d = 1$  bestätigt.

Für  $d > 1$  betrachten wir zunächst den Fall  $d \geq n$ , wo

$$p_{d,n} = 1 \tag{2.31}$$

erfüllt ist. Hierfür nutzen wir in (2.29) die Gleichung

$$\sum_{k=0}^{n-1} \binom{n-1}{k} = 2^{n-1} \tag{2.32}$$

(siehe beispielsweise [19]) und  $\binom{n-1}{k} = 0$  für  $k > n-1$ . Um zu sehen, wieso die Wahrscheinlichkeit 1 ist, seien  $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^d$  die  $n \leq d$  zufälligen Punkte. Wegen der Gleichverteilung sind diese mit Wahrscheinlichkeit 1 linear unabhängig. Wir betrachten dann  $\mathbf{y} = \sum_{i \in [n]} \mathbf{x}_i$  und erhalten einen Vektor, der  $\mathbf{y}^T \mathbf{x}_i > 0$  für alle  $i \in [n]$  erfüllt, sodass alle  $\mathbf{x}_i$  mit Wahrscheinlichkeit 1 auf einer Halbkugel liegen.

Im Folgenden soll noch

$$p_{d,n} = \frac{1}{2} (p_{d,n-1} + p_{d-1,n-1}) \tag{2.33}$$

für  $d > 1$ ,  $n > 2$  und  $d < n$  gezeigt werden. Da die Formel (2.29) die Beziehung (2.33) erfüllt (wie wir in der nächsten Gleichungskette nachrechnen), ist der Beweis dann abgeschlossen. Es gilt, wenn wir mit (2.29) in (2.33) einsetzen:

$$\frac{1}{2} (p_{d,n-1} + p_{d-1,n-1}) \tag{2.34}$$

$$= \frac{1}{2} \left( 2^{-(n-1)+1} \sum_{k=0}^{d-1} \binom{(n-1)-1}{k} + 2^{-(n-1)+1} \sum_{k=0}^{(d-1)-1} \binom{(n-1)-1}{k} \right) \tag{2.35}$$

$$= 2^{-n+1} \left( \sum_{k=0}^{d-1} \binom{n-2}{k} + \sum_{k=0}^{d-2} \binom{n-2}{k} \right) \tag{2.36}$$

$$= 2^{-n+1} \left( \sum_{k=0}^{d-1} \binom{n-2}{k} + \sum_{k=1}^{d-1} \binom{n-2}{k-1} \right) \tag{2.37}$$

$$= 2^{-n+1} \left( \underbrace{\binom{n-2}{0}}_{=1} + \sum_{k=1}^{d-1} \left( \binom{n-2}{k} + \binom{n-2}{k-1} \right) \right) \tag{2.38}$$

$$= 2^{-n+1} \left( \underbrace{\binom{n-1}{0}}_{=1} + \sum_{k=1}^{d-1} \binom{n-1}{k} \right) = 2^{-n+1} \sum_{k=0}^{d-1} \binom{n-1}{k} = p_{d,n}. \tag{2.39}$$

Hier haben wir

$$\binom{n-2}{k} + \binom{n-2}{k-1} = \binom{n-1}{k} \quad (2.40)$$

ausgenutzt, siehe beispielsweise [19].

Es seien  $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^d$  die  $n$  zufällig verteilten Punkte und  $\mathbf{y} \in \mathbb{R}^d$  ein Vektor, der zu keinem Punkt orthogonal ist. Wir betrachten den  $n$ -dimensionalen Vektor  $\mathbf{s}_\mathbf{y} = (\text{sign}(\mathbf{y}^\top \mathbf{x}_1), \text{sign}(\mathbf{y}^\top \mathbf{x}_2), \dots, \text{sign}(\mathbf{y}^\top \mathbf{x}_n))$  der Vorzeichen der Innenprodukte zwischen  $\mathbf{y}$  und allen Punkten  $\mathbf{x}_i$ . Es handelt sich bei  $\mathbf{s}_\mathbf{y}$  dann um einen zufälligen Punkt in  $\mathcal{S} = \{-1, +1\}^n$ . Wir sagen ein bestimmtes  $\mathbf{s} \in \mathcal{S}$  *kommt vor*, falls ein  $\mathbf{y}$  mit  $\mathbf{s}_\mathbf{y} = \mathbf{s}$  existiert. Wir schreiben  $A_\mathbf{s}$  für dieses Ereignis und definieren  $I_\mathbf{s}$  als die Indikatorfunktion, die den Wert 1 annimmt, falls  $A_\mathbf{s}$  eintritt, und die andernfalls den Wert 0 annimmt. Wir können allgemein den  $j$ ten Eintrag von  $\mathbf{s}$  ändern, indem wir  $\mathbf{x}_j$  durch  $-\mathbf{x}_j$  ersetzen. Da  $\mathbf{x}_j$  und  $-\mathbf{x}_j$  mit gleicher Wahrscheinlichkeit auftreten, sind alle  $A_\mathbf{s}$  gleich wahrscheinlich. Mit  $\mathbf{s}_0 = (1, 1, \dots, 1)$  können wir  $p_{d,n} = \mathbb{P}(A_{\mathbf{s}_0})$  schreiben (alle  $\mathbf{x}_i$  zeigen in Richtung  $\mathbf{y}$  und sind damit auf einer Halbkugel).

Wir können nun über  $Q_{d,n} = \sum_{\mathbf{s} \in \mathcal{S}} I_\mathbf{s}$  zählen, wie viele unterschiedliche  $\mathbf{s}$  vorkommen und davon den Erwartungswert bestimmen:

$$\mathbb{E}(Q_{d,n}) = \sum_{\mathbf{s} \in \mathcal{S}} \mathbb{E}(I_\mathbf{s}) = \sum_{\mathbf{s} \in \mathcal{S}} \mathbb{P}(A_\mathbf{s}) = 2^n p_{d,n}. \quad (2.41)$$

Wir wollen im weiteren Verlauf zeigen, dass  $Q_{d,n}$  nicht von den  $\mathbf{x}_1, \dots, \mathbf{x}_n$  abhängt. Wir nutzen im restlichen Beweis, dass wegen der Gleichverteilung jeweils  $d$  der  $n$  Punkte mit Wahrscheinlichkeit 1 linear unabhängig sind. Es sei  $h_i$  die über  $\mathbf{x}_i$  als Normalenvektor definierte Hyperebene. Das Hyperebenenarrangement zu allen  $n$  Vektoren zerlegt dann  $\mathbb{R}^d$  in  $Q_{d,n}$  zusammenhängende Bereiche. Es gibt  $Q_{d,n}$  Bereiche, da jeder Vektor  $\mathbf{y}$  in einem fest gewählten Bereich zum gleichen  $\mathbf{s}_\mathbf{y}$  führt.

Es gilt jetzt die Anzahl der Bereiche zu zählen. Dafür entfernen wir hypothetisch eine beliebige Hyperebene, beispielsweise  $h_n$ . Die verbleibenden  $n-1$  Hyperebenen führen per Definition zu  $Q_{d,n-1}$  Bereichen. Es gibt dabei genau zwei Arten von Bereichen: (i) jene  $Q^{(i)}$  Bereiche, durch welche zuvor  $h_n$  verlief, und (ii) jene  $Q^{(ii)}$  Bereiche, durch welche  $h_n$  zuvor nicht verlief. Es gilt  $Q_{d,n-1} = Q^{(i)} + Q^{(ii)}$ . Fügen wir die Hyperebene  $h_n$  wieder hinzu, bleiben die Bereiche der Art (ii) unverändert und die Bereiche der Art (i) werden in zwei Bereiche aufgeteilt. Das führt zu  $Q_{d,n} = 2Q^{(i)} + Q^{(ii)}$ . Wir schließen insgesamt auf

$$Q_{d,n} = Q_{d,n-1} + Q^{(i)}. \quad (2.42)$$

Wir zeigen unten noch  $Q^{(i)} = Q_{d-1,n-1}$ . Wenn wir das in (2.42) einsetzen und anschließend noch (2.41) beachten, erhalten wir

$$2^n p_{d,n} = 2^{n-1} p_{d,n-1} + 2^{n-1} p_{d-1,n-1}, \quad (2.43)$$

was dann (2.33) bestätigt und den Beweis abschließt.

Zu  $Q^{(i)}$ : Für jedes  $j \in [n-1]$  handelt es sich bei  $h_j \cap h_n$  um eine Hyperebene im  $(d-1)$ -dimensionalen Raum  $h_n$ , sodass jeweils  $d-1$  der die Hyperbenen  $h_j \cap h_n$

definierenden Normalenvektoren linear unabhängig sind. Wir sind also in einer ähnlichen Situation wie zuvor – mit dem Unterschied, dass wir  $n - 1$  Hyperebenen  $\{h_j \cap h_n\}_{j \in [n-1]}$  in der  $(d - 1)$ -dimensionalen Hyperebene  $h_n$  anstelle von  $n$  Hyperebenen  $\{h_j\}_{j \in [n]}$  im  $d$ -dimensionalen  $\mathbb{R}^d$  betrachten. Die Anzahl der Bereiche in  $h_n$  ist daher  $Q_{d-1, n-1}$ . Diese  $Q_{d-1, n-1}$  Bereiche in  $h_n$  sind nun aber die Schnitte der  $Q^{(i)}$  Bereiche der Art (i) mit  $h_n$  und wir erhalten  $Q^{(i)} = Q_{d-1, n-1}$ .  $\square$

**Korollar 2.14** (zu Satz 2.13). *Die Wahrscheinlichkeit, dass  $n = 2d$  Punkte auf einer Halbkugel liegen, ist*

$$p_{d, 2d} = 2^{-2d+1} \sum_{k=0}^{d-1} \binom{2d-1}{k} = \frac{1}{2}. \quad (2.44)$$

*Beweis.* Unter Ausnutzung der Gleichung

$$\binom{2d-1}{k} = \binom{2d-1}{2d-1-k} \quad (2.45)$$

(siehe beispielsweise [19]) stellen wir

$$\sum_{k=0}^{2d-1} \binom{2d-1}{k} = \sum_{k=0}^{d-1} \binom{2d-1}{k} + \sum_{k=d}^{2d-1} \binom{2d-1}{k} \quad (2.46)$$

$$= \sum_{k=0}^{d-1} \binom{2d-1}{k} + \sum_{k=d}^{2d-1} \binom{2d-1}{2d-1-k} \quad (2.47)$$

$$= \sum_{k=0}^{d-1} \binom{2d-1}{k} + \sum_{k=0}^{d-1} \binom{2d-1}{k} = 2 \sum_{k=0}^{d-1} \binom{2d-1}{k} \quad (2.48)$$

fest. Zusammen mit der bekannten Beziehung

$$\sum_{k=0}^{2d-1} \binom{2d-1}{k} = 2^{2d-1} \quad (2.49)$$

(vergleiche (2.32)) schließen wir nun auf (2.44).  $\square$

Wir betrachten in diesem Kontext dann das Ereignis, dass alle Punkte auf einer Halbkugel liegen, als Misserfolg und wir betrachten das Gegenereignis als Erfolg. Falls die Wahrscheinlichkeit, dass ein Erfolg eintritt,  $p$  ist und wir  $N$  unabhängige Zufallsexperimente durchführen, können wir mit der nachfolgenden Bemerkung 2.15 die Wahrscheinlichkeit berechnen, dass mindestens oder höchstens  $k \leq N$  Erfolge beobachtet werden.

**Bemerkung 2.15.** *Es seien  $X_1, \dots, X_N$  unabhängige Zufallsvariablen, die jeweils mit Wahrscheinlichkeit  $p \in [0, 1]$  den Wert 1 („Erfolg“) und mit Wahrscheinlichkeit  $1 - p$  den Wert 0 („Misserfolg“) annehmen. Die Wahrscheinlichkeit, dass wir bei genau  $k \in \{0, 1, \dots, N\}$  dieser Zufallsvariablen den Wert 1 beobachten ist [19]*

$$\mathbb{P} \left( \sum_{i=1}^N X_i = k \right) = \binom{N}{k} p^k (1-p)^{N-k}. \quad (2.50)$$

Wir benötigen außerdem ein Hilfsmittel, um die Wahrscheinlichkeit abzuschätzen, dass mehrere Ereignisse gleichzeitig eintreten. Zu diesem Zweck ist es vorteilhaft, die Wahrscheinlichkeit abschätzen zu können, dass mindestens eines der Ereignisse eintritt:

**Bemerkung 2.16** (union bound). *Es seien  $E_1, \dots, E_m$  Ereignisse, die mit den Wahrscheinlichkeiten  $\mathbb{P}(E_1), \dots, \mathbb{P}(E_m)$  eintreten. Dann gilt [19]*

$$\mathbb{P} \left( \bigcup_{i \in [m]} E_i \right) \leq \sum_{i \in [m]} \mathbb{P}(E_i). \quad (2.51)$$

Zuletzt zitieren wir noch das Gesetz der großen Zahlen, welches in manchen Beweisen zum Einsatz kommt:

**Bemerkung 2.17** (Gesetz der großen Zahlen, law of large numbers). *Für jedes  $N \in \mathbb{N}$  seien  $X_1, \dots, X_N$  unabhängige und identisch verteilte Zufallsvariablen, die jeweils den Mittelwert  $\mu$  und die Varianz  $\sigma$  haben. Dann gilt für jedes beliebige  $\varepsilon > 0$  [19]*

$$\lim_{N \rightarrow \infty} \mathbb{P} \left( \left| \frac{1}{N} \sum_{i=1}^N X_i - \mu \right| < \varepsilon \right) = 1. \quad (2.52)$$



deterministische Version	stochastische Version
Radon	Satz 3.1
Tverberg	Satz 3.4, Satz 3.5
Radon mit Toleranz	Satz 3.7
Tverberg mit Toleranz	Satz 3.2, Satz 3.6, Korollar 3.9

Tabelle 1: Überblick über die Ergebnisse in Abschnitt 3.

### 3 Stochastische Radon- und Tverberg-Sätze

Wir stellen in diesem Abschnitt stochastische Radon- und Tverberg-Sätze mit und ohne Toleranz vor. Tabelle 1 verschafft einen Überblick.

#### 3.1 In der Literatur bekannte Aussagen

Die zwei folgenden Sätze stellen stochastische Radon- und Tverberg-Aussagen dar, welche bereits in der Literatur bekannt sind. Satz 3.1 berechnet die Wahrscheinlichkeit, dass eine zufällige Zuordnung (Definition 2.11) zu einer Radon-Zerlegung führt, bemerkenswerterweise exakt. Die weiteren Sätze in Abschnitt 3 liefern untere oder obere Schranken an entsprechende Wahrscheinlichkeiten. Die Quelle für Satz 3.1 ist [6].

**Satz 3.1** (Cover). *Für eine stetige Wahrscheinlichkeitsverteilung  $D$  auf  $\mathbb{R}^d$  gilt*

$$\mathbb{P}(\mathcal{R}_{2,n,D} \text{ ist Radon}) = 1 - \frac{1}{2^{n-1}} \sum_{k=0}^{d-1} \binom{n-1}{k}. \quad (3.1)$$

*Im Speziellen gilt*

$$\mathbb{P}(\mathcal{R}_{2,2d,D} \text{ ist Radon}) = 1/2. \quad (3.2)$$

*Außerdem gilt für jedes  $\varepsilon > 0$  und jede Folge  $(D_i)_{i \in \mathbb{N}}$  von stetigen Wahrscheinlichkeitsverteilungen  $D_i$  auf  $\mathbb{R}^i$ :*

$$\lim_{i \rightarrow \infty} \mathbb{P}(\mathcal{R}_{2,(1+\varepsilon)2i,D_i} \text{ ist Radon}) = 1 \quad (3.3)$$

*und*

$$\lim_{i \rightarrow \infty} \mathbb{P}(\mathcal{R}_{2,(1-\varepsilon)2i,D_i} \text{ ist Radon}) = 0. \quad (3.4)$$

*Beweis von Gleichung (3.2).* Wir setzen  $n = 2d$  in (3.1) ein und erhalten

$$\mathbb{P}(\mathcal{R}_{2,2d,D} \text{ ist Radon}) = 1 - \frac{1}{2^{2d-1}} \sum_{k=0}^{d-1} \binom{2d-1}{k}, \quad (3.5)$$

was zusammen mit Korollar 2.14 zu (3.2) führt.  $\square$

Eine Verallgemeinerung dieses Ergebnisses ist der folgende Satz 3.2 aus [20]. Er ist hier zunächst zur Vollständigkeit aufgelistet und begegnet uns beim Thema Separierbarkeit (siehe Abschnitt 4.2) erneut. Wir diskutieren den Phasenübergang, der in Satz 3.1 von Gleichung (3.4) zu Gleichung (3.3) zu beobachten ist, in Abschnitt 3.2 nach Satz 3.4.

**Satz 3.2** (Soberón). *Es seien  $k, t, d, m \in \mathbb{N}$  und  $\varepsilon > 0$ . Eine zufällige Zuordnung (Definition 2.11) von  $k$  Punkten im  $\mathbb{R}^d$  in  $m$  Farben ist mit einer Wahrscheinlichkeit von mindestens  $1 - \varepsilon$  eine Tverberg-Zerlegung mit Toleranz  $t$ , falls*

$$t + 1 \leq k/m - \sqrt{\frac{1}{2} \left[ (d+1)(m-1)k \ln(km) + k \ln \left( \frac{1}{\varepsilon} \right) \right]} \quad (3.6)$$

erfüllt ist.

Bei den nachfolgend dargestellten Aussagen in Abschnitt 3.2 und in Abschnitt 3.3 handelt es sich um Radon- und Tverberg-Sätze, die erstmals in [10] bewiesen wurden.

### 3.2 Zufällige Gleichverteilung

Wir konzentrieren uns in diesem Abschnitt auf das Modell der zufälligen Gleichverteilung (Definition 2.10). In Abschnitt 3.3 liegt der Fokus dann auf dem Modell der zufälligen Zuordnung (Definition 2.11). Wir beginnen mit einer Abschätzung der Wahrscheinlichkeit, dass im Modell der zufälligen Gleichverteilung (Definition 2.10) eine Tverberg-Zerlegung (mit Toleranz  $t = 0$ ) entsteht. Im Beweis nutzen wir folgende Überlegung:

**Proposition 3.3.** *Es seien  $m, n \in \mathbb{N}$  und es sei  $D$  eine Wahrscheinlichkeitsverteilung auf  $\mathbb{R}^d$ , die um den Ursprung  $\mathbf{0} \in \mathbb{R}^d$  in dem Sinne ausgeglichen ist, dass jede Hyperebene durch  $\mathbf{0}$  die Verteilung  $D$  in zwei Mengen gleichen Maßes aufteilt. Weiter bezeichne  $\mathcal{E}_{m,n,D} = \bigcup_{i \in [m]} \mathcal{S}_i = \{\mathbf{x}_k\}_{k=1}^{mn}$  eine zufällige Gleichverteilung wie in Definition 2.10, wobei  $\mathcal{S}_i = \{\mathbf{x}_{(i-1)n+j}\}_{j=1}^n$  die  $n$  Punkte der Farbe  $i \in [m]$  enthalte und keiner der Punkte gleich  $\mathbf{0}$  sei. Es sei  $0 < r < \min_{k \in [mn]} \|\mathbf{x}_k\|$ . Wir projizieren jeden Punkt  $\mathbf{x}_k$ ,  $k \in [mn]$ , mittels  $\tilde{\mathbf{x}}_k = r\mathbf{x}_k/\|\mathbf{x}_k\|$  auf die Oberfläche der um  $\mathbf{0}$  zentrierten  $d$ -dimensionalen Kugel mit Radius  $r$  und schreiben  $\tilde{\mathcal{E}}_{m,n,D} = \bigcup_{i \in [m]} \tilde{\mathcal{S}}_i = \{\tilde{\mathbf{x}}_k\}_{k=1}^{mn}$  für die Menge der so entstehenden Punkte, wobei wieder  $\tilde{\mathcal{S}}_i$  nur Punkte der Farbe  $i \in [m]$  enthalte. Dann gilt:*

*Der Ursprung  $\mathbf{0} \in \mathbb{R}^d$  ist genau dann ein Tverberg-Punkt von  $\mathcal{E}_{m,n,D}$ , wenn  $\mathbf{0}$  ein Tverberg-Punkt von  $\tilde{\mathcal{E}}_{m,n,D}$  ist.*

*Beweis.* „ $\Rightarrow$ “: Wegen  $\mathbf{0} \in \bigcap_{i \in [m]} \text{conv}(\mathcal{S}_i)$  finden wir Koeffizienten  $\{\lambda_k\}_{k \in [mn]}$  mit  $0 \leq \lambda_k \leq 1$ ,  $k \in [mn]$ , und mit

$$\mathbf{0} = \sum_{j \in [n]} \lambda_{(i-1)n+j} \mathbf{x}_{(i-1)n+j} \quad \text{und} \quad \sum_{j \in [n]} \lambda_{(i-1)n+j} = 1 \quad (3.7)$$

für jedes  $i \in [m]$ . Wir definieren  $\alpha_i = \sum_{j \in [n]} \lambda_{(i-1)n+j} \|\mathbf{x}_{(i-1)n+j}\|$  für  $i \in [m]$  und

$$\tilde{\lambda}_{(i-1)n+j} = \lambda_{(i-1)n+j} \frac{\|\mathbf{x}_{(i-1)n+j}\|}{\alpha_i} \quad (3.8)$$

für  $i \in [m]$  und  $j \in [n]$ . Dann gilt

$$1 = \frac{\sum_{j \in [n]} \lambda_{(i-1)n+j} \|\mathbf{x}_{(i-1)n+j}\|}{\alpha_i} = \sum_{j \in [n]} \lambda_{(i-1)n+j} \frac{\|\mathbf{x}_{(i-1)n+j}\|}{\alpha_i} = \sum_{j \in [n]} \tilde{\lambda}_{(i-1)n+j}, \quad (3.9)$$

wobei außerdem  $0 \leq \tilde{\lambda}_k \leq 1$  für  $k \in [mn]$  erfüllt ist. Wir schreiben nun für  $i \in [m]$

$$\mathbf{0} = \sum_{j \in [n]} \lambda_{(i-1)n+j} \mathbf{x}_{(i-1)n+j} \quad (3.10)$$

$$= \sum_{j \in [n]} \lambda_{(i-1)n+j} \frac{\|\mathbf{x}_{(i-1)n+j}\|}{r} \frac{r}{\|\mathbf{x}_{(i-1)n+j}\|} \mathbf{x}_{(i-1)n+j} \quad (3.11)$$

$$= \frac{\alpha_i}{r} \sum_{j \in [n]} \lambda_{(i-1)n+j} \frac{\|\mathbf{x}_{(i-1)n+j}\|}{\alpha_i} \tilde{\mathbf{x}}_{(i-1)n+j} \quad (3.12)$$

$$= \frac{\alpha_i}{r} \sum_{j \in [n]} \tilde{\lambda}_{(i-1)n+j} \tilde{\mathbf{x}}_{(i-1)n+j}. \quad (3.13)$$

Nach einer Multiplikation mit  $\frac{r}{\alpha_i}$  auf beiden Seiten schließen wir auf  $\mathbf{0} \in \bigcap_{i \in [m]} \text{conv}(\tilde{\mathcal{S}}_i)$ .

„ $\Leftarrow$ “: Die Argumentation verläuft sehr ähnlich. Wegen  $\mathbf{0} \in \bigcap_{i \in [m]} \text{conv}(\tilde{\mathcal{S}}_i)$  finden wir Koeffizienten  $\{\tilde{\lambda}_k\}_{k \in [mn]}$  mit  $0 \leq \tilde{\lambda}_k \leq 1$ ,  $k \in [mn]$ , und mit

$$\mathbf{0} = \sum_{j \in [n]} \tilde{\lambda}_{(i-1)n+j} \tilde{\mathbf{x}}_{(i-1)n+j} \quad \text{und} \quad \sum_{j \in [n]} \tilde{\lambda}_{(i-1)n+j} = 1 \quad (3.14)$$

für jedes  $i \in [m]$ . Wir definieren  $\tilde{\alpha}_i = \sum_{j \in [n]} \frac{\tilde{\lambda}_{(i-1)n+j}}{\|\mathbf{x}_{(i-1)n+j}\|}$  für  $i \in [m]$  und

$$\lambda_{(i-1)n+j} = \frac{\tilde{\lambda}_{(i-1)n+j}}{\|\mathbf{x}_{(i-1)n+j}\| \tilde{\alpha}_i} \quad (3.15)$$

für  $i \in [m]$  und  $j \in [n]$ . Dann gilt

$$1 = \frac{\sum_{j \in [n]} \frac{\tilde{\lambda}_{(i-1)n+j}}{\|\mathbf{x}_{(i-1)n+j}\|}}{\tilde{\alpha}_i} = \sum_{j \in [n]} \frac{\tilde{\lambda}_{(i-1)n+j}}{\|\mathbf{x}_{(i-1)n+j}\| \tilde{\alpha}_i} = \sum_{j \in [n]} \lambda_{(i-1)n+j}, \quad (3.16)$$

wobei außerdem  $0 \leq \lambda_k \leq 1$  für  $k \in [mn]$  erfüllt ist. Wir schreiben nun für  $i \in [m]$

$$\mathbf{0} = \sum_{j \in [n]} \tilde{\lambda}_{(i-1)n+j} \tilde{\mathbf{x}}_{(i-1)n+j} \quad (3.17)$$

$$= \sum_{j \in [n]} \tilde{\lambda}_{(i-1)n+j} \frac{r}{\|\mathbf{x}_{(i-1)n+j}\|} \frac{\|\mathbf{x}_{(i-1)n+j}\|}{r} \tilde{\mathbf{x}}_{(i-1)n+j} \quad (3.18)$$

$$= r \tilde{\alpha}_i \sum_{j \in [n]} \frac{\tilde{\lambda}_{(i-1)n+j}}{\|\mathbf{x}_{(i-1)n+j}\| \tilde{\alpha}_i} \mathbf{x}_{(i-1)n+j} \quad (3.19)$$

$$= r \tilde{\alpha}_i \sum_{j \in [n]} \lambda_{(i-1)n+j} \mathbf{x}_{(i-1)n+j}, \quad (3.20)$$

woraus wir letztlich auf  $\mathbf{0} \in \bigcap_{i \in [m]} \text{conv}(\mathcal{S}_i)$  schließen.  $\square$

**Satz 3.4** (Theorem 3.2). *Es seien  $m, n \in \mathbb{N}$  und es sei  $D$  eine Wahrscheinlichkeitsverteilung auf  $\mathbb{R}^d$ , die um einen Punkt  $\mathbf{p} \in \mathbb{R}^d$  in dem Sinne ausgeglichen ist, dass jede Hyperebene durch  $\mathbf{p}$  die Verteilung  $D$  in zwei Mengen gleichen Maßes aufteilt. Weiter bezeichne  $\mathcal{E}_{m,n,D}$  eine zufällige Gleichverteilung wie in Definition 2.10. Dann gelten*

$$\left(1 - \frac{1}{2^{n-1}} \sum_{k=0}^{d-1} \binom{n-1}{k}\right)^m \leq \mathbb{P}(\mathcal{E}_{m,n,D} \text{ ist Tverberg}) \quad (3.21)$$

und

$$\mathbb{P}(\mathcal{E}_{m,n,D} \text{ ist Tverberg}) \leq (2(1 - 2^{-n})^m - (1 - 2^{-n+1})^m)^d. \quad (3.22)$$

*Beweis der unteren Schranke (3.21).* Wir können  $\mathbf{p} = \mathbf{0}$  annehmen, da wir gegebenenfalls eine Verschiebung aller Punkte um  $\mathbf{p}$  durchführen können. Dann projizieren wir alle Punkte wie in Proposition 3.3 beschrieben auf eine Kugeloberfläche. Wir zeigen (3.21), indem wir die Wahrscheinlichkeit, dass der Ursprung  $\mathbf{0} \in \mathbb{R}^d$  ein Tverberg-Punkt ist (siehe auch Proposition 3.3), nach unten abschätzen. Wir nehmen im Folgenden an, dass keiner der zufällig gewählten Punkte in  $\mathcal{E}_{m,n,D}$  der Ursprung ist. Falls dem doch so ist, verbessert dieser Punkt lediglich die Schranke.

Nachdem die Projektion durchgeführt wurde, können wir annehmen, dass alle Punkte mit einer Gleichverteilung auf der Kugeloberfläche realisiert wurden (als Folgerung aus der Ausgeglichenheit der Verteilung um  $\mathbf{0}$ ). Der Ursprung ist ein Tverberg-Punkt, falls die konvexe Hülle jeder Farbe den Ursprung enthält. Falls alle Punkte einer Farbe auf einer Halbkugel liegen, enthält deren konvexe Hülle den Ursprung nicht. Wir wollen also, dass die Punkte von keiner Farbe auf einer Halbkugel liegen. Die Wahrscheinlichkeit, dass  $n$  Punkte auf einer Halbkugel liegen, ist nach Satz 2.13 durch

$$\frac{1}{2^{n-1}} \sum_{k=0}^{d-1} \binom{n-1}{k} \quad (3.23)$$

gegeben. Wenn also die Punkte aller  $m$  Farben gleichzeitig jeweils *nicht* auf einer Halbkugel liegen sollen, ist die Wahrscheinlichkeit dafür durch die linke Seite in (3.21) gegeben.  $\square$

*Beweis der oberen Schranke (3.22).* Im Folgenden spielen Münzwurfwahrscheinlichkeiten eine Rolle. Wir berechnen diese im Voraus, um dann darauf zurückzugreifen. Es werfen  $m$  Personen jeweils  $n$  faire Münzen. Es sei  $K$  das Ereignis, dass mindestens eine Person  $n$ -mal Kopf beobachtet und es sei  $Z$  das Ereignis, dass mindestens eine Person  $n$ -mal Zahl beobachtet. Wir benötigen die Wahrscheinlichkeit, dass beide Ereignisse eintreten:

$$\mathbb{P}(K \cap Z) = \mathbb{P}(K) + \mathbb{P}(Z) - \mathbb{P}(K \cup Z) = \mathbb{P}(K) + \mathbb{P}(Z) - \left(1 - \mathbb{P}\left((K \cup Z)^c\right)\right). \quad (3.24)$$

Die Wahrscheinlichkeit, dass bei  $n$  Münzwürfen  $n$ -mal Kopf beobachtet wird, ist  $2^{-n}$ . Die Wahrscheinlichkeit, dass dies bei *keiner* der  $m$  Personen, die jeweils  $n$  Münzen werfen, passiert, ist dann  $(1 - 2^{-n})^m$ . Das Gegenereignis dazu ist  $K$ , weshalb

$$\mathbb{P}(K) = 1 - (1 - 2^{-n})^m = \mathbb{P}(Z) \quad (3.25)$$

gilt. Die Wahrscheinlichkeit, dass eine Person  $n$ -mal Kopf oder  $n$ -mal Zahl beobachtet, ist  $2 \cdot 2^{-n} = 2^{-n+1}$ . Die Wahrscheinlichkeit, dass dies bei *keiner* der  $m$  Personen, die jeweils  $n$  Münzen werfen, passiert, ist dann

$$\mathbb{P}\left((K \cup Z)^c\right) = (1 - 2^{-n+1})^m. \quad (3.26)$$

Daraus folgt schließlich nach Einsetzen:

$$\mathbb{P}(K \cap Z) = (1 - (1 - 2^{-n})^m) + (1 - (1 - 2^{-n})^m) - (1 - (1 - 2^{-n+1})^m) \quad (3.27)$$

$$= 1 - 2(1 - 2^{-n})^m + (1 - 2^{-n+1})^m. \quad (3.28)$$

Nach dieser Vorüberlegung kommen wir zum Hauptbeweis. Wir nehmen wie zuvor  $\mathbf{p} = \mathbf{0}$  an und konzentrieren uns zunächst auf  $d = 1$ . Um eine obere Schranke für  $\mathbb{P}(\mathcal{E}_{m,n,D} \text{ ist Tverberg})$  zu finden, leiten wir eine untere Schranke an die Wahrscheinlichkeit für das Gegenereignis her. Es sei also  $E$  das Ereignis, dass sich nicht alle konvexen Hüllen der Farben schneiden. Im eindimensionalen ( $d = 1$ ) Fall tritt dieses Ereignis gewiss ein, wenn alle Punkte mindestens einer Farbe negativ und gleichzeitig alle Punkte mindestens einer Farbe positiv sind. Da der Ursprung  $D$  in zwei Mengen gleichen Maßes einteilt, ist sowohl die Wahrscheinlichkeit einen negativen Punkt zu realisieren als auch die Wahrscheinlichkeit einen positiven Punkt zu realisieren jeweils  $1/2$ . Das Ereignis, dass alle Punkte mindestens einer Farbe negativ und gleichzeitig alle Punkte mindestens einer Farbe positiv sind, lässt sich deshalb wie folgt formulieren: Wenn  $m$  Personen  $n$  faire Münzen werfen, soll mindestens eine Person  $n$ -mal Kopf beobachten (das Ereignis  $K$ ) und mindestens eine Person soll  $n$ -mal Zahl beobachten (das Ereignis  $Z$ ). Es gilt nun

$$\mathbb{P}(\mathcal{E}_{m,n,D} \text{ ist Tverberg}) = 1 - \mathbb{P}(E) \leq 1 - \mathbb{P}(K \cap Z) \quad (3.29)$$

und mit (3.28) erhalten wir insgesamt

$$\mathbb{P}(\mathcal{E}_{m,n,D} \text{ ist Tverberg}) \leq 2(1 - 2^{-n})^m - (1 - 2^{-n+1})^m, \quad (3.30)$$

womit der Fall  $d = 1$  behandelt ist.

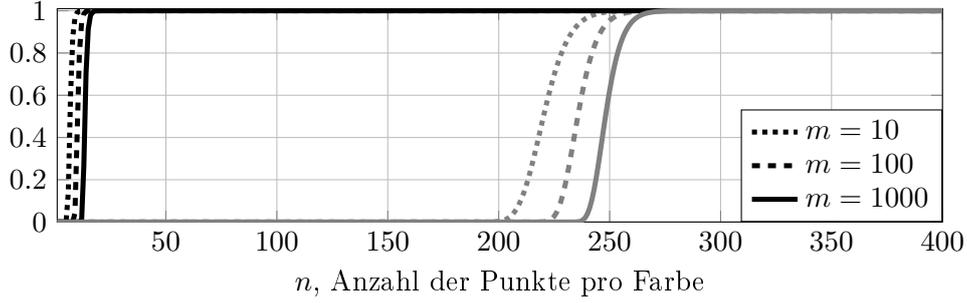


Abbildung 1: Die untere (grau, Ungleichung (3.21)) und obere (schwarz, Ungleichung (3.22)) Schranke von Satz 3.4 ausgewertet für  $d = 100$  bei  $m = 10$  (gepunktet),  $m = 100$  (gestrichelt) und  $m = 1000$  (durchgezogen) Farben.

Für  $d > 1$  sei  $p_i(\mathbf{x})$  die Projektion eines Punktes  $\mathbf{x} \in \mathbb{R}^d$  auf die  $i$ te Koordinatenachse ( $i \in [d]$ ). Da allgemein die zur  $i$ ten Achse orthogonale Hyperebene die Verteilung  $D$  in zwei Mengen gleichen Maßes aufteilt (wegen der Annahme  $\mathbf{p} = \mathbf{0}$ ), liegt  $p_i(\mathbf{x})$  jeweils mit Wahrscheinlichkeit  $1/2$  auf der einen Seite der Hyperebene oder auf der anderen Seite. Wir sind also wieder im Münzwurffall angekommen und eine Tverberg-Zerlegung muss die Eigenschaft haben, dass kein Farbpaar nach der Projektion auf die  $d$  Koordinatenachsen durch den Ursprung getrennt wird. Diese  $d$  Ereignisse sind unabhängig, weshalb wir schließlich die gewünschte Schranke (3.22) erhalten.  $\square$

Wir werten die untere und obere Schranke von Satz 3.4 in Abbildung 1 für  $d = 100$  Dimensionen bei unterschiedlicher Anzahl  $m$  an Farben aus. In diesem Beispiel liegen die beiden Schranken für kleine  $n$  und für große  $n$  jeweils nahe beieinander (mit Werten nahe bei 0 beziehungsweise nahe bei 1). Bei einem passenden Verhältnis von  $m$  und  $n$  liegen demnach scharfe Schranken vor. Ein Vergleich der Kurven für unterschiedliche Anzahl  $m$  an Farben führt zu der interessanten Beobachtung, dass die Anzahl  $n$  der Punkte pro Farbe nicht um den gleichen Faktor wie  $m$  erhöht werden muss, um in einem Bereich scharfer Schranken zu landen. Der nachfolgende Satz, welcher aus Satz 3.4 folgt, zeigt, dass wir allgemein einen Phasenübergang von einem fast unmöglichen zu einem fast sicheren Ereignis beobachten können, wenn  $m$  und  $n$  im richtigen Verhältnis erhöht werden. Wir folgern aus dem Beweis im Speziellen, dass die Schranken von Satz 3.4 asymptotisch ( $m \rightarrow \infty$ ) scharf sind.

Die Gleichungen (3.4) und (3.3) in Satz 3.1 zeigen, dass auch im Modell der zufälligen Zuordnung (Definition 2.11) bei einer bestimmten Anzahl an Punkten ein Übergang von einem fast unmöglichen zu einem fast sicheren Ereignis stattfindet. Es ist bemerkenswert, dass Satz 3.5 im Gegensatz zu Satz 3.1 von der Dimension unabhängig ist. In diesem Sinne ist Satz 3.5 mit Satz 3.2 verwandt.

**Satz 3.5** (Theorem 3.3). *Es sei  $D$  eine Wahrscheinlichkeitsverteilung auf  $\mathbb{R}^d$ , die um einen Punkt  $\mathbf{p} \in \mathbb{R}^d$  in dem Sinne ausgeglichen ist, dass jede Hyperebene durch  $\mathbf{p}$  die Verteilung  $D$  in zwei Mengen gleichen Maßes aufteilt. Wir betrachten die Folge zufällig*

gleichverteilter (Definition 2.10) Punktmengen  $\mathcal{E}_{m,f(m),D}$ , bei welchen  $m \in \mathbb{N}$  gilt und  $n = f(m)$  von  $m$  abhängt. Dann ist  $\mathcal{E}_{m,f(m),D}$  mit hoher Wahrscheinlichkeit Tverberg, falls  $f(m) \gg \log_2(m)$ , und  $\mathcal{E}_{m,f(m),D}$  ist mit hoher Wahrscheinlichkeit nicht Tverberg, falls  $f(m) \ll \log_2(m)$ .

*Beweis.* Es sei  $\varepsilon > 0$  fest vorgegeben. Wir beweisen die beiden Aussagen, indem wir zwei Grenzwerte berechnen. Zum einen setzen wir  $n = (1 + \varepsilon) \log_2(m)$  in die untere Schranke (3.21) von Satz 3.4 ein und zeigen deren Konvergenz gegen 1 (für  $m \rightarrow \infty$ ). Zum anderen setzen wir  $n = (1 - \varepsilon) \log_2(m)$  in die obere Schranke (3.22) von Satz 3.4 ein und zeigen deren Konvergenz gegen 0 (für  $m \rightarrow \infty$ ).

*Grenzwert der unteren Schranke.* Als Vorbereitung berechnen wir zuerst den Grenzwert von

$$\frac{(1 + \varepsilon) \log_2(m)}{m^\varepsilon} \quad (3.31)$$

für  $m \rightarrow \infty$ . Hierfür stellen wir zunächst  $\lim_{m \rightarrow \infty} (1 + \varepsilon) \log_2(m) = \infty$  für den Zähler und  $\lim_{m \rightarrow \infty} m^\varepsilon = \infty$  für den Nenner fest. Wir wenden nun die Regel von L'Hospital an, um den Grenzwert als

$$\lim_{m \rightarrow \infty} \frac{(1 + \varepsilon) \log_2(m)}{m^\varepsilon} = \lim_{m \rightarrow \infty} \frac{(1 + \varepsilon) \frac{1}{m \ln(2)}}{\varepsilon m^{\varepsilon-1}} = \lim_{m \rightarrow \infty} \frac{1 + \varepsilon}{\ln(2) \varepsilon m^\varepsilon} = 0 \quad (3.32)$$

zu bestimmen.

Wir zeigen als Nächstes, dass  $\mathcal{E}_{m,f(m),D}$  mit hoher Wahrscheinlichkeit Tverberg ist, falls  $f(m) > \log_2(m)$  gilt. Da die Voraussetzungen von Satz 3.4 erfüllt sind, können wir  $n = (1 + \varepsilon) \log_2(m)$  in (3.21) einsetzen und erhalten

$$\mathbb{P}(\mathcal{E}_{m,n,D} \text{ ist Tverberg}) \geq \left( 1 - \frac{1}{2^{(1+\varepsilon) \log_2(m)-1}} \sum_{k=0}^{d-1} \binom{n-1}{k} \right)^m \quad (3.33)$$

$$= \left( 1 - 2^{-(1+\varepsilon) \log_2(m)+1} \sum_{k=0}^{d-1} \binom{n-1}{k} \right)^m \quad (3.34)$$

$$= \left( 1 - 2m^{-(1+\varepsilon)} \sum_{k=0}^{d-1} \binom{n-1}{k} \right)^m. \quad (3.35)$$

Mit der bekannten Abschätzung

$$\binom{n}{k} \leq \left( \frac{en}{k} \right)^k \quad (3.36)$$

können wir wegen  $k \leq d$  eine Konstante  $K$  finden, sodass

$$2 \sum_{k=0}^{d-1} \binom{n-1}{k} \leq K n^d \quad (3.37)$$

gilt. Wir setzen das in (3.35) ein und erhalten

$$\left(1 - Kn^d m^{-(1+\varepsilon)}\right)^m \leq \mathbb{P}(\mathcal{E}_{m,n,D} \text{ ist Tverberg}). \quad (3.38)$$

Wir zeigen nun, dass die linke Seite für  $m \rightarrow \infty$  gegen 1 konvergiert. Dazu sei ein beliebiges  $\delta > 0$  fest vorgegeben. Wegen (3.32) gibt es ein  $M$ , sodass

$$Kn^d m^{-\varepsilon} < \delta \quad \text{für alle } m \geq M \quad (3.39)$$

für  $n = (1 + \varepsilon) \log_2(m)$  erfüllt ist. Damit ergibt sich  $(1 - Kn^d m^{-(1+\varepsilon)})^m > (1 - \delta m^{-1})^m$  für alle  $m \geq M$ . Das erlaubt uns, den Grenzwert wie folgt abzuschätzen:

$$\lim_{m \rightarrow \infty} (1 - Kn^d m^{-(1+\varepsilon)})^m \geq \lim_{m \rightarrow \infty} (1 - \delta m^{-1})^m = e^{-\delta}. \quad (3.40)$$

Da wir  $\delta > 0$  beliebig klein machen können, konvergiert also  $\mathbb{P}(\mathcal{E}_{m,n,D} \text{ ist Tverberg})$  mit (3.38) gegen 1 für  $m \rightarrow \infty$ .

*Grenzwert der oberen Schranke.* Wir zeigen als Nächstes, dass  $\mathcal{E}_{m,f(m),D}$  mit hoher Wahrscheinlichkeit nicht Tverberg ist, falls  $f(m) < \log_2(m)$  gilt. Wir wählen  $n = (1 - \varepsilon) \log_2(m)$  und setzen in (3.22) ein:

$$\mathbb{P}(\mathcal{E}_{m,n,D} \text{ ist Tverberg}) \leq (2(1 - m^{-1+\varepsilon})^m - (1 - 2m^{-1+\varepsilon})^m)^d. \quad (3.41)$$

Es sei  $\gamma > 0$  beliebig gewählt. Es gibt ein  $M$ , sodass sowohl  $m^{-1+\varepsilon} > \gamma m^{-1}$  als auch  $2m^{-1+\varepsilon} > \gamma m^{-1}$  für  $m \geq M$  erfüllt sind. Das sehen wir, indem wir beide Seiten der Ungleichungen mit  $m$  multiplizieren und  $\lim_{m \rightarrow \infty} m^\varepsilon = \infty$  bedenken. Der erste Summand  $2(1 - m^{-1+\varepsilon})^m$  in (3.41) ist deshalb für  $m \geq M$  kleiner als  $2(1 - \gamma m^{-1})^m$  und der zweite Summand  $(1 - 2m^{-1+\varepsilon})^m$  ist kleiner als  $(1 - \gamma m^{-1})^m$ . Nun gilt  $\lim_{m \rightarrow \infty} (1 - \gamma m^{-1})^m = e^{-\gamma}$ . Da  $\gamma$  beliebig groß gewählt werden kann, können wir beide Summanden beliebig klein machen. Deshalb konvergiert die rechte Seite in (3.41) für  $m \rightarrow \infty$  gegen 0, was zu zeigen war.  $\square$

Im Speziellen schließen wir aus dem Beweis von Satz 3.5, dass die Schranken in Satz 3.4 bezüglich  $m$  asymptotisch scharf sind. In Abbildung 2 sehen wir ein numerisches Beispiel für diese Aussage. Der nächste Satz bringt nun das Thema Toleranz ins Spiel.

**Satz 3.6** (Theorem 3.4). *Es seien  $m, n \in \mathbb{N}$  und es sei  $D$  eine Wahrscheinlichkeitsverteilung auf  $\mathbb{R}^d$ , die um einen Punkt  $\mathbf{p} \in \mathbb{R}^d$  in dem Sinne ausgeglichen ist, dass jede Hyperebene durch  $\mathbf{p}$  die Verteilung  $D$  in zwei Mengen gleichen Maßes aufteilt. Es sei  $\mathcal{E}_{m,n,D}$  eine zufällige Gleichverteilung wie in Definition 2.10. Dann gilt*

$$\mathbb{P}(\mathcal{E}_{m,n,D} \text{ ist Tverberg mit Toleranz } t) \geq \left(1 - 2^{-\lfloor n/2d \rfloor} \sum_{i=0}^t \binom{\lfloor n/2d \rfloor}{i}\right)^m. \quad (3.42)$$

*Beweis.* Wie schon in anderen Beweisen können wir  $\mathbf{p} = \mathbf{0}$  annehmen. Es sei  $\mathcal{S}$  die Menge aller Punkte einer fest gewählten Farbe. Die Menge enthält gemäß Modellannahme  $n$  Punkte. Wir zerlegen  $\mathcal{S}$  in  $\lfloor n/2d \rfloor$  Teilmengen  $\mathcal{S}_1, \dots, \mathcal{S}_{\lfloor n/2d \rfloor}$ , welche alle jeweils mindestens  $2d$  Punkte enthalten.

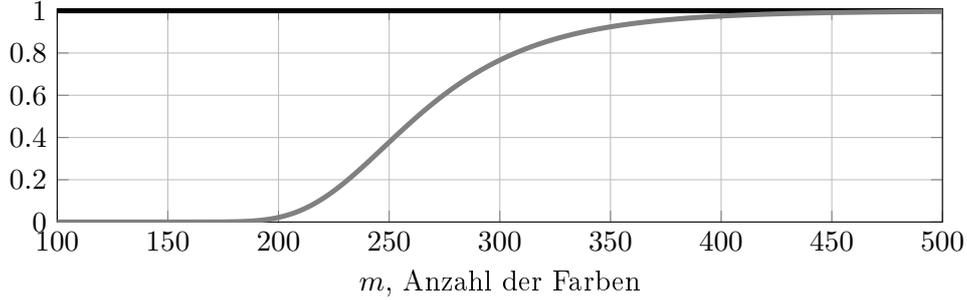


Abbildung 2: Die untere (grau, Ungleichung (3.21)) und obere (schwarz, Ungleichung (3.22)) Schranke von Satz 3.4 ausgewertet für  $d = 100$  bei  $n = 30 \log_2(m)$  Punkten pro Farbe.

Nach Korollar 2.14 ist die Wahrscheinlichkeit, dass  $2d$  Punkte auf einer Halbkugel liegen,  $1/2$ . Das wiederum impliziert, dass die konvexe Hülle von  $2d$  Punkten den Ursprung mit einer Wahrscheinlichkeit von  $1/2$  enthält. Da jede Menge  $\mathcal{S}_i$  mindestens  $2d$  Punkte enthält, enthält also jede konvexe Hülle  $\text{conv}(\mathcal{S}_i)$  den Ursprung mit einer Wahrscheinlichkeit von mindestens  $1/2$ . Wir haben insgesamt  $\lfloor n/2d \rfloor$  Mengen  $\mathcal{S}_i$  und für jede dieser Mengen gilt die gerade hergeleitete Wahrscheinlichkeit. Wir bezeichnen „die konvexe Hülle enthält den Ursprung“ als „Erfolg“ und „die konvexe Hülle enthält den Ursprung nicht“ als „Misserfolg“. Da alle Punkte unabhängig voneinander generiert werden, liegt die Situation in Bemerkung 2.15 (mit  $p = 1/2$ ) vor. Wir können also die Wahrscheinlichkeit, dass weniger als  $t + 1$  der konvexen Hüllen  $\text{conv}(\mathcal{S}_i)$  den Ursprung enthalten durch  $(1/2)^{\lfloor n/2d \rfloor} \sum_{i=0}^t \binom{\lfloor n/2d \rfloor}{i}$  nach oben beschränken.

Wenn mindestens  $t + 1$  der  $\text{conv}(\mathcal{S}_i)$  den Ursprung enthalten, enthält die konvexe Hülle  $\text{conv}(S \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_t\})$  den Ursprung für beliebige  $\mathbf{x}_1, \dots, \mathbf{x}_t \in S$ . Das liegt daran, dass bei beliebiger Wahl der  $t$  Punkte  $\mathbf{x}_1, \dots, \mathbf{x}_t$  stets eine der  $t + 1$  konvexen Hüllen  $\text{conv}(\mathcal{S}_i)$  unverändert ist (eine Menge  $\mathcal{S}_i$  enthält also weiterhin alle ihre Punkte) und deshalb weiterhin den Ursprung enthält. Deshalb enthält  $\text{conv}(S \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_t\})$  mit Wahrscheinlichkeit

$$1 - (1/2)^{\lfloor n/2d \rfloor} \sum_{i=0}^t \binom{\lfloor n/2d \rfloor}{i} \quad (3.43)$$

den Ursprung. Da diese Überlegung für jede der  $m$  Farben unabhängig gilt, folgt die Aussage des Satzes.  $\square$

### 3.3 Zufällige Zuordnung

Wir kommen zu Aussagen zum Modell der zufälligen Zuordnung (Definition 2.11). Der nachfolgende Satz 3.7 kann als eine um Toleranz erweiterte Version von Satz 3.1 betrachtet werden.

**Satz 3.7** (Theorem 3.5). *Es sei  $k \in \mathbb{N}$ . Wenn  $D$  eine stetige Wahrscheinlichkeitsverteilung auf  $\mathbb{R}^d$  und  $\mathcal{R}_{2,k,D}$  eine zufällige Zuordnung wie in Definition 2.11 ist, dann gilt*

$$\mathbb{P}(\mathcal{R}_{2,k,D} \text{ ist Radon mit Toleranz } t) \geq 1 - 2^{-\lfloor k/(2d+2) \rfloor} \sum_{i=0}^t \binom{\lfloor k/(2d+2) \rfloor}{i}. \quad (3.44)$$

*Im Speziellen gilt*

$$\mathbb{P}(\mathcal{R}_{2,k,D} \text{ ist Radon mit Toleranz } \lfloor k/(4d+4) \rfloor) \geq 1/2. \quad (3.45)$$

*Beweis.* Wir teilen die  $k$  Punkte beliebig auf  $\lfloor k/(2d+2) \rfloor$  disjunkte Mengen auf, welche dann jeweils mindestens  $2d+2$  Punkte enthalten. Mit (3.2) schneiden sich die konvexen Hüllen der beiden Farben (wir haben  $m=2$ ) in einer gegebenen Menge mit einer Wahrscheinlichkeit von mindestens  $1/2$ . Wir betrachten dieses Ereignis als einen „Erfolg“. Ähnlich wie im Beweis von Satz 3.6 wollen wir mindestens  $t+1$  Erfolge haben, damit  $t$  Punkte entfernt werden können und sich die konvexen Hüllen immer noch schneiden. Die Wahrscheinlichkeit dafür ist durch  $1 - 2^{-\lfloor k/(2d+2) \rfloor} \sum_{i=0}^t \binom{\lfloor k/(2d+2) \rfloor}{i}$  gegeben (vergleiche (3.43)) und erklärt (3.44).

Um (3.45) zu sehen, setzen wir zunächst

$$2^{\lfloor k/(2d+2) \rfloor} = \sum_{i=0}^{\lfloor k/(2d+2) \rfloor} \binom{\lfloor k/(2d+2) \rfloor}{i} = 2^{\lfloor k/(4d+4) \rfloor} \sum_{i=0}^{\lfloor k/(2d+2) \rfloor} \binom{\lfloor k/(2d+2) \rfloor}{i} \quad (3.46)$$

an (vergleiche die Argumentation im Beweis von Korollar 2.14). Wenn wir  $t = \lfloor k/(4d+4) \rfloor$  in (3.44) einsetzen, erhalten wir

$$\mathbb{P}(\mathcal{R}_{2,k,D} \text{ ist Radon mit Toleranz } \lfloor k/(4d+4) \rfloor) \geq 1 - 2^{-\lfloor k/(2d+2) \rfloor} \sum_{i=0}^{\lfloor k/(4d+4) \rfloor} \binom{\lfloor k/(2d+2) \rfloor}{i}. \quad (3.47)$$

Wegen (3.46) ist die rechte Seite gleich  $1/2$ , was (3.45) erklärt.  $\square$

Um den Fall der zufälligen Zuordnung mit Toleranz  $t \geq 0$  zu behandeln, ist sowohl das Ergebnis des folgenden Satzes als auch dessen Notation nötig.

**Satz 3.8** (Erdős, Renyi). *Wir geben eine Kugel nach der anderen und unabhängig voneinander mit einer Wahrscheinlichkeit von jeweils  $1/u$  in eine von  $u$  Urnen. Es bezeichne  $N_K(u)$  die Anzahl an Kugeln, die nötig ist, um mindestens  $K$  Kugeln in jeder Urne zu sehen. Dann gilt für jedes  $x \in \mathbb{R}$ :*

$$\lim_{u \rightarrow \infty} \mathbb{P}(N_K(u) < u(\ln(u) + (K-1)\ln(\ln(u)) + x)) = \exp\left(-\frac{e^{-x}}{(K-1)!}\right). \quad (3.48)$$

Statt Kugeln in Urnen zu geben, generieren wir Punkte und färben diese ein. Wir sind daran interessiert, mittels Beobachtung 2.12 bereits für das Modell der zufälligen Gleichverteilung (Definition 2.10) bekannte Aussagen auf das Modell der zufälligen Zuordnung (Definition 2.11) zu übertragen. Laut Beobachtung 2.12 brauchen wir deshalb die Wahrscheinlichkeit dafür, dass wir im Modell der zufälligen Zuordnung (Definition 2.11) mindestens  $n$  Punkte jeder Farbe haben. Im folgenden Korollar 3.9 soll die für dieses Ereignis benötigte Anzahl an Punkten (die Zahl  $N_n(m)$ ) kleiner als die Anzahl an vorhandenen Punkten (die Zahl  $k$ ) sein. Wir verwenden also die Wahrscheinlichkeit für  $N_n(m) \leq k$ .

**Korollar 3.9** (Corollary 3.6). *Es seien  $n, m \in \mathbb{N}$ . Wir bezeichnen mit  $N_n(m)$  die Größe aus Satz 3.8. Es sei  $D$  eine Wahrscheinlichkeitsverteilung auf  $\mathbb{R}^d$ , die um einen Punkt  $\mathbf{p} \in \mathbb{R}^d$  in dem Sinne ausgeglichen ist, dass jede Hyperebene durch  $\mathbf{p}$  die Verteilung  $D$  in zwei Mengen gleichen Maßes aufteilt.*

1. *Dann gilt*

$$\mathbb{P}(\mathcal{R}_{m,k,D} \text{ ist Tverberg mit Toleranz } t) \geq \mathbb{P}(N_n(m) \leq k) \left( 1 - 2^{-\lfloor n/2d \rfloor} \sum_{i=0}^t \binom{\lfloor n/2d \rfloor}{i} \right)^m. \quad (3.49)$$

2. *Für den Fall ohne Toleranz gilt*

$$\mathbb{P}(\mathcal{R}_{m,k,D} \text{ ist Tverberg}) \geq \mathbb{P}(N_n(m) \leq k) \left( 1 - \frac{1}{2^{n-1}} \sum_{i=0}^{d-1} \binom{n-1}{i} \right)^m. \quad (3.50)$$

3. *Es sei  $(\mathcal{R}_{m,f(m),D})_{m \in \mathbb{N}}$  eine Folge von zufällig zugeordneten (Definition 2.11) Punkt-mengen  $\mathcal{R}_{m,f(m),D}$ , bei welchen die Anzahl  $f(m)$  der Punkte von  $m$  abhängt. Falls  $f(m) \gg m \log_2(m) \ln(\ln(m))$ , ist  $\mathcal{R}_{m,f(m),D}$  mit hoher Wahrscheinlichkeit Tverberg.*

*Beweis von Aussage 1.* Das folgt aus Beobachtung 2.12 zusammen mit Satz 3.6. □

*Beweis von Aussage 2.* Das folgt aus Beobachtung 2.12 zusammen mit Satz 3.4. □

*Beweis von Aussage 3.* Nach Satz 3.5 ist eine zufällige Gleichverteilung (Definition 2.10) von Punkten in  $m$  Farben, wobei jede Farbe mehr als  $\log_2(m)$  Punkte hat, mit hoher Wahrscheinlichkeit Tverberg. Nach Beobachtung 2.12 ist dann auch eine zufällige Zuordnung (Definition 2.11) mit hoher Wahrscheinlichkeit Tverberg, falls mit hoher Wahrscheinlichkeit mehr als  $\log_2(m)$  Punkte jeder Farbe vorhanden sind. Wir wollen also sicherstellen, dass wir  $f(m)$  (die Anzahl der Punkte im aktuellen Korollar) groß genug wählen, um mit hoher Wahrscheinlichkeit mehr als  $\log_2(m)$  Punkte pro Farbe zu beobachten.

Das erfolgt über Satz 3.8. Die Anzahl an Urnen entspricht der Anzahl an Farben ( $u = m$ ) und wir wollen mindestens  $K = \log_2(m)(1 + \varepsilon)$  Punkte pro Farbe (Kugeln

pro Urne) beobachten. Letztlich soll also  $N_K(m) < f(m)$  mit hoher Wahrscheinlichkeit eintreten. Das Argument der Wahrscheinlichkeit in Satz 3.8 lautet mit der Notation des Korollars:

$$N_K(m) < m(\ln(m) + (\log_2(m)(1 + \varepsilon) - 1) \ln(\ln(m)) + x) \quad (3.51)$$

$$= m \ln(m) + m(\log_2(m)(1 + \varepsilon) - 1) \ln(\ln(m)) + mx. \quad (3.52)$$

Der Term  $m \log_2(m)(1 + \varepsilon) \ln(\ln(m))$  wächst mit  $m$  am schnellsten und die rechte Seite

$$\exp\left(-\frac{e^{-x}}{(K-1)!}\right) \quad (3.53)$$

in Satz 3.8 geht für  $K = \log_2(m)(1 + \varepsilon)$  schnell gegen 1. Deswegen erreichen wir mit  $f(m) \gg m \log_2(m) \ln(\ln(m))$  das Ziel.  $\square$

## 4 Anwendungsbeispiele

Wir präsentieren drei Anwendungsbeispiele, bei welchen Tverberg-Zerlegungen eine wichtige Rolle spielen. Mit den in Abschnitt 3 vorgestellten stochastischen Aussagen zur Tverberg-Zerlegung lassen sich nun entsprechende stochastische Aussagen zu den Anwendungsbeispielen machen.

### 4.1 Zentralwert

Das erste Anwendungsbeispiel betrifft die Berechnung von Zentralwerten. Es handelt sich bei einem Zentralwert um die Verallgemeinerung des Medians auf höhere Dimensionen.

**Definition 4.1** (Zentralwert, centerpoint). *Ein Zentralwert einer Punktmenge  $\mathcal{S} \subset \mathbb{R}^d$  mit  $|\mathcal{S}| = n$  ist ein Punkt  $\mathbf{p} \in \mathbb{R}^d$  mit der Eigenschaft, dass jeder Halbraum, der  $\mathbf{p}$  enthält, auch mindestens  $\frac{n}{d+1}$  Punkte von  $\mathcal{S}$  enthält.*

Ein Zusammenhang zwischen Tverberg-Punkten und Zentralwerten kann wie folgt hergestellt werden: Ein Tverberg-Punkt einer  $m$ -farbigen Punktmenge mit  $m = \lceil \frac{n}{d+1} \rceil$  ist ein Zentralwert. Das liegt letztlich daran, dass jeder Halbraum, der den Tverberg-Punkt enthält, mindestens einen Punkt jeder Farbe enthalten muss. Denn, gäbe es eine Farbe, die nicht im Halbraum liegt, könnte der Tverberg-Punkt keine Konvexkombination der Punkte dieser Farbe sein und damit läge kein Tverberg-Punkt vor.

Eine effiziente Methode, Tverberg-Punkte zu berechnen, liefert also eine Methode, Zentralwerte zu bestimmen. Allerdings scheint es weder zur Berechnung von Tverberg-Punkten noch zur Berechnung von Zentralwerten effiziente Methoden zu geben. Es gibt jedoch schnelle Methoden, um Tverberg-Punkte approximativ zu berechnen, falls eine Tverberg-Zerlegung vorliegt. Das folgende Korollar 4.2 zeigt, wie  $n$ -elementige Punktmen- gen, die aus einer passenden Verteilung erzeugt wurden, eingefärbt werden können, um mit hoher Wahrscheinlichkeit eine Tverberg-Zerlegung zu erhalten. Es kommen letztlich wieder die beiden aus Abschnitt 2.4 bekannten Modelle zur Erzeugung zufälliger Punkt- mengen zum Einsatz. Sobald eine Tverberg-Zerlegung vorliegt, kann ein Algorithmus zur approximativen Berechnung eines Tverberg-Punktes benutzt werden, um schließlich einen Zentralwert der Menge zu erhalten.

**Korollar 4.2** (Corollary 1.3). *Es sei  $D$  eine Wahrscheinlichkeitsverteilung auf  $\mathbb{R}^d$ , die um einen Punkt  $\mathbf{p} \in \mathbb{R}^d$  in dem Sinne ausgeglichen ist, dass jede Hyperebene durch  $\mathbf{p}$  die Verteilung in zwei Mengen gleichen Maßes aufteilt. Es werden  $n$ -elementige Punktmen- gen gemäß  $D$  erzeugt.*

1. *Wenn  $\delta > 0$  fest vorgegeben ist und die Punktmen- gen gemäß einer zufälligen Gleich- verteilung (Definition 2.10) eingefärbt werden, wobei  $m = (1 - \delta) \binom{n}{\log_2(n)}$  Farben benutzt werden, erhalten wir mit gegen 1 konvergierender Wahrscheinlichkeit (bei  $n \rightarrow \infty$ ) eine Tverberg-Zerlegung.*
2. *Wenn  $\delta > 0$  fest vorgegeben ist und die Punktmen- gen gemäß einer zufälligen Zuord- nung (Definition 2.11) eingefärbt werden, wobei  $m = (1 - \delta) \binom{n}{\log_2(n) \ln(\ln(n))}$  Farben*

benutzt werden, erhalten wir mit gegen 1 konvergierender Wahrscheinlichkeit (bei  $n \rightarrow \infty$ ) eine Tverberg-Zerlegung.

*Beweis von Aussage 1.* Satz 3.5 gibt die Anzahl  $m$  an Farben vor und besagt, wie groß die Anzahl der Punkte pro Farbe gewählt werden muss, um eine Tverberg-Zerlegung zu erhalten. Im aktuellen Korollar ist dagegen die Gesamtzahl  $n$  der Punkte vorgegeben und wir wollen die Anzahl  $m$  der Farben klein genug wählen, um Satz 3.5 anwenden zu können. Das führt zur Wahl  $m = (1 - \delta) \left( \frac{n}{\log_2(n)} \right)$ .  $\square$

*Beweis von Aussage 2.* Die Wahl  $m = (1 - \delta) \left( \frac{n}{\log_2(n) \ln(\ln(n))} \right)$  führt wegen  $m < n$  zu  $m < (1 - \delta) \left( \frac{n}{\log_2(m) \ln(\ln(m))} \right)$ . Demnach ist  $m \log_2(m) \ln(\ln(m)) < (1 - \delta)n < n$  erfüllt. Die Anzahl  $n$  der Punkte ist also bei obiger Wahl von  $m$  groß genug, um Korollar 3.9 (Aussage 3) anwenden zu können.  $\square$

Korollar 4.2 erklärt also zum Beispiel, dass der Rechenaufwand, um (mit hoher Wahrscheinlichkeit) eine  $m$ -Tverberg-Zerlegung mit  $m = \frac{n}{\log_2(n)}$  Farben zu erhalten, lediglich darin besteht,  $n$  Punkte gemäß der zufälligen Gleichverteilung (Definition 2.10) zu erzeugen und einzufärben. Eine ähnliche Aussage gilt im Fall der zufälligen Zuordnung (Definition 2.11). Wir fassen das in den letzten zwei Zeilen von Tabelle 2 zusammen.

Die verbleibenden drei Zeilen von Tabelle 2 machen ähnliche Aussagen bei anderen Methoden. Es handelt sich bei Tabelle 2 um eine Reproduktion von [10, Table 5.2]. Leider bleibt diese Tabelle bei den Autoren unkommentiert. Wir ergänzen, dass es sich bei „PPAD“ um die Komplexitätsklasse „Polynomial Parity Arguments on Directed graphs“ handelt und diese Aussage in [12] gezeigt wird.

Methode	Anzahl der Farben	Zeitkomplexität
Tverberg [22]	$\lfloor (n + 1)/d + 1 \rfloor$	PPAD
Mulzer, Werner [13]	$n/(4d + 1)^3$	$d^{O(\ln(d))}n$
Rolnick, Soberón [15]	$n/((d(d + 1)^2))$ mit Fehlerwahrscheinlichkeit $\varepsilon$	schwach polynomial in $n, d, \ln(1/\varepsilon)$
zufällige Gleichverteilung	$O\left(\frac{n}{\log_2(n)}\right)$	$O(n)$
zufällige Zuordnung	$O\left(\frac{n}{\log_2(n) \ln(\ln(n))}\right)$	$O(n)$

Tabelle 2: Die Zeitkomplexität zur Berechnung approximativer Tverberg-Zerlegungen bei  $n$  Punkten, die aus einer passenden Verteilung generiert wurden.

## 4.2 Separierbarkeit

Die Konzepte einer Tverberg-Zerlegung mit Toleranz und der Begriff der Separierbarkeit einer Menge stehen in einem engen Zusammenhang. Die in Abschnitt 3 hergeleiteten Aussagen zur Tverberg-Zerlegung führen deswegen zu Aussagen über Separierbarkeit, die wir im Folgenden vorstellen. Letztlich geht es bei allen Aussagen darum, dass eine zufällige Zuordnung (Definition 2.11) einerseits schwer separierbare Punktmengen und andererseits möglichst tolerante Tverberg-Zerlegungen erzeugt. Die Autoren von [10] definieren leider nicht, was sie unter einer separierbaren Menge verstehen. Wir entscheiden uns hier für die strikte Version der Separierbarkeit (siehe beispielsweise [1]).

**Definition 4.3.** *Es sei  $\mathcal{X}$  eine zweifarbige Punktmenge in  $\mathbb{R}^d$ . Wir nennen  $\mathcal{X}$  separierbar, falls es eine durch  $\mathbf{a} \in \mathbb{R}^d$  und  $a_0 \in \mathbb{R}$  definierte Hyperebene*

$$h = \{\mathbf{x} \in \mathbb{R}^d : \mathbf{x}^T \mathbf{a} = a_0\} \quad (4.1)$$

*gibt, welche die zwei Farbmengen trennt. Das heißt, alle Punkte  $\mathbf{x}$  der einen Farbe erfüllen  $\mathbf{x}^T \mathbf{a} < a_0$  und alle Punkte  $\mathbf{y}$  der anderen Farbe erfüllen  $\mathbf{y}^T \mathbf{a} > a_0$ .*

**Bemerkung 4.4.** *Es seien  $\mathbf{x}_1, \dots, \mathbf{x}_{N_1}, \mathbf{x}_{N_1+1}, \dots, \mathbf{x}_{N_1+N_2}$  Punkte im  $\mathbb{R}^d$ , die wir auf zwei Mengen  $\mathcal{X}_1 = \{\mathbf{x}_1, \dots, \mathbf{x}_{N_1}\}$  und  $\mathcal{X}_2 = \{\mathbf{x}_{N_1+1}, \dots, \mathbf{x}_{N_1+N_2}\}$  aufteilen. Es sei  $h$  eine Hyperebene mit  $\mathbf{a} \in \mathbb{R}^d$  und  $a_0 \in \mathbb{R}$ , welche  $\mathcal{X}_1$  und  $\mathcal{X}_2$  über*

$$\mathbf{x}_i^T \mathbf{a} < a_0 \quad \text{und} \quad \mathbf{x}_j^T \mathbf{a} > a_0 \quad (4.2)$$

*für  $i \in [N_1]$  und  $j \in [N_2] \setminus [N_1]$  trennt. Wir betrachten mit  $\mathbf{x} = \sum_{i \in [N_1]} \lambda_i \mathbf{x}_i$  für  $\lambda_i \geq 0$  und  $\sum_{i \in [N_1]} \lambda_i = 1$  ein beliebiges Element aus  $\text{conv}(\mathcal{X}_1)$ . Dann folgt*

$$\mathbf{x}^T \mathbf{a} = \sum_{i \in [N_1]} \lambda_i \mathbf{x}_i^T \mathbf{a} < \sum_{i \in [N_1]} \lambda_i a_0 = a_0. \quad (4.3)$$

*Analog sehen wir auch*

$$\mathbf{y}^T \mathbf{a} > a_0 \quad (4.4)$$

*für beliebige  $\mathbf{y} \in \text{conv}(\mathcal{X}_2)$ , woraus wir insgesamt auf  $\text{conv}(\mathcal{X}_1) \cap \text{conv}(\mathcal{X}_2) = \emptyset$  schließen. Wenn zwei endliche Mengen durch eine Hyperebene getrennt werden, schneiden sich die zugehörigen konvexen Hüllen also nicht.*

*Es gilt auch die umgekehrte Aussage: Aus  $\text{conv}(\mathcal{X}_1) \cap \text{conv}(\mathcal{X}_2) = \emptyset$  folgt, dass es eine Hyperebene gibt, welche  $\mathcal{X}_1$  und  $\mathcal{X}_2$  trennt. Diese Aussage gilt hier, weil sowohl  $\mathcal{X}_1$  als auch  $\mathcal{X}_2$  endliche Mengen sind. Das impliziert die Kompaktheit von  $\text{conv}(\mathcal{X}_1)$  und  $\text{conv}(\mathcal{X}_2)$ , woraus sich dann die strikte Separierbarkeit folgern lässt, siehe [1, Abschnitt 2.4].*

Falls eine gegebene Menge nicht separierbar ist, ist eine naheliegende Frage, wie viele Punkte entfernt (oder an eine andere Stelle bewegt) werden müssen, um doch eine separierbare Menge zu erhalten.

**Definition 4.5.** Für eine zweifarbige Punktmenge  $\mathcal{X}$  in  $\mathbb{R}^d$  bezeichne  $M$  die minimale Anzahl an Punkten, die bewegt werden müssen, um  $\mathcal{X}$  separierbar zu machen. Dann definieren wir

$$\text{PertSEP}_0(\mathcal{X}) = \frac{M}{|\mathcal{X}|}. \quad (4.5)$$

Anschaulich bedeutet ein größeres  $\text{PertSEP}_0(\mathcal{X})$  eine schwerere Separierbarkeit von  $\mathcal{X}$ . Gleichzeitig führt das zu einer toleranteren Radon-Zerlegung. Der folgende Satz 4.6 präzisiert diese Anschauung. Der Beweis nutzt an mehreren Stellen Bemerkung 4.4.

**Satz 4.6.** Es sei  $\mathcal{X} = \mathcal{X}_1 \cup \mathcal{X}_2$  mit  $|\mathcal{X}| = n$  eine Radon-Zerlegung. Wir betrachten  $\mathcal{X}$  als zweifarbige Punktmenge, wobei alle Punkte in  $\mathcal{X}_1$  in eine Farbe und alle Punkte in  $\mathcal{X}_2$  in eine andere Farbe eingefärbt seien. Die Toleranz von  $\mathcal{X}$  ist genau dann  $t$ , wenn gilt:

$$\text{PertSEP}_0(\mathcal{X}) = \frac{t+1}{n}. \quad (4.6)$$

**Bemerkung 4.7.** Es handelt sich bei Satz 4.6 in etwa um [10, Theorem 1.2]. Ein Unterschied ist, dass [10, Theorem 1.2] die Rückrichtung (“ $\Leftarrow$ ”) nicht formuliert. Es fehlt also die Aussage, dass aus Separierbarkeit auch Toleranz folgt. Diese Aussage ist aber für einen der nachfolgenden Beweise essenziell. Wir orientieren uns deshalb bei der Formulierung von Satz 4.6 an [4, Theorem 3]. Es handelt sich hierbei letztlich um die gleiche Aussage in etwas anderer Sprechweise. Außerdem ergänzen wir das “+ 1” im Zähler von (4.6) (vergleiche [4, Theorem 3]).

*Beweis von Satz 4.6.* „ $\Rightarrow$ “: Es sei zunächst  $\mathcal{X}$  eine genau  $t$ -tolerante Radon-Zerlegung. Das heißt, die Radon-Zerlegung ist zwar  $t$ -tolerant, aber es gibt  $t+1$  Punkte  $\mathbf{x}_1, \dots, \mathbf{x}_{t+1}$  in  $\mathcal{X}$ , sodass  $\text{conv}(\mathcal{X}_1 \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_{t+1}\}) \cap \text{conv}(\mathcal{X}_2 \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_{t+1}\}) = \emptyset$  gilt. Im Speziellen gibt es also eine Hyperebene  $h$ , die  $\mathcal{X}_1 \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_{t+1}\}$  und  $\mathcal{X}_2 \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_{t+1}\}$  trennt. Alle Punkte auf einer Seite von  $h$  haben also eine Farbe und alle Punkte auf der anderen Seite haben eine andere Farbe. Wir können nun jeden der Punkte  $\mathbf{x}_1, \dots, \mathbf{x}_{t+1}$  auf die seiner Farbe entsprechende Seite von  $h$  bewegen. Das führt zu einer neuen Menge  $\mathcal{X}'$ , die mittels  $h$  separierbar ist. Wir schließen deshalb, dass höchstens  $t+1$  Punkte bewegt werden müssen, um  $\mathcal{X}$  separierbar zu machen:  $\text{PertSEP}_0(\mathcal{X}) \leq \frac{t+1}{n}$ .

Wir zeigen nun, dass auch mindestens  $t+1$  Punkte bewegt werden müssen, um Separierbarkeit zu erreichen. Dafür nehmen wir an, dass es schon ausreicht,  $t$  Punkte zu bewegen, um Separierbarkeit zu erreichen. Es seien also  $\mathbf{x}_1, \dots, \mathbf{x}_t$  Punkte in  $\mathcal{X}$ , die bewegt werden müssen, um  $\mathcal{X}$  separierbar zu machen. Dann gibt es eine Hyperebene, die  $\mathcal{X}_1 \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_t\}$  und  $\mathcal{X}_2 \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_t\}$  trennt. Daraus folgt aber  $\text{conv}(\mathcal{X}_1 \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_t\}) \cap \text{conv}(\mathcal{X}_2 \setminus \{\mathbf{x}_1, \dots, \mathbf{x}_t\}) = \emptyset$ . Das steht im Widerspruch zur  $t$ -Toleranz der Radon-Zerlegung. Wir haben also auch  $\text{PertSEP}_0(\mathcal{X}) \geq \frac{t+1}{n}$ .

„ $\Leftarrow$ “: Es gelte  $\text{PertSEP}_0(\mathcal{X}) = \frac{t+1}{n}$ . Da mindestens  $t+1$  Punkte entfernt werden müssen, um  $\mathcal{X}$  separierbar zu machen, gilt für beliebige  $\mathcal{C} \subset \mathcal{X}$  mit  $|\mathcal{C}| \leq t$ , dass der Schnitt aus  $\text{conv}(\mathcal{X}_1 \setminus \mathcal{C})$  und  $\text{conv}(\mathcal{X}_2 \setminus \mathcal{C})$  nichtleer ist. Die Toleranz ist demnach mindestens  $t$ . Andererseits kann  $\mathcal{X}$  nicht  $(t+1)$ -tolerant sein. Denn dann wäre  $\text{conv}(\mathcal{X}_1 \setminus \mathcal{C}) \cap \text{conv}(\mathcal{X}_2 \setminus \mathcal{C}) \neq \emptyset$  für beliebige Mengen  $\mathcal{C} \subset \mathcal{X}$  mit  $|\mathcal{C}| = t+1$  erfüllt.

Das würde jedoch implizieren, dass  $\mathcal{X}_1 \setminus \mathcal{C}$  und  $\mathcal{X}_2 \setminus \mathcal{C}$  für kein  $\mathcal{C}$  durch eine Hyperebene getrennt werden kann, was aber bei  $\text{PertSEP}_0(\mathcal{X}) = \frac{t+1}{n}$  schon der Fall sein muss.  $\square$

Es sei  $\mathcal{X} = \mathcal{X}_1 \cup \mathcal{X}_2$  wie in Satz 4.6. Da entweder  $\mathcal{X}_1$  oder  $\mathcal{X}_2$  höchstens  $|\mathcal{X}|/2$  Punkte enthält und außerdem höchstens  $\min(|\mathcal{X}_1|, |\mathcal{X}_2|)$  Punkte bewegt werden müssen, um  $\mathcal{X}$  separierbar zu machen, ist  $1/2$  eine obere Schranke an  $\text{PertSEP}_0(\mathcal{X})$ . Das folgende Korollar 4.9 zeigt, dass diese Schranke im Modell der zufälligen Zuordnung (Definition 2.11) asymptotisch (mit der Anzahl der generierten Punkte) erreicht wird.

**Beobachtung 4.8.** *Wir setzen bei festen  $m, d \in \mathbb{N}$  und  $\delta > 0$  die Toleranz  $t = (1-\delta)k/m$  in Satz 3.2 ein und erhalten die Bedingung*

$$\delta k/m \geq \sqrt{\frac{1}{2} \left[ (d+1)(m-1)k \ln(km) + k \ln\left(\frac{1}{\varepsilon}\right) \right]}. \quad (4.7)$$

Da die linke Seite bezüglich  $k$  schneller wächst als die rechte, finden wir für beliebiges  $\varepsilon > 0$  stets ein  $k$ , sodass die Bedingung (4.7) erfüllt ist. Deshalb konvergiert die Wahrscheinlichkeit, dass eine zufällige Zuordnung (Definition 2.11) von  $k$  Punkten in  $m$  Farben eine Toleranz von mindestens  $(1-\delta)k/m$  hat, für  $k \rightarrow \infty$  gegen 1.

Gleichzeitig muss bei einer Aufteilung von  $k$  Punkten in  $m$  Farben eine Farbmenge dabei sein, die höchstens  $k/m$  Punkte enthält (wenn jede der  $m$  Farbmengen mehr als  $k/m$  Punkte hätte, gäbe es mehr als  $k$  Punkte). Die Toleranz kann deshalb im Allgemeinen nicht größer als  $k/m$  sein. Die vorangehende Überlegung sagt also, dass die Toleranz asymptotisch ( $k \rightarrow \infty$ ) größtmöglich ist. Das kann nun mit Satz 4.6 zu folgendem Korollar kombiniert werden (beachte  $m = 2$ ).

**Korollar 4.9** (Corollary 2.4). *Es sei  $D$  eine Verteilung auf  $\mathbb{R}^d$  und  $(\mathcal{R}_{2,k,D})_{k \in \mathbb{N}}$  eine Folge von zufällig zugeordneten (Definition 2.11) Punktmengen. Für jedes beliebige  $\varepsilon > 0$  gilt  $|\text{PertSEP}_0(\mathcal{R}_{2,k,D}) - 1/2| < \varepsilon$  mit hoher Wahrscheinlichkeit.*

Diese Aussage wird im folgenden Satz 4.10 auf den Fall von mehr als  $m = 2$  Farben erweitert:  $\text{PertSEP}_0$  ist für jedes Farbpaar asymptotisch größtmöglich.

**Satz 4.10** (Theorem 3.1). *Es sei  $\varepsilon > 0$  vorgegeben. Für jede Verteilung  $D$  auf  $\mathbb{R}^d$  und jede Folge  $(\mathcal{R}_{m,k,D})_{k \in \mathbb{N}}$  von zufällig zugeordneten (Definition 2.11) Punktmengen  $\mathcal{R}_{m,k,D} = \{\mathcal{X}_1, \dots, \mathcal{X}_m\}$  gilt*

$$\lim_{k \rightarrow \infty} \left( \min_{\mathcal{X}_i, \mathcal{X}_j \in \mathcal{R}_{m,k,D}} \text{PertSEP}_0(\mathcal{X}_i \cup \mathcal{X}_j) \right) = 1/2 \quad (4.8)$$

mit hoher Wahrscheinlichkeit.

*Beweis.* Es seien  $m, d, \delta$  fest vorgegeben. Mit  $T(k)$  bezeichnen wir das Ereignis, dass eine zufällige Zuordnung (Definition 2.11) von  $k$  Punkten im  $\mathbb{R}^d$  in  $m$  Farben eine Toleranz von mindestens  $(1-\delta)k/m$  hat. Nach Satz 3.2 (siehe auch Beobachtung 4.8) ist

$\lim_{k \rightarrow \infty} \mathbb{P}(T(k)) = 1$ . Zu einem festen  $\zeta > 0$  existiert deshalb ein  $\tilde{k}$ , sodass  $\mathbb{P}(T(\tilde{k})) > 1 - \zeta$  erfüllt ist.

Für fest vorgegebene  $m$  und  $\varepsilon$  sei  $E_i(k)$  das Ereignis, dass eine zufällige Zuordnung (Definition 2.11) von  $k$  Punkten in  $m$  Farben zwischen  $(1 - \varepsilon)k/m$  und  $(1 + \varepsilon)k/m$  Punkte der Farbe  $i \in [m]$  hat. Nach Bemerkung 2.17 (Gesetz der großen Zahlen) gilt

$$\lim_{k \rightarrow \infty} \mathbb{P}(E_i(k)) = 1 \quad \text{für jedes } i \in [m]. \quad (4.9)$$

Wir können nun die Wahrscheinlichkeit dafür, dass alle Ereignisse gleichzeitig eintreten, ebenfalls abschätzen. Es gilt nämlich mit Bemerkung 2.16

$$\mathbb{P}\left(\bigcap_{i \in [m]} E_i(k)\right) = 1 - \mathbb{P}\left(\left(\bigcap_{i \in [m]} E_i(k)\right)^{\complement}\right) = 1 - \mathbb{P}\left(\bigcup_{i \in [m]} E_i(k)^{\complement}\right) \quad (4.10)$$

$$\geq 1 - \sum_{i \in [m]} \mathbb{P}\left(E_i(k)^{\complement}\right) = 1 - \sum_{i \in [m]} (1 - \mathbb{P}(E_i(k))) \quad (4.11)$$

$$= 1 - m + \sum_{i \in [m]} \mathbb{P}(E_i(k)) \quad (4.12)$$

Wegen (4.9) konvergiert (4.12) für  $k \rightarrow \infty$  gegen 1. Die Wahrscheinlichkeit, dass alle Ereignisse  $E_1(k), \dots, E_m(k)$  gleichzeitig eintreten, konvergiert also für  $k \rightarrow \infty$  ebenfalls gegen 1. Deshalb existiert für festes  $\varepsilon_2 > 0$  ein  $k' \in \mathbb{N}$ , sodass alle Ereignisse  $E_1(k'), \dots, E_m(k')$  mit einer Wahrscheinlichkeit von mindestens  $1 - \varepsilon_2$  gleichzeitig eintreten.

Mit einer ähnlichen Argumentation wie in (4.10) bis (4.12) schließen wir, dass alle Ereignisse  $T(k)$  und  $E_1(k), \dots, E_m(k)$  mit einer Wahrscheinlichkeit von mindestens  $1 - \zeta - \varepsilon_2$  gleichzeitig eintreten, wenn  $k \geq \max(\tilde{k}, k')$  gewählt wird. Das heißt, jede Farbmenge hat höchstens  $(1 + \varepsilon)k/m$  Punkte und gleichzeitig hat  $\mathcal{R}_{m,k,D}$  eine Toleranz von mindestens  $(1 - \delta)k/m$ .

Als Folge davon gilt mit einer Wahrscheinlichkeit von mindestens  $1 - \zeta - \varepsilon_2$ , dass jedes Farbpaar  $\mathcal{X}_i, \mathcal{X}_j \in \mathcal{R}_{m,k,D}$  in Summe höchstens  $2(1 + \varepsilon)k/m$  Punkte hat. Da die Toleranz von  $\mathcal{X}_i \cup \mathcal{X}_j$  durch die Toleranz von  $\bigcup_{i \in [m]} \mathcal{X}_i$  nach unten beschränkt wird, ist jedes Farbpaar eine Radon-Zerlegung mit einer Toleranz von mindestens  $(1 - \delta)k/m$ . Dank Satz 4.6 gilt nun mit einer Wahrscheinlichkeit von mindestens  $1 - \zeta - \varepsilon_2$ :

$$\text{PertSEP}_0 \geq \frac{(1 - \delta)\frac{k}{m} + 1}{2(1 + \varepsilon)\frac{k}{m}} = \frac{1 - \delta + \frac{m}{k}}{2(1 + \varepsilon)}. \quad (4.13)$$

Einerseits gilt  $\lim_{k \rightarrow \infty} \frac{m}{k} = 0$  und andererseits waren  $\delta, \zeta, \varepsilon, \varepsilon_2$  beliebig gewählt. Damit ist der Beweis abgeschlossen.  $\square$

### 4.3 Maximum-Likelihood-Schätzung

Wir beschäftigen uns auch beim letzten Anwendungsbeispiel mit farbigen Punktmengen. Es sei angemerkt, dass im Kontext der nachfolgend vorgestellten logistischen Regression

in der Regel eher von „Klassen“ als von „Farben“ gesprochen wird. Wir benutzen jedoch die bisher eingeführten Bezeichnungen und sprechen weiterhin von Farben.

Wir beschreiben in Abschnitt 4.3.1 zunächst ein Farberkennungsproblem (ein Klassifizierungsproblem) und zwei Lösungsstrategien für dieses Problem. Beide Lösungsstrategien bauen auf der binären logistischen Regression auf, welche das Thema von Abschnitt 4.3.2 ist. Der Zusammenhang zu Tverberg-Sätzen wird klar, sobald wir geklärt haben, wie die Regressionskoeffizienten der binären logistischen Regression typischerweise gewählt werden. Abschnitt 4.3.3 stellt diesen Zusammenhang genauer dar.

### 4.3.1 Farberkennung bei mehreren Farben

Es sei  $\mathcal{X} \times \mathcal{Y} \subset \mathbb{R}^d \times [m]$  eine  $N$ -elementige und  $m$ -farbigen Punktmenge, deren Elemente wir als  $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)$  schreiben. Für die nachfolgende Darstellung ist es instruktiv, sich unter  $\mathcal{X} \times \mathcal{Y}$  einen bekannten Datensatz wie CIFAR-10 [9] vorzustellen. In Fall von CIFAR-10 ist jeder Punkt  $\mathbf{x} \in \mathcal{X}$  ein  $32 \times 32$  Pixel großes Bild ( $n = 32^2$ ), welches genau eines der folgenden zehn Objekte darstellt: Flugzeug, Auto, Vogel, Katze, Hirsch, Hund, Frosch, Pferd, Schiff, Lastwagen. Diese zehn Objekte entsprechen  $m = 10$  Farben. Demnach besteht jedes Element  $(\mathbf{x}, y) \in \mathcal{X} \times \mathcal{Y}$  also aus einem Bild  $\mathbf{x}$  und der zugehörigen Farbe  $y$ .

Wir sind nun an einer Strategie interessiert, die bei Vorgabe eines neuen Bildes  $\mathbf{x} \in \mathbb{R}^d$ , das nicht in  $\mathcal{X}$  liegt und deshalb noch keiner Farbe zugeordnet ist, entscheidet, welche der zehn Farben vorliegt. Eine andere Formulierung wäre, dass die Strategie das im Bild dargestellte Objekt erkennen soll. Wir beschreiben in Abschnitt 4.3.2, wie dieses Problem im Fall von zwei Farben mithilfe der binären logistischen Regression gelöst werden kann. Mit diesem Wissen können wir den Fall  $m > 2$  mit einer der beiden Strategien „Eins-gegen-Alle“ oder „Eins-gegen-Eins“ behandeln.

*Eins-gegen-Alle:* Mit der in Abschnitt 4.3.2 beschriebenen logistischen Regression finden wir  $m$  Funktionen  $f_1, \dots, f_m$  mit der folgenden Eigenschaft. Die Funktion  $f_i, i \in [m]$ , gibt zu einem gegebenen Punkt  $\mathbf{x}$  die Wahrscheinlichkeit an, dass  $\mathbf{x}$  die Farbe  $i$  hat. Die Funktion  $f_i$  hat demnach die Aufgabe, zwischen einer Farbe  $i$  und allen anderen Farben in  $[m] \setminus \{i\}$  zu unterscheiden (daher „Eins-gegen-Alle“). Wir werten alle  $m$  Funktionen aus und ordnen dem Punkt  $\mathbf{x}$  eine Farbe  $y$  zu, die zur höchsten Wahrscheinlichkeit führt:  $f_y(\mathbf{x}) = \max_{i \in [m]} f_i(\mathbf{x})$ .

*Eins-gegen-Eins.* Mit der in Abschnitt 4.3.2 beschriebenen logistischen Regression finden wir zu jedem der  $\binom{m}{2}$  möglichen Farbpaare  $(i, j)$  mit  $i, j \in [m]$  und  $i \neq j$  eine Funktion, die entscheidet, ob ein gegebener Punkt  $\mathbf{x}$  die Farbe  $i$  oder die Farbe  $j$  hat. Die Funktion hat demnach die Aufgabe, zwischen einer Farbe  $i$  und einer anderen Farbe  $j$  zu unterscheiden (daher „Eins-gegen-Eins“). Wir werten alle  $\binom{m}{2}$  Funktionen aus und ordnen dem Punkt  $\mathbf{x}$  eine Farbe  $y$  zu, für welche die  $\binom{m}{2}$  Funktion am häufigsten entschieden haben.

### 4.3.2 Binäre logistische Regression

Wir betrachten eine Möglichkeit, das Farberkennungsproblem im Fall von zwei Farben zu lösen. Es liegen  $N_1$  Punkte  $\mathbf{x}_1, \dots, \mathbf{x}_{N_1} \in \mathbb{R}^d$  der Farbe 1 und  $N_2$  Punkte  $\mathbf{x}_{N_1+1}, \dots, \mathbf{x}_{N_1+N_2} \in \mathbb{R}^d$  der Farbe 2 vor. Die logistische Regression führt Regressionskoeffizienten  $\boldsymbol{\beta} \in \mathbb{R}^{d+1}$  ein und modelliert die Wahrscheinlichkeit, dass der Punkt  $\mathbf{x}_i$ ,  $i \in [N_1 + N_2]$ , die Farbe 1 hat als [8]

$$\mathbb{P}_{\boldsymbol{\beta}}(y_i = 1 | \mathbf{x}_i) = \frac{1}{1 + \exp(-\beta_0 - \mathbf{x}_i^T \boldsymbol{\beta}')}, \quad (4.14)$$

wobei wir  $\boldsymbol{\beta} = (\beta_0, \boldsymbol{\beta}'^T)^T$  mit  $\beta_0 \in \mathbb{R}$  und  $\boldsymbol{\beta}' \in \mathbb{R}^d$  schreiben. Es gilt dann entsprechend  $\mathbb{P}_{\boldsymbol{\beta}}(y_i = 2 | \mathbf{x}_i) = 1 - \mathbb{P}_{\boldsymbol{\beta}}(y_i = 1 | \mathbf{x}_i)$ . Um die Regressionskoeffizienten  $\boldsymbol{\beta}$  zu bestimmen, betrachten wir die sogenannte Log-Likelihood-Funktion [8]:

$$l(\boldsymbol{\beta}) = \sum_{i=1}^{N_1} \ln(\mathbb{P}_{\boldsymbol{\beta}}(y_i = 1 | \mathbf{x}_i)) + \sum_{j=N_1+1}^{N_1+N_2} \ln(\mathbb{P}_{\boldsymbol{\beta}}(y_j = 2 | \mathbf{x}_j)). \quad (4.15)$$

Jeder Maximierer  $\hat{\boldsymbol{\beta}}$  dieser Funktion wird Maximum-Likelihood-Schätzung von  $\boldsymbol{\beta}$  genannt. Die Intuition ist, die Regressionskoeffizienten so zu wählen, dass die Wahrscheinlichkeit, den  $N_1 + N_2$  gegebenen Punkten ihre zugehörige Farbe korrekt zuzuordnen, maximiert wird. Falls ein Maximierer  $\hat{\boldsymbol{\beta}}$  existiert, sagen wir auch „die Maximum-Likelihood-Schätzung zum gegebenen Farbpaar existiert“.

Sobald die Maximum-Likelihood-Schätzung  $\hat{\boldsymbol{\beta}}$  bestimmt ist, können wir für einen neuen Punkt  $\mathbf{x}$  die Wahrscheinlichkeit, dass er die Farbe  $y = 1$  hat, bestimmen:  $\mathbb{P}_{\hat{\boldsymbol{\beta}}}(y = 1 | \mathbf{x})$ . Bei der in Abschnitt 4.3.1 beschriebenen Strategie Eins-gegen-Alle sind wir nur an dieser Wahrscheinlichkeit interessiert. Bei der Strategie Eins-gegen-Eins sind wir an einer Entscheidung interessiert: Wir entscheiden uns für die Farbe  $y = 1$ , falls  $\mathbb{P}_{\hat{\boldsymbol{\beta}}}(y = 1 | \mathbf{x}) \geq 1/2$  gilt, und für die Farbe 2 andernfalls.

Es gibt einen Zusammenhang zu Tverberg-Sätzen, weil die Maximum-Likelihood-Schätzung  $\hat{\boldsymbol{\beta}}$  existiert, falls

$$\text{conv}(\{\mathbf{x}_1, \dots, \mathbf{x}_{N_1}\}) \cap \text{conv}(\{\mathbf{x}_{N_1+1}, \dots, \mathbf{x}_{N_1+N_2}\}) \neq \emptyset \quad (4.16)$$

gilt. Die konvexen Hüllen der beiden Farben müssen sich also schneiden. Da wir in Abschnitt 3 Wahrscheinlichkeiten für das Eintreten eines solchen Ereignisses bestimmt haben, können wir Aussagen über die Wahrscheinlichkeit, dass die Maximum-Likelihood-Schätzung existiert, machen.

### 4.3.3 Existenz der Maximum-Likelihood-Schätzung

Das folgende Korollar 4.11 macht eine Aussage darüber, wann die Maximum-Likelihood-Schätzung im Modell der zufälligen Zuordnung (Definition 2.11) existiert. Wir konzentrieren uns auf die Strategie Eins-gegen-Eins, da dadurch auch die Strategie Eins-gegen-Alle abgedeckt ist (siehe Bemerkung 4.13).

Um das Farberkennungsproblem zu einer  $m$ -farbigen Punktmenge zu lösen, muss nach Abschnitt 4.3.1 die Maximum-Likelihood-Schätzung zu jedem Farbpaar existieren. Nach der Bemerkung um (4.16) existiert die Maximum-Likelihood-Schätzung zu einem gegebenen Farbpaar, falls sich die zugehörigen konvexen Hüllen schneiden. Es soll deshalb das Ereignis eintreten, dass sich die konvexen Hüllen aller Farbpaare jeweils schneiden. Diese Bedingung ist auf jeden Fall erfüllt, wenn eine  $m$ -Tverberg-Zerlegung vorliegt. Korollar 4.11 gibt eine Antwort auf die Frage, wie viele Punkte nötig sind, damit dieses Ereignis mit hoher Wahrscheinlichkeit eintritt.

Die erste Aussage von Korollar 4.11 behandelt den Fall einer Folge von Punktfolgen, bei der sowohl die Anzahl der Farben als auch die Anzahl der Punkte mit dem Folgenindex wächst. Für die zweite Aussage ist die Anzahl der Farben fest und nur die Anzahl der Punkte wächst mit dem Folgenindex.

**Korollar 4.11** (Corollary 1.1). *Es sei  $\varepsilon > 0$  vorgegeben und es sei  $D$  eine Wahrscheinlichkeitsverteilung auf  $\mathbb{R}^d$ , die um einen Punkt  $\mathbf{p} \in \mathbb{R}^d$  in dem Sinne ausgeglichen ist, dass jede Hyperebene durch  $\mathbf{p}$  die Verteilung  $D$  in zwei Mengen gleichen Maßes aufteilt.*

1. *Es sei  $(\mathcal{R}_{m,f(m),D})_{m \in \mathbb{N}}$  eine Folge  $m$ -farbiger zufällig zugeordneter (Definition 2.11) Punktfolgen  $\mathcal{R}_{m,f(m),D}$  mit  $f(m)$  Elementen. Es bezeichne  $X_m$  das Ereignis, dass die Maximum-Likelihood-Schätzung zu jedem Farbpaar in  $\mathcal{R}_{m,f(m),D}$  existiert. Dann tritt  $X_m$  mit hoher Wahrscheinlichkeit ein sofern*

$$f(m) \gg (1 + \varepsilon)m \log_2(m) \ln(\ln(m)). \quad (4.17)$$

2. *Es seien  $m, \delta > 0$  vorgegeben und es sei  $(\mathcal{R}_{m,g(t),D})_{t \in \mathbb{N}}$  eine Folge  $m$ -farbiger zufällig zugeordneter (Definition 2.11) Punktfolgen  $\mathcal{R}_{m,g(t),D}$  mit  $g(t)$  Elementen. Es bezeichne  $Y_t$  das Ereignis, dass die Maximum-Likelihood-Schätzung zu jedem Farbpaar in  $\mathcal{R}_{m,g(t),D}$  sogar dann existiert, wenn  $t - 1$  beliebige Punkte entfernt werden. Dann tritt  $Y_t$  mit hoher Wahrscheinlichkeit ein sofern*

$$g(t) \gg (1 + \delta)m(t + 1). \quad (4.18)$$

**Bemerkung 4.12.** *Der Grund, weshalb Aussage 2 von Korollar 4.11 leicht anders formuliert ist als Aussage 2 von [10, Corollary 1.1], ist das “+ 1” im Zähler von (4.6), siehe auch Bemerkung 4.7. Da die Beweise der Aussagen von Korollar 4.11 auf Aussagen zurückgreifen, die für ausgeglichene Verteilungen gelten, haben wir außerdem auch Korollar 4.11 für ausgeglichene Verteilungen formuliert.*

**Bemerkung 4.13.** *In beiden Fällen geht es um die Strategie Eins-gegen-Eins (vergleiche Abschnitt 4.3.1). Die Schranken in Korollar 4.11 gelten jedoch auch für die Strategie Eins-gegen-Alle (vergleiche ebenfalls Abschnitt 4.3.1). Das liegt letztlich daran, dass sich bei der Strategie Eins-gegen-Eins die konvexen Hüllen aller Farbpaare jeweils schneiden müssen. Das impliziert im Speziellen jedoch, dass man eine beliebige Farbe betrachten kann, und garantiert ist, dass sich deren konvexe Hülle mit der konvexen Hülle aller verbleibenden Punkte (der anderen Farben) schneidet. Und damit ist die Strategie Eins-gegen-Alle abgedeckt.*

*Beweis von Aussage 1.* Nach Korollar 3.9 (Aussage 3) sind die Punktmenge  $\mathcal{R}_{m,f(m),D}$  mit hoher Wahrscheinlichkeit  $m$ -Tverberg-Zerlegungen. Da die Maximum-Likelihood-Schätzung für jedes Farbpaar einer  $m$ -Tverberg-Zerlegung existiert, existiert sie mit hoher Wahrscheinlichkeit also auch für jedes Farbpaar der Menge  $\mathcal{R}_{m,f(m),D}$ .  $\square$

*Beweis von Aussage 2.* Wir schreiben  $\mathcal{R}_{m,g(t),D} = \{\mathcal{X}_1, \dots, \mathcal{X}_m\}$ . Wir müssen dafür sorgen, dass jedes Farbpaar  $\mathcal{X}_i, \mathcal{X}_j$  der Punktmenge  $\mathcal{R}_{m,g(t),D}$  mit hoher Wahrscheinlichkeit eine  $t$ -tolerante Radon-Zerlegung ist. (Denn dann können problemlos  $t-1$  Punkte entfernt werden und die Maximum-Likelihood-Schätzung zu jedem Farbpaar existiert trotzdem.) Nach Satz 4.10 gilt

$$\lim_{t \rightarrow \infty} \left( \min_{\mathcal{X}_i, \mathcal{X}_j \in \mathcal{R}_{m,g(t),D}} \text{PertSEP}_0(\mathcal{X}_i \cup \mathcal{X}_j) \right) = 1/2 \quad (4.19)$$

mit hoher Wahrscheinlichkeit. Für beliebige  $\delta, \varepsilon > 0$  finden wir deshalb ein  $N \in \mathbb{N}$ , sodass

$$\min_{\mathcal{X}_i, \mathcal{X}_j \in \mathcal{R}_{m,g(t),D}} \text{PertSEP}_0(\mathcal{X}_i \cup \mathcal{X}_j) \geq \frac{1 - \delta}{2} \quad (4.20)$$

mit Wahrscheinlichkeit  $1 - \frac{\varepsilon}{2}$  für alle  $t \geq N$  erfüllt ist. Nach Satz 4.6 stellt dann jedes Paar  $\mathcal{X}_i, \mathcal{X}_j$  mit Wahrscheinlichkeit  $1 - \frac{\varepsilon}{2}$  eine Radon-Zerlegung mit Toleranz  $t' \geq (|\mathcal{X}_i| + |\mathcal{X}_j|)^{\frac{1-\delta}{2}} - 1$  dar. Es verbleibt noch zu argumentieren, dass mit hoher Wahrscheinlichkeit  $t' \geq t$  gilt. Mit einer ähnlichen Argumentation wie in (4.10) bis (4.12) gibt es ein  $N' \geq N$ , sodass für  $t \geq N'$  mit Wahrscheinlichkeit  $1 - \frac{\varepsilon}{2}$  jedes  $\mathcal{X}_i$  mindestens  $\frac{|\mathcal{R}_{m,g(t),D}|}{m(1+\delta)}$  Punkte enthält. Die Wahl  $|\mathcal{R}_{m,g(t),D}| = g(t) \gg (1+\delta)m(t+1)$  führt also mit Wahrscheinlichkeit  $1 - \varepsilon$  zu

$$t' \geq (|\mathcal{X}_i| + |\mathcal{X}_j|) \frac{1 - \delta}{2} - 1 \geq \left( \frac{|\mathcal{R}_{m,g(t),D}|}{m(1+\delta)} + \frac{|\mathcal{R}_{m,g(t),D}|}{m(1+\delta)} \right) \frac{1 - \delta}{2} - 1 \quad (4.21)$$

$$\gg (t + 1 + t + 1) \frac{1 - \delta}{2} - 1 = (t + 1)(1 - \delta) - 1. \quad (4.22)$$

Die Aussage folgt, weil  $\delta$  und  $\varepsilon$  beliebig waren.  $\square$

Der Grund, weshalb wir bei der Existenz der Maximum-Likelihood-Schätzung in Aussage 2 von Korollar 4.11 auch daran interessiert sind, möglicherweise Punkt entfernen zu können, ist, dass manche Punkte eines gegebenen Datensatzes als Ausreißer (outlier) betrachtet werden können, welche wir möglicherweise zunächst entfernen wollen, bevor das Farberkennungsproblem (Klassifizierungsproblem) gelöst werden soll. So könnte es etwa im Beispiel des Datensatzes CIFAR-10 aus Abschnitt 4.3.1 sein, dass auf manchen Bildern nicht eindeutig zu erkennen ist, welches Objekt dargestellt ist (zum Beispiel Hund oder Katze). Es könnte dann vorteilhaft sein, diese Bilder zu entfernen bevor die Maximum-Likelihood-Schätzung bestimmt wird. Tabelle 3 fasst die Ergebnisse von Abschnitt 4.3 zusammen.

deterministische Version	stochastische Version	wahrscheinliche Existenz der Maximum-Likelihood Schätzung
Radon	Satz 3.1	Abschnitt 4.3.2
Tverberg	Satz 3.4, Satz 3.5	Korollar 4.11 (Aussage 1)
Radon mit Toleranz	Satz 3.7	Korollar 4.11 (Aussage 2)
Tverberg mit Toleranz	Satz 3.2, Satz 3.6, Korollar 3.9	Korollar 4.11 (Aussage 2)

Tabelle 3: Überblick über die Ergebnisse in Abschnitt 4.3. Die stochastischen Versionen in der mittleren Spalte können benutzt werden, um Wahrscheinlichkeiten dafür anzugeben, dass Ereignisse eintreten, welche zur Existenz der Maximum-Likelihood-Schätzung (rechte Spalte) führen.



## 5 Schluss

Wir haben zwei Modelle zur Erzeugung zufälliger Punktmengen studiert und untersucht, unter welchen Umständen dabei Tverberg-Zerlegungen entstehen. Das Modell der zufälligen Zuordnung (Definition 2.11) hat bei den vorgestellten Anwendungsbeispielen eine größere Rolle gespielt. Die Beweise der zugehörigen stochastischen Tverberg-Sätze bauen jedoch bei mehr als zwei Farben auf den Sätzen zum Modell der zufälligen Gleichverteilung (Definition 2.10) auf.

Die beiden Modelle sind zwar allgemein für stetige Wahrscheinlichkeitsverteilungen definiert, aber es war bei den meisten Sätzen wichtig, die Punkte gemäß einer ausgeglichenen Verteilung zu generieren, um dann in Beweisen über gleich wahrscheinliche Ereignisse argumentieren zu können. Es ist eine offene Frage, ob und wie sich die Sätze auf andere Verteilungen erweitern lassen.

Die beiden Modelle haben außerdem gemeinsam, dass erst alle Punkte generiert werden, um ihnen im Anschluss unabhängig Farben zuzuordnen. Es ist eine weitere offene Frage, ob und wie sich die präsentierten Ergebnisse auf den Fall erweitern lassen, dass die Farben nicht unabhängig verteilt werden.



## Literatur

- [1] M. Bazaraa, H. Sherali, C. Shetty, *Nonlinear programming: Theory and applications*, Wiley, 2006.
- [2] I. Bárány, *A generalization of Carathéodory's theorem*, Discrete Math., 1982.
- [3] I. Bárány, P. Soberón, *Tverberg's theorem is 50 years old: A survey*, Bull. of Amer. Math. Soc., 2018.
- [4] S. Bereg, M. Haghpanah, *Algorithms for Radon partitions with tolerance*, In: Algorithms and Discrete Applied Mathematics, Springer, 2020.
- [5] S. Boyd, L. Vandenberghe, *Convex optimization*, Cambridge University Press, 2004.
- [6] T. Cover, *Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition*, IEEE Trans. on Electron. Comput., 1964.
- [7] U. Faigle, W. Kern, G. Still, *Algorithmic principles of mathematical programming*, Springer, 2022.
- [8] D. Kleinbaum, M. Klein, *Logistic regression: A self-learning text*, Springer, 2010.
- [9] A. Krizhevsky, *Learning multiple layers of features from tiny images*, Technical Report, 2009.
- [10] J. De Loera, T. Hogan, *Stochastic Tverberg theorems with applications in multiclass logistic regression, separability, and centerpoints of data*, SIAM J. Math. Data Sci., 2020.
- [11] J. Matoušek, *Lectures on discrete geometry*, Springer, 2002.
- [12] F. Meunier, W. Mulzer, P. Sarrabezolles, Y. Stein, *The rainbow at the end of the line — A PPAD formulation of the colorful Carathéodory theorem with applications*, preprint, <https://arxiv.org/abs/1608.01921>, 2016.
- [13] W. Mulzer, D. Werner, *Approximating Tverberg points in linear time for any fixed dimension*, Discrete Comput. Geom., 2013.
- [14] J. Radon, *Mengen konvexer Körper, die einen gemeinsamen Punkt enthalten*, Math. Ann., 1921.
- [15] D. Rolnick, P. Soberón, *Algorithms for Tverberg's theorem via centerpoint theorems*, preprint, <https://arxiv.org/abs/1601.03083>, 2016.
- [16] S. Roman, *Advanced linear algebra*, Springer, 2005.
- [17] J. Roudneff, *Partitions of points into simplices with  $k$ -dimensional intersection. Part I: The conic Tverberg's theorem*, Europ. J. Combinatorics, 2001.

- [18] K. Sarkaria, *Tverberg's theorem via number fields*, Israel J. Math., 1992.
- [19] G. Schay, *Introduction to probability with statistical applications*, Birkhäuser, 2007.
- [20] P. Soberón, *Robust Tverberg and colourful Carathéodory results via random choice*, Combin. Probab. Comput., 2018.
- [21] P. Soberón, R. Strausz, *A generalization of Tverberg's theorem*, Discrete Comput. Geom., 2012.
- [22] H. Tverberg, *A generalization of Radon's theorem*, J. London Math. Soc., 1966.
- [23] H. Tverberg, *A generalization of Radon's theorem II*, Bull. Austral. Math. Soc., 1981.
- [24] H. Tverberg, S. Vrećica, *On generalizations of Radon's theorem and the ham sandwich theorem*, European J. Combin., 1993.
- [25] J. Wendel, *A problem in geometric probability*, Math. Scandinavia, 1962.
- [26] M. Zvager'skii, *An elementary proof of Tverberg's theorem*, J. of Math. Sci., 2008.