

# Statistische Mechanik des Zahlenaufteilungsproblems

Heiko Bauke

Heiko.Bauke@physik.uni-magdeburg.de

JAHRE **50** JAHRE



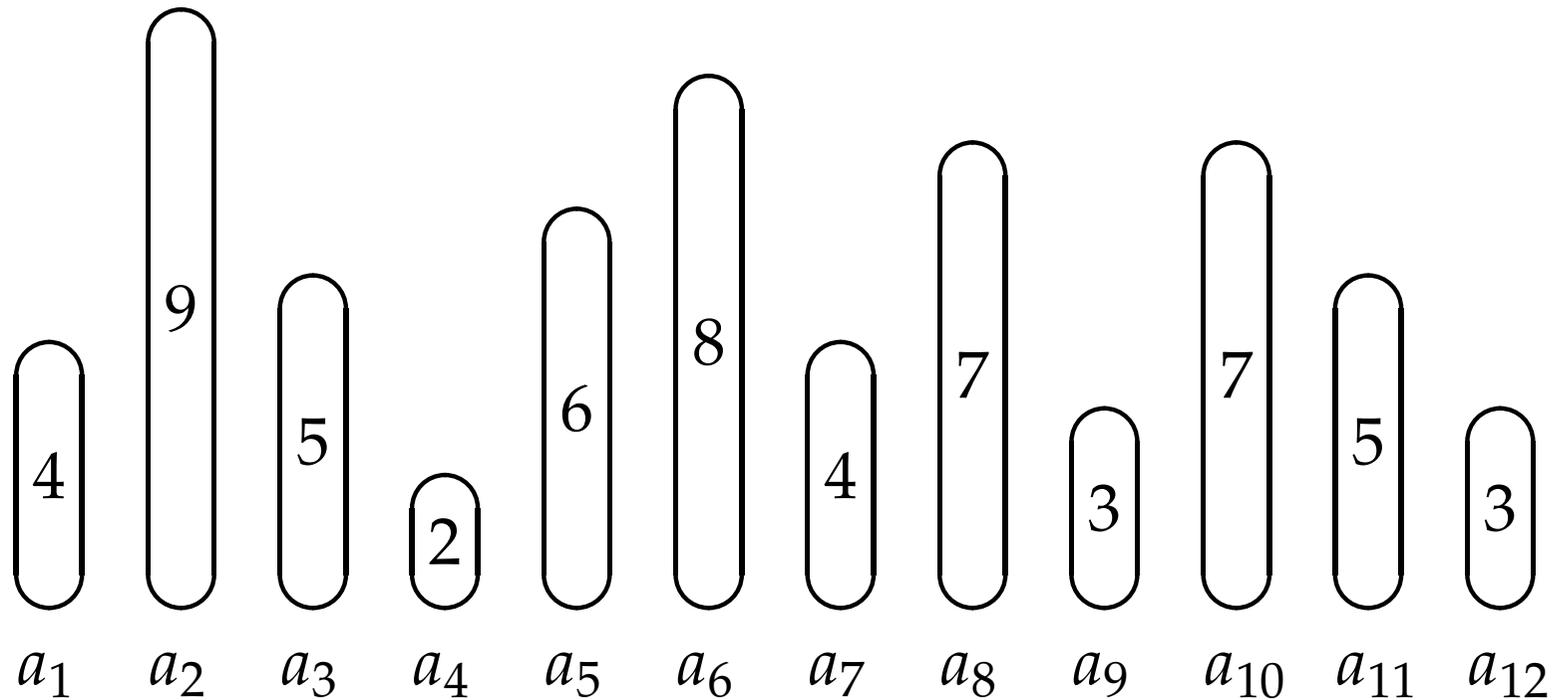
1953 - 2003

50 JAHRE HOCHSCHULSTANDORT

1993 - 2003

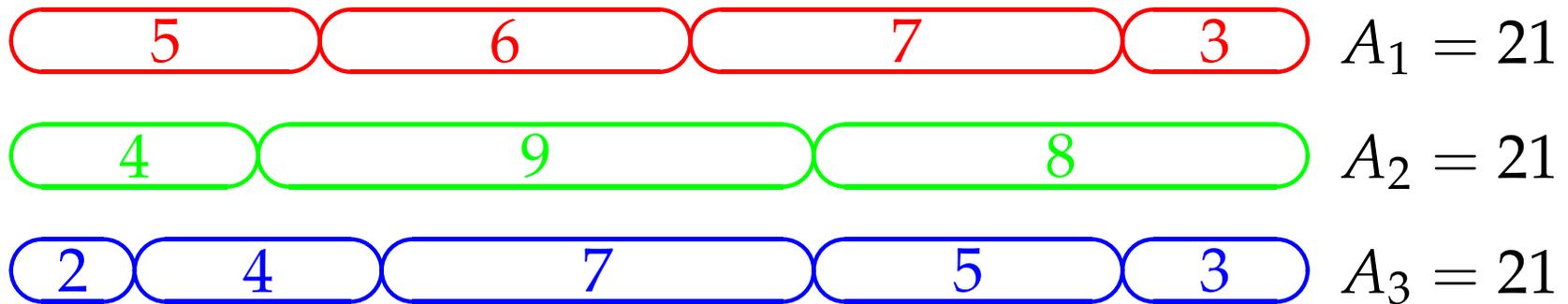
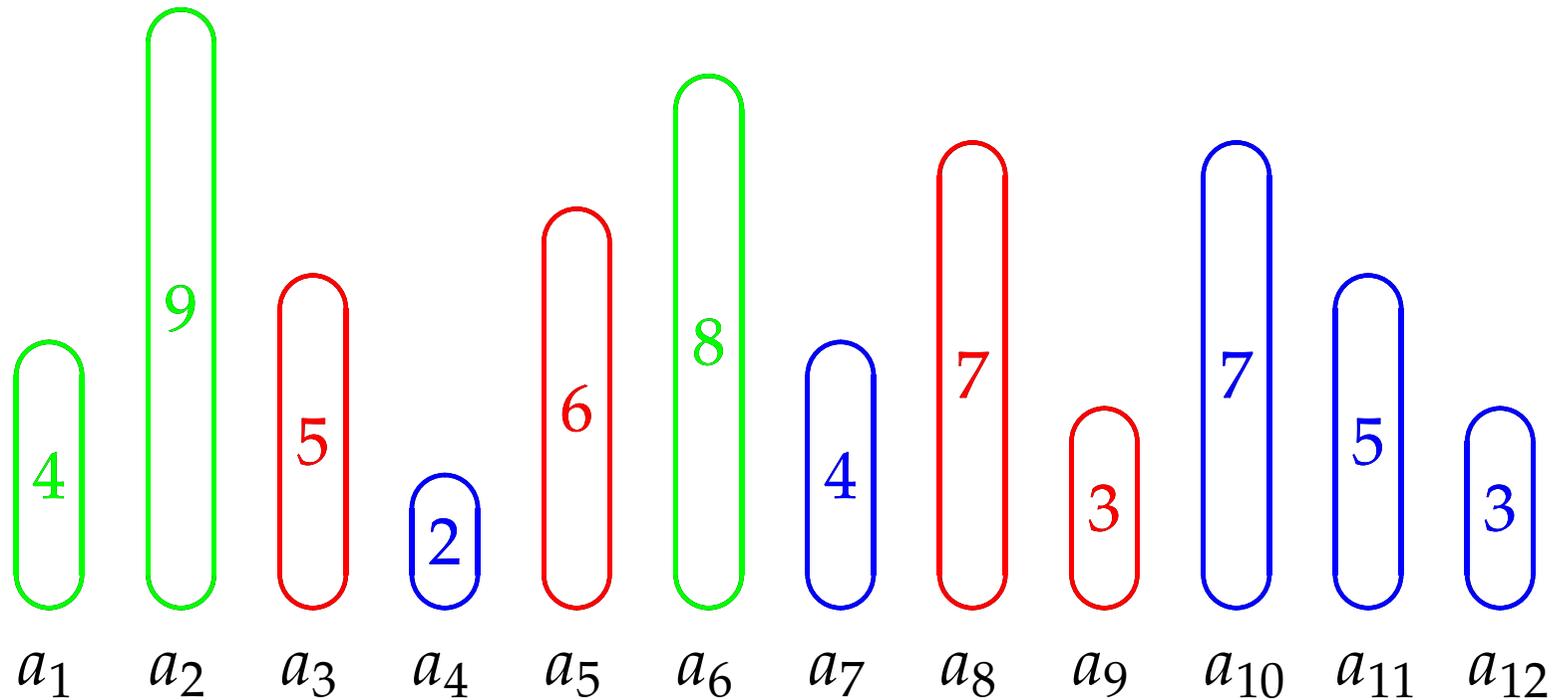
10 JAHRE OTTO-VON-GUERICKE-UNIVERSITÄT

# Das Zahlenaufteilungsproblem (NPP)

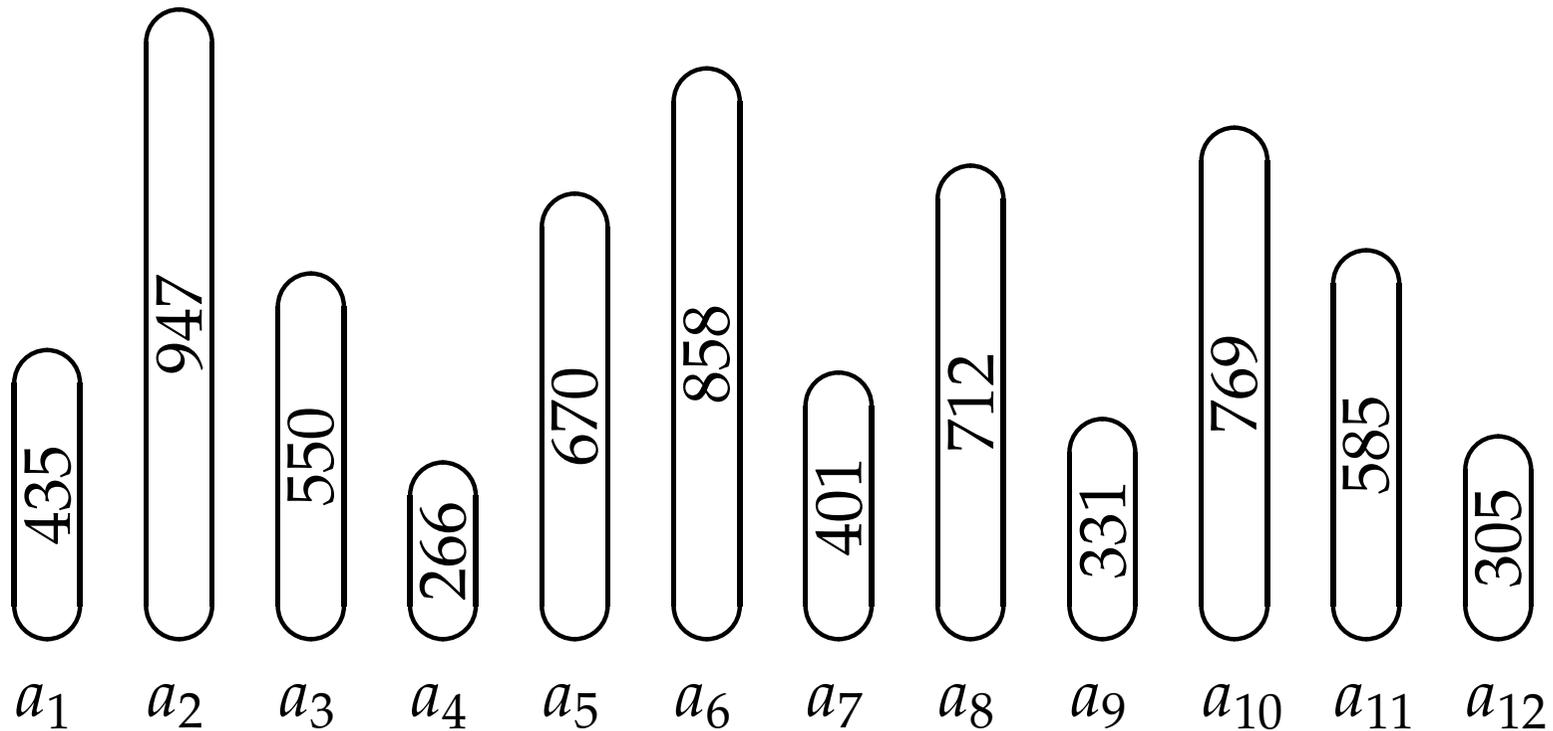


Lassen sich die Gewichte gleichmäßig (perfekt)  
auf 3 Mengen aufteilen?

# Das Zahlenaufteilungsproblem (NPP)

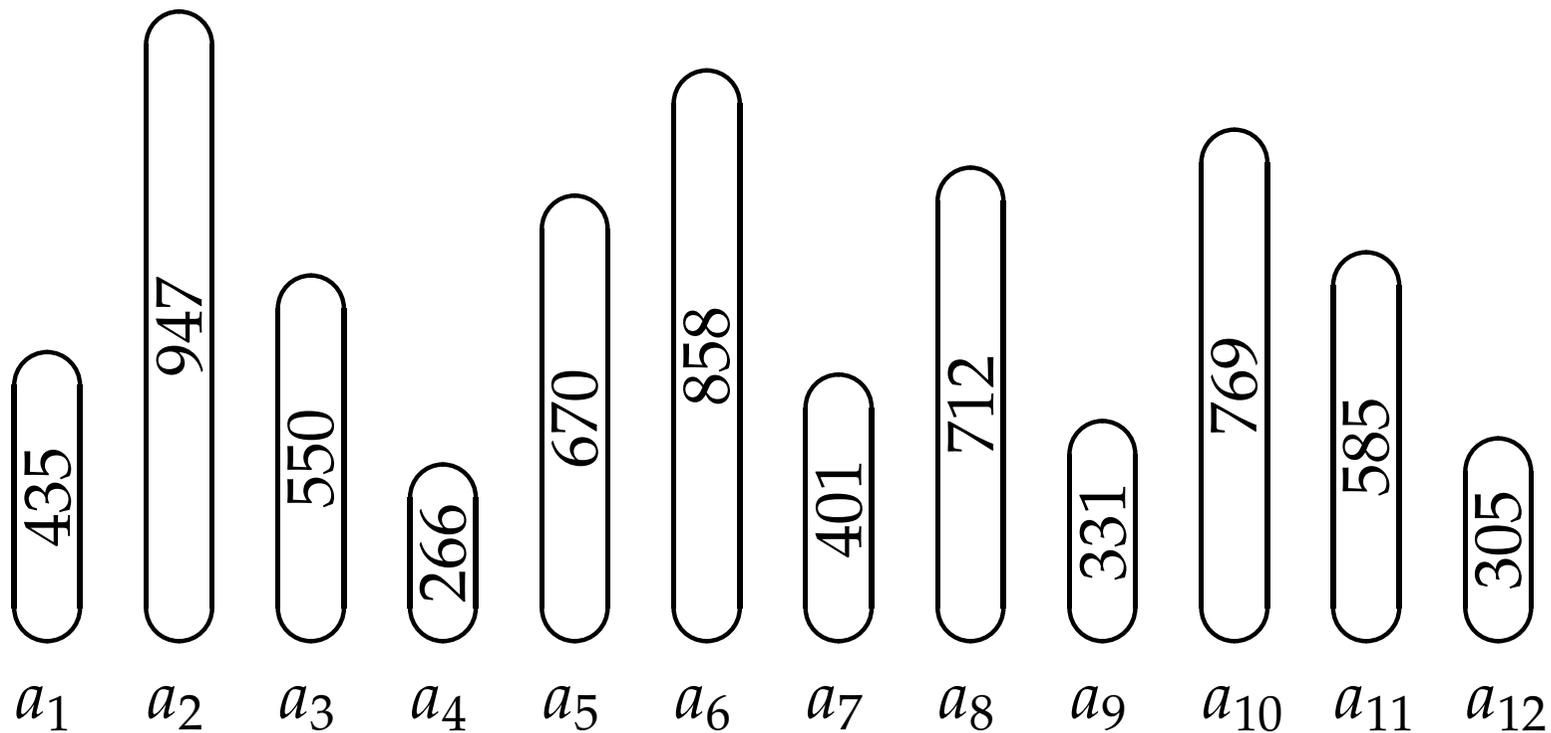


# Das Zahlenaufteilungsproblem (NPP)



Lassen sich auch diese Gewichte gleichmäßig  
(perfekt) auf 3 Mengen aufteilen?

# Das Zahlenaufteilungsproblem (NPP)



Lassen sich auch diese Gewichte gleichmäßig  
(perfekt) auf 3 Mengen aufteilen?

**Nein !**



# Das Zahlenaufteilungsproblem (NPP)



- $\mathcal{NP}$ -vollständiges Entscheidungsproblem

# Das Zahlenaufteilungsproblem (NPP)



- $\mathcal{NP}$ -vollständiges Entscheidungsproblem
- Rechenzeit wächst im **ungünstigsten** Fall exponentiell

# Das Zahlenaufteilungsproblem (NPP)



- $\mathcal{NP}$ -vollständiges Entscheidungsproblem
- Rechenzeit wächst im **ungünstigsten** Fall exponentiell
- einfach, falls Probleminstanz perfekte Partition besitzt

# Das Zahlenaufteilungsproblem (NPP)



- $\mathcal{NP}$ -vollständiges Entscheidungsproblem
- Rechenzeit wächst im **ungünstigsten** Fall exponentiell
- einfach, falls Problem Instanz perfekte Partition besitzt
- Wechsel in algorithmischer Komplexität  
⇒ **Phasenübergang**

# Das Zahlenaufteilungsproblem (NPP)



- $\mathcal{NP}$ -vollständiges Entscheidungsproblem
- Rechenzeit wächst im **ungünstigsten** Fall exponentiell
- einfach, falls Probleminstanz perfekte Partition besitzt
- Wechsel in algorithmischer Komplexität  
⇒ **Phasenübergang**
- Untersuchung des **typischen** Falls mittels statistischer Mechanik

# Warum ist NPP interessant?



Vom CS-Standpunkt:

- $\mathcal{NP}$ -vollständiges Entscheidungsproblem bzw.  $\mathcal{NP}$ -hartes Optimierungsproblem

# Warum ist NPP interessant?



Vom CS-Standpunkt:

- $\mathcal{NP}$ -vollständiges Entscheidungsproblem bzw.  $\mathcal{NP}$ -hartes Optimierungsproblem
- keine gute Heuristik bekannt



# Warum ist NPP interessant?

Vom CS-Standpunkt:

- $\mathcal{NP}$ -vollständiges Entscheidungsproblem bzw.  $\mathcal{NP}$ -hartes Optimierungsproblem
- keine gute Heuristik bekannt

Vom physikalischen Standpunkt:

- NPP auf frustrierten Antiferromagneten abbildbar (Spinglas)



# Warum ist NPP interessant?

Vom CS-Standpunkt:

- $\mathcal{NP}$ -vollständiges Entscheidungsproblem bzw.  $\mathcal{NP}$ -hartes Optimierungsproblem
- keine gute Heuristik bekannt

Vom physikalischen Standpunkt:

- NPP auf frustrierten Antiferromagneten abbildbar (Spinglas)
- mathematisch viel einfacher als andere Spinglasmodelle, keine Replica-Methode notwendig



# Warum ist NPP interessant?

Vom CS-Standpunkt:

- $\mathcal{NP}$ -vollständiges Entscheidungsproblem bzw.  $\mathcal{NP}$ -hartes Optimierungsproblem
- keine gute Heuristik bekannt

Vom physikalischen Standpunkt:

- NPP auf frustrierten Antiferromagneten abbildbar (Spinglas)
- mathematisch viel einfacher als andere Spinglasmodelle, keine Replica-Methode notwendig

Praktische Anwendungen:

- Scheduling, Chip-Design, Public-Key-Kryptographie

# Das Zahlenaufteilungsproblem (NPP)



Geben seien:

- $N$  zufällige Gewichte  $a_1, a_2, \dots, a_N$  mit  $a_i \in \mathbb{N}$  und  $q$  „Töpfe“

# Das Zahlenaufteilungsproblem (NPP)



Geben seien:

- $N$  zufällige Gewichte  $a_1, a_2, \dots, a_N$  mit  $a_i \in \mathbb{N}$  und  $q$  „Töpfe“

Aufgabe:

- Verteile die Gewichte  $a_1, a_2, \dots, a_N$  auf  $q$  disjunkte Teilmengen, so dass die Summen der Gewichte  $A_1, A_2, \dots, A_q$  in den Teilmengen möglichst gleich groß sind.



# Das Zahlenaufteilungsproblem (NPP)

Geben seien:

- $N$  zufällige Gewichte  $a_1, a_2, \dots, a_N$  mit  $a_i \in \mathbb{N}$  und  $q$  „Töpfe“

Aufgabe:

- Verteile die Gewichte  $a_1, a_2, \dots, a_N$  auf  $q$  disjunkte Teilmengen, so dass die Summen der Gewichte  $A_1, A_2, \dots, A_q$  in den Teilmengen möglichst gleich groß sind.

Problem:

- Wieviele **perfekte Partitionen** existieren?

$$\text{perfekte Partition} \iff A_j - \left\lfloor \frac{1}{q} \sum_{i=1}^N a_i \right\rfloor = \begin{cases} 1 & j \leq r \\ 0 & j > r \end{cases}$$

$$\text{mit } r \equiv \sum_{i=1}^N a_i \pmod{q}$$



- $q$  Teilmengen  $\hat{=}$   $q$  POTTS-Vektoren  $S_q = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_q\}$   
(Symmetrie!)



- $q$  Teilmengen  $\hat{=}$   $q$  POTTS-Vektoren  $S_q = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_q\}$   
(Symmetrie!)
- Partition  $\hat{=}$  walk auf  $(q - 1)$ -dim. Gitter in  $q$  Richtungen



- $q$  Teilmengen  $\hat{=}$   $q$  POTTS-Vektoren  $S_q = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_q\}$   
(Symmetrie!)
- Partition  $\hat{=}$  walk auf  $(q - 1)$ -dim. Gitter in  $q$  Richtungen
- Schrittweite: Gewicht  $a_i$



- $q$  Teilmengen  $\hat{=}$   $q$  POTTS-Vektoren  $S_q = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_q\}$   
(Symmetrie!)
- Partition  $\hat{=}$  walk auf  $(q - 1)$ -dim. Gitter in  $q$  Richtungen
- Schrittweite: Gewicht  $a_i$
- Richtung:  $\vec{s}_i \in S_q = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_q\}$



- $q$  Teilmengen  $\hat{=}$   $q$  POTTS-Vektoren  $S_q = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_q\}$  (Symmetrie!)
- Partition  $\hat{=}$  walk auf  $(q - 1)$ -dim. Gitter in  $q$  Richtungen
- Schrittweite: Gewicht  $a_i$
- Richtung:  $\vec{s}_i \in S_q = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_q\}$
- Kostenfunktion  $\hat{=}$  Abstand von Ursprung nach  $N$  Schritten

$$H(\vec{s}_1, \vec{s}_2, \dots, \vec{s}_N) = \left| \sum_{i=1}^N \vec{s}_i a_i \right| = \sqrt{\sum_{j=1}^N \sum_{i=1}^N \vec{s}_i a_i a_j \vec{s}_j}$$



- $q$  Teilmengen  $\hat{=}$   $q$  POTTS-Vektoren  $S_q = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_q\}$  (Symmetrie!)
- Partition  $\hat{=}$  walk auf  $(q - 1)$ -dim. Gitter in  $q$  Richtungen
- Schrittweite: Gewicht  $a_i$
- Richtung:  $\vec{s}_i \in S_q = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_q\}$
- Kostenfunktion  $\hat{=}$  Abstand von Ursprung nach  $N$  Schritten

$$H(\vec{s}_1, \vec{s}_2, \dots, \vec{s}_N) = \left| \sum_{i=1}^N \vec{s}_i a_i \right| = \sqrt{\sum_{j=1}^N \sum_{i=1}^N \vec{s}_i a_i a_j \vec{s}_j}$$

- Spin-HAMILTON-Funktion vom MATTIS-Typ, aber:



- $q$  Teilmengen  $\hat{=}$   $q$  POTTS-Vektoren  $S_q = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_q\}$  (Symmetrie!)
- Partition  $\hat{=}$  walk auf  $(q - 1)$ -dim. Gitter in  $q$  Richtungen
- Schrittweite: Gewicht  $a_i$
- Richtung:  $\vec{s}_i \in S_q = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_q\}$
- Kostenfunktion  $\hat{=}$  Abstand von Ursprung nach  $N$  Schritten

$$H(\vec{s}_1, \vec{s}_2, \dots, \vec{s}_N) = \left| \sum_{i=1}^N \vec{s}_i a_i \right| = \sqrt{\sum_{j=1}^N \sum_{i=1}^N \vec{s}_i a_i a_j \vec{s}_j}$$

- Spin-HAMILTON-Funktion vom MATTIS-Typ, aber:
- Stärke der Wechselwirkung zufällig, nicht das Vorzeichen  $\Rightarrow$  starke Frustration



NPP

statistische Physik



## NPP

- $q$  Teilmengen:  
 $s_i \in \{1, 2, \dots, q\}$

## statistische Physik

- $q$  POTTIS-Spins:  
 $\vec{s}_i \in \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_q\}$



## NPP

- $q$  Teilmengen:  
 $s_i \in \{1, 2, \dots, q\}$
- Gewichte  $a_i$

## statistische Physik

- $q$  POTTS-Spins:  
 $\vec{s}_i \in \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_q\}$
- Spin-Spin-Kopplungen  
 $J_{i,j} = -a_i a_j$



## NPP

- $q$  Teilmengen:  
 $s_i \in \{1, 2, \dots, q\}$
- Gewichte  $a_i$
- „Unausgeglichenheit“

## statistische Physik

- $q$  POTTS-Spins:  
 $\vec{s}_i \in \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_q\}$
- Spin-Spin-Kopplungen  
 $J_{i,j} = -a_i a_j$
- Energie der Spinkonfiguration

## NPP

- $q$  Teilmengen:  
 $s_i \in \{1, 2, \dots, q\}$
- Gewichte  $a_i$
- „Unausgeglichenheit“
- beste Partition der Gewichte

## statistische Physik

- $q$  POTTS-Spins:  
 $\vec{s}_i \in \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_q\}$
- Spin-Spin-Kopplungen  
 $J_{i,j} = -a_i a_j$
- Energie der Spinkonfiguration
- Grundzustand



## NPP

- $q$  Teilmengen:  
 $s_i \in \{1, 2, \dots, q\}$
- Gewichte  $a_i$
- „Unausgeglichenheit“
- beste Partition der Gewichte
- perfekte Partition existiert

## statistische Physik

- $q$  POTTS-Spins:  
 $\vec{s}_i \in \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_q\}$
- Spin-Spin-Kopplungen  
 $J_{i,j} = -a_i a_j$
- Energie der Spinkonfiguration
- Grundzustand
- Grundzustangsenergie  $\cong 0$

## NPP

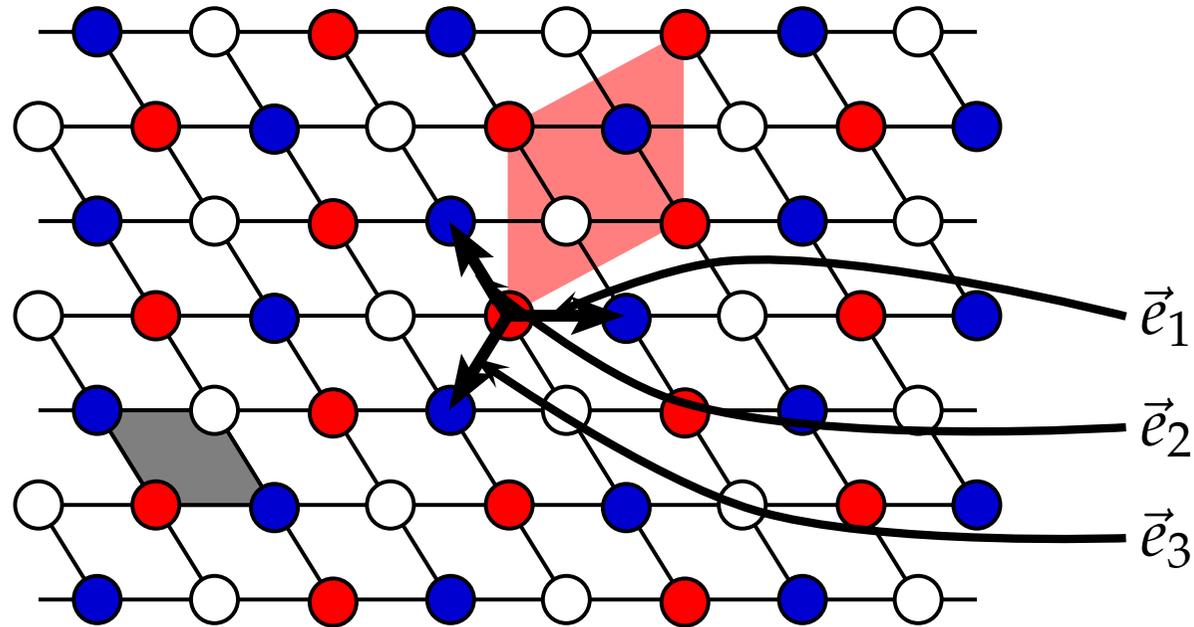
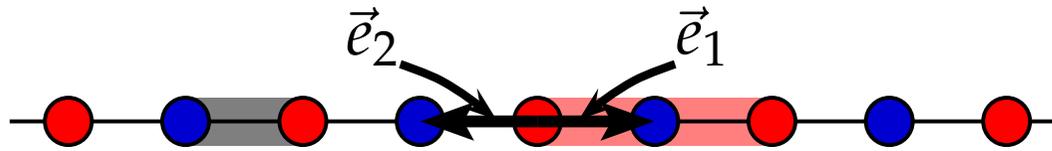
- $q$  Teilmengen:  
 $s_i \in \{1, 2, \dots, q\}$
- Gewichte  $a_i$
- „Unausgeglichenheit“
- beste Partition der Gewichte
- perfekte Partition existiert
- keine perfekte Partition

## statistische Physik

- $q$  POTTS-Spins:  
 $\vec{s}_i \in \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_q\}$
- Spin-Spin-Kopplungen  
 $J_{i,j} = -a_i a_j$
- Energie der Spinkonfiguration
- Grundzustand
- Grundzustangsenergie  $\cong 0$
- Grundzustangsenergie  $> 0$

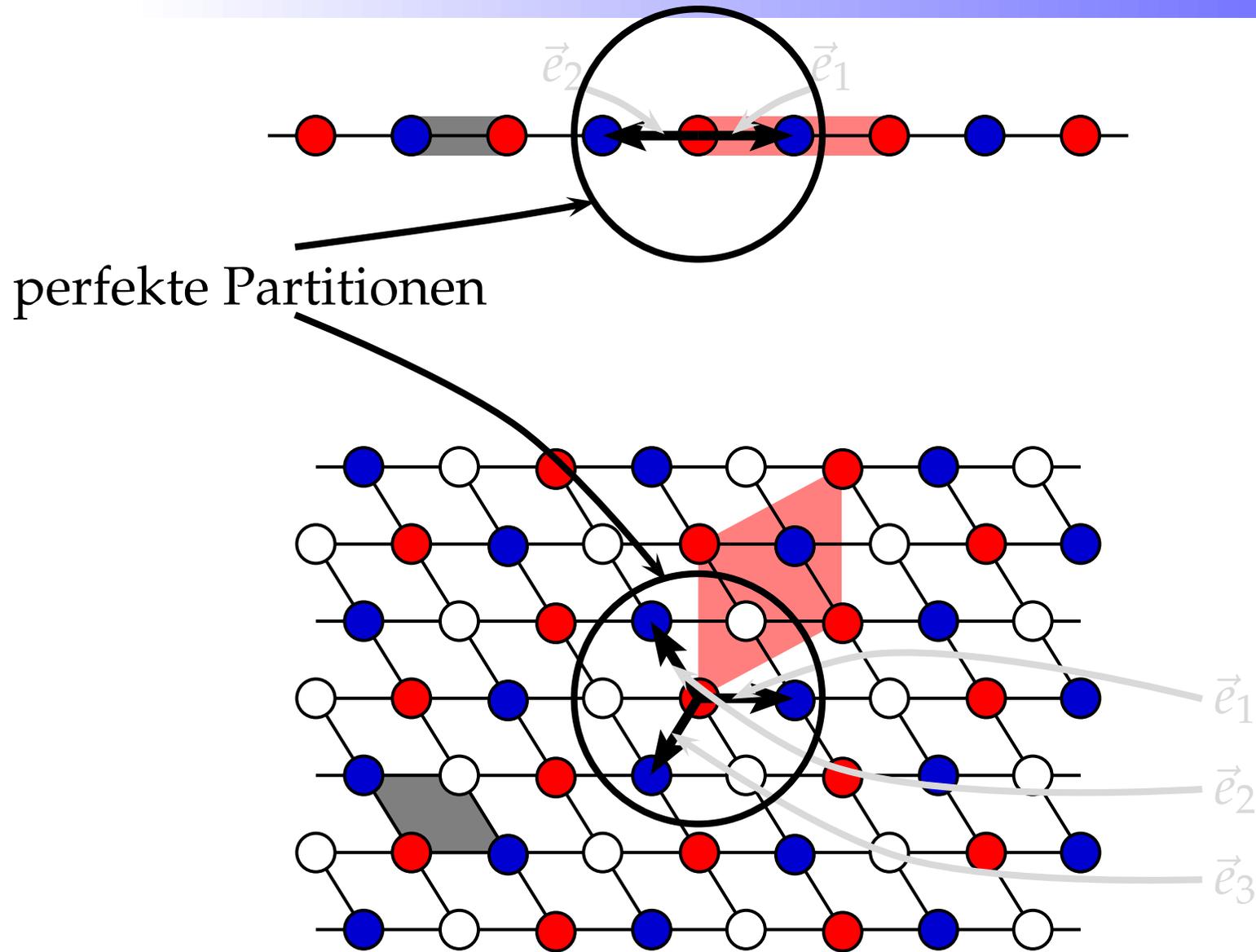


# Mikrokanonische Zustandssumme



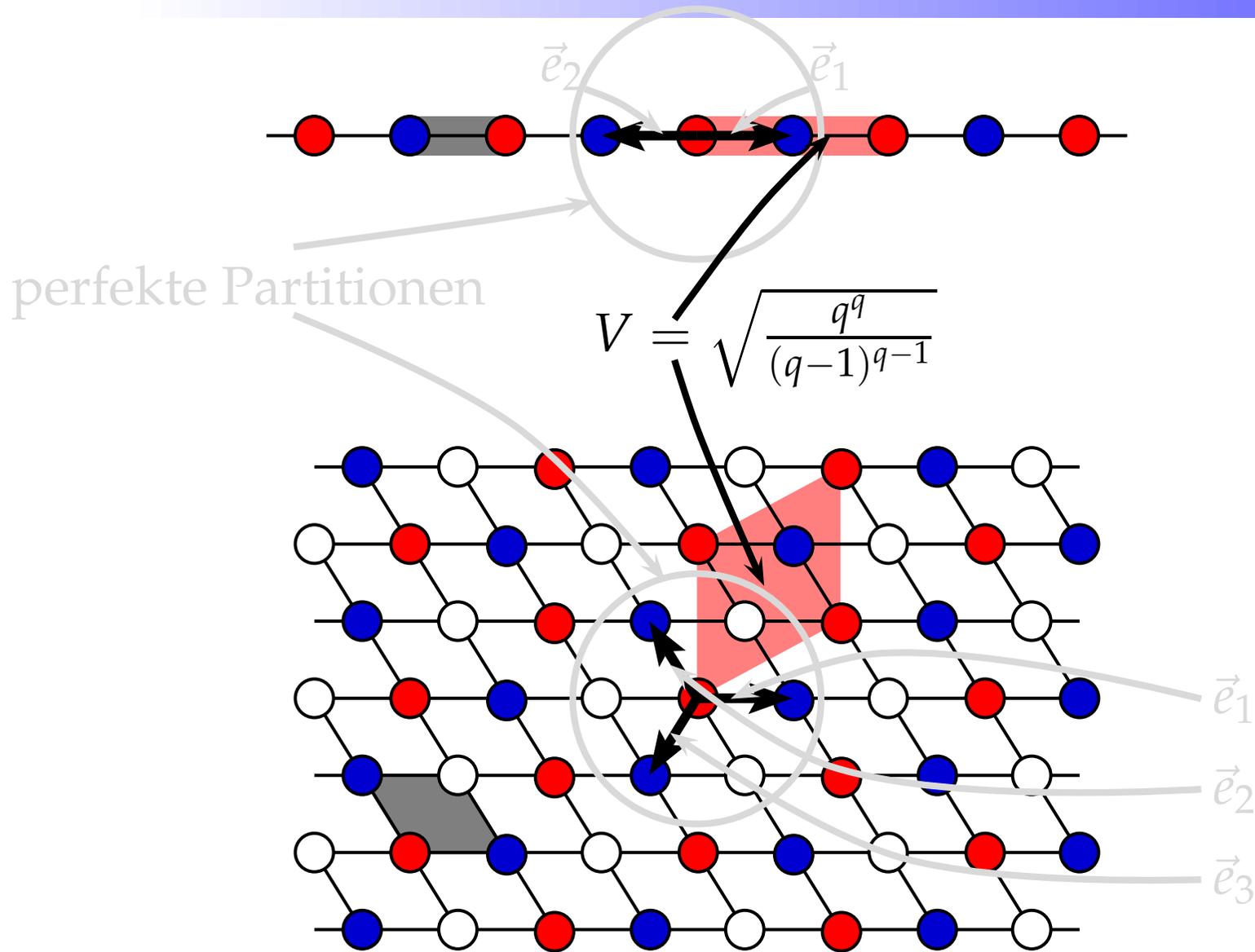


# Mikrokanonische Zustandssumme





# Mikrokanonische Zustandssumme



# Mikrokanonische Zustandssumme



$$\Omega(\vec{D}) = \sum_{\substack{(\vec{s}_1, \vec{s}_2, \dots, \vec{s}_N) \\ \vec{s}_i \in \mathcal{S}_q}} \delta \left( \vec{D} - \sum_{j=1}^N a_j \vec{s}_j \right)$$



$$\begin{aligned}\Omega(\vec{D}) &= \sum_{\substack{(\vec{s}_1, \vec{s}_2, \dots, \vec{s}_N) \\ \vec{s}_i \in S_q}} \delta \left( \vec{D} - \sum_{j=1}^N a_j \vec{s}_j \right) \\ &= q^N P \left( \vec{d} = \left( \frac{\vec{D}}{\sqrt{N}} \right) \right) = q^N \int_{\frac{v}{\sqrt{N^{q-1}}}} p(\vec{d}) d^{q-1} d\end{aligned}$$



$$\begin{aligned}\Omega(\vec{D}) &= \sum_{\substack{(\vec{s}_1, \vec{s}_2, \dots, \vec{s}_N) \\ \vec{s}_i \in S_q}} \delta \left( \vec{D} - \sum_{j=1}^N a_j \vec{s}_j \right) \\ &= q^N P \left( \vec{d} = \left( \frac{\vec{D}}{\sqrt{N}} \right) \right) = q^N \int_{\frac{V}{\sqrt{N^{q-1}}}} p(\vec{d}) d^{q-1} d\end{aligned}$$

$p(\vec{d})$  ist  $(q - 1)$ -dimensionale Normalverteilung

$$p(\vec{d}) \approx \frac{1}{\left( \frac{2\pi}{q-1} \langle a^2 \rangle \right)^{q-1}} e^{-\frac{\vec{d}^2}{\frac{2}{q-1} \langle a^2 \rangle}}$$



- Zahl der Partionen mit Vektor  $\vec{D}$  normalverteilt

$$\Omega(\vec{D}) = \frac{q^N q^{\frac{q}{2}}}{(2\pi N \langle a^2 \rangle)^{\frac{q-1}{2}}} e^{-\frac{(q-1)\vec{D}^2}{2N \langle a^2 \rangle}}$$



- Zahl der Partionen mit Vektor  $\vec{D}$  normalverteilt

$$\Omega(\vec{D}) = \frac{q^N q^{\frac{q}{2}}}{(2\pi N \langle a^2 \rangle)^{\frac{q-1}{2}}} e^{-\frac{(q-1)\vec{D}^2}{2N \langle a^2 \rangle}}$$

- Zahl der perfekten Partitionen (zur Graphik)

$$\Omega_{\text{perfekt},r} = \binom{q}{r} \frac{q^N q^{\frac{q}{2}}}{(2\pi N \langle a^2 \rangle)^{\frac{q-1}{2}}}, \quad r \equiv \sum_{i=1}^N a_i \pmod{q}$$



- Zahl der Partionen mit Vektor  $\vec{D}$  normalverteilt

$$\Omega(\vec{D}) = \frac{q^N q^{\frac{q}{2}}}{(2\pi N \langle a^2 \rangle)^{\frac{q-1}{2}}} e^{-\frac{(q-1)\vec{D}^2}{2N \langle a^2 \rangle}}$$

- Zahl der perfekten Partitionen (zur Graphik)

$$\Omega_{\text{perfekt},r} = \binom{q}{r} \frac{q^N q^{\frac{q}{2}}}{(2\pi N \langle a^2 \rangle)^{\frac{q-1}{2}}}, \quad r \equiv \sum_{i=1}^N a_i \pmod{q}$$

- Zahl der perfekten Partitionen im Mittel

$$\begin{aligned} \log_q \Omega_{\text{perfekt}} &= N \left( \underbrace{1 + \frac{\log_q \frac{2^q - 1}{q}}{N}}_{\kappa_c} - \frac{\log_q \frac{(2\pi N)^{q-1}}{3^{q-1} q^q}}{2N} - \underbrace{\frac{\log_q 3^{q-1} \langle a^2 \rangle^{q-1}}{2N}}_{\kappa} \right) \\ &= N(\kappa_c - \kappa) \end{aligned}$$

# Bedeutung des Parameters $\kappa$



- Seien die Gewichte gleichverteilte Zufallszahlen aus  $[1, a_{\max}]$ .



# Bedeutung des Parameters $\kappa$

- Seien die Gewichte gleichverteilte Zufallszahlen aus  $[1, a_{\max}]$ .

- $$\langle a^2 \rangle = \frac{1}{a_{\max}} \sum_{i=1}^{a_{\max}} i^2 \approx \frac{1}{3} a_{\max}^2$$



# Bedeutung des Parameters $\kappa$

- Seien die Gewichte gleichverteilte Zufallszahlen aus  $[1, a_{\max}]$ .

- $$\langle a^2 \rangle = \frac{1}{a_{\max}} \sum_{i=1}^{a_{\max}} i^2 \approx \frac{1}{3} a_{\max}^2$$

- mit  $a_{\max} = q^b$

$$\kappa = \frac{\log_q 3^{q-1} \langle a^2 \rangle^{q-1}}{2N} = \frac{(q-1)b}{N}$$



# Bedeutung des Parameters $\kappa$

- Seien die Gewichte gleichverteilte Zufallszahlen aus  $[1, a_{\max}]$ .

- $$\langle a^2 \rangle = \frac{1}{a_{\max}} \sum_{i=1}^{a_{\max}} i^2 \approx \frac{1}{3} a_{\max}^2$$

- mit  $a_{\max} = q^b$

$$\kappa = \frac{\log_q 3^{q-1} \langle a^2 \rangle^{q-1}}{2N} = \frac{(q-1)b}{N}$$

- $\kappa$  ein Maß für die numerische Auflösung der Gewichte



# Bedeutung des Parameters $\kappa$

- Seien die Gewichte gleichverteilte Zufallszahlen aus  $[1, a_{\max}]$ .

- $$\langle a^2 \rangle = \frac{1}{a_{\max}} \sum_{i=1}^{a_{\max}} i^2 \approx \frac{1}{3} a_{\max}^2$$

- mit  $a_{\max} = q^b$

$$\kappa = \frac{\log_q 3^{q-1} \langle a^2 \rangle^{q-1}}{2N} = \frac{(q-1)b}{N}$$

- $\kappa$  ein Maß für die numerische Auflösung der Gewichte
- $\kappa < \kappa_c \Rightarrow$  viele perfekte Lösungen



# Bedeutung des Parameters $\kappa$

- Seien die Gewichte gleichverteilte Zufallszahlen aus  $[1, a_{\max}]$ .

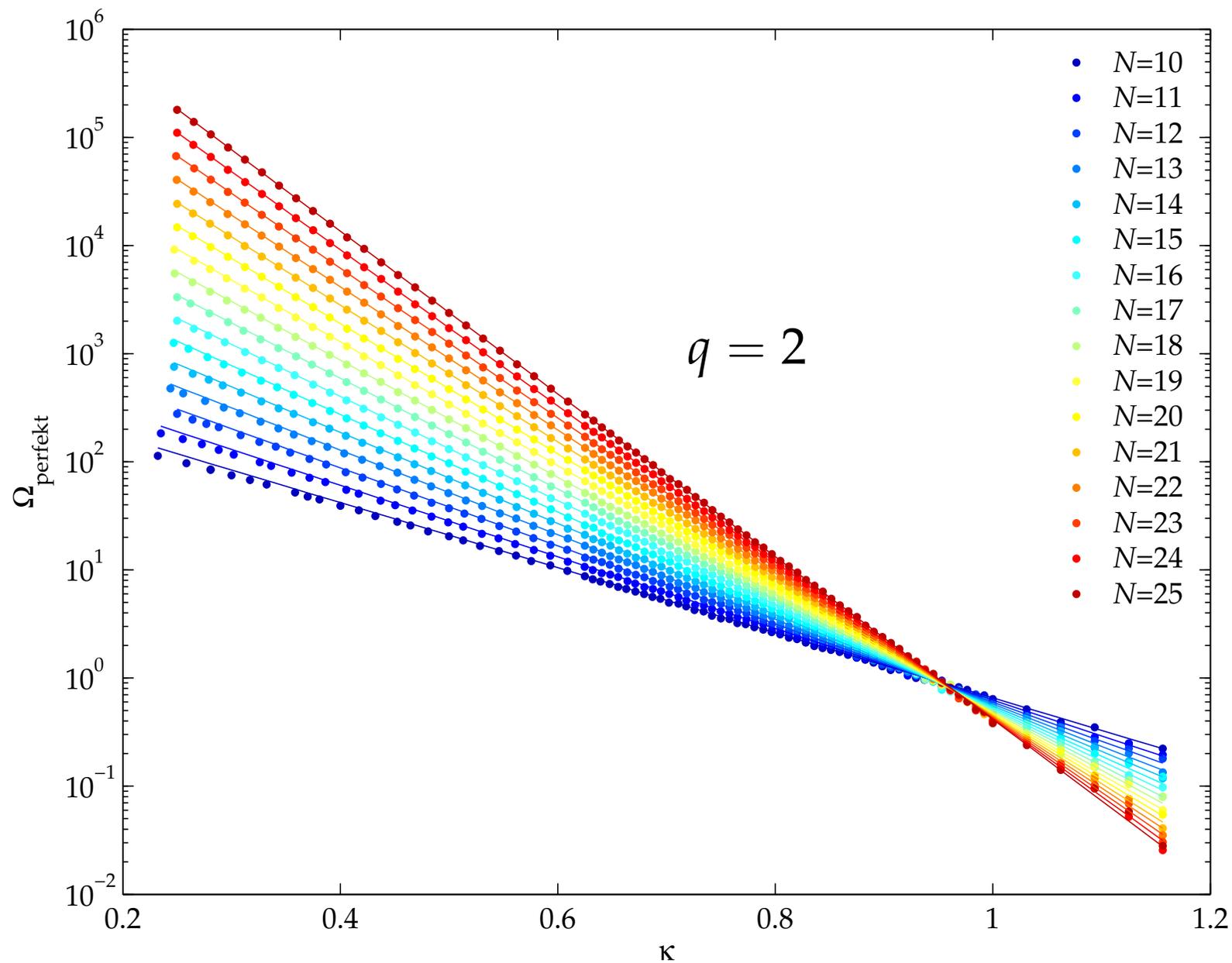
- $$\langle a^2 \rangle = \frac{1}{a_{\max}} \sum_{i=1}^{a_{\max}} i^2 \approx \frac{1}{3} a_{\max}^2$$

- mit  $a_{\max} = q^b$

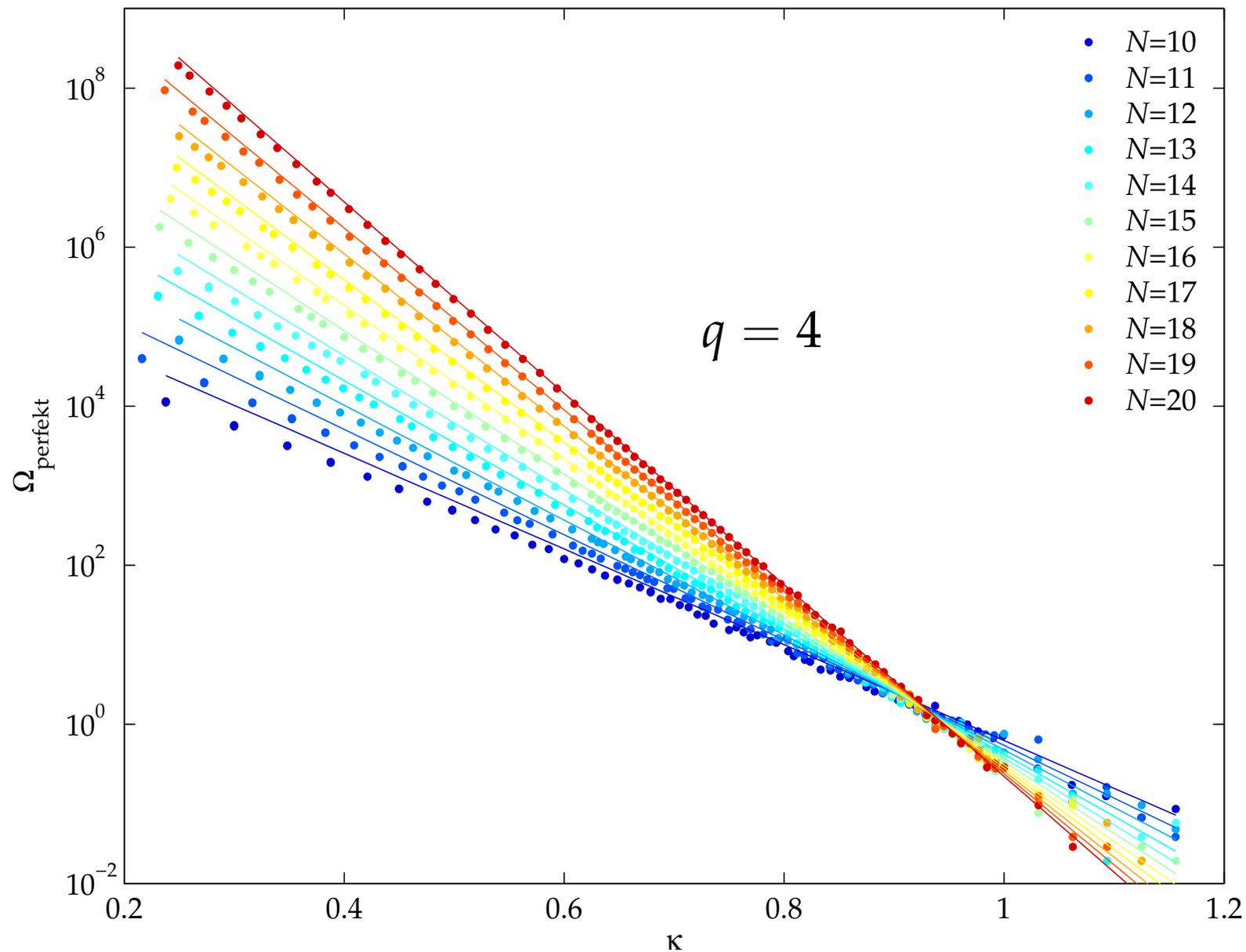
$$\kappa = \frac{\log_q 3^{q-1} \langle a^2 \rangle^{q-1}}{2N} = \frac{(q-1)b}{N}$$

- $\kappa$  ein Maß für die numerische Auflösung der Gewichte
- $\kappa < \kappa_c \Rightarrow$  viele perfekte Lösungen
- $\kappa > \kappa_c \Rightarrow$  keine perfekten Lösungen (siehe vorn)

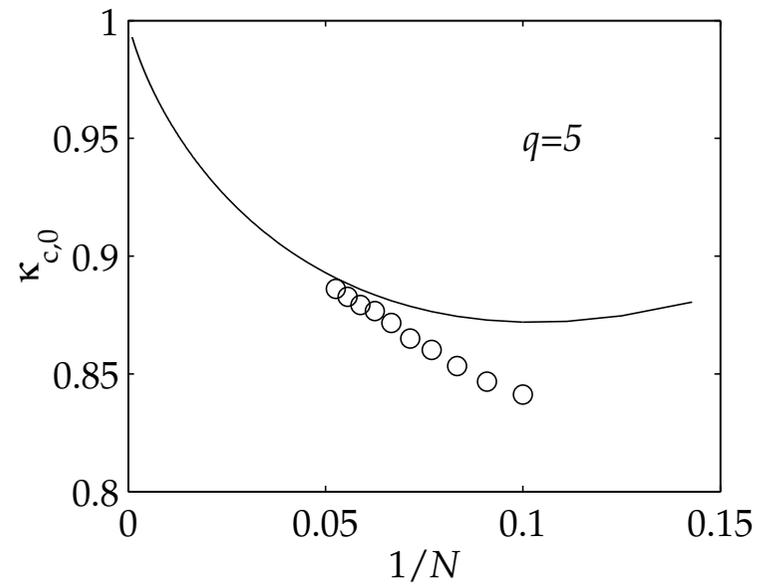
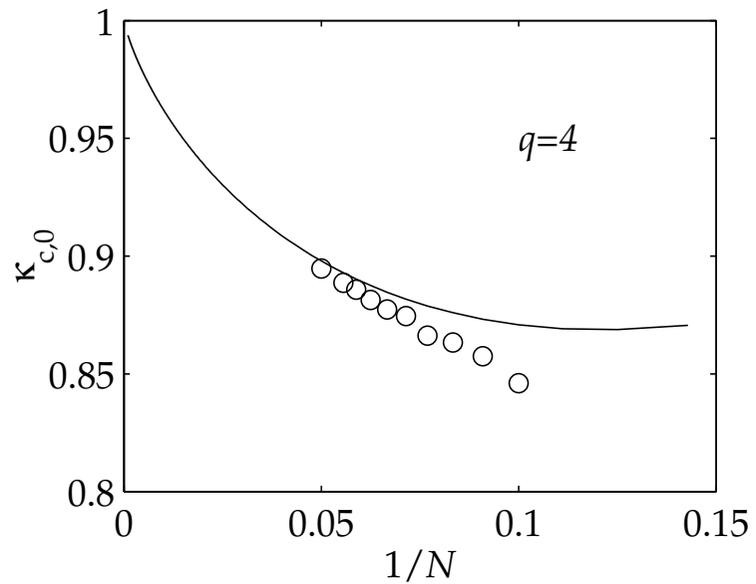
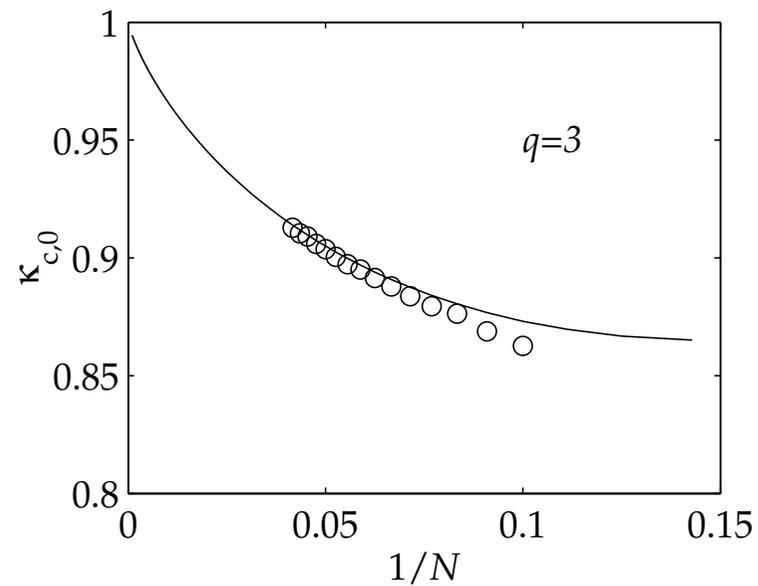
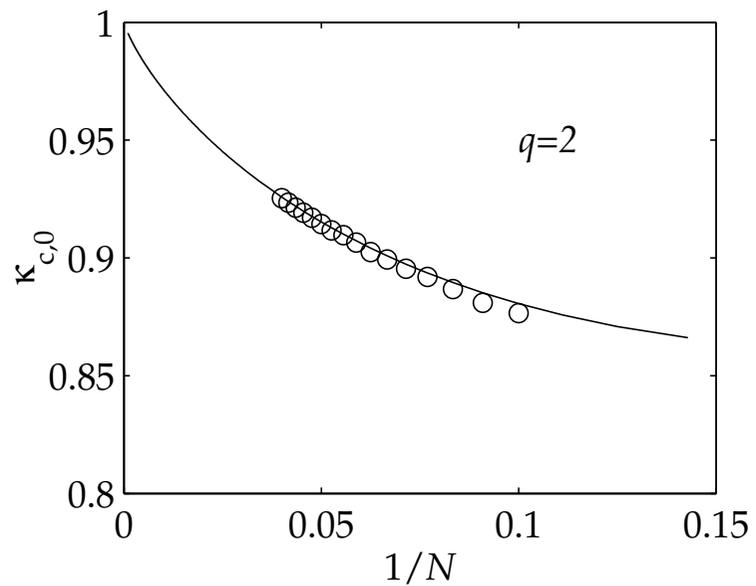
# Zahl der perfekten Partitionen



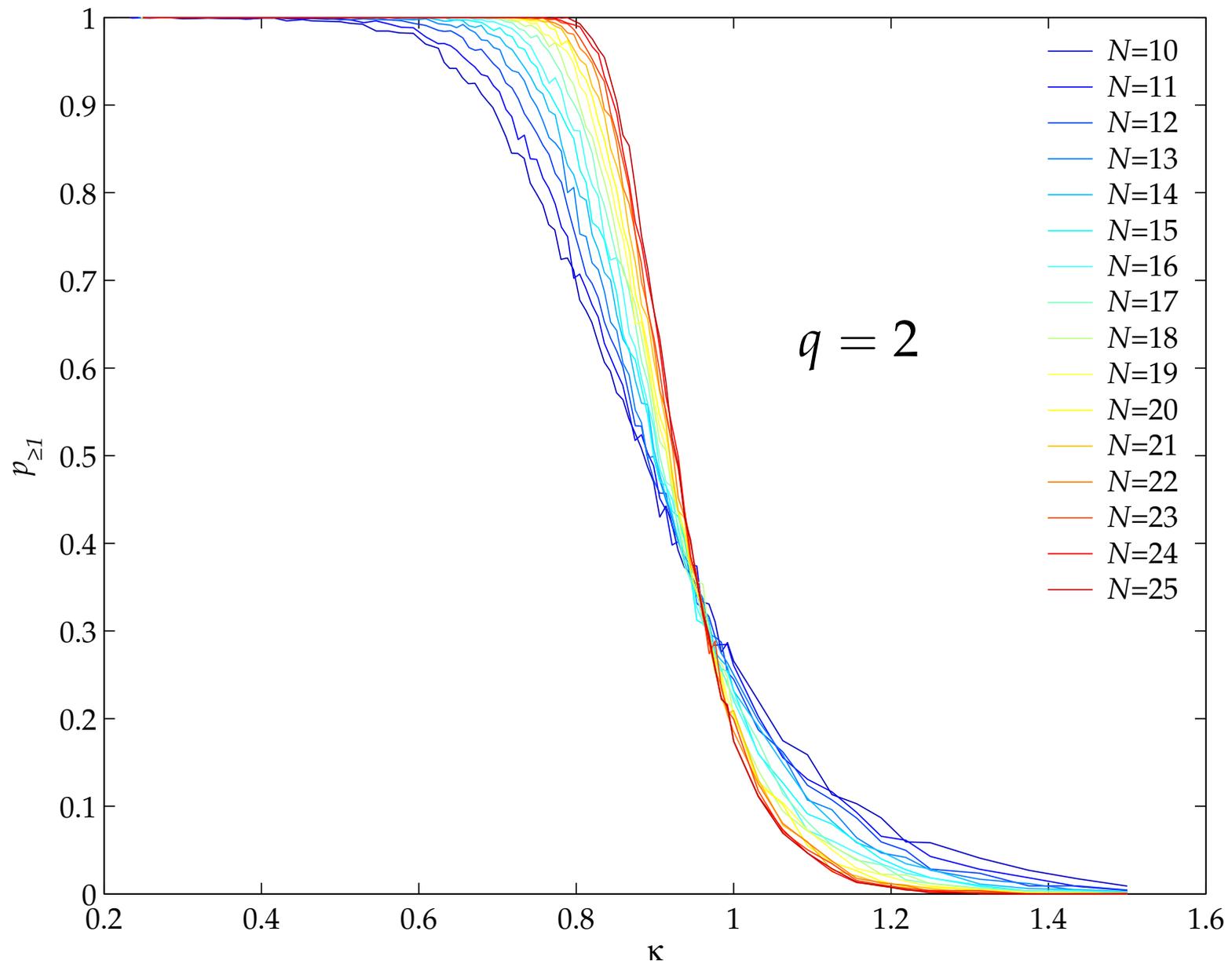
# Zahl der perfekten Partitionen



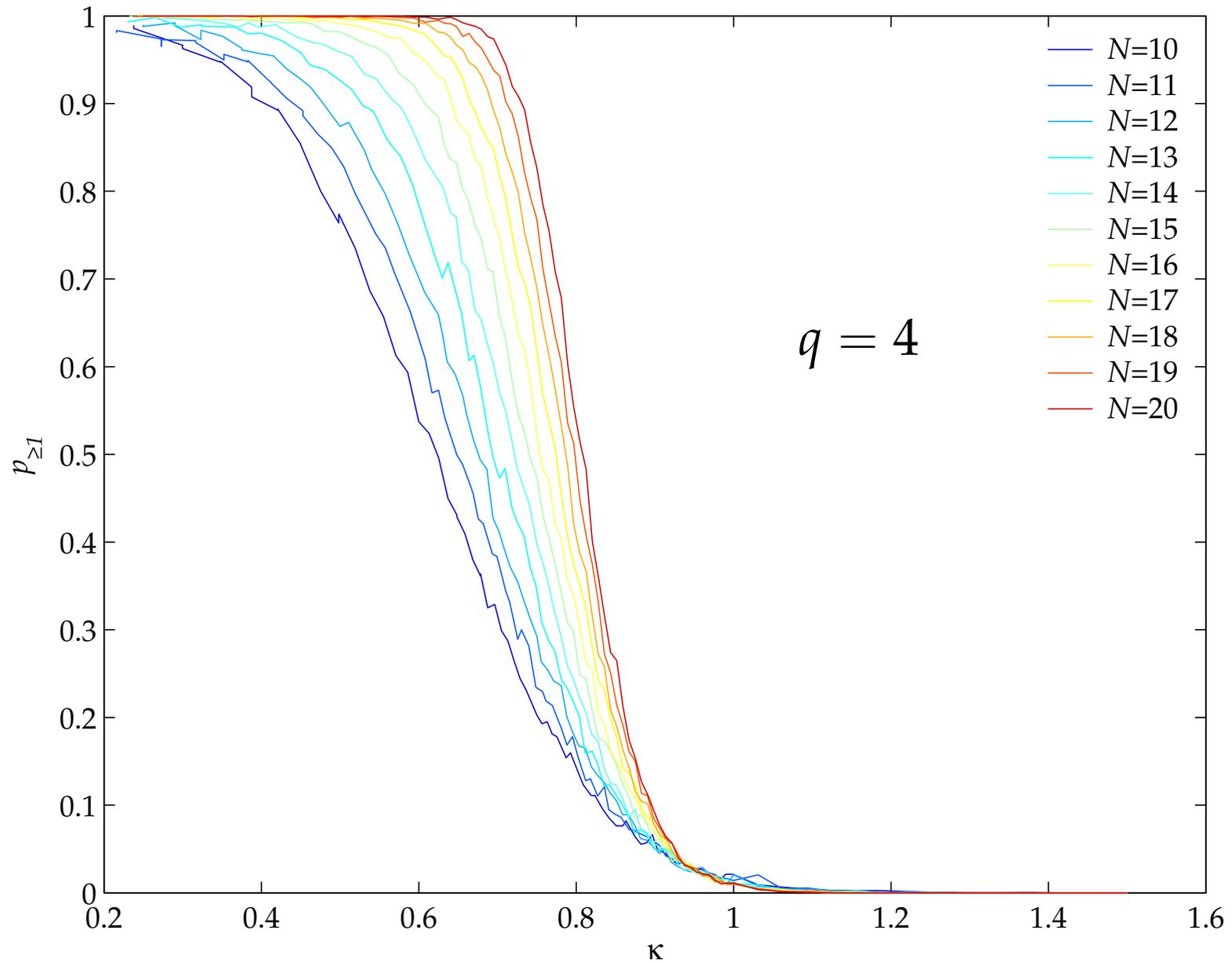
# Kritischer Wert von $\kappa$



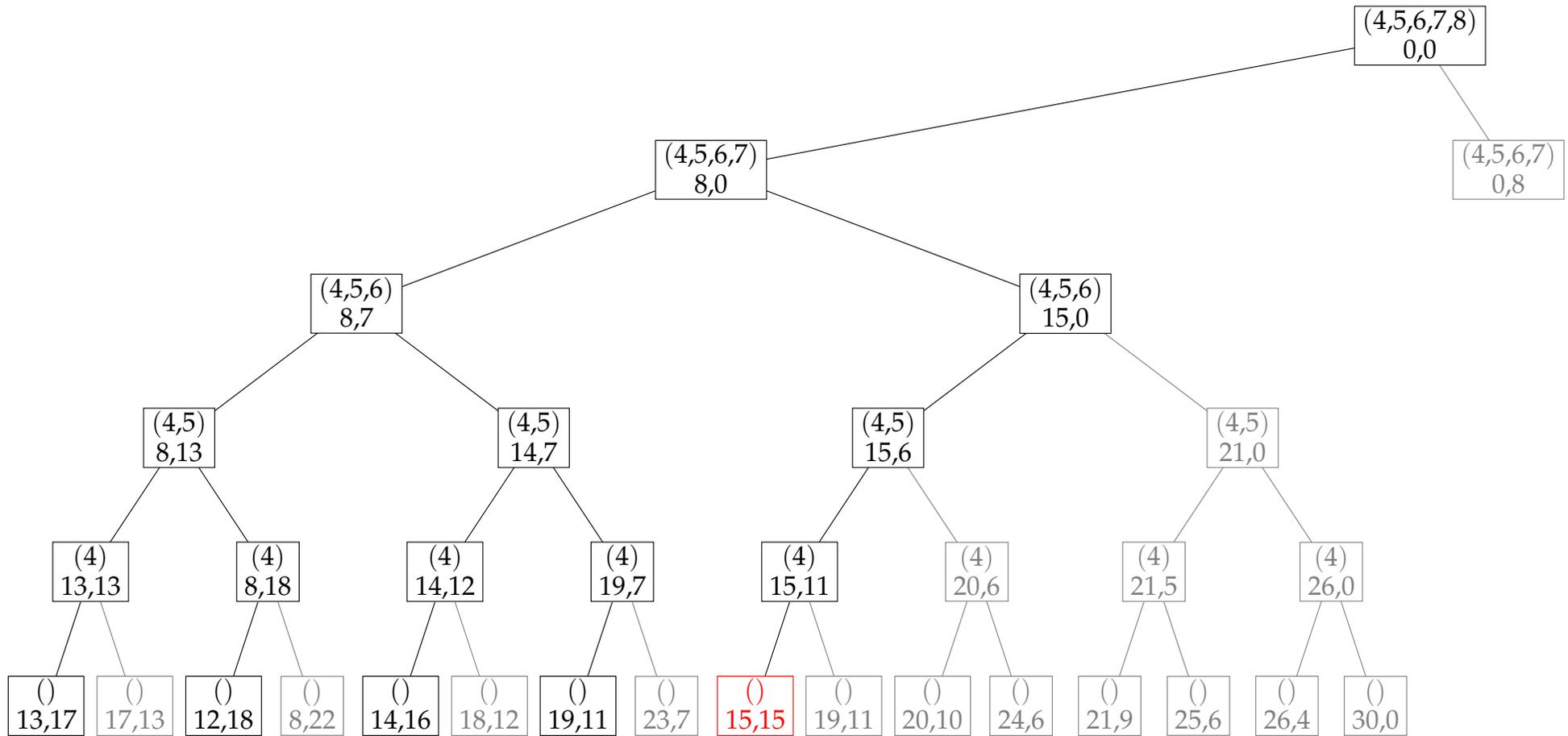
# Phasenübergang



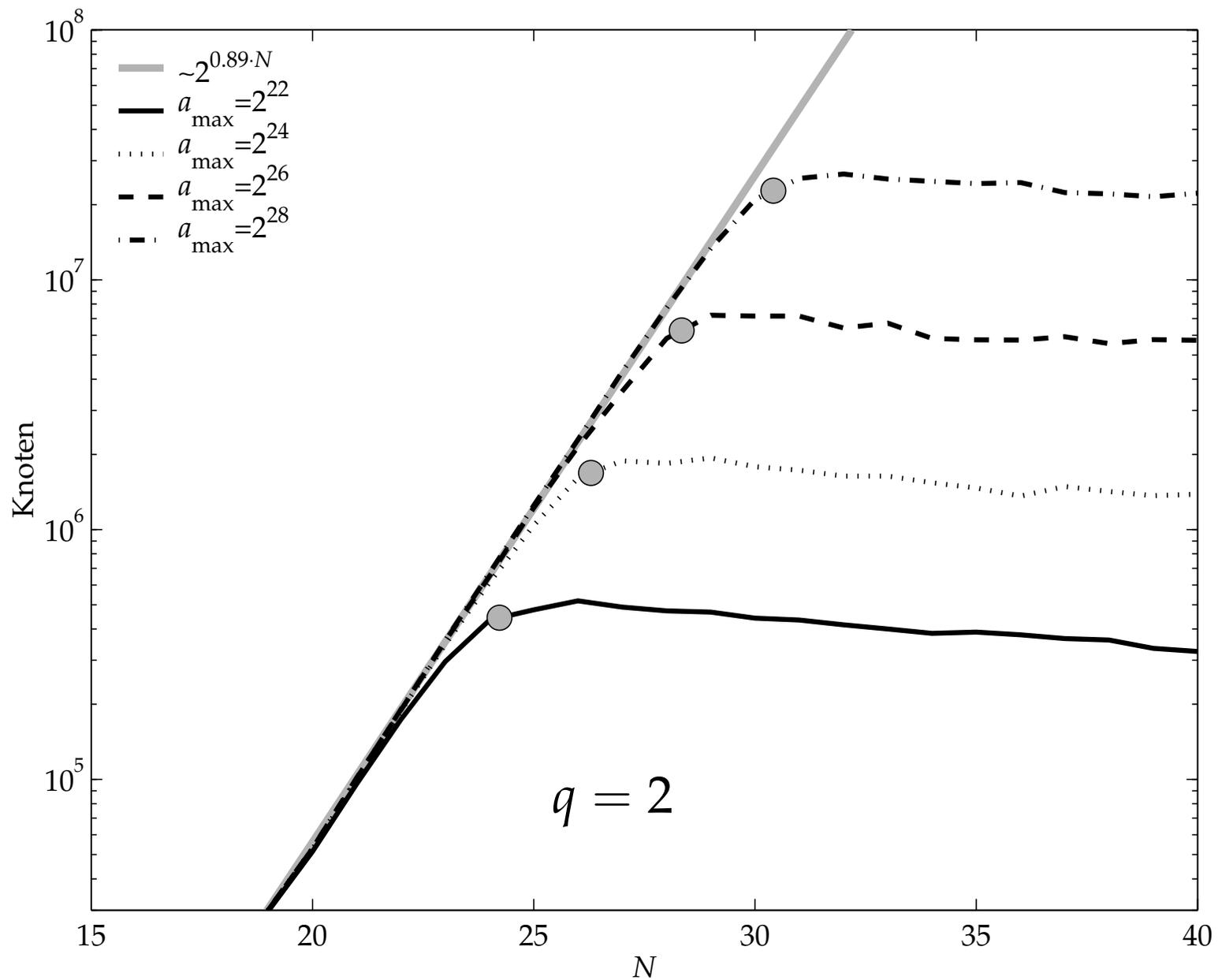
# Phasenübergang



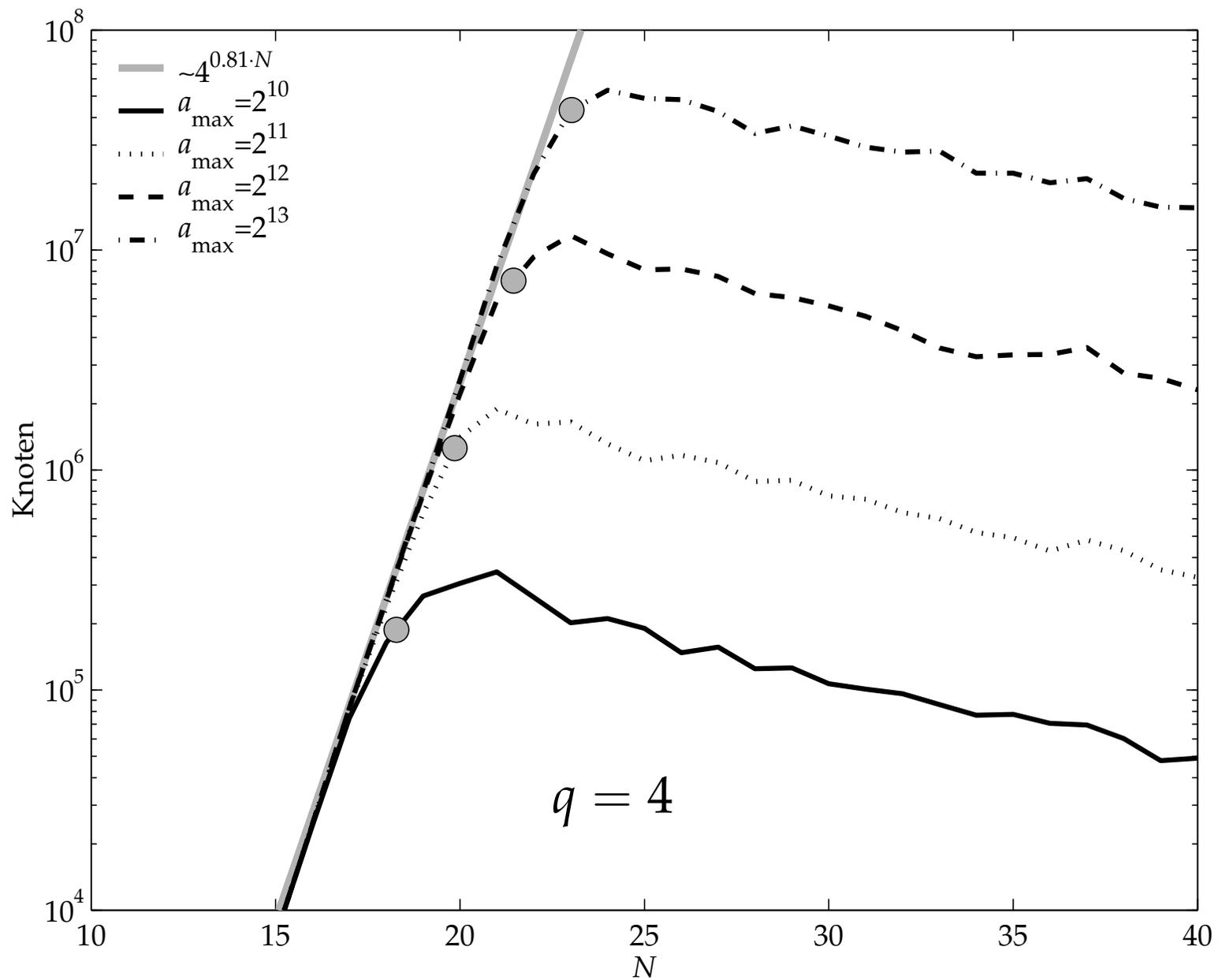
# Greedy-Algorithmus ( $q = 2$ )



# Dynamik von Algorithmen



# Dynamik von Algorithmen



# Zusammenfassung



- Abbildung des Zahlenaufteilungsproblems auf ein Spinsystem

# Zusammenfassung



- Abbildung des Zahlenaufteilungsproblems auf ein Spinsystem
- Spinvariablen sind Potts-Vektoren

# Zusammenfassung



- Abbildung des Zahlenaufteilungsproblems auf ein Spinsystem
- Spinvariablen sind POTTS-Vektoren
- Beobachtung eines Phasenübergangs



- Abbildung des Zahlenaufteilungsproblems auf ein Spinsystem
- Spinvariablen sind Potts-Vektoren
- Beobachtung eines Phasenübergangs
- Ordnungsparameter: Wahrscheinlichkeit eine perfekte Partition zu finden

- Abbildung des Zahlenaufteilungsproblems auf ein Spinsystem
- Spinvariablen sind POTTS-Vektoren
- Beobachtung eines Phasenübergangs
- Ordnungsparameter: Wahrscheinlichkeit eine perfekte Partition zu finden
- Kontrollparameter: numerische Auflösung der Gewichte

- Abbildung des Zahlenaufteilungsproblems auf ein Spinsystem
- Spinvariablen sind Potts-Vektoren
- Beobachtung eines Phasenübergangs
- Ordnungsparameter: Wahrscheinlichkeit eine perfekte Partition zu finden
- Kontrollparameter: numerische Auflösung der Gewichte
- Phasenübergang beeinflusst Laufzeitverhalten von Algorithmen



- Abbildung des Zahlenaufteilungsproblems auf ein Spinsystem
- Spinvariablen sind POTTS-Vektoren
- Beobachtung eines Phasenübergangs
- Ordnungsparameter: Wahrscheinlichkeit eine perfekte Partition zu finden
- Kontrollparameter: numerische Auflösung der Gewichte
- Phasenübergang beeinflusst Laufzeitverhalten von Algorithmen
- gute Übereinstimmung von analytischen und numerischen Ergebnissen



- Abbildung des Zahlenaufteilungsproblems auf ein Spinsystem
- Spinvariablen sind POTTS-Vektoren
- Beobachtung eines Phasenübergangs
- Ordnungsparameter: Wahrscheinlichkeit eine perfekte Partition zu finden
- Kontrollparameter: numerische Auflösung der Gewichte
- Phasenübergang beeinflusst Laufzeitverhalten von Algorithmen
- gute Übereinstimmung von analytischen und numerischen Ergebnissen
- signifikante finite-size-Effekte



- Stephan Mertens
- Andreas Engel
- Brian Hayes, *The Easiest Hard Problem*,  
in *American Scientist*, March-April 2002
- PRL, vol. 90, nr. 15, 158701  
<http://de.arxiv.org/abs/cond-mat/0208081>
- Diplomarbeit  
<http://www.uni-magdeburg.de/bauke>