

# **Evolutionäre Algorithmen für das Graph-Bisection-Problem**

Gero Greiner

8. Februar 2008



# Inhaltsverzeichnis

<b>1 Grundlagen</b>	<b>5</b>
1.1 Das Graph-Bisection-Problem . . . . .	5
1.1.1 Bekannte Resultate . . . . .	6
1.2 Randomisierte Suchheuristiken . . . . .	7
1.2.1 Multikriterielle Optimierung . . . . .	9
1.3 Branch-and-Bound-Algorithmen . . . . .	11
1.4 Der A*-Algorithmus und IDA* . . . . .	12
1.5 Analytische Hilfsmittel . . . . .	13
1.5.1 Mutationswahrscheinlichkeiten und Erwartungswerte . . . . .	13
1.5.2 Die Chernoff-Ungleichungen . . . . .	14
1.5.3 Die Stirling-Formel . . . . .	14
1.5.4 Random Walk auf Plateaus . . . . .	15
1.5.5 Das Gamblers-Ruin-Problem . . . . .	16
<b>2 Algorithmen für GBP</b>	<b>17</b>
2.1 Ein optimales Lösungsverfahren . . . . .	17
2.2 Analysierte Suchheuristiken . . . . .	21
2.2.1 RLS . . . . .	21
2.2.2 (1+1)-EA . . . . .	21
2.2.3 SEMO . . . . .	22
<b>3 Theoretische Analyse</b>	<b>25</b>
3.1 Zwei-Cliquen-Graph . . . . .	25
3.1.1 Der Zwei-Cliquen-Graph mit dem global SEMO . . . . .	25
3.1.2 Der Zwei-Cliquen-Graph mit dem (1+1)-EA . . . . .	34
3.2 Ringe . . . . .	36
3.2.1 RLS auf Ringen . . . . .	37
3.2.2 (1+1)-EA auf Ringen . . . . .	39
3.3 Graphen mit lokalen Optima . . . . .	42
3.3.1 Der Vier-Cliquen-Graph . . . . .	42
3.3.2 Der $2n$ - $n$ -Cliquen-Graph . . . . .	57
3.3.3 Der Clique-Ring-Graph . . . . .	65
<b>4 Experimente</b>	<b>75</b>
4.1 Der Zwei-Cliquen-Graph . . . . .	75
4.2 Ringe . . . . .	78
4.3 Der Vier-Cliquen-Graph . . . . .	80
4.4 Der $2n$ - $n$ -Cliquen-Graph . . . . .	82
4.5 Der Clique-Ring-Graph . . . . .	84
<b>5 Zusammenfassung</b>	<b>87</b>

5.1	Ausblick . . . . .	88
-----	--------------------	----

# 1 Grundlagen

## 1.1 Das Graph-Bisection-Problem

Gegenstand dieser Diplomarbeit ist die Analyse von evolutionären Algorithmen für das Graph-Bisection-Problem (GBP). Eine Bisektion eines Graphen  $G = (V, E)$  mit gerader Anzahl an Knoten  $|V| = 2n$  für  $n \in \mathbb{N}$  ist eine Partition der Knotenmenge  $V$  in zwei gleich große Mengen. Diesen weisen wir willkürlich die Bezeichnungen *links* und *rechts* zu und nennen sie entsprechend  $L$  und  $R$ . Der Wert einer Bisektion entspricht dann der Anzahl der Kanten, die zwischen den Mengen  $L$  und  $R$  verlaufen.

**Definition 1.1.** Sei  $G = (V, E)$  ein ungerichteter Graph mit  $V = \{v_1, \dots, v_{2n}\}$  für  $n \in \mathbb{N}$ . Die Zeichenkette  $x \in \{1, 2\}^{2n}$  wird als eine Bisektion von  $G$  bezeichnet, falls für die Mengen

$$L(x) := \{v_i \in V \mid x_i = 1\}$$

und

$$R(x) := \{v_i \in V \mid x_i = 2\}$$

gilt:  $|L(x)| = |R(x)|$ . Die Funktion

$$\begin{aligned} \text{GBP}_G : \{1, 2\}^{2n} &\rightarrow \{0, \dots, |E|\} \\ x &\mapsto \sum_{\{v_i, v_j\} \in E} (1 - \delta_{x_i, x_j}), \end{aligned}$$

mit dem Kronecker-Delta

$$\delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}$$

beschreibt den Wert einer Bisektion  $x$ .

**Definition 1.2.** Ziel des Graph-Bisection-Problem ist es eine Bisektion  $x^{opt}$  mit Wert

$$\text{bis}(G) := \min_{x \in \mathcal{S}} \text{GBP}_G(x),$$

für  $\mathcal{S} := \{x \mid x \in \{1, 2\}^{2n}, |L(x)| = |R(x)|\}$  zu finden<sup>1</sup>.

Die Menge der Kanten einer Bisektion  $x$ , die zwischen den Mengen  $L(x)$  und  $R(x)$  verlaufen, bezeichnen wir auch als *Schnitt*, die Kanten selbst als *Schnittkanten*. Für jede Bisektion  $x$  existiert eine äquivalente Bisektion  $x'$  mit  $L(x') = R(x)$  und  $R(x') = L(x)$  mit gleichem Funktionswert  $\text{GBP}_G(x') = \text{GBP}_G(x)$ . Diese Bezeichnen wir als *Spiegelbild* von  $x$ . Wir weisen im Folgenden jedoch nicht immer explizit auf die Existenz dieser Bisektion hin. Wird eine spezielle Bisektion, wie ein (lokales) Optimum,

<sup>1</sup>Da wir später eine ternäre Kodierung einführen, kodieren wir die Bisektion hier mit  $\{1, 2\}^{2n}$  anstelle der konventionellen Kodierung durch  $\{0, 1\}^{2n}$ .

beschrieben, so existiert entsprechend auch immer das Spiegelbild mit vertauschten Schnittseiten und gleichen Eigenschaften.

Das Problem, die bzw. eine minimale Bisektion eines Graphen zu finden, tritt bei vielen Netzwerkproblemen auf, deren Lösungsalgorithmen auf Divide-and-Conquer-Paradigmen beruhen [5]. Vor allem das Placement-Problem für VLSI-Schaltkreise und Routing-Probleme sind oft genannte Beispiele. Ziel des Placement-Problems ist die Anordnung von Schaltkreiselementen innerhalb eines Mikrochips, so dass die Wege zwischen den Elementen minimiert und somit die Kommunikationsgeschwindigkeit maximiert wird. Eine Approximationsmethode ist die rekursive Partitionierung der Schaltkreiselemente [20].

### 1.1.1 Bekannte Resultate

Das Graph-Bisection-Problem ist bekanntermaßen  $\mathcal{NP}$ -vollständig [11]. Somit kann unter der Hypothese  $\mathcal{NP} \neq \mathcal{P}$  nicht für alle Graphen die optimale Lösung in Polynomialzeit gefunden werden. Von Bui und Jones [4] wurde außerdem gezeigt, dass eine Approximation der minimalen Bisektionen von Graphen mit  $n$  Knoten mit der Lösung  $\text{OPT} + n^{2-\epsilon}$  für alle  $\epsilon > 0$   $\mathcal{NP}$ -schwer ist, wobei  $\text{OPT}$  den Wert einer minimalen Bisektion bezeichnet. Es ist also  $\text{GBP} \notin \mathcal{APX}$ , falls  $\mathcal{NP} \neq \mathcal{P}$  ist. Sogar für Graphen mit einem maximalen Grad von 3 ist das Erzeugen eines Schnitts der Größe  $\text{OPT} + n^{1/2-\epsilon}$  bereits  $\mathcal{NP}$ -schwer.

Für einige Graphklassen, wie Bäume und planare Graphen mit einer minimalen Bisektion von  $\mathcal{O}(\log n)$ , sind dagegen effiziente Algorithmen bekannt [5, 6]. Für andere Graphen werden in der Praxis meist Heuristiken eingesetzt, die jedoch keine Güte der Lösung garantieren. Die wahrscheinlich bekannteste Heuristik ist der Kernighan-Lin-Algorithmus [16], welcher durch gezieltes Vertauschen zweier Knotenmengen eine vorhandene Bisektion iterativ verbessert. Häufig verwendet wird auch Simulated Annealing bzw. der Metropolis-Algorithmus [17]. Nach einer Analyse in [15] liefert Simulated Annealing im Durchschnitt zwar bessere Ergebnisse als Kernighan-Lin, jedoch wurde in [2] die Überlegenheit von Kernighan-Lin unter dem Einsatz einer Methode zur Verdichtung des Graphen gegenüber Simulated Annealing experimentell bestätigt. Es ist allerdings bekannt, dass der Kernighan-Lin-Algorithmus auf bestimmten Graphklassen, wie beispielsweise Leiter-Graphen, generell sehr schlechte Ergebnisse liefert [2].

Zur Anwendung von evolutionären Algorithmen auf dem auch als Graph-Bisectioning bezeichneten GBP, existieren einige empirische Arbeiten für bestimmte Graphklassen, wie Zufallsgraphen, Gitter oder Bäume [6, 19]. Leider gibt es jedoch kaum theoretische Analysen über das Verhalten eines evolutionären Algorithmus für GBP.

Für einige andere Suchheuristiken auf verschiedenen Klassen von Zufallsgraphen wurde dagegen erfolgreich gezeigt, dass diese mit großer Wahrscheinlichkeit über die Graphen dieser Klasse eine minimale Bisektion finden. Dazu wurden meist solche Klassen von Zufallsgraphen verwendet, bei denen sich die minimale Bisektion deutlich von einem zufälligen Schnitt durch den Graphen abhebt. So wurde von Bui et al. [3] das Modell  $\mathcal{G}_{nmb}$  verwendet, bei dem für einen Graph mit  $n$  Knoten und  $m$  Kanten ein minimaler Schnitt von  $b$  bekannt ist. Von Jerrum et al. [14] wurde gezeigt, dass der Metropolis-Algorithmus in polynomieller Zeit mit hoher Wahrscheinlichkeit eine minimale Bisektion für  $p-q = \Omega(n^{-1/6+\epsilon})$  findet, wobei  $b = qn^2/4$  und  $m = (p+q)n^2/4$  gilt. Für die etwas weniger bekannte, lokal arbeitende Heuristik Go-With-The-Winners

[1] wurde ebenso eine hohe Wahrscheinlichkeit für das Erzeugen einer minimalen Bisektion eines Graphen mit  $\sqrt{mn^2/\ln n} < m/2 - b < \sqrt{m \ln n}$  bewiesen [8]. Für eine randomisierte lokale Suche zeigen Carson et al. in [7], dass diese für  $p - q = \Omega(n^{-1/4} \ln^4 n)$  mit hoher Wahrscheinlichkeit in polynomieller Zeit gegen die minimale Bisektion konvergiert.

Ein solches Ergebnis konnte im Rahmen dieser Arbeit für einen evolutionären Algorithmus nicht gezeigt werden. Stattdessen werden wir uns mit mehreren Graphklassen befassen, für die die behandelten Heuristiken entweder mit hoher Wahrscheinlichkeit in polynomieller Zeit ein Optimum erreichen oder exponentielle Rechenzeit benötigen um ein lokales Optima zu verlassen.

## 1.2 Randomisierte Suchheuristiken

Der Begriff *Heuristik* ist im Bereich der Algorithmen nicht fest definiert. Als Heuristiken werden jedoch häufig Methoden bezeichnet, die nur eine Hoffnung auf das erfolgreiche Lösen eines Problems bieten, aber keine Garantie für die Qualität einer Lösung oder die Laufzeit geben. Der Begriff *randomisierte Suchheuristik* wird meist für Heuristiken verwendet, die zur Optimierung einer beliebigen Funktion

$$f : S \rightarrow Z$$

auf einem Suchraum  $S$  eingesetzt werden und dabei zufällige Entscheidungen treffen. Dazu muss mindestens eine Halbordnung auf  $Z$  existieren, um den Begriff der Optimierung zu definieren.

Für viele theoretische Optimierungsprobleme ist  $Z = \mathbb{R} \cup \{-\infty, +\infty\}$  und es wird die gewohnte Totalordnung  $\leq$  verwendet. Es gilt dann, je nach Problemstellung, die Funktion  $f$  zu maximieren oder zu minimieren. Da wir in dieser Arbeit mit dem Graph-Bisektion-Problem ein Minimierungsproblem behandeln, ist mit dem Begriff *Optimieren* stets das Minimieren einer Funktion gemeint. Es wird also ein  $s \in S$  gesucht, für das  $f(s)$  minimal ist. Die Funktion  $f$  wird als *Fitnessfunktion* bezeichnet, der Funktionswert eines Suchpunktes  $s \in S$  als *Fitness*.

Eine Suchheuristik schreibt nun zum einen vor, wie ein erster Suchpunkt  $x_0$  gewählt wird, zum anderen wie ein nächster Suchpunkt  $x_{t+1}$  gewählt wird, unter der Voraussetzung, dass die Suchpunkte  $x_0, \dots, x_t$  ausgewertet wurden. Unabhängig von der Komplexität einer Funktionsauswertung geben wir dem nötigen Zeitintervall die Größe 1 und sehen somit die Anzahl der Funktionsauswertungen als Maß für die Zeit an. Da solche iterativen Heuristiken häufig keine fixe Rechenzeit haben, wird normalerweise zur Reduzierung der Komplexität nur ein Teil der bereits ausgewerteten Suchpunkte zur Bestimmung des nächsten Suchpunkts benutzt. Die Menge dieser Suchpunkte bezeichnen wir als *Population*.

Generell verwenden randomisierte Suchheuristiken nur sehr wenig problemspezifisches Wissen. Meist wird nur eine Nachbarschaftsrelation oder Distanzfunktion der Elemente des Suchraums vorausgesetzt und anhand des Vergleichs der Fitness einiger Suchpunkte das weitere Verhalten beeinflusst. In diesem Fall liegt für den Algorithmus ein *Black-Box-Szenario* vor, da nur der Funktionswert für eine bestimmte Eingabe, jedoch kein Wissen über die Funktion selbst verwendet wird.

Zu diesen randomisierten Suchheuristiken zählen unter anderem die randomisierte lokale Suche (RLS), der Metropolis-Algorithmus und Simulated Annealing (SA), sowie evolutionäre Algorithmen (EA). Diese Algorithmen arbeiten mit der Annahme, dass im Suchraum bezüglich einer Distanzfunktion nahe liegende Punkte auch eine ähnliche Fitness aufweisen. Wird ein Suchpunkt mit guter Fitness gefunden, so vermutet man weitere Punkte mit ähnlicher bzw. besserer Fitness in der Nähe. Auf dieser Annahme beruht die Konstruktion eines *Mutationsoperators*, welcher aus einem Suchpunkt  $x$  einen anderen Punkt  $x'$  erzeugt, welcher im Erwartungswert in der Nähe von  $x$  liegt. Ist der maximale Abstand von  $x'$  zu  $x$  im Suchraum eingeschränkt, so spricht man von einer lokalen Suche, andernfalls von einer globalen.

Bei evolutionären Algorithmen wird der Prozess der Evolution auf den Suchpunkten nachgebildet, was im Sinne der Kreuzung zweier Lebewesen zu einem *Crossoveroperator*, auch als *Rekombination* bekannt, führt. Dabei geht man davon aus, dass gute Suchpunkte besondere Vorzüge haben, welche kombiniert zu einem noch besseren Suchpunkt führen. Für den Einsatz eines Crossoveroperators ist selbstverständlich eine Population von mindestens zwei Suchpunkten notwendig, um diesen sinnvoll einsetzen zu können. Rekombination und Mutation fasst man unter dem Begriff der *Reproduktion* zusammen. Ein Suchpunkt wird im Kontext der Evolution auch als *Individuum* bezeichnet, die Individuen, die als Eingabe der Reproduktion dienen, als *Eltern* und die so erzeugten als *Kinder*.

Als dritter Operator der evolutionären Algorithmen dient die *Selektion*, welche dazu verwendet wird, eine Menge von Individuen aus der Population auszuwählen. Sie findet zum einen ihren Einsatz in der Auswahl der Eltern für die Reproduktion, zum anderen zur Reduktion der Populationsgröße.

Der Basisalgorithmus eines evolutionären Algorithmus besteht dann aus der Wiederholung der Schritte *Auswertung*, *Selektion der Population*, *Selektion zur Reproduktion* und *Reproduktion* bis ein Abbruchkriterium eintritt. Die Population des  $i$ -ten Schleifendurchlaufs wird dabei auch als  $i$ -te *Generation* bezeichnet. Zu Beginn wird die Population eines EA durch zufälliges Erzeugen von Individuen nach einer gegebenen Wahrscheinlichkeitsverteilung über dem Suchraum, meist der Gleichverteilung, gefüllt.

Für die Anzahl der Individuen in der Population  $\mu$ , die Anzahl erzeugter Kinder pro Schleifendurchlauf  $\lambda$  und die Anzahl der Eltern zur Erzeugung eines Individuums  $\rho$  wird ein EA kurz als  $(\mu/\rho + \lambda)$ -EA bzw.  $(\mu/\rho, \lambda)$ -EA bezeichnet, abhängig davon, ob eine Plus- oder Komma-Selektion bei der Erstellung der nächsten Generation verwendet wird. Bei der Plus-Selektion werden die Individuen der alten Population, sowie die neu erzeugten Kinder der Reproduktion in einer Population zusammengefasst, auf welche dann die Selektion angewendet wird, um die Population der nächsten Generation zu erzeugen. Im Gegensatz dazu wird bei der Komma-Selektion die neue Population nur aus den Kindern erstellt, während alle Eltern verworfen werden. Dazu muss notwendigerweise  $\lambda \geq \mu$  gelten, damit die Populationsgröße nicht abnimmt. Sinnvollerweise sollte jedoch  $\lambda > \mu$  gelten, damit die Suche zielgerichtet sein kann. Die Selektion für die Erzeugung der nächsten Population erfolgt in der Regel fitnessbasiert, da es sich sonst um eine ungerichtete randomisierte Suche handelt. Plus- und Komma-Selektion sind Formen der Schnitt Selektion, bei der die Individuen einer Population absteigend (bzw. aufsteigend im Falle eines Maximierungsproblems) nach deren Fitness sortiert werden und die neue Population aus den ersten  $\mu$  Individuen dieser Ordnung besteht.

Im Rahmen dieser Arbeit werden wir uns von den hier vorgestellten Algorithmen nur mit dem (1+1)-EA beschäftigen, welcher entsprechend eine Population eines einzelnen Individuums  $x$  verwaltet und pro Iteration genau ein Kind  $x'$  durch Mutation erstellt. Dieses Kind bildet die neue Population, d.h. es wird genau dann akzeptiert, wenn seine Fitness nicht schlechter ist als die des Elters. Andernfalls wird  $x'$  verworfen (s. Algorithmus 1). Als Repräsentation des Suchraums wird für die meisten Analysen außerhalb dieser Arbeit ein Bitstring  $\{0,1\}^n$  verwendet. Die Distanzfunktion auf diesem Suchraum ist dann meist der Hamming-Abstand:

**Definition 1.3.** Der Hamming-Abstand  $H : A^n \rightarrow \mathbb{N}$  zweier  $n$ -stelliger Zeichenketten  $x, y \in A^n$  ist definiert durch

$$H(x, y) = \sum_{i=1}^n (1 - \delta_{x_i, y_i})$$

mit dem Kronecker-Delta aus Definition 1.1.

Ein EA hat grundsätzlich einen globalen Mutationsoperator. Meist wird dabei jedes Bit mit Wahrscheinlichkeit  $1/n$  geflippt.

---

#### Algorithmus 1 (1+1)-EA auf $\{0,1\}^n$

---

```

Wähle  $x \in \{0,1\}^n$  zufällig gemäß Gleichverteilung
repeat
  Erstelle  $x' := x$ 
  for  $1 \leq i \leq n$  do
    Setze  $x'_i := 1 - x_i$  mit Wahrscheinlichkeit  $1/n$ 
  end for
  if  $f(x') \leq f(x)$  then
     $x := x'$ 
  end if
until Abbruchkriterium

```

---

Bei der entsprechenden lokalen Variante des (1+1)-EA, der RLS, wird dagegen ein Mutationsoperator verwendet, der immer genau ein Bit verändert, welches nach Gleichverteilung gewählt wird.

Als Abbruchkriterium dient meist entweder eine Grenze für die Zahl der Generationen oder eine Schranke für die Verbesserung der Fitness. Zum Zwecke der Analysierbarkeit gehen wir in dieser Arbeit davon aus, dass der Algorithmus so lange läuft, bis ein optimales Individuum erzeugt wurde. Im Falle der RLS kann dies zu einer unendlichen Laufzeit führen, wenn das Individuum nicht durch die lokalen Mutationen erreicht werden kann.

### 1.2.1 Multikriterielle Optimierung

Evolutionäre Algorithmen können zwar auf Funktionen  $f : S \rightarrow \mathbb{R}$  zu guten Ergebnissen führen, jedoch gilt es in vielen realen Problemen mehrere Parameter zu optimieren, die möglicherweise in Konkurrenz zueinander stehen. So möchte man beispielsweise bei einem Motor für Verbrauch und Kraft möglichst gute Werte finden, obwohl sich dies

für gewöhnlich widerspricht. Für eine solche Funktion  $f : S \rightarrow Z$  existiert dann unter Umständen keine totale Ordnung auf  $Z$  und ein gewöhnlicher EA nach Algorithmus 1 lässt sich nicht anwenden, da nicht-vergleichbare Elemente in  $Z$  existieren. Sogenannte *multiobjective evolutionary algorithms* (MOEA) arbeiten dagegen auf  $f$  falls eine Halbordnung  $\preceq$  auf  $Z$  existiert. Anhand dieser Ordnung lassen sich dann auch die Suchpunkte in  $S$  vergleichen.

**Definition 1.4.** Sei  $\preceq$  eine Halbordnung auf den Elementen einer Menge  $Z$  und  $f : S \rightarrow Z$ . Dann ist die Quasiordnung  $\preceq_f$  auf  $S$  gegeben durch

$$x \preceq_f y \Leftrightarrow f(x) \preceq f(y).$$

Da normalerweise für  $\preceq_f$  die Antisymmetrie nicht gegeben ist, ist  $\preceq_f$  nur eine Quasiordnung.

Für zwei Elemente  $x, y \in Z$  gilt:

- $x$  dominiert  $y$ , falls  $x \prec_f y$  mit  $x \prec_f y \Leftrightarrow a \preceq_f y \wedge y \not\preceq_f x$ .
- $x$  dominiert  $y$  schwach, falls  $x \preceq_f y$ .

Die Menge der optimalen Elemente für eine Halbordnung  $\preceq$  kann dann mehrere nicht-vergleichbare Elemente beinhalten.

**Definition 1.5.** Für die Halbordnung  $\preceq$  einer Menge  $Z$  wird  $x^* \in S$  als pareto-optimal bezeichnet, wenn  $x^*$  von keinem anderen  $x \in S$  dominiert wird. Die Menge der pareto-optimalen Suchpunkte

$$P^* := \{x^* \in S \mid \nexists x \in X : x \prec_f x^*\}$$

wird als Pareto-Menge bezeichnet, die Menge  $F^* = f(P^*)$  als Pareto-Front.

Ziel eines MOEA ist es genügend pareto-optimale Punkte im Suchraum zu finden um die Pareto-Front abzudecken. Dazu sollte ein MOEA, im Gegensatz zu dem (1+1)-EA, eine Population verwalten um alle nicht-vergleichbaren Elemente beizubehalten. Ansonsten würde eine Entscheidung getroffen, welches Individuum präferiert wird. Nach endlicher Zeit sollte dann die Population mit der Pareto-Front übereinstimmen.

In [18] stellen Laumanns et al. zwei lokale multikriterielle evolutionäre Algorithmen vor (SEMO und FEMO). Der SEMO (simple evolutionary multi-objective optimizer) verwaltet eine Population  $\mathbb{P}$  von sich nicht dominierenden Individuen und wählt in jeder Iteration ein Individuum  $x$ , um durch Mutation ein Kind  $x'$  zu erstellen. Wird dieses  $x'$  von keinem  $y \in \mathbb{P}$  dominiert, so wird es in  $\mathbb{P}$  aufgenommen. Dabei werden alle Individuen  $y' \in \mathbb{P}$  für die  $x' \preceq y'$  gilt, die also von  $x'$  dominiert werden, aus der Population entfernt. Giel [12] erweitert den SEMO zu einer globalen Variante durch Verwendung eines globalen Mutationsoperators und unterscheidet dann den lokalen und den globalen SEMO. Der globale SEMO wird in Algorithmus 2 beschrieben. Der lokale SEMO verwendet dagegen einen zur RLS analogen Mutationsoperator, in dem genau ein Bit nach Gleichverteilung gewählt und geflippt wird.

Zu Analyse Zwecken sehen wir wieder von der Existenz eines Abbruchkriteriums ab und bestimmen die erwartete Zeit, bis für die Population  $\mathbb{P}$  des SEMO gilt  $f(\mathbb{P}) = F^*$ . Dann ist die Population eine Untermenge der Pareto-Menge  $P^*$ .

**Algorithmus 2** Globaler SEMO auf  $\{0, 1\}^n$ 


---

```

Wähle  $x \in \{0, 1\}^n$  zufällig gemäß Gleichverteilung
 $\mathbb{P} \leftarrow x$ 
repeat
  Wähle  $x \in \mathbb{P}$  gemäß Gleichverteilung
  Erstelle  $x' := x$ 
  for  $1 \leq i \leq n$  do
    Setze  $x'_i := 1 - x_i$  mit Wahrscheinlichkeit  $1/n$ 
  end for
  if  $\nexists z \in \mathbb{P} : z \prec_f x'$  then
     $\mathbb{P} \leftarrow \{z \in \mathbb{P} \mid x' \not\prec_f z\} \cup \{x'\}$ 
  end if
until Abbruchkriterium

```

---

### 1.3 Branch-and-Bound-Algorithmen

Branch-and-Bound-Algorithmen sind eine weitere Form heuristischer Methoden, die zur Gewinnung von optimalen Ergebnissen auf einer Funktion  $f : S \rightarrow \mathbb{R}$  eingesetzt werden. Dabei wird der Suchraum rekursiv in Teilmengen zerlegt, welche durch Knoten eines Suchbaums repräsentiert werden (s. Abbildung 1.1). Die Blätter des Baumes entsprechen dann den Elementen des Suchraums, also den Lösungen. Im Falle des Graph-Bisection-Problems wäre beispielsweise eine Zerlegung sinnvoll, bei der an einem Knoten der Tiefe  $t$  im Suchbaum nur Partitionen betrachtet werden, in denen die bzgl. einer vorher definierten Ordnung ersten  $t$  Knoten fest den Mengen  $L$  und  $R$  zugeteilt sind.

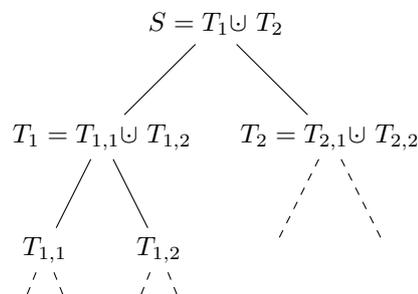


Abbildung 1.1: Suchbaum eines Branch-and-Bound-Algorithmus

Im Verlauf des Algorithmus wird eine obere Schranke  $s^u \geq \min_{x \in S} f(x)$  für den Wert einer optimalen Lösung verwaltet (z.B. der Wert der besten bisher gefundenen Lösung) und mit einer unteren Schranke  $s_T^l \leq \min_{x \in T} f(x)$  für die Elemente der Teilmenge  $T$  am Knoten  $k$  im Suchbaum verglichen. Ist  $s_T^l > s^u$ , kann der Zweig zu  $k$  im Baum abgeschnitten werden, da keine optimale Lösung in  $T$  enthalten sein kann. Die Berechnung der unteren Schranken muss entsprechend problemspezifisch durchgeführt werden, somit bezeichnet Branch-and-Bound nur den Oberbegriff für spezifische Algorithmen, die nach diesem Paradigma arbeiten.

Zu Beginn kann entweder mittels einer Heuristik eine obere Schranke für das Optimum erzeugt werden, oder es wird mit einer Schranke von  $+\infty$  begonnen. Anschließend wird der Suchbaum nach einem Verfahren wie Depth-First-Search oder Breadth-First-Search durchlaufen, wobei auch der Wert der unteren Schranken der Kinder eines Knotens  $k$  zur Reihenfolge der Abarbeitung der Kinder herangezogen werden kann. Man bezeichnet den Algorithmus dann entsprechend als DFSBnB oder BFSBnB.

Durch das Abtrennen der Äste im Suchbaum kann verhindert werden, dass alle Punkte im Suchraum ausgewertet werden. Je nach Problemstellung, verwendeter Methode zur Berechnung unterer Schranken und Art der Partitionierung des Suchraums kann dies von der Auswertung nur weniger Suchpunkte bis zur vollständigen Auswertung variieren. Das Verfahren hat also keine fixe Laufzeit, ist jedoch optimal. Da aber ab dem Erreichen des ersten Blattes stets der Wert einer besten Lösung verwaltet wird, kann das Verfahren auch vorzeitig abgebrochen werden und liefert dann eine Approximation.

Der Suchbaum selbst wird normalerweise nicht explizit als solcher im Speicher erstellt, da sonst eine exponentielle Speicherlast vorliegen würde. Stattdessen wird ein Knoten nur dann erstellt, wenn er untersucht wird und kann wieder verworfen werden, wenn keine Berechnungen der Kinder mehr anfallen.

## 1.4 Der A\*-Algorithmus und IDA\*

Der A\*-Algorithmus ist ein aus dem Kontext des Kürzeste-Wege-Problems bekannter Algorithmus. Jedem Knoten  $k$  wird auf dem Weg zwischen einem Start- und einem Zielknoten ein Funktionswert  $f(k) = g(k) + h(k)$  zugewiesen, der eine untere Schranke für die Kosten eines Weges über  $k$  angibt. Die Funktion ist zusammengesetzt aus den Kosten  $g(k)$ , die vom Startknoten zum Knoten  $k$  anfallen und der Funktion  $h(k)$ , die eine untere Schranke für nötigen Kosten zum Zielknoten beschreibt. A\* verwaltet eine Liste, in der die *bekannt*en Knoten nach deren Funktionswert  $f(k)$  sortiert gespeichert werden. Es wird dann in jeder Iteration der Knoten mit minimalem  $f(k)$  aus der Liste entfernt und dessen unbekannte Nachbarknoten, nach Berechnung von  $f$ , in die Liste aufgenommen. Zu Beginn befindet sich in der Liste nur der Startknoten. Im einfachsten Fall ist  $h(x) = 0$ , dann entspricht A\* dem Dijkstra-Algorithmus.

Angewandt auf einen Suchbaum, wie in Abschnitt 1.3, dient der Algorithmus der Suche optimaler Lösungen. Die Funktion  $g(k)$  eines Knotens  $k$  mit der Teilmenge  $T$  beschreibt dann die Kosten, die für alle  $x \in T$  auf jeden Fall anfallen. Eine untere Schranke für die noch notwendigen Kosten, um ein Individuum über die Eigenschaften von  $T$  hinaus zu vervollständigen, wird mit  $h(x)$  berechnet. Sobald A\* ein Blatt im Suchbaum zur Liste der bekannten Knoten hinzufügt, so ist mit diesem eine optimale Lösung gefunden.

Ein großer Nachteil von A\* ist der exponentielle Speicherbedarf. Die Liste der bekannten Knoten kann sehr schnell auf die Größenordnung des Suchraums anwachsen. Abhilfe schafft unter anderem Iterative-Deepening A\* (IDA\*) mit einem, in der Tiefe des Suchbaums, linearen Speicherbedarf. IDA\* verhält sich wie eine Folge von DFS-Läufen, wobei ein Schwellwert  $C_i$  verwaltet wird. Die Tiefensuche einer Iteration verfolgt dabei nur Knoten mit  $f(k) \leq C_i$  weiter. Wird ein Kind  $k$  mit  $f(k) > C_i$  gefunden, so wird  $k$  einer Menge  $S_i$  zugefügt. Zu Beginn ist  $C_0 := f(r)$ , für die Wurzel

$r$  des Suchbaums. Am Ende der  $i$ -ten Iteration wird der nächste Schwellwert durch  $C_{i+1} := \min_{x \in S_i} f(x)$  berechnet, mit der Menge der im Lauf untersuchten Knoten  $S_i$ , welche den Schwellwert  $C_i$  überschreiten. Durch die vielfachen DFS-Läufe werden von IDA\* im Gegensatz zu A\* Knoten mehrfach ausgewertet. Auch hier entspricht das erste Blatt, welches den Schwellwert nicht überschreitet, einer optimalen Lösung.

## 1.5 Analytische Hilfsmittel

### 1.5.1 Mutationswahrscheinlichkeiten und Erwartungswerte

Für eine Reihe von unabhängigen Zufallszahlen  $X_1, \dots, X_n$  mit  $P(X_i = 1) = p$ ,  $P(X_i = 0) = 1 - p$ , so genannte Bernoulli-Experimente, sei  $X = \sum_i X_i$ . Man nennt  $X$  dann binomialverteilt nach  $B(n, p)$  und es ist

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Für den Erwartungswert und die Varianz von  $X$  gilt

$$\begin{aligned} E(X) &= np \\ \text{Var}(X) &= np(1-p). \end{aligned}$$

Eine Mutation des (1+1)-EA aus Algorithmus 1 kann als Resultat solcher Bernoulli-Experimente mit Wahrscheinlichkeit  $p = 1/n$  betrachtet werden, da jedes Bit des Individuums unabhängig mit Wahrscheinlichkeit  $1/n$  geflippt wird. Die Anzahl der in einer Mutation geflippten Bits  $X$  ist damit binomialverteilt nach  $B(n, 1/n)$ . Im Erwartungswert wird damit pro Mutation genau ein Bit mutiert. Der Erwartungswert für die Dauer, bis eine bestimmte Mutation mit der Wahrscheinlichkeit  $p$  eintritt ist gleich  $1/p$ .

Wir benötigen häufig eine untere Schranke für die Wahrscheinlichkeit genau  $m$  bestimmte Zeichen der Repräsentation zu mutieren. Diese Wahrscheinlichkeit lässt sich abschätzen durch

$$\frac{1}{n^m} \left(1 - \frac{1}{n}\right)^{n-m} \geq \frac{1}{en^m},$$

denn es gilt

$$\left(1 - \frac{1}{n}\right)^{n-1} \geq e^{-1} \geq \left(1 - \frac{1}{n}\right)^n.$$

In Kapitel 3.3.1 werden wir einen (1+1)-EA mit geringerer Mutationswahrscheinlichkeit  $i/n$  für ein  $i < 1$  betrachten. In diesem Fall wird die Wahrscheinlichkeit einen Knoten nicht zu mutieren steigen und es gilt

$$\begin{aligned} \left(1 - \frac{i}{n}\right)^{n-1} &\stackrel{(*)}{\geq} \left(1 - \frac{i}{n}\right)^{n-i} \\ &\geq \left(\left(1 - \frac{i}{n}\right)^{n/i-1}\right)^i \\ &\geq e^{-i}, \end{aligned}$$

wobei in  $(*)$   $i \leq 1$  und  $(1 - i/n) < 1$  verwendet wurde.

### 1.5.2 Die Chernoff-Ungleichungen

Für eine binomialverteilte Zufallsvariable  $X$  wurde von Herman Chernoff eine obere Schranke für die Wahrscheinlichkeit einer Abweichung von  $X$  von seinem Erwartungswert in beide Richtungen bestimmt. Für eine Abweichung vom Erwartungswert nach oben gilt für jedes  $\delta > 0$

$$P[X \geq (1 + \delta) \cdot E(X)] \leq \left( \frac{e^\delta}{(1 + \delta)^{1+\delta}} \right)^{E(X)} \leq \exp\left(-\frac{\min\{\delta, \delta^2\}}{3} E(X)\right), \quad (1.1)$$

für eine Abweichung nach unten gilt für  $\delta \in [0, 1]$

$$P[X \leq (1 - \delta) \cdot E(X)] \leq \left( \frac{e^{-\delta}}{(1 - \delta)^{1-\delta}} \right)^{E(X)} \leq \exp\left(-\frac{\delta^2}{2} E(X)\right). \quad (1.2)$$

Ein Beweis für Gleichung 1.2, sowie Gleichung 1.1 für  $0 \leq \delta \leq 1$  befindet sich in [13]. Der Fall  $\delta > 1$  kann leicht gezeigt werden, indem folgendes Lemma für die zweite Ungleichung in (1.1) verwendet wird.

**Lemma 1.6.** Für  $\delta \geq 1$  gilt

$$(1 + \delta) \ln(1 + \delta) \geq \delta + \frac{1}{3}\delta. \quad (1.3)$$

*Beweis.* Für  $\delta = 1$  ist  $2 \ln 2 > 1, 38 > 4/3$ . Die Ableitung beider Seiten nach  $\delta$  führt zu

$$1 + \ln(1 + \delta) \geq \frac{4}{3}$$

und es ist  $\ln(1 + \delta) \geq \ln(2) > 1/3$  für  $\delta \geq 1$ . Somit steigt die linke Seite von Gleichung 1.3 schneller als die rechte und da die Gleichung für  $\delta = 1$  erfüllt ist, gilt sie für alle  $\delta > 1$ .  $\square$

### 1.5.3 Die Stirling-Formel

Mit der Stirling-Formel werden untere und obere Schranken für die Fakultät  $n!$  berechnet. Es gilt

$$\sqrt{2\pi n}^{n+1/2} e^{-n+\frac{1}{12n+1}} < n! < \sqrt{2\pi n}^{n+1/2} e^{-n+\frac{1}{12n}}.$$

Damit werden in [21] die folgenden unteren und oberen Schranken für Binomialkoeffizienten gezeigt:

$$\binom{mn}{n} \geq \frac{m^{m(n-1)+1}}{(m-1)^{(m-1)(n-1)}} n^{-\frac{1}{2}} \quad (1.4)$$

$$\binom{mn}{n} \leq \frac{1}{\sqrt{2\pi}} n^{-\frac{1}{2}} \frac{m^{mn+\frac{1}{2}}}{(m-1)^{(m-1)n+\frac{1}{2}}} \quad (1.5)$$

### 1.5.4 Random Walk auf Plateaus

Haben im Suchraum mehrere bezüglich einer Nachbarschaftsbeziehung in Relation stehende Knoten die gleiche Fitness, so spricht man von einem Plateau. Bei großen Plateaus gibt es daher keinen Hinweis auf eine „Richtung“ im Suchraum, die zu einer Verbesserung führen kann. Deshalb wird das Verhalten des (1+1)-EA auf einem Plateau im englischen auch als *Random Walk* bezeichnet, denn der (1+1)-EA bewegt sich mit gleicher Wahrscheinlichkeit in jede Richtung, solange der Rand des Plateaus nicht erreicht wurde. In Kapitel 3.2 und 3.3.3 werden wir dieses Verhalten auf Ringen analysieren. Dazu benötigen wir das folgende Lemma:

**Lemma 1.7.** *Für eine nach  $B(N, 1/2)$ , mit  $N = cn^2$ , binomialverteilte Zufallsvariable  $X$  existieren Konstanten  $c$  und  $q > 0$ , so dass gilt:*

$$P(X \geq \frac{1}{2}cn^2 + \frac{1}{2}n) \geq q.$$

*Beweis.* Nach der Verteilung von  $X$  gilt

$$P(X = k) = \binom{N}{k} 2^{-N},$$

was für  $k = N/2$  maximal ist. Nach Gleichung 1.5 gilt dann

$$P(X = N/2) \leq \sqrt{\frac{2}{\pi}} N^{-1/2}.$$

Da die Verteilung zu beiden Seiten von  $N/2$  abnimmt gilt für eine Abweichung von  $N/2$  um höchstens  $n/2$

$$P\left(\left|X - \frac{N}{2}\right| \leq \frac{1}{2}n\right) \leq \frac{1}{2}n \cdot \sqrt{\frac{2}{\pi}} N^{-1/2} = \sqrt{\frac{2}{\pi c}},$$

da  $N = cn^2$  ist. Die Verteilung von  $X$  ist symmetrisch um  $N/2$ . Somit gilt für die Wahrscheinlichkeit einer Abweichung vom Maximum um mehr als  $n/2$  in eine Richtung

$$P\left(X \geq \frac{1}{2}N + \frac{1}{2}n\right) \geq \frac{1}{2} - \frac{1}{\sqrt{2\pi c}} =: q.$$

Es ist dann  $q > 0$  für alle  $c > 2/\pi$ . □

Dieses Lemma gilt für mit Wahrscheinlichkeit  $p = 1/2$  binomialverteilte Zufallsvariablen. Da wir häufig obere und untere Schranken für Wahrscheinlichkeiten verwenden, werden die entsprechenden Zufallsvariablen oft nicht exakt binomialverteilt sein, sondern die Einzelexperimente liefern mit einer Wahrscheinlichkeit von mindestens  $1/2$  ein für uns günstiges Ergebnis. Für eine solche Zufallsvariable  $X' = \sum_i X_i$  mit den Einzelexperimenten  $X_i$ , für die gilt  $P(X_i = 0) \leq 1/2$  und  $P(X_i = 1) \geq 1/2$ , ist jedoch stets

$$P(X' \geq k) \geq P(X \geq k)$$

und somit gilt das Lemma auch für  $X'$ .

### 1.5.5 Das Gamblers-Ruin-Problem

Das Gamblers-Ruin-Problem beschreibt die Situation eines Spielers beim Glücksspiel. Dabei gewinnt oder verliert er mit Wahrscheinlichkeit  $p$  bzw.  $q$  einen Euro. Mit einem Startguthaben  $z$  soll der Spieler gegen einen Gegner mit Guthaben  $a - z$  antreten. Nun stellt sich die Frage wie hoch die Wahrscheinlichkeit für den Spieler ist, das Guthaben auf  $a$  zu steigern, bevor der Ruin des Spielers eintritt. Feller zeigt in [9], dass folgende Lemmata gelten:

**Lemma 1.8.** *Für  $r = q/p$  ist die Wahrscheinlichkeit das Guthaben  $a$  zu erreichen, bevor der Spieler auf das Guthaben  $b$  sinkt gleich*

$$\frac{r^{z-b} - r^{a-b}}{1 - r^{a-b}}.$$

**Lemma 1.9.** *Die erwartete Zeit bis das Guthaben des Spielers entweder  $a$  oder  $b$  erreicht, ist gleich*

$$\frac{z-b}{q-p} - \frac{a-b}{q-p} \cdot \frac{1 - (q/p)^{z-b}}{1 - (q/p)^{a-b}}.$$

Übertragen werden kann dies auf das Verhalten eines Algorithmus, der bei einem Random Walk eines Individuums mit Wahrscheinlichkeit  $p$  einen Parameter  $P$  um 1 erhöht, mit Wahrscheinlichkeit  $q$  um 1 senkt und sonst konstant hält. Es lässt sich dann Abschätzen, mit welcher Wahrscheinlichkeit  $P = a$  vor  $P = 0$  gilt und somit z.B. der Rand eines Plateaus erreicht ist. Außerdem kann dann die erwartete Laufzeit berechnet werden.

## 2 Algorithmen für GBP

### 2.1 Ein optimales Lösungsverfahren

Von Felner werden in [10] vier Methoden zur Berechnung unterer Schranken für Teilmengen des Suchraums für das Graph-Bisection-Problem vorgestellt. Diese werden sowohl für DFSBnB als auch IDA\* (s. Kapitel 1.3 und 1.4) zur Suche nach optimalen Lösungen eingesetzt. Für eine vorher festgelegte Reihenfolge  $v_{i_1}, \dots, v_{i_n}$  der Knoten des Graphen  $G = (V, E)$  wird an einem Punkt der Tiefe  $t$  im Suchbaum entschieden, ob  $v_{i_t}$  in die Menge  $L$  oder  $R$  eingeteilt wird. Somit erhalten wir einen binären Baum. In der Funktion  $f(k) = g(k) + h(k)$ , für einen Knoten  $k$  im Suchbaum, beschreibt  $g(k)$  die Kanten zwischen den Knoten von  $G$ , die bereits in die Mengen  $L$  oder  $R$  eingeteilt wurden.

**Definition 2.1.** Für einen Graphen  $G = (V, E)$  mit  $V = \{v_1, \dots, v_n\}$  sei  $x \in \{0, 1, 2\}^n$  und die Mengen  $L(x)$  und  $R(x)$ , nach Definition 1.1 gegeben durch

$$\begin{aligned}L(x) &= \{v_i \in V \mid x_i = 1\}, \\R(x) &= \{v_i \in V \mid x_i = 2\}.\end{aligned}$$

Dann ist  $\mathcal{S}^* := \{x \mid x \in \{0, 1, 2\}^n, |L(x)| \leq n, |R(x)| \leq n\}$  die Menge der partiellen Knoteneinteilungen und  $\mathcal{S} := \{x \mid x \in \{1, 2\}^{2n}, |L(x)| = |R(x)|\}$  die Menge der vollständigen Knoteneinteilungen.

Wir repräsentieren damit im Folgenden einen Knoten im Suchbaum durch eine ternäre Zeichenkette  $x \in \mathcal{S}^*$ , wobei  $x_i = 1$  genau dann gilt, wenn  $v_i \in L$  und  $x_i = 2$  genau dann gilt, wenn  $v_i \in R$ . Für einen Knoten  $v_i$  des Graphen  $G$ , der noch keiner Menge zugeteilt wurde gilt  $x_i = 0$ . Diese Knoten bezeichnen wir als *freie Knoten*.

**Definition 2.2.** Die Menge  $\mathcal{S}(x) \subseteq \mathcal{S}$  ist definiert durch

$$y \in \mathcal{S}(x) \Leftrightarrow L \subseteq L(y), R \subseteq R(y),$$

für die am Knoten von  $x$  vorliegende partielle Einteilung in  $L$  und  $R$ .

Dann lässt sich  $g(k)$  beschreiben durch die entsprechende Zeichenkette  $x$  am Knoten  $k$  mit

$$g(x) = \sum_{\{v,w\} \in E, v \in L, w \in R} 1.$$

Als einfachste Variante stellt Felner  $f_0(x) = g(x)$  vor, die er als Uniform-Cost-Search (UCS) bezeichnet. Bei der Berechnung von  $g(x)$  muss nur zu den Kosten des Elternknotens im Suchbaum die Zahl der Kanten des neu zugeteilten Knotens  $v \in V$

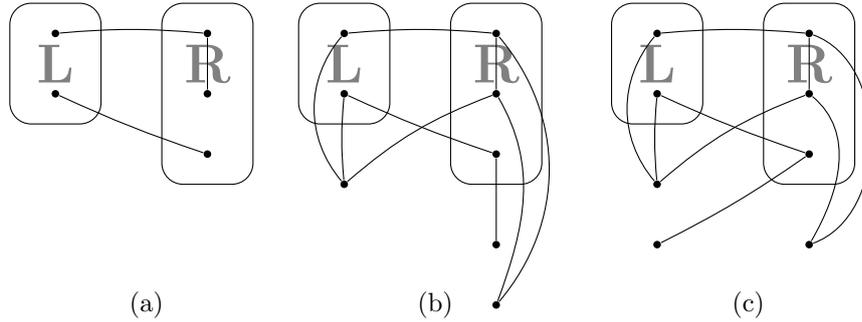


Abbildung 2.1: Die von Felner vorgestellten unteren Schranken für eine partielle Zuweisung der Knoten zu  $L$  und  $R$ . Für diesen Graphen ist  $f_0(x) = 2$  (a),  $f_1(x) = 3$  (b) und  $f_2(x) = 4$  (c).

in die Mengen  $L$  bzw.  $R$  addiert werden. Somit ist die Komplexität der Berechnung von  $f_0(x)$  im Grad  $d_v$  des neu eingeteilten Knotens  $v$  beschränkt.

Die zweite Schranke  $f_1(x) = g(x) + h_1(x)$  verwendet eine einfache Heuristik  $h_1(x)$ , um eine untere Schranke der Kosten zu erstellen. Für einen Knoten  $v$  in  $G$  sei  $d(v, M)$  mit  $M \subseteq V$  die Anzahl der Kanten zwischen  $v$  und Knoten der Menge  $M$ . Es ist dann

$$h_1(x) = \sum_{v_i \in V, x_i = 0} \min\{d(v_i, L), d(v_i, R)\}.$$

Dies können wir als Zuteilung der Knoten mit  $x_i = 0$  zu Mengen  $L'$  und  $R'$  verstehen, so dass eine minimale Anzahl an Kanten zwischen  $L'$  und  $R'$ , sowie zwischen  $R'$  und  $L$  verläuft (s. Abbildung 2.1 (b)). Die Schranke ist offensichtlich korrekt, da jede andere Einteilung der Knoten zu einem schlechteren Schnitt führt. Dabei wird jedoch nicht beachtet, dass bei GBP die Anzahl der Knoten auf beiden Schnittseiten gleich sein muss. Die Schranke gilt also auch für eine relaxierte Variante von GBP.

Für diese Schranke kann für jeden Knoten  $w$  der Wert  $d(w, L)$  und  $d(w, R)$  gespeichert werden und im Falle der Zuweisung eines Knotens  $v$  der entsprechende Wert für die Nachbarknoten von  $v$  geändert werden. Die Komplexität für  $f_1(x)$  ist also ebenso im Grad  $d_v$  des neu eingeteilten Knotens  $v$  beschränkt.

Bei der Heuristik  $h_2(x)$  der dritten Funktion  $f_2(x) = g(x) + h_2(x)$  wird die Beschränkung der Größen von  $L$  und  $R$  dagegen berücksichtigt. Für jeden freien Knoten  $v_i$  mit  $x_i = 0$  wird der Wert  $N_L(v_i) = d(v_i, L) - d(v_i, R)$  berechnet. Anschließend werden die Knoten absteigend nach ihrer Wertigkeit sortiert und die ersten  $n/2 - |L|$  der Menge  $L'$  zugeordnet, die verbleibenden  $n/2 - |R|$  der Menge  $R'$ . Es ist dann

$$h_2(x) = \sum_{x_i = 0, v_i \in L'} d(v_i, R) + \sum_{x_i = 0, v_i \in R'} d(v_i, L).$$

$N_L(v)$  kann als Maß für den Vorteil verstanden werden den Knoten  $v$  der Menge  $L$  statt  $R$  zuzuordnen. Jede andere Partition der Knoten kann durch Vertauschungen zweier Knotenmengen  $L'' \subseteq L'$  und  $R'' \subseteq R'$  erreicht werden. Ein Tausch der Mengen

führt zu einer Änderung des Wertes der Bisektion von

$$\sum_{v \in L''} (d(v, L) - d(v, R)) + \sum_{w \in R''} (d(w, R) - d(w, L)) = \sum_{v \in L''} N_L(v) - \sum_{w \in R''} N_L(w).$$

Durch die Sortierung der Knoten gilt  $\sum_{v \in L''} N_L(v) - \sum_{w \in R''} N_L(w) \geq 0$  und somit ist  $f_2(x)$  eine untere Schranke für alle Suchpunkte in  $\mathcal{S}(x)$ .

Auch hier können in Zeit  $\mathcal{O}(d_v)$  für den Grad  $d_v$  des neu einsortierten Knotens  $v$  die Werte  $d(w, L)$ ,  $d(w, R)$  für freie Knoten geändert werden. Zusätzlich muss eine Sortierung der freien Knoten nach  $N_L(w)$  verwaltet werden. Hierzu verwendet Felner einen Bucketsort, bei dem jeder freie Knoten in einem Bucket entsprechend seinem Funktionswert  $N_L$  enthalten ist. Eine Umsortierung ist nur notwendig für freie Knoten  $w$ , für die  $\{v, w\} \in E$  ist. Also ist die Änderung von  $f_2$  in  $\mathcal{O}(d_v)$  berechenbar.

Für eine Einteilung der Knoten in  $L$ ,  $L'$ ,  $R$  und  $R'$  entsprechend  $f_2$  klassifiziert Felner die Kanten von  $G$  folgendermaßen (vgl. Abbildung 2.2):

1. Typ I: Kanten, die innerhalb der Mengen  $L \cup L'$  bzw.  $R \cup R'$  verlaufen und damit nicht zum Schnitt beitragen.
2. Typ II: Kanten, welche zwischen den Mengen  $L$  und  $R$  verlaufen.
3. Typ III: Kanten zwischen  $L'$  und  $R$ , sowie  $R'$  und  $L$ .
4. Typ IV: Kanten, die zwischen den Mengen  $L'$  und  $R'$  verlaufen.

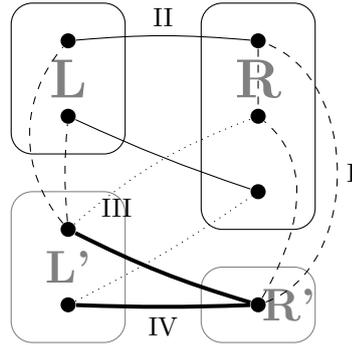


Abbildung 2.2: Die verschiedenen Klassen der Kanten einer partiellen Bisektion. Gestrichelt sind Kanten vom Typ I, durchgezogen solche vom Typ II, die gepunkteten Kanten entsprechen Typ III und die fett gedruckten Typ IV.

Die bisherigen Schranken bieten eine Abschätzung für Kanten der Typen I, II und III. In seiner vierten Heuristik  $h_3(x)$  bezieht Felner die Kanten vom Typ IV mit ein. Zunächst erfolgt eine Einordnung der freien Knoten in  $L'$  und  $R'$  wie für  $h_2$ . In der nach  $N_L(x)$  absteigend sortierten Liste der freien Knoten  $v_{s_1}, \dots, v_{s_m}$  sind  $v_{s_i} \in L'$  mit  $i = n/2 - |L|$  und  $v_{s_{i+1}} \in R'$  die Knoten, die bei einem Seitenwechsel die wenigsten neuen Kanten zur Bisektion hinzufügen. Wir bezeichnen  $v_{s_i}$  als  $v_{L'}$  und  $v_{s_{i+1}}$  als  $v_{R'}$ . Jede Bisektion, in der ein Knoten  $v$  anders eingeordnet ist hat also, wenn die Balance erhalten bleiben soll, mindestens

$$N(v) = \begin{cases} N_L(v) - N_L(v_{R'}) & \text{falls } v \in L' \\ N_L(v_{L'}) - N_L(v) & \text{falls } v \in R' \end{cases}$$

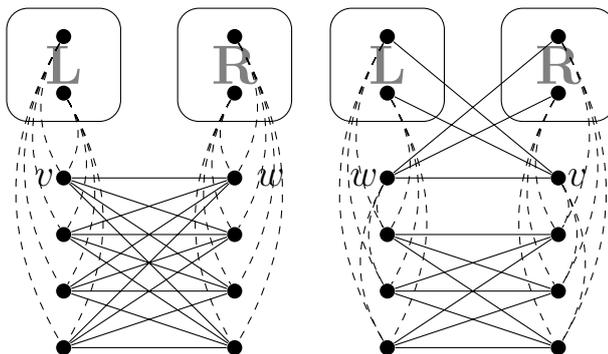


Abbildung 2.3: Ein Gegenbeispiel für die von Felner verwendete Heuristik  $h_3$ . Links: Die Einteilung der Knoten nach  $h_2$  mit einem Wert von  $f_3(x) = 16$ . Es ist  $N_L(v) = 2$  und  $N_L(w) = -2$  und damit  $N(v) = N(w) = 4$ , so dass alle Kanten Typ IV in  $h_3$  eingehen. Rechts: Eine Bisektion, in der zwei Knoten aus  $L'$  und  $R'$  vertauscht wurden, mit Wert 14.

zusätzliche Kanten vom Typ III. Für einen Knoten  $v \in L'$  mit adjazenten Kanten vom Typ IV ist nach Felner  $N(v)$  eine obere Schranke für die Anzahl der Kanten vom Typ IV, die wir hinzunehmen dürfen, ohne  $\min_{y \in \mathcal{S}^*(x)} \text{GBP}_G(y)$  zu überschreiten. So wird jedem Knoten  $v$  ein Kontingent von  $N(v)$  möglichen adjazenten Kanten vom Typ IV zugeschrieben und es soll eine maximale Anzahl der Kanten vom Typ IV gewählt werden, ohne dass ein Knoten sein Kontingent überschreitet. Dies ist eine Generalisierung des Maximal-Matching-Problems auf dem bipartiten Graphen  $G' = (L' \cup R', E_{IV})$  mit der Menge der Kanten vom Typ IV  $E_{IV}$ . Die Anzahl der so gezählten Kanten ist  $h_3(x)$  und es ist  $f_3(x) = g(x) + h_2(x) + h_3(x)$ .

Für diese Schranke lässt sich jedoch ein einfaches Gegenbeispiel konstruieren, in dem der Untergraph der Knoten  $L' \cup R'$  ein vollständiger bipartiter Graph ist (s. Abbildung 2.3). Felner spricht jedem Knoten ein Kontingent von  $N(x)$  zu. Tatsächlich können so aber die Knoten  $v_{L'}$  und  $v_{R'}$  jeweils Kanten vom Typ IV einbringen in Höhe der Kosten, die ein Tausch der beiden bringt, also unter Umständen doppelt so viel wie erlaubt.

Die Komplexität von  $f_3$  setzt sich jetzt aus der von  $f_2$ , sowie der Berechnung des generalisierten Matchings auf dem Untergraph der freien Knoten zusammen. Dieses kann über die Berechnung eines maximalen Flusses gelöst werden. Der Graph für das Flussproblem ist dann der Untergraph der freien Knoten mit zusätzlich einer Quelle und einer Senke, wobei von der Quelle zu jedem freien Knoten  $v \in L'$  eine Kante mit Kapazität  $N(v)$  führt, sowie von jedem Knoten  $w \in R'$  eine Kante mit Kapazität  $N(w)$  zur Senke.

Wir werden uns im Folgenden immer wieder auf diese Funktionen beziehen. Taucht also die Bezeichnung  $f_i$  in den späteren Analysen auf, so wird damit Bezug auf die hier vorgestellten Funktionen genommen.

## 2.2 Analyalisierte Suchheuristiken

Wir stellen nun die drei randomisierten Suchheuristiken vor, die wir auf einigen Instanzen des Graph-Bisection-Problems in den Kapiteln 3 und 4 analysieren werden. Wie in Kapitel 1.2 beschrieben gehen wir in den Analysen davon aus, dass ein Abbruchkriterium niemals erfüllt wird und die Algorithmen für immer in der Schleife verweilen. Wir messen dann die erwartete Zeit bis ein Optimum erreicht wurde. Für die ersten beiden Heuristiken, welche eine Populationsgröße von 1 haben, bezeichnen wir die beiden Mengen der Bisektion des aktuellen Individuums mit  $L$  und  $R$ , anstelle von  $L(x)$  und  $R(x)$ , wenn keine Verwechslungsgefahr besteht.

### 2.2.1 RLS

Das einfachste hier analysierte heuristische Suchverfahren ist die RLS. Als Repräsentation der Lösung für einen Graphen  $G = (V, E)$  mit  $|V| = n$  verwenden wir die in Definition 1.2 vorgestellte Menge  $\mathcal{S} = \{x \mid x \in \{1, 2\}^n, |L(x)| = |R(x)|\}$ . Jedes Element in  $\mathcal{S}$  lässt sich für eine  $n$ -stellige Permutation  $\pi \in \Pi_n$  darstellen durch  $\pi(1^{n/2}2^{n/2})$ . Zur Generierung des initialen Individuums der RLS wird deshalb eine  $n$ -stellige Permutation  $\pi \in \Pi_n$  gleichverteilt gewählt und  $x = \pi(1^{n/2}2^{n/2})$  erzeugt. Als Nachbarn von  $x$  verwenden wir Suchpunkte, die im Hamming-Abstand 2 zu  $x$  stehen, d.h. Zeichenketten  $x' \in \mathcal{S}$ , die sich nur in zwei Zeichen unterscheiden. Für diese zwei Zeichen  $x_i$  und  $x_j$  gilt zwangsläufig  $x_i \neq x_j$ , da sonst die Balance von  $L$  und  $R$  verletzt ist. In einer Mutation werden also zwei Knoten  $v_i$  und  $v_j$  gewählt und vertauscht, wobei  $v_i \in L$  und  $v_j \in R$  gilt. Nach der Mutation gilt dann  $x'_i = 2$  und  $x'_j = 1$ . Die verwendete RLS ist in Algorithmus 3 beschrieben.

---

#### Algorithmus 3 RLS für GBP

---

```

Wähle  $x = \pi(1^{n/2}2^{n/2})$  gleichverteilt
repeat
   $x' := x$ 
  Wähle gleichverteilt  $1 \leq i, j \leq n$  mit  $x_i = 1, x_j = 2$  und setze  $x'_i := 2, x'_j := 1$ 
  if  $\text{GBP}_G(x') \leq \text{GBP}_G(x)$  then
     $x := x'$ 
  end if
until Abbruchkriterium

```

---

### 2.2.2 (1+1)-EA

Der simple Mutationsoperator der RLS kann leicht in lokalen Optima gefangen werden, aus dem ein Entkommen unmöglich ist. Ein (1+1)-EA erlaubt dagegen die Mutation beliebig vieler Knoten. Da beim Graph-Bisection-Problem allerdings stets die gleiche Anzahl an Knoten auf beiden Seiten der Bisektion mutiert werden sollte, um eine zulässige Lösung zu erhalten, werden wir den Mutationsoperator des klassischen (1+1)-EA hier anpassen. Wie in 1.5.1 beschrieben ist die Anzahl der durch eine Standardmutation mutierten Zeichen einer Zeichenkette der Länge  $n$  binomialverteilt nach  $B(n, 1/n)$ . Wir wollen nun jeden Knoten aus der Menge  $L$  mit Wahrscheinlichkeit  $1/|L|$

unabhängig voneinander mutieren und anschließend die gleiche Anzahl an Knoten der Menge  $R$  ändern. Für einen Graphen  $G = (V, E)$  mit  $|V| = n$  wird deshalb in Algorithmus 4 die Zahl der zu vertauschenden Knoten  $M$  vorher nach der Binomialverteilung  $B(n/2, 2/n)$  gewählt. Danach werden aus den Mengen  $L$  und  $R$  je  $M$  Knoten gleichverteilt gewählt und vertauscht. Damit ergibt sich eine Wahrscheinlichkeit einer bestimmten  $m$ -Zeichen-Mutation von

$$P(M = m) \binom{n/2}{m}^{-2} = \left(\frac{2}{n}\right)^m \left(1 - \frac{2}{n}\right)^{n/2-m} \binom{n/2}{m}^{-1}.$$

Insbesondere entspricht die Wahrscheinlichkeit eines Tausches von zwei bestimmten Knoten der Wahrscheinlichkeit der Mutation dieser Knoten durch die vorgestellte RLS, wobei jedoch gleichzeitig andere Knoten mutiert werden können.

Wir bezeichnen mit  $p := 1/|L|$  die Mutationswahrscheinlichkeit von Algorithmus 4. Für eine geänderte Mutationswahrscheinlichkeit  $p$  wird  $M$  gemäß  $B(n/2, p)$  gewählt und es werden wieder in  $L$  und  $R$  je  $M$  Knoten gewählt, die vertauscht werden. Damit ergibt sich eine Wahrscheinlichkeit einer bestimmten  $m$ -Zeichen-Mutation von

$$p^m (1 - p)^{n/2-m} \binom{n/2}{m}^{-1}.$$

---

#### Algorithmus 4 (1+1)-EA für GBP

---

Wähle  $\pi \in \Pi_n$  gemäß Gleichverteilung und erzeuge  $x := \pi(1^{n/2}2^{n/2})$   
**repeat**  
     $x' := x$   
    Wähle  $M \sim B(n/2, p)$   
    Wähle  $M$  Zeichen  $x_{s_1}, \dots, x_{s_M}$  aus  $L$ , sowie  $m$  Zeichen  $x_{s_{M+1}}, \dots, x_{s_{2M}}$  aus  $R$   
    gemäß Gleichverteilung  
    Setze  $x'_{s_i} := 3 - x_{s_i}$  für alle  $s_i$   
    **if**  $\text{GBP}_G(x') \leq \text{GBP}_G(x)$  **then**  
         $x := x'$   
    **end if**  
**until** Abbruchkriterium

---

### 2.2.3 SEMO

Ein anderer Ansatz, welcher sich an dem von Felner vorgestellten Algorithmus aus Abschnitt 2.1 orientiert, ist der im Folgenden beschriebene multikriterielle Algorithmus. Dieser arbeitet, im Gegensatz zu den beiden vorher vorgestellten Suchheuristiken, auf einer ternären Repräsentation der Knoteneinteilung, entsprechend der Knoten des Suchbaums in Abschnitt 2.1.

Wir verwenden als Suchraum die Menge  $\mathcal{S}^*$  aus Definition 2.1 mit den folgenden Bezeichnungen:

**Definition 2.3.** Für eine partielle Knoteneinteilung  $x \in \mathcal{S}^*$  wird ein Knoten  $v_i$  mit  $x_i \in \{1, 2\}$  als eingeteilt, ein Knoten mit  $x_i = 0$  als freier Knoten bezeichnet.

Zur Berechnung einer oberen Schranke für die Qualität einer solchen Knoteneinteilung sollen die von Felner verwendeten Funktionen  $f_0, \dots, f_3$  aus Kapitel 2.1 dienen. Da diese generell eine untere Schranke für die Bisektion berechnen, kann der Funktionswert nicht als Fitness eines (1+1)-EA verwendet werden. Dieser würde höchstwahrscheinlich im Punkt  $0^n$  enden. Daher verwenden wir einen multikriteriellen Ansatz auf der folgenden Funktion:

**Definition 2.4.** Für einen Graphen  $G = (V, E)$  mit  $|V| = n$  und die in Kapitel 2.1 definierte Funktion  $f_i$ ,  $i \in \{0, 1, 2, 3\}$  sei

$$\begin{aligned} \text{GBP}_G^i : \{0, 1, 2\}^n &\rightarrow \{0, \dots, |E|, \infty\} \times \{0, \dots, n\} \\ x &\mapsto \begin{cases} (f_i(x), \text{zeros}(x)) & \text{für } x \in \mathcal{S}^* \\ (\infty, \text{zeros}(x)) & \text{sonst} \end{cases} \end{aligned}$$

mit

$$\text{zeros}(x) := \sum_{i=1}^n \delta_{x_i, 0},$$

wobei  $\delta_{i,j}$  wieder dem Kronecker-Delta entspricht.

Zur Optimierung von  $\text{GBP}_G^i$  benutzen wir dann die folgende Ordnung der Elemente in  $\{0, \dots, |E|, \infty\} \times \{0, \dots, n\}$ :

**Definition 2.5.** Die Halbordnung  $\preceq$  auf  $\mathbb{R}^m$  ist definiert durch

$$x \preceq y \Leftrightarrow \forall 1 \leq i \leq m : x_i \leq y_i, \quad (2.1)$$

für  $x, y \in \mathbb{R}^m$ .

Der SEMO aus Kapitel 1.2.1 wird dann in seiner globalen wie auch der lokalen Variante zur Optimierung von  $\text{GBP}_G^i$  eingesetzt. Initialisiert werden beide Algorithmen mit einer Population eines gültigen Individuums  $x \in \mathcal{S}^*$ . Beim global SEMO wird bei einer Mutation jedes  $x_i$  mit Wahrscheinlichkeit  $1/n$  zufällig gleichverteilt gewählt und ebenfalls gleichverteilt in einen der beiden Werte  $x_i + 1 \pmod 3$  oder  $x_i + 2 \pmod 3$  geändert. Der local SEMO mutiert nur ein einzelnes Zeichen eines Individuums. Dieses wird gleichverteilt gewählt und mit gleicher Wahrscheinlichkeit in  $x_i + 1 \pmod 3$  oder  $x_i + 2 \pmod 3$  verändert. Wird durch eine Mutation ein Individuum außerhalb des Suchraums  $\mathcal{S}^*$  erzeugt, so wird dieses mit einer Fitness von  $\infty$  im ersten Argument des Ergebnistupels bestraft.

Eine optimale Bisektion des Graphen lässt sich dann in der Pareto-Menge  $P^*$  finden. Da es sich um ein Minimierungsproblem handelt, existiert in jedem Falle in der Pareto-Front  $F_i^*$  von  $\text{GBP}_G^i$  ein Element  $(f_i(x), 0)$  und dementsprechend ein  $x$  mit  $\text{zeros}(x) = 0$  und minimalem  $\text{GBP}_G(x)$  in der Pareto-Menge  $P_i^*$ .

Die Größe der Population ist für den hier beschriebenen Algorithmus folgendermaßen beschränkt:

**Lemma 2.6.** Für die Population  $\mathbb{P}$  des global SEMO, sowie des local SEMO auf  $\text{GBP}_G^i$ ,  $i = 0, 1, 2, 3$  eines Graphen  $G = (V, E)$  mit  $|V| = n$  gilt stets  $|\mathbb{P}| \leq n + 1$ .

*Beweis.* Die Funktion  $\text{zeros}$  kann nur Werte von 0 (alle Knoten sind eingeteilt) bis  $n$  (alle Knoten sind frei) annehmen. Nach Definition 1.4 von  $\preceq_{\text{GBP}_G^i}$  dominieren sich Individuen mit gleichem Wert von  $\text{zeros}$  mindestens schwach. Somit kann für jeden Wert  $z \in \{0, \dots, n\}$  maximal ein Individuum  $x$  mit  $\text{zeros}(x) = z$  in  $\mathbb{P}$  enthalten sein und es ergibt sich  $|\mathbb{P}| \leq n + 1$ .  $\square$

Ein Vorteil der ternären Repräsentation ist die Möglichkeit neuer Mutationen und das Erreichen anderer Punkte im Suchraum durch neu eröffnete Pfade. Übertragen auf einen Branch-and-Bound-Algorithmus kann der SEMO mögliche Fehler beim Einordnen der Knoten durch das Entfernen selbiger wieder rückgängig machen und sogar die Reihenfolge der Einordnung frei wählen. Gleichmaßen stellt die Vergrößerung des Suchraums des SEMO aber eine Redundanz bezüglich des tatsächlichen Suchraums von GBP dar. Während  $\mathcal{S}$  für einen Graphen mit  $n$  Knoten eine Größe von  $\binom{n}{n/2} = \Theta(2^n / \sqrt{n})$  hat, hat der Suchraum  $\{0, 1, 2\}^n$  des SEMO eine Größe von  $3^n$ . Es existiert jedoch keine allgemeine Theorie, die besagt, wann und ob eine Vergrößerung des Suchraumes eines evolutionären Algorithmus über den eigentlichen Suchraum hinaus vorteilhaft ist.

Im Laufe dieser Diplomarbeit bezeichnen wir sowohl die disjunkte Aufteilung von  $V$  in  $L$  und  $R$  als Bisektion, als auch das zugehörige Individuum  $x$ . Weiterhin beziehen sich die Bezeichnungen RLS, (1+1)-EA und local bzw. global SEMO im weiteren Verlauf der Arbeit auf die hier beschriebenen Algorithmen zur Anwendung auf GBP.

## 3 Theoretische Analyse

### 3.1 Zwei-Cliquen-Graph

Für die erste Analyse der vorgestellten Algorithmen betrachten wir in diesem Kapitel die Bisektion eines sehr einfachen zusammenhängenden Graphen. Diesen werden wir für den global SEMO, sowie den (1+1)-EA analysieren.

**Definition 3.1.** Ein Graph  $G = (V, E)$  wird als Zwei-Cliquen-Graph bezeichnet, wenn sich  $G$  aus zwei disjunkten Cliques  $C_1, C_2$ ,  $|C_1| = |C_2| = n$ , sowie einer einzelnen Verbindungskante (s. Abbildung 3.1) zusammensetzt. Es ist also  $V = C_1 \cup C_2$  und  $E = \{\{v, w\} \mid v, w \in C_k, k \in \{1, 2\}\} \cup \{v_{C_1}, v_{C_2}\}$  für zwei Knoten  $v_{C_1} \in C_1, v_{C_2} \in C_2$ .

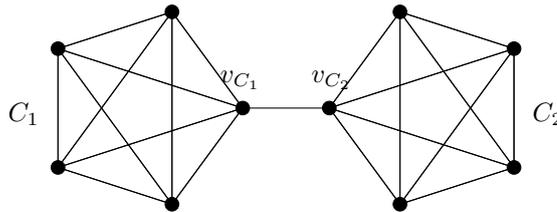


Abbildung 3.1: Ein Zwei-Cliquen-Graph der Größe  $|V| = 10$ .

#### 3.1.1 Der Zwei-Cliquen-Graph mit dem global SEMO

**Lemma 3.2.** Die Pareto-Front  $F_3^*$  von  $\text{GBP}_G^3$ , mit einem Zwei-Cliquen-Graphen  $G = (V, E)$  mit  $|V| = 2n$ , ist  $F_3^* = \{(0, 2n), (1, 0)\}$ , die Größe der Pareto-Menge  $P_3^*$  ist  $|P_3^*| = 3$ .

*Beweis.* Für die Individuen  $x^L = (1^n 2^n)$  und  $x^R = (2^n 1^n)$ , bei denen also jede Clique genau einer Schnittseite zugeordnet ist, gilt  $\text{GBP}_G^3(x^L) = \text{GBP}_G^3(x^R) = (1, 0)$ . Da der Graph zusammenhängend ist, gibt es keine vollständige Einteilung der Knoten, mit weniger Schnittkanten. Weiterhin ist  $\text{GBP}_G^3(x^0) = (0, 2n)$  für  $x^0 = 0^{2n}$ .

Wir betrachten nun ein beliebiges  $x = (x_1, \dots, x_{2n})$  mit  $\text{zeros}(x) = 2n - 1$ . Sei dafür o.B.d.A  $x_i = 1$  für ein  $v_i \in C_1$  und  $x_j = 0$  für alle  $j \neq i$ . Für die zusammengesetzte Funktion  $f_3$  ist somit  $g(x) = 0$ , da kein weiterer Knoten eingeordnet ist. Zur Berechnung von  $h_2$  werden die verbliebenen  $n - 1$  Knoten entsprechend der Differenz der Anzahl ihrer Kanten in die Mengen  $L(x) = \{v_i\}$  und  $R(x) = \emptyset$  absteigend sortiert und in die Mengen  $L'$  und  $R'$  eingeteilt.

Für den Fall  $v_i = v_{C_1}$  sind alle Knoten  $v_j \in (C_1 \setminus \{v_i\}) \cup v_{C_2}$  mit einer adjazenten Kante in die Menge  $L$  gleichwertig, während alle  $v_k \in C_2 \setminus \{v_{C_2}\}$  keine Kante in die Menge  $L$  aufweisen. Da nur  $n - 1$  Knoten in die Menge  $L'$  eingeteilt werden können, muss einer der Knoten  $(C_1 \setminus \{v_i\}) \cup \{v_{C_2}\}$  in  $R'$  eingeordnet werden. Damit schneidet eine Kante vom Typ III (Schnittkanten zwischen  $L' \cup R'$  und  $L \cup R$ , s. Kapitel 2.1) und es ergibt sich  $h_2(x) = 1$ , also ist  $f_3(x) \geq 1$ .

Falls  $v_i \neq v_{C_1}$  gilt, ist  $L' = C_1 \setminus \{v_i\}$  und  $R' = C_2$ . Dann existiert keine Kante vom Typ III im Schnitt und es ist  $h_2(x) = 0$ . Für alle Knoten  $v_j \in L'$  ist  $N(v_j, v_{R'}) = 1$  und für alle  $v_k \in R'$  ist  $N(v_j, v_{L'}) = 1$ , mit den Knoten  $v_{R'}, v_{L'}$  mit maximalem bzw. minimalem  $N_L$ -Wert aus  $L'$  bzw.  $R'$ . Daher kann für jeden Knoten genau eine adjazente Kante vom Typ IV (Kanten zwischen  $L'$  und  $R'$ ) in die Berechnung von  $h_3$  einfließen und es wird die Kante  $\{v_{C_1}, v_{C_2}\}$  auf jeden Fall in  $h_3$  gefunden und mitgezählt. Daher gilt:  $f_3(x) = 1$ .

Also ist  $\text{GBP}_G^3(x) = (1, 2n - 1)$  und  $x$  wird dominiert von  $x^L$  und  $x^R$ . Da wir nicht sicher wissen ob  $f_3$  eine untere Schranke für alle Schnitte mit  $v_i \in L(x)$  bietet, müssen wir auch andere Bisektionen betrachten. Dazu erstellen wir ein  $y$  aus  $x$ , indem wir eine unbestimmte Anzahl von freien Knoten in  $L(y)$  und  $R(y)$  einordnen. Dabei unterscheiden wir die folgenden drei Fälle:

1. Für ein  $i \in \{1, 2\}$  gilt  $C_i \cap L(y) \neq \emptyset$  und  $C_i \cap R(y) \neq \emptyset$ . Dann verläuft mindestens eine Kante zwischen den Mengen  $L(y)$  und  $R(y)$ . Es gilt also  $x^L \preceq_{\text{GBP}_G^3} y$ .
2. Auf einer Seite befindet sich mindestens je ein eingeordneter Knoten aus  $C_1$  und aus  $C_2$ . Sei dies  $L(y)$  gemäß der Wahl der Position von  $v_i$ . Dann besitzt jeder Knoten in  $R'$  auch eine Kante zu einem der Knoten in  $L(x)$ . Diese Kanten vom Typ III führen zu  $h_2(x) \geq 1$ .
3. Es befinden sich keine zwei Knoten aus  $C_1$  und  $C_2$  auf der selben Schnittseite. Da wir o.B.d.A davon ausgehen, dass  $v_i \in L(y)$  ist, sind also alle Knoten aus  $C_1$  in  $L(y)$ , alle aus  $C_2$  in  $R(y)$ . Dann kann sich der Wert  $N_L(v_{R'})$  höchstens verkleinern und der Wert  $N_L(v_{L'})$  höchstens erhöhen. Die Kante  $\{v_{C_1}, v_{C_2}\}$ , sofern nicht als Kante vom Typ III enthalten, wird also genauso in  $h_3$  gefunden.

Damit ist  $f_3(x)$  mit  $\text{zeros}(x) = n - 1$  eine untere Schranke für alle Schnitte im Suchbaum unterhalb von  $x$  und für alle Schnitte  $y$  mit  $\text{zeros}(y) \leq n - 1$  ist die Funktion  $f_3(y) \geq 1$  und damit  $x^L \prec_{\text{GBP}_G^3} y$ . Die Pareto-Front besteht also aus  $\{(0, 2n), (1, 0)\}$  und die Pareto-Menge aus  $x^0, x^L$  und  $x^R$ .  $\square$

**Theorem 3.3.** *Die erwartete Anzahl an Funktionsauswertungen des global SEMO bis zum Erreichen der Pareto-Front  $F_3^*$  für  $\text{GBP}_G^3$  auf einem Zwei-Cliquen-Graph  $G = (V, E)$  mit  $|V| = 2n$  ist  $\Omega(2^{0,83n} n^{2n})$ .*

*Beweis.* Mit Wahrscheinlichkeit  $1/(3^{2n})$  ist das initiale Individuum  $x^{\text{init}} = x^L$  mit  $\text{GBP}_G^3(x^L) = (1, 0)$ . Wie in Lemma 3.2 gezeigt, werden alle  $x \in \mathcal{S}^* \setminus \{0^{2n}\}$  von  $x^L$  dominiert. Die Population  $\mathbb{P} = \{x^L\}$  wird also nur erweitert, wenn  $x^0 = 0^{2n}$  erzeugt wird. Dazu müssen allerdings alle Knoten in  $x^L$  gleichzeitig mutiert werden. Dies geschieht mit Wahrscheinlichkeit

$$P(x^0 \mid x^L) \leq \frac{1}{(2n)^{2n}} \cdot \frac{1}{2^{2n}},$$

also in erwarteter Zeit  $2^{4n}n^{2n}$ . Insgesamt ist damit die erwartete Anzahl an Funktionsauswertungen bis zum Erreichen der Pareto-Front

$$E(T) \geq \frac{1}{3^{2n}} \cdot 2^{4n}n^{2n} = 2^{2n(2-\log 3)}n^{2n} = \Omega(2^{0,83n}n^{2n}).$$

□

Die gezeigte erwartete Laufzeit gilt jedoch nur für die Zeit, bis die komplette Pareto-Front vom global SEMO erfasst wurde. Das Theorem 3.3 hat damit insbesondere keine Aussagekraft über die Dauer, bis eine optimale Bisektion von  $G$  gefunden wurde und zeigt nur die mit Felters Heuristik  $h_3$  (s. Kapitel 2.1) verbundenen Schwierigkeiten in der Analyse des multikriteriellen Ansatzes.

Es zeigte sich außerdem, dass für alle  $i = 1, 2, 3$  die Funktion  $f_i$  in nur einzelnen signifikanten Eigenschaften der Zeichenkette  $x \in \mathcal{S}^*$ , wie der Anzahl „korrekt“ eingeordneter Knoten in  $L$  und  $R$ , nicht monoton ist. Es spielt immer eine Vielzahl an Parametern in  $f_i$  eine Rolle, so dass sich eine Analyse für  $\text{GBP}_G^1, \text{GBP}_G^2$  und  $\text{GBP}_G^3$  als äußerst schwierig erweist. Wir beschränken uns daher auf die Analyse der einfachsten Variante  $\text{GBP}_G^0$ .

**Lemma 3.4.** *Für die Pareto-Front  $F_0^*$  auf  $\text{GBP}_G^0$  mit dem Zwei-Cliquen-Graphen  $G = (V, E)$  mit  $|V| = 2n$  gilt  $F_0^* = \{(0, 1), (1, 0)\}$ , für die Pareto-Menge  $P_0^*$  gilt  $|P_0^*| = 6$ .*

*Beweis.* Für  $x^L = (1^n 2^n)$  und  $x^R = (2^n 1^n)$  gilt  $\text{GBP}_G^0(x^L) = \text{GBP}_G^0(x^R) = (1, 0)$ . Für  $\text{zeros}(x) = 0$  sind dies die beiden Individuen mit minimalem  $f_0$ .

Eine Einteilung  $x$  mit  $\text{zeros}(x) = 1$  kann nur dann keine Schnittkanten enthalten, wenn entweder gilt  $C_1 \cap L(x) = C_2 \cap R(x) = \emptyset$  oder  $C_1 \cap R(x) = C_2 \cap L(x) = \emptyset$ . In jedem anderen Fall gäbe es zwei Knoten einer Clique, welche verschiedenen Seiten der Bisektion zugeordnet sind und damit mindestens eine Schnittkante. Falls nun einer der beiden Knoten  $v_{C_1}$  oder  $v_{C_2}$  keiner Menge zugeordnet wird, so taucht auch die Kante  $\{v_{C_1}, v_{C_2}\}$  nicht in den von  $f_0$  gezählten Schnittkanten auf. Da von  $x^L$  oder  $x^R$  ausgehend jeweils die Möglichkeit besteht  $v_{C_1}$  oder  $v_{C_2}$  aus der Einteilung zu entfernen, existieren genau vier Einteilungen mit  $\text{GBP}_G^0(x) = (0, 1)$ . Da kein  $y$  mit  $\text{GBP}_G^0(y) = (0, 0)$  existieren kann, weil der Zwei-Cliquen-Graph zusammenhängend ist, ist  $F_0^* = \{(0, 1), (1, 0)\}$ . Mit diesen vier Individuen und den beiden Individuen  $x^L$  und  $x^R$  ist  $|P_0^*| = 6$ . □

**Theorem 3.5.** *Die erwartete Anzahl der Funktionsauswertungen, bis die Population  $\mathbb{P}$  des global SEMO die Pareto-Front für  $\text{GBP}_G^0$  mit dem Zwei-Cliquen-Graphen  $G = (V, E)$  für  $|V| = 2n$  abdeckt, ist  $\mathcal{O}(n^2 \log n)$ .*

*Beweis.* Der Beweis gliedert sich in drei Phasen. Zunächst soll ein Individuum  $x^{\text{voll}}$  mit  $\text{zeros}(x^{\text{voll}}) = 0$  erzeugt werden. Danach wird dieses Individuum so lange durch Austausch von Knoten optimiert, bis eines der optimalen Individuen  $x^L$  oder  $x^R$  erzeugt wurde. In der dritten Phase soll die Zeit bis zur Erzeugung eines zweiten pareto-optimalen Individuums  $x^{\text{par}}$  mit  $\text{GBP}_G^0(x^{\text{par}}) = (0, 1)$  abgeschätzt werden.

**Erste Phase** Für die erste Phase definieren wir auf dem Suchraum  $\mathcal{S}^*$  eine Partition in die Mengen  $\mathbb{X}_0^I, \dots, \mathbb{X}_{2n}^I$  mit  $x \in \mathbb{X}_i^I$  genau dann, wenn  $\text{zeros}(x) = i$ . Sei  $x^{\text{init}} \in \mathbb{X}_l^I$  für  $0 \leq l \leq 2n$  das initiale Individuum. Wir bestimmen dann die erwartete Zeit bis ein  $x^{\text{voll}} \in \mathbb{X}_0^I$  erzeugt wird. Die Größe der Population  $\mathbb{P}$  des global SEMO lässt sich nach Lemma 2.6 durch  $|\mathbb{P}| \leq 2n + 1$  abschätzen. Da während der ersten Phase jedoch noch kein  $x \in \mathbb{X}_0^I$  in der Population enthalten ist, gilt  $|\mathbb{P}| \leq 2n$ . Die Wahrscheinlichkeit, aus einem  $x^i \in \mathbb{X}_i^I \cap \mathbb{P}$  ein  $x^{i-1} \in \mathbb{X}_{i-1}^I$  zu erzeugen, lässt sich dann durch die Wahrscheinlichkeit abschätzen,  $x^i = (x_1^i, \dots, x_{2n}^i)$  aus der Population zu wählen und genau eines der  $x_j^i = 0$ , von denen es insgesamt  $i$  gibt, beliebig zu mutieren:

$$\mathbb{P}(x^{i-1} \mid x^i) \geq \frac{1}{2n} \cdot i \cdot \frac{1}{2n} \left(1 - \frac{1}{2n}\right)^{2n-1} \geq \frac{i}{4en^2}.$$

Der Erwartungswert der Zeit  $T_i^I$  bis zum Eintreten dieses Falles ist also durch  $\mathbb{E}(T_i^I) \leq 4en^2/i$  beschränkt. Im schlimmsten Falle ist  $x^{\text{init}} \in \mathbb{X}_{2n}^I$  und die erwartete Laufzeit bis zur Beendigung der ersten Phase beträgt

$$\mathbb{E}(T^I) \leq \sum_{i=1}^{2n} \frac{4en^2}{i} = 4en^2 \sum_{i=1}^{2n} \frac{1}{i} = 4en^2 H_{2n} = \mathcal{O}(n^2 \log n),$$

mit der  $2n$ -ten harmonischen Zahl  $H_{2n}$ .

Ein Individuum  $x'$  kann ein anderes Individuum  $x$  nur dann (schwach) dominieren und damit aus der Population entfernen, wenn keines der beiden Argumente des Ergebnistupels  $\text{GBP}_G^0(x')$  größer ist als das entsprechende in  $\text{GBP}_G^0(x)$ . Daher kann  $x$  nur von  $x'$  dominiert werden, wenn  $\text{zeros}(x') \leq \text{zeros}(x)$  gilt. Sollte also  $x^i \in \mathbb{X}_i^I$  aus der Population entfernt werden, bevor  $x^{i-1} \in \mathbb{X}_{i-1}^I$  daraus generiert werden konnte, so muss es ein  $x' \in \mathbb{X}_j^I \cap \mathbb{P}$  mit  $j \leq i$  geben. Wir können dem Ziel also höchstens näher kommen.

**Zweite Phase** Die zweite Phase beschreibt nun die eigentliche Optimierungsphase. Dabei sollen die folgenden Definitionen helfen:

**Definition 3.6.** Sei die Funktion  $d_M : \mathcal{S}^* \rightarrow \mathbb{Z}$  mit  $M \subseteq V$  gegeben durch  $d_M(x) := |M \cap L(x)| - |M \cap R(x)|$ . Es kann dann  $d_M(x)$  als Maßstab für die Zugehörigkeit einer Knotenmenge  $M$  zur Menge  $L(x)$  gegenüber  $R(x)$  gesehen werden.

Ein Knoten  $v \in V$  in einer Bisektion  $x$ , für die  $d_{C_1}(x) \geq d_{C_2}(x)$  gilt, wird als falscher Knoten bzgl.  $x$  bezeichnet, wenn entweder  $v \in C_1 \cap R(x)$  oder  $v \in C_2 \cap L(x)$  gilt (s. Abbildung 3.2). Für den Fall  $d_{C_1}(x) < d_{C_2}(x)$  gilt analog:  $v$  ist ein falscher Knoten, wenn entweder  $v \in L(x) \cap C_1$  oder  $v \in R(x) \cap C_2$  gilt.

Für eine vollständige Knoteneinteilung  $x$ , also  $\text{zeros}(x) = 0$ , mit  $d_{C_1}(x) \neq d_{C_2}(x)$  gilt damit insbesondere:  $v \in C_j$  ist genau dann ein falscher Knoten, wenn entweder  $v \in L(x)$  und  $|L(x) \cap C_j| < n/2$  oder  $v \in R(x)$  und  $|R(x) \cap C_j| < n/2$  gilt.

Die Anzahl der falschen Knoten einer Knoteneinteilung wird durch

$$\text{wrongs}(x) := \min\{|L(x) \cap C_1| + |R(x) \cap C_2|, |L(x) \cap C_2| + |R(x) \cap C_1|\}$$

berechnet. Die Menge der falschen Knoten ist also eine minimale Menge von Knoten, welche mutiert werden müssen, um ein Individuum  $x'$  mit  $f_0(x') \leq 1$  zu erzeugen. Knoten, die bezüglich  $x$  nicht als falsch gelten, werden als korrekt bezeichnet.

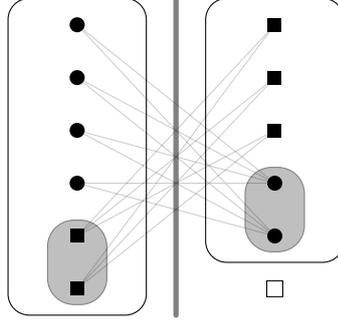


Abbildung 3.2: Darstellung einer Bisektion  $x$  mit einem freien Knoten (weiß). Die Form der Knoten soll die Zugehörigkeit zu den jeweiligen Cliques darstellen. Die vier markierten Knoten sind dann die falschen Knoten bzgl.  $x$ . Wird aus  $x$  eine vollständige Knoteneinteilung  $x'$  durch Hinzunahme des weißen Knotens erzeugt, so zeigt sich außerdem, dass  $\text{wrongs}(x')$  gerade ist.

**Lemma 3.7.** *Für eine vollständige Knoteneinteilung  $x \in \mathbb{X}_0^I$  ist die Anzahl der falschen Knoten  $w := \text{wrongs}(x)$  gerade. Es gilt außerdem, dass  $L(x)$  und  $R(x)$  je  $w/2$  falsche Knoten enthalten.*

*Beweis.* Falls  $d_{C_1}(x) \geq d_{C_2}(x)$ , so ist  $(C_1 \cap R(x)) \cup (C_2 \cap L(x))$  die Menge der falschen Knoten. Es gilt dann

$$|C_1 \cap R(x)| = |C_1 \setminus L(x)| = n - |C_1 \cap L(x)| = |L(x) \setminus C_1| = |C_2 \cap L(x)|,$$

also enthalten  $L(x)$  und  $R(x)$  gleich viele falsche Knoten. Daraus folgt insbesondere, dass  $w$  gerade ist.

Für  $d_{C_1}(x) < d_{C_2}(x)$  lässt sich dies analog zeigen. □

Für ein Individuum  $x^{\text{voll}} \in \mathbb{X}_0^I$  mit vollständiger Knoteneinteilung kann dann  $f_0(x^{\text{voll}})$  mit der Anzahl der falschen Knoten  $w = \text{wrongs}(x)$  und der Kenntnis über die Position der Verbindungskante berechnet werden: Auf jeder Seite des Schnittes befinden sich nach Lemma 3.7  $w/2$  falsche Knoten, welche jeweils durch eine Kante mit jedem der  $n - w/2$  korrekten Knoten der anderen Schnittseite verbunden sind. Also ist

$$f_0(x^{\text{voll}}) = 2 \cdot \frac{w}{2} \cdot \left(n - \frac{w}{2}\right) + r(x^{\text{voll}}) = \frac{w}{2} (2n - w) + r(x^{\text{voll}}), \quad (3.1)$$

mit

$$r(x) = \begin{cases} 1 & \text{falls } v_{C_1} \in L(x) \wedge v_{C_2} \in R(x) \text{ oder } v_{C_1} \in R(x) \wedge v_{C_2} \in L(x) \\ 0 & \text{sonst.} \end{cases} \quad (3.2)$$

Da nicht mehr als die Hälfte der Knoten falsch eingeteilt sein können, gilt  $w \leq n$ . Damit ist  $f_0(x) - r(x)$  monoton fallend in der Anzahl der falschen Knoten  $w$ . Für  $n \geq 2$  gilt sogar, dass  $f_0$  streng monoton fallend ist: Da  $w$  immer gerade ist, muss

sich der Wert von  $w$  um mindestens 2 ändern. Für zwei Individuen  $x$  und  $x'$  mit  $\text{zeros}(x) = \text{zeros}(x') = 0$  und  $\text{wrongs}(x') = \text{wrongs}(x) + 2 =: w + 2$  gilt dann

$$\begin{aligned}
f_0(x') &= \frac{w+2}{2} (2n - (w+2)) + r(x') \\
&\stackrel{(1)}{\geq} \frac{w+2}{2} (2n - (w+2)) \\
&= \frac{w}{2} (2n - w) + (2n - 2) \\
&\stackrel{(2)}{>} \frac{w}{2} (2n - w) + 1 \\
&\stackrel{(3)}{\geq} \frac{w}{2} (2n - w) + r(x) \\
&= f_0(x),
\end{aligned}$$

wobei  $0 \leq r(x) \leq 1$  in <sup>(1)</sup> und <sup>(3)</sup> verwendet wurde und  $n \geq 2$  für <sup>(2)</sup>. Für  $n = 1$  existiert nur eine Bisektion und ihr Spiegelbild. Es gibt also keine zwei Bisektionen mit unterschiedlichem  $w$ .

Ein Individuum  $x^{\text{voll}} \in \mathbb{X}_0^I$  kann dann nur durch ein anderes Individuum mit vollständiger Knoteneinteilung und weniger falschen Knoten dominiert werden.

In der zweiten Phase verwenden wir die disjunkten Teilmengen  $\mathbb{X}_n^{II}, \dots, \mathbb{X}_0^{II}$  des Suchraums. Dabei gilt  $x \in \mathbb{X}_i^{II}$  genau dann, wenn  $\text{zeros}(x) \leq 1$  und  $\text{wrongs}(x) = i$  ist. Aus einem Individuum  $x^i \in \mathbb{X}_i^{II}$  mit  $\text{zeros}(x^i) = 0$  kann dann durch Mutation eines Knotens ein  $x^{i-1} \in \mathbb{X}_{i-1}^{II}$  erstellt werden. Die Wahrscheinlichkeit  $x^i$  aus der Population zu wählen und genau einen der bezüglich  $x^i$  falschen Knoten zu mutieren und zu einem freien Knoten zu machen ist mindestens

$$P(x^{i-1} | x^i) \geq \frac{1}{2n+1} \cdot i \cdot \frac{1}{2n} \left(1 - \frac{1}{2n}\right)^{2n-1} \cdot \frac{1}{2} \geq \frac{i}{4en(2n+1)}, \quad (3.3)$$

da die Größe der Population jetzt nur noch durch  $|\mathbb{P}| \leq 2n+1$  beschränkt ist. Die erwartete Zeit, bis eine solche Mutation eintritt, ist  $E(T_i^{II}) \leq 4en(2n+1)/i$ .

Da von jedem falschen Knoten mindestens eine Kante zu einem korrekten Knoten der anderen Schnittseite existiert, ist  $f_0(x^{i-1}) < f_0(x^i)$ . Also ist  $x^i \not\prec_{\text{GBP}_G^0} x^{i-1}$  und  $x^{i-1}$  wird in die Population aufgenommen, wenn kein  $y \not\prec_{\text{GBP}_G^0} x^{i-1}$  in der Population existiert. Da  $\text{zeros}(x^{i-1}) = 1$  ist, muss für ein solches  $\text{zeros}(y) \leq 1$  gelten. Wir zeigen nun, dass im Falle, dass  $x^{i-1}$  nicht akzeptiert wird,  $\mathbb{P} \cap \bigcup_{l=i-1}^0 \mathbb{X}_l^{II} = \emptyset$  gilt. Dann sind wir aber schon einen Schritt weiter.

**Lemma 3.8.** *Für jede Knoteneinteilung  $x$  mit genau einem freien Knoten  $v$  kann die Fitness  $f_0(x)$  durch die Anzahl der falschen Knoten  $w = \text{wrongs}(x)$  wie folgt berechnet werden:*

$$f_0(x) = \frac{w}{2} (2n - w - 1) + r(x). \quad (3.4)$$

*Beweis.* Sei o.B.d.A.  $|L(x)| = n - 1$ . Falls  $w$  gerade ist, so existieren auf beiden Seiten des Schnittes gleich viele falsche Knoten: Sei  $x'$  die Knoteneinteilung mit  $L(x') = L(x) \cup \{v\}$  und  $R(x') = R(x)$ . Dann muss  $v$  ein korrekter Knoten bezüglich  $x'$  sein. Andernfalls wäre die Zahl der falschen Knoten in einer vollständigen Einteilung

ungerade, was Lemma 3.7 widerspricht. In dieser vollständigen Einteilung  $x'$  müssen dann die falschen Knoten gleichmäßig auf  $L(x')$  und  $R(x')$  aufgeteilt sein, also ebenso in  $x$ .

Es gibt damit  $w/2$  falsche Knoten in  $L(x)$  und  $n - w/2$  korrekte Knoten in  $R(x)$ , zwischen denen paarweise je eine Kante verläuft. In  $R(x)$  existieren ebenso  $w/2$  falsche Knoten, während  $n - w/2 - 1$  korrekte Knoten in  $L(x)$  existieren. Dies ergibt eine Summe der Schnittkanten von

$$\begin{aligned} f_0(x) &= \frac{w}{2} \left( n - \frac{w}{2} \right) + \frac{w}{2} \left( n - \frac{w}{2} - 1 \right) + r(x) \\ &= \frac{w}{2} (2n - w - 1) + r(x), \end{aligned}$$

mit  $r(x)$  aus Gleichung 3.2, was also einen Zusatz von 1 liefert, falls die Verbindungskante  $\{v_{C_1}, v_{C_2}\}$  in der Bisektion enthalten ist.

Für ungerades  $w$  kann ebenso ein  $x'$  mit  $L(x') = L(x) \cup \{v\}$  und  $R(x') = R(x)$  konstruiert werden, welches nach Lemma 3.7 eine gerade Anzahl falscher Knoten enthält, die gleichmäßig auf die beide Schnittseiten verteilt sind. Da  $w$  ungerade ist, muss  $v$  in  $x'$  ein falscher Knoten sein und es ist  $\text{wrongs}(x') = w + 1$ . Nach Konstruktion von  $x'$  ist  $R(x') = R(x)$  und somit sind  $(w + 1)/2$  falsche Knoten in  $R(x)$  enthalten. Es bleiben dann noch  $(w - 1)/2$  falsche Knoten, welche zu  $L(x)$  gehören. Die Gesamtzahl der Schnittkanten berechnet sich dann zu

$$\begin{aligned} f_0(x) &= \frac{w + 1}{2} \left( n - \frac{w - 1}{2} - 1 \right) + \frac{w - 1}{2} \left( n - \frac{w + 1}{2} \right) + r(x) \\ &= \left( \frac{w + 1}{2} + \frac{w - 1}{2} \right) \left( n - \frac{w + 1}{2} \right) + r(x) \\ &= \frac{w}{2} (2n - w - 1) + r(x). \end{aligned}$$

□

**Lemma 3.9.** *Ein Individuum  $x$  mit  $\text{zeros}(x) = 1$  und  $w = \text{wrongs}(x)$  kann nur durch ein Individuum  $x'$  mit  $\text{wrongs}(x') \leq w$  dominiert werden, oder es gilt  $r(x) \geq r(x')$  und  $\text{wrongs}(x) = 2n - 1$ .*

*Beweis.* Ein solches Individuum  $x$  kann nur von einem Individuum  $x'$  mit  $\text{zeros}(x') \leq 1$  dominiert werden. Wir betrachten die beiden Fälle  $\text{zeros}(x') = 1$  und  $\text{zeros}(x') = 0$  getrennt. Beginnend mit dem ersten Fall soll nun gezeigt werden, dass  $f_0$  auf der Menge  $\mathbb{X}_1^I$ , welche alle Individuen  $x'$  mit  $\text{zeros}(x') = 1$  beinhaltet, in  $\text{wrongs}(x)$  monoton steigend ist.

Per definitionem ist  $\text{wrongs}(x)$  das Minimum aus  $|L(x) \cap C_1| + |R(x) \cap C_2|$  und  $|L(x) \cap C_2| + |R(x) \cap C_1|$ . Die Summe beider Argumente ergibt die Anzahl aller eingeordneter Knoten, also  $2n - 1$  für  $x$ . Somit muss eine der beiden Zahlen kleiner oder gleich der Hälfte der Anzahl eingeordneter Knoten sein. Es ist also  $\text{wrongs}(x) \leq$

$\lfloor (2n-1)/2 \rfloor = n-1$ . Für ein  $x'$  mit  $\text{wrongs}(x') = w-1$  gilt dann

$$\begin{aligned}
f_0(x') &= \frac{w-1}{2}(2n - (w-1) - 1) + r(x') \\
&\stackrel{(1)}{\leq} \frac{w-1}{2}(2n - (w-1) - 1) + 1 \\
&= \frac{w}{2}(2n-w) - \frac{1}{2}(2n-w) + 1 \\
&\stackrel{(2)}{\leq} \frac{w}{2}(2n-w) - \frac{1}{2}(2n - (n-1)) + 1 \\
&= \frac{w}{2}(2n-w) - \frac{1}{2}(n+1) + 1 \\
&\stackrel{(3)}{\leq} \frac{w}{2}(2n-w) - \frac{1}{2}(w+2) + 1 \\
&= \frac{w}{2}(2n-w-1) \\
&\stackrel{(4)}{\leq} \frac{w}{2}(2n-w-1) + r(x) \\
&= f_0(x),
\end{aligned}$$

wobei in <sup>(2)</sup> und <sup>(3)</sup>  $\text{wrongs}(x') \leq n-1$  verwendet wurde und in <sup>(1)</sup> und <sup>(4)</sup>  $0 \leq r(x) \leq 1$ . Die Gleichheit von  $f(x')$  und  $f(x)$  gilt nur, wenn  $r(x') = 1$  und  $r(x) = 0$  ist, sonst ist entweder Ungleichung <sup>(1)</sup> und <sup>(4)</sup> echt kleiner, und weiterhin gilt  $w = 2n-1$ , denn sonst sind die Ungleichungen <sup>(2)</sup> und <sup>(3)</sup> echt kleiner. Für  $r(x') \leq r(x)$  oder  $w < 2n-1$  ist somit  $f(x') < f(x)$  und es ist  $x \not\prec_{\text{GBP}^0} x'$ .

Es bleibt zu zeigen, dass  $x$  nicht von einem  $y \in \mathbb{X}_0^I$  (den Individuen mit  $\text{zeros}(y) = 0$ ) dominiert wird, für das  $\text{wrongs}(y) > \text{wrongs}(x)$  gilt. Dies lässt sich einfach zeigen, da jedes  $y$  mit  $\text{wrongs}(y) = w+1$  eine schlechtere Fitness hat und  $f_0$  auf  $\mathbb{X}_0^I$  wie oben gezeigt monoton in  $w$  ist: Mit Gleichung 3.1 gilt

$$\begin{aligned}
f_0(y) &= \frac{w+1}{2}(2n - (w+1)) + r(y) \\
&\stackrel{(5)}{\geq} \frac{w+1}{2}(2n - (w+1)) \\
&= \frac{w}{2}(2n-w-1) + \frac{1}{2}(2n-w-1) \\
&\stackrel{(6)}{\geq} \frac{w}{2}(2n-w-1) + \frac{1}{2}(2n - (n-1) - 1) \\
&= \frac{w}{2}(2n-w-1) + \frac{n}{2} \\
&\stackrel{(7)}{>} \frac{w}{2}(2n-w-1) + 1 \\
&\stackrel{(8)}{\geq} \frac{w}{2}(2n-w-1) + r(x) \\
&= f_0(x),
\end{aligned}$$

für  $n > 2$ , wobei in <sup>(5)</sup> und <sup>(8)</sup> wieder  $0 \leq r(x) \leq 1$  und in <sup>(6)</sup>  $w \leq n-1$  verwendet wurde. Die Ungleichung <sup>(7)</sup> gilt dabei zwar nur für  $n > 2$ , falls aber  $r(x) \leq r(y)$  oder  $w < 2n-1$

gilt, so ist auch  $f_0(y) < f_0(x)$  für  $n \leq 2$ . Durch die strenge Monotonie von  $f_0$  bezüglich  $w$  auf  $\mathbb{X}_0^I$  gilt damit  $f_0(y') > f_0(x)$  für alle  $y'$  mit  $\text{wrongs}(y') > \text{wrongs}(y)$ .  $\square$

Aus einem Individuum  $x^i \in \mathbb{X}_i^{II}$  mit  $\text{zeros}(x^i) = 1$  kann ein  $x^{i-1} \in \mathbb{X}_{i-1}^{II}$  erzeugt werden, indem ein falscher Knoten  $v_j$  auf die andere Schnittseite verschoben wird und damit zu einem korrekten Knoten wird. Die Mutation ändert also  $x_j := 3 - x_j$ . Dies ist natürlich nur für Knoten  $v_j$  möglich, die auf der Schnittseite liegen, die  $n/2$  Knoten enthält, da andernfalls kein Platz für einen Wechsel vorhanden ist. Davon gibt es  $\lceil i/2 \rceil$  solcher Knoten. Die Wahrscheinlichkeit zur Erzeugung eines solchen Individuums beträgt damit

$$\mathbb{P}(x^{i-1} | x^i) \geq \frac{1}{2n+1} \cdot \left\lceil \frac{i}{2} \right\rceil \frac{1}{2n} \left(1 - \frac{1}{2n}\right)^{2n-1} \frac{1}{2} \geq \frac{i}{8en(2n+1)}.$$

Nach Lemma 3.9 gilt  $f_0(x^{i-1}) \leq f_0(x^i)$ . Die Gleichheit  $f_0(x^{i-1}) = f_0(x^i)$  tritt nur genau dann ein, wenn gilt  $r(x^{i-1}) = 1$ ,  $r(x^i) = 0$  und  $w = n - 1$ . Aus  $r(x^i) = 0$  folgt, dass  $v_{C_1}$  und  $v_{C_2}$  in  $x^i$  beide auf der selben Schnittseite liegen. Da die beiden Knoten unterschiedlichen Cliques angehören bedeutet dies, dass einer der beiden Knoten ein falscher Knoten ist, der Andere ein korrekter. Im Spezialfall  $w = n - 1$  führt also die Mutation dieses falschen Knotens zu einem Individuum  $x^{i-1}$  mit gleicher Fitness wie  $x^i$ . Eine Mutation der anderen  $\lceil (n-2)/2 \rceil$  falschen Knoten führt dagegen zu einem Individuum mit  $x^{i-1} \prec_{\text{GBP}_G^0} x^i$ .

Der Fall  $w = n$  tritt nur für ein  $x$  mit  $\text{zeros}(x) = 0$  auf und betrifft somit nicht das Lemma 3.9. Das Verhalten dazu beschreibt Gleichung 3.3.

Wird  $x^i$  von einem  $y^j \in \mathbb{X}_j^{II}$  mit  $j \leq i$  und  $\text{zeros}(y^j) = 0$  dominiert, so kann aus diesem wieder, durch Entfernen eines falschen Knotens entsprechend Gleichung 3.3, in erwarteter Zeit  $4en(2n+1)/j$  ein  $x^{j-1} \in \mathbb{X}_{j-1}^{II}$  erzeugt werden. Damit ergibt sich eine erwartete Zeitspanne von  $\mathbb{E}(T_i^{II}) \leq 8en(2n+1)/i$ , um von einer Population mit  $\mathbb{X}_i^{II} \cap \mathbb{P} \neq \emptyset$  eine Population mit  $\mathbb{X}_{i-1}^{II} \cap \mathbb{P} \neq \emptyset$  zu erreichen, mit  $1 \leq i \leq n$ ,  $i \neq n-1$ . Für die erwartete Gesamtzeit bis zur Generierung eines Individuums  $x \in \mathbb{X}_0^{II}$  ohne falsche Knoten, jedoch mit einem nicht zugeordneten Knoten, gilt also

$$\begin{aligned} \mathbb{E}(T^{II}) &\leq \sum_{i=1}^{n-2} \frac{8en(2n+1)}{i} + \frac{8en(2n+1)}{n-2} + \frac{4en(2n+1)}{n} \\ &= 8en(2n+1) \sum_{i=1}^{n-2} \frac{1}{i} + \frac{8en(2n+1)}{n-2} + 4e(2n+1) \\ &= 8en(2n+1)H_{n-2} + \frac{8en(2n+1)}{n-2} + 4e(2n+1) \\ &= \mathcal{O}(n^2 \log n), \end{aligned}$$

mit der  $(n-2)$ -ten harmonischen Zahl  $H_{n-2}$ . Aus  $x \in \mathbb{X}_0^{II}$  mit  $\text{zeros}(x) = 1$  kann dann ein  $x^{\text{opt}} \in \{x^L, x^R\}$  durch Mutation des freien Knotens mit Wahrscheinlichkeit

$$\mathbb{P}(x^{\text{opt}} | x) \geq \frac{1}{2n+1} \cdot \frac{1}{2n} \cdot \left(1 - \frac{1}{2n}\right)^{2n-1} \cdot \frac{1}{2} \geq \frac{1}{4en(2n+1)},$$

also in einer erwarteter Zeit von höchstens  $4en(2n+1) = \mathcal{O}(n^2)$  erstellt werden.

**Dritte Phase** Ist ein optimales Individuum  $x^{\text{opt}}$  erst einmal gefunden, so kann durch Mutation von einem der beiden Knoten  $v_{C_1}$  und  $v_{C_2}$  ein Knoten  $x^{\text{par}}$  mit  $\text{GBP}_G^0(x^{\text{par}}) = (0, 1)$  erzeugt werden. Da nun das pareto-optimale Individuum  $x^{\text{opt}}$  Teil der Population  $\mathbb{P}$  ist, kann es nur ein zweites Individuum  $x$  mit  $\text{GBP}_G^0(x) = (0, \text{zeros}(x))$  geben. Somit gilt  $|P| \leq 2$ . Die Wahrscheinlichkeit  $x^{\text{par}}$  zu erzeugen ist dann mindestens

$$\frac{1}{2} \cdot 2 \cdot \frac{1}{2n} \left(1 - \frac{1}{2n}\right)^{2n-1} \frac{1}{2} \leq \frac{1}{4en}.$$

Die erwartete Zeit für die letzte Phase ist damit durch  $4en = \mathcal{O}(n)$  beschränkt.

Insgesamt ergibt sich also eine erwartete Laufzeit, die durch  $\mathcal{O}(n^2 \log n)$  nach oben beschränkt ist.  $\square$

### 3.1.2 Der Zwei-Cliquen-Graph mit dem (1+1)-EA

**Theorem 3.10.** *Die erwartete Anzahl der Funktionsauswertungen des (1+1)-EA für  $\text{GBP}_G$  auf dem Zwei-Cliquen-Graph  $G = (V, E)$  mit  $|V| = 2n$ , bis zum Erzeugen eines Optimums, ist  $\mathcal{O}(n^2)$ .*

*Beweis.* Wir verwenden dazu die Funktion  $\text{wrongs} : \mathcal{S}^* \rightarrow \mathbb{N}$  aus Definition 3.6 des Theorems 3.5. Da für  $x \in \mathcal{S} \subset \mathcal{S}^*$  gilt  $\text{zeros}(x) = 0$  und  $f_0(x) = \text{GBP}_G(x)$ , können wir außerdem Gleichung 3.1 zur Berechnung der Fitness verwenden:

$$\text{GBP}_G(x) = \frac{w}{2} (2n - w) + r(x)$$

mit  $r(x)$  entsprechend Gleichung 3.2 aus Theorem 3.5.

Wir bilden nun die Partition  $\mathbb{X}_0, \dots, \mathbb{X}_n$  auf  $\mathcal{S}$  mit  $x \in \mathbb{X}_i$  genau dann, wenn  $\text{wrongs}(x) = i$ . Wie wir aus Lemma 3.7 wissen, ist  $\text{wrongs}(x)$  für alle  $x \in \mathcal{S}$  gerade, da  $\mathcal{S}$  der Menge der vollständigen Knoteneinteilungen entspricht. Dann gilt für die Fitness einer Knoteneinteilung  $x'$  mit  $\text{wrongs}(x') = \text{wrongs}(x) + 2 =: w + 2$

$$\begin{aligned} \text{GBP}_G(x') &= \frac{w+2}{2} (2n - (w+2)) + r(x') \\ &\stackrel{(1)}{\geq} \frac{w}{2} (2n - (w+2)) \\ &= \frac{w}{2} (2n - w) + 2n - 2 \\ &\stackrel{(2)}{>} \frac{w}{2} (2n - w) + 1 \\ &\stackrel{(3)}{\geq} \frac{w}{2} (2n - w) + r(x) \\ &= \text{GBP}_G(x). \end{aligned}$$

In <sup>(1)</sup> und <sup>(3)</sup> wurde wieder  $0 \leq r(x) \leq 1$  verwendet. Die Ungleichung <sup>(2)</sup> gilt zwar nur für  $n > 3/2$ , da aber  $n \in \mathbb{N}$  gilt, reicht die schwächere Bedingung  $n > 1$  aus. Für  $n = 1$  gibt es jedoch keine zwei Individuen  $x'$  und  $x$  mit  $\text{wrongs}(x') \neq \text{wrongs}(x)$ . Also ist  $\text{GBP}_G$  streng monoton fallend in  $\text{wrongs}$ .

Hat der (1+1)-EA ein Individuum  $x$  mit  $\text{wrongs}(x) = w$  erzeugt, so werden nur noch Individuen  $x'$  mit  $\text{wrongs}(x') \leq w$  akzeptiert. Aus einem Individuum  $x \in \mathbb{X}_w$  kann durch Mutation eines der  $w^2$  möglichen Paare falscher Knoten dann ein  $x' \in \mathbb{X}_{w-1}$  erzeugt werden. Dies geschieht mit Wahrscheinlichkeit

$$P(x' | x) \geq w^2 \cdot \frac{1}{n} \left(1 - \frac{1}{n}\right)^{n-1} \cdot \binom{n}{1}^{-1} \geq \frac{w^2}{e n^2},$$

also in einer erwarteten Zeit von höchstens  $E(T_w) \leq e n^2 / w^2$ .

Im schlimmsten Fall ist das initiale Individuum aus  $\mathbb{X}_n$ . Dann ergibt sich eine erwartete Gesamtzeit, bis ein  $x^{opt}$  mit  $\text{wrongs}(x^{opt}) = 0$  erzeugt wurde, von

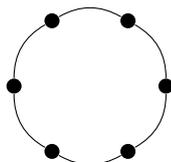
$$E(T) \leq \sum_{i=1}^{n/2} \frac{e n^2}{(2i)^2} \leq e n^2 \sum_{j=1}^n \frac{1}{j^2} \leq \frac{e \pi n^2}{6} = \mathcal{O}(n^2).$$

□

## 3.2 Ringe

Bei der nächsten Analyse widmen wir uns einer Klasse von zyklischen Graphen, den sogenannten Ringen.

**Definition 3.11.** Ein Graph  $G = (V, E)$  mit  $V = \{v_0, \dots, v_{n-1}\}$  wird als Ring bezeichnet, wenn gilt  $E = \{\{v_i, v_j\} \mid j \equiv i + 1 \pmod{n}\}$ .



In der Fitnesslandschaft von  $\text{GBP}_G$  eines Rings existieren viele Plateaus (vgl. Kapitel 1.5.4). Wir werden zeigen, dass der (1+1)-EA in der Lage ist, diese Plateaus mit einem Random Walk zu überwinden. Da der global SEMO aber mit einer Population arbeitet, in der häufig mehrere Individuen mit geringem Hamming-Abstand zur gleichen Zeit existieren, ist ein klassischer Random Walk eines Individuums nicht möglich. Auch wenn ein einzelnes Individuum der Population eine große Strecke durch einen Random Walk zurücklegt, so ist die Wahrscheinlichkeit hoch, dass dieses Individuum von einem andere Individuum schwach dominiert wird, welches einen geringeren Abstand zu dem Rest der Population hat. Daher ist eine Analyse des Verhaltens des global SEMO sehr schwierig und wir werden in diesem Kapitel darauf verzichten.

Im Gegensatz zu dem Zwei-Cliquen-Graph gibt es beim Ring nicht nur eine einzige optimale Einteilung der Knoten, sondern jeder Schnitt durch den Ring, der nur zwei Kanten schneidet ist optimal. Weiterhin führen aufgrund der Plateaus nur wenige Mutationen zu einer Verbesserung. Ein suboptimaler Schnitt kann von der RLS nur verbessert werden, wenn ein einzelner Knoten im Schnitt separiert ist von seinen Nachbarknoten. Ansonsten existiert keine Mutation, die zu einer Verbesserung der Fitness führt (s. Abbildung 3.3). Der (1+1)-EA verbessert mit einem Fehler von  $\mathcal{O}(1/n)$  die Bisektion nur, wenn separierte Blöcke aus nicht mehr als vier Knoten existieren.

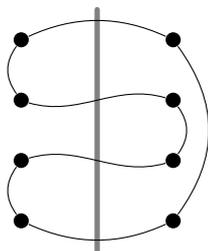


Abbildung 3.3: Beispiel eines Plateaus der Fitnessfunktion auf dem Ring: Keine Mutation, die nur zwei Knoten tauscht, führt zu einer Verbesserung. Erst nach einer weiteren Mutation kann eine Senkung der Fitness erzielt werden.

Zur besseren Anschaulichkeit färben wir die Knoten nach ihrer jeweiligen Einteilung in der Bisektion. Ein Knoten  $v$  sei grün gefärbt, wenn gilt  $v \in l$ , ansonsten rot. Eine beliebige Bisektion des Ringes teilt die Knoten dann in *Blöcke* gleicher Farbe. Dieser Begriff soll auch im Folgenden die maximalen zusammenhängenden, gleichfarbigen Teile des Graphen bezeichnen. Die Länge eines Blocks ist dann durch die Anzahl der enthaltenen Knoten gegeben. Kanten, welche zwei Blöcke verschiedener Farbe teilen, sollen im Folgenden als *Trennstelle* bezeichnet werden (s. Abbildung 3.4). Man beobachtet, dass immer eine gerade Anzahl an Trennstellen vorliegt. Die Anzahl der Trennstellen einer Bisektion ist äquivalent zu ihrer Fitness:

$$\text{GBP}_G(x) = \sum_{i=0}^{n-1} (1 - \delta_{x_i, x_{(i+1) \bmod n}}),$$

mit dem Kronecker-Delta aus Definition 1.1.

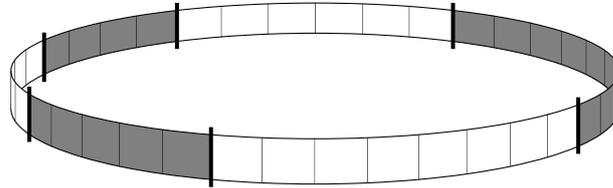


Abbildung 3.4: Ein Ring der Größe  $n = 38$  mit sechs Trennstellen.

### 3.2.1 RLS auf Ringen

**Theorem 3.12.** Für einen Ring  $G = (V, E)$  der Größe  $|V| = n$  ist die erwartete Zeit bis zur Erzeugung eines optimalen Individuums durch die RLS auf  $\text{GBP}_G$  beschränkt durch  $\mathcal{O}(n^4)$ .

*Beweis.* Die Fitness der Bisektion kann nur verbessert werden, indem mindestens zwei Trennstellen entfernt werden. Nach dem obigen Farbmodell wählt eine Mutation zwei Knoten unterschiedlicher Farbe und vertauscht die Farbe der beiden Knoten. Eine Verringerung der Trennstellen erfolgt nur, wenn ein aus einem Knoten bestehender Block die Farbe flippt, während ein anderer Knoten farblich flippt, welcher keine zusätzlichen Trennstellen einführt. Es muss also ein Block bestehend aus einem Knoten existieren.

**Lemma 3.13.** Für eine Bisektion  $x$  mit  $i \geq 4$  Blöcken wird von der RLS in erwarteter Zeit  $\mathcal{O}(n^4/i^3)$  ein Block der Länge eins erzeugt.

*Beweis.* Bei  $i$  Blöcken gibt es einen kürzeste Block der Größe  $l \leq \lfloor n/i \rfloor$ . Im Folgenden sollen nur die Mutationen betrachtet werden, die die Größe dieses Blocks verändern. Diese bezeichnen wir als *relevante* Mutationen. Wir können davon ausgehen, dass alle Blöcke eine Länge von mehr als einem Knoten haben, sonst ist das Ziel bereits erreicht. Es ist dann nur möglich Knoten, die an Trennstellen grenzen, zu mutieren. Das Flippen von Knoten im Inneren eines Blocks wird nicht von der RLS akzeptiert, da dabei

zusätzliche Trennstellen entstehen und damit die Fitness steigt. Es gibt vier Knoten, die an die beiden Trennstellen des kleinsten Blocks grenzen und zu einer Änderung der Größe führen können. Sollten die beiden ausgewählten Knoten einer Mutation beide auf die vier angrenzenden Knoten fallen, so führt dies entweder zu einer Verschiebung des Blocks (je ein Knoten an einer Trennstelle) oder zu einer Verschlechterung der Fitness (zwei Knoten an einer Trennstelle) und damit zum Verwerfen dieser Bisektion. Eine Änderung der Größe erfolgt also lediglich, wenn genau einer der vier Knoten geflippt wird, während der zweite zu flippende Knoten außerhalb des Blocks und der beiden angrenzenden Steine liegt (s. Abbildung 3.5).

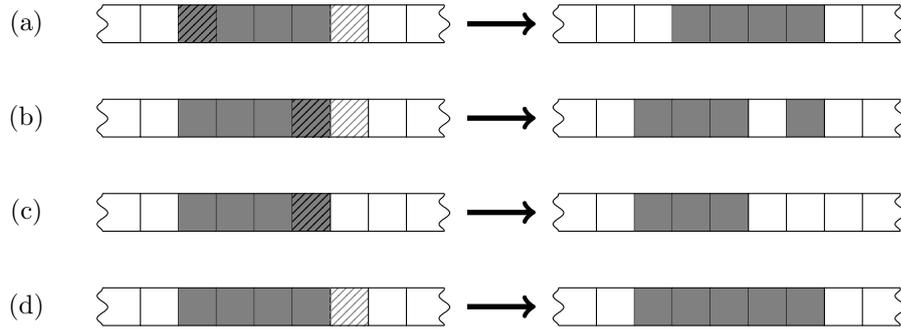


Abbildung 3.5: Mutationen der RLS, bei der Nachbarknoten der Trennstellen mutiert werden (schraffierte Knoten werden mutiert): (a) zeigt eine Verschiebung des Blocks, (b) das Einfügen neuer Trennstellen, (c) und (d) die Verkleinerung bzw. Vergrößerung eines Blocks, wobei ein zweiter Knoten an einer beliebigen anderen Trennstelle geflippt wird.

Mit Wahrscheinlichkeit mindestens  $1/2$  wird bei einer solchen Mutation der Block verkleinert. Nach Lemma 1.5.4 schrumpft ein Block der Größe  $l$  nach einer erwarteten Anzahl solcher Mutationen von  $\mathcal{O}(l^2)$  auf Größe 1.

Die Wahrscheinlichkeit einer relevanten Mutation lässt sich folgendermaßen abschätzen: Eine Mutation der RLS von zwei bestimmten verschiedenfarbigen Knoten hat eine Wahrscheinlichkeit von  $(2/n)^2$ . Es existieren nun vier angrenzende Knoten, welche in einer relevanten Mutation gewählt werden können. Als weiterer zu flippender Knoten muss dann ein Knoten anderer Farbe gewählt werden, welcher an eine Trennstelle grenzt, die nicht zum kleinsten Block gehört. Es existieren  $i/2$  Blöcke anderer Farbe, die mindestens eine solche Trennstelle aufweisen. Insgesamt ergibt sich damit eine untere Schranke für die Wahrscheinlichkeit einer relevanten Mutation von

$$4 \cdot \frac{i}{2} \cdot \left(\frac{2}{n}\right)^2 = \frac{8i}{n^2}$$

und damit eine erwartete Dauer von  $\mathcal{O}(l^2(n^2/i)) = \mathcal{O}(n^4/i^3)$  bis ein Block der Größe 1 erzeugt wurde.  $\square$

**Lemma 3.14.** *Für eine Bisektion  $x$  mit einem Block der Länge 1 gilt: Mit Wahrscheinlichkeit  $1/3$  werden in  $\mathcal{O}(l \cdot n)$  Schritten zwei Blöcke von der RLS aus der Bisektion entfernt.*

*Beweis.* Auch hier betrachten wir die relevanten Mutationen. Für einen Block der Länge 1 gibt es nur noch drei Knoten, die bei einer Mutation die Größe des Blocks ändern. Wieder betrachten wir nur die 2-Zeichen Mutationen, die keine neuen Trennstellen hinzufügen. Diese beinhalten die oben genannten relevanten Mutationen. Hinzu kommt jetzt allerdings noch der Fall des Flippens des einzelnen Knotens sowie eines beliebigen anderen Knotens, welcher nicht an eine Trennstelle grenzt. In diesem Falle wird der kleinste Block entfernt und gleichzeitig ein neuer Block der Größe 1 erzeugt. Es ändert sich also nichts und wir können uns wieder mit den relevanten Mutationen für diesen Block befassen. Da nur noch drei Knoten an den Trennstellen eines kürzesten Blocks liegen, haben diese nun eine Wahrscheinlichkeit von mindestens  $6/(ln)$ . Mit Wahrscheinlichkeit  $1/3$  wird dabei der kleinste Block entfernt.  $\square$

Nach einer erwarteten Anzahl von drei der Random Walks aus Lemma 3.13 wird ein Block damit entfernt. Für eine Bisektion mit  $i$  Blöcken wird also ein Block der Länge  $l \leq \lfloor n/i \rfloor$  in erwarteter Zeit  $\mathcal{O}(l^3 n) = \mathcal{O}(n^4/i^3)$  entfernt und die Anzahl der Blöcke sinkt um zwei, da die Nachbarblöcke verschmelzen.

Eine Bisektion hat im schlechtesten Fall genau  $n$  Blöcke. Die erwartete Zeit bis zum Erreichen einer optimalen Lösung mit nur zwei Blöcken ergibt sich also zu

$$c \cdot \sum_{i=2}^{n/2} \frac{n^4}{(2i)^3} \leq c \cdot \sum_{i=4}^n \frac{n^4}{i^3} = \mathcal{O}(n^4),$$

für eine Konstante  $c$ .  $\square$

### 3.2.2 (1+1)-EA auf Ringen

**Theorem 3.15.** *Der (1+1)-EA auf  $\text{GBP}_G$  findet eine optimale Partition auf einem Ring  $G = (V, E)$  der Größe  $|V| = n$  in erwarteter Zeit  $\mathcal{O}(n^4)$ .*

*Beweis.* Aus Lemma 3.13 wissen wir, dass in erwarteter Zeit  $\mathcal{O}(l^3 \cdot n)$  ein Block der Länge  $l$  auf Länge 1 geschrumpft wird, wenn nur 2-Zeichen-Mutationen zugelassen werden. Wir zeigen nun, dass auch unter der Verwendung von Mehrbitmutationen ein Block mit Wahrscheinlichkeit je  $1/2$  verkleinert oder vergrößert wird.

**Lemma 3.16.** *Für eine Partition  $x$  mit  $i \geq 4$  Blöcken wird in erwarteter Zeit  $\mathcal{O}(n^4/i^3)$  ein Block der Länge maximal 4 vom (1+1)-EA erzeugt.*

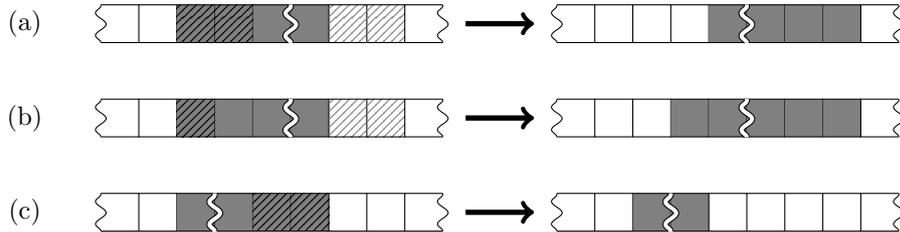
*Beweis.* Sei der kleinste Block von Länge  $l > 4$ , sonst ist das Ziel bereits erreicht. Weiterhin gilt wieder  $l \leq \lfloor n/i \rfloor$ . Wir betrachten nur die Mutationen, die einen oder eine Reihe von Knoten mutieren, welche an eine der Trennstellen des kürzesten Blocks grenzt. Wird dabei die Größe des Blocks verändert, so bezeichnen wir diese Mutation wieder als *relevant*. Dabei beobachten wir, dass Mutationen, die Knoten auf beiden Seiten einer Trennstellen mutieren nicht akzeptiert werden, falls nicht der angrenzende Block dadurch verschwindet (vgl. Abbildung 3.5 (b)). Sonst wird die Zahl der Blöcke weiter erhöht und damit die Fitness des so erzeugten Individuums verschlechtert. Demnach können mutierte Knoten nur rechts oder nur links einer Trennstelle sein, wenn nicht der gesamte Block entfernt wird. In Zeit  $\mathcal{O}(n^4)$  kommt mit einer Fehlerwahrscheinlichkeit von  $\mathcal{O}(1/n)$  keine Mutationen vor, bei der mehr als vier Knoten an den Trennstellen des

kürzesten Blocks mutiert werden. Wir vernachlässigen demnach Mutationen von mehr als acht Zeichen, da sie die Größe des Blocks mit der angegebenen Fehlerwahrscheinlichkeit nicht ändern. Da gilt  $l > 4$  kann der kürzeste Block also nicht entfernt werden und es existieren nur Mutationen, welche die Trennstellen verschieben.

Wir untersuchen nun den Einfluss der möglichen relevanten Mutationen und zeigen, dass diese mit Wahrscheinlichkeit je  $1/2 l$  vergrößern und verkleinern:

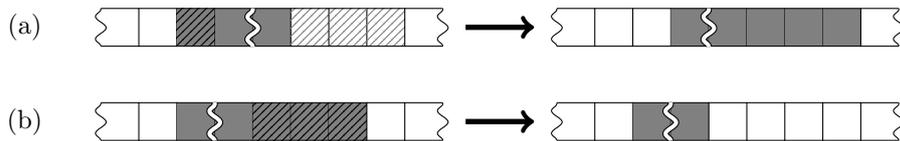
- 2-Zeichen-Mutationen: Für 2-Zeichen-Mutationen wurde bereits im Rahmen des Beweises für RLS gezeigt, dass diese die Größe des kürzesten Blocks mit Wahrscheinlichkeit  $1/2$  verkleinern.

- 4-Zeichen-Mutationen:



Falls wie in (a) alle vier mutierten Knoten an die Trennstellen des kürzesten Blocks grenzen, so wird eine Verschiebung des Blocks erreicht. Fälle bei denen pro Trennstelle höchstens ein angrenzendes Zeichen mutiert wird, entsprechen einer 2-Zeichen-Mutation, bei der weitere Knoten mutiert werden, die jedoch nichts an der Länge des kürzesten Blocks ändern. Falls an beiden Trennstellen des Blocks gleichzeitig Knoten mutiert werden, so entspricht dies der Folge von zwei bestimmten 2-Zeichen-Mutation. Auch hier wird dann mit Wahrscheinlichkeit mindestens  $1/2$  der Block verkleinert. Betrachten wir also die Fälle (b) und (c) bei denen an einer Trennstelle des Blocks zwei Zeichen der Repräsentation mutiert werden. Falls wie in (c) nur diese zwei Zeichen am Block mutiert werden, so wird der Block mit Wahrscheinlichkeit  $1/2$  um zwei Zeichen verkleinert und ebenso mit Wahrscheinlichkeit  $1/2$  um zwei Zeichen vergrößert, je nachdem ob die mutierten Zeichen Teil des Blocks sind oder außerhalb liegen. Für eine Mutation wie in (b) wird der Block entweder um zwei Zeichen verkürzt und um eines verlängert oder umgekehrt um zwei Zeichen verlängert und um eines verkürzt. Der Block verschiebt sich damit um ein Zeichen und wird zusätzlich mit gleicher Wahrscheinlichkeit um ein Zeichen vergrößert oder verkleinert.

- 6-Zeichen-Mutationen: Mutationen, die weniger als eine Reihe von drei Knoten an eine Trennstelle angrenzend mutieren, können wie 4-Zeichen-Mutationen behandelt werden. Damit bleiben die folgenden drei Fälle, für die die Aussage gezeigt werden muss:



Fall (a) verschiebt den Block um einen Knoten und führt mit gleicher Wahrscheinlichkeit zu einer Verlängerung oder Verkürzung des Blocks um zwei Zeichen. Fall (b) führt dagegen nur zu einer Verlängerung bzw. Verkürzung um drei Zeichen.

Diese geschehen wieder mit gleicher Wahrscheinlichkeit.

- 8-Zeichen-Mutationen: Wieder verweisen wir zur Behandlung solcher Mutationen bei denen weniger als vier Nachbarknoten einer Trennstelle mutiert werden auf die vorherigen Punkte. Dann verbleibt nur noch der Fall bei dem genau vier an eine Trennstelle grenzende Knoten mutiert werden.



Auch dies führt mit gleicher Wahrscheinlichkeit zu einer Verlängerung oder Verkürzung des Blocks.

Ignorieren wir Mutationen, die den Block nur verschieben, nicht aber dessen Größe ändern, wird ein Block mit Wahrscheinlichkeit mindestens  $1/2$  verkleinert. Die Wahrscheinlichkeit, dass tatsächlich mindestens die Hälfte solcher relevanter Mehr-Zeichen-Mutationen die Größe des Blocks verkleinern, ist damit beschränkt durch  $1/2$ . Wenn dies der Fall ist, wird nach  $\mathcal{O}(l^2)$  relevanter 2-Zeichen-Mutationen ein Block nach Lemma 1.5.4 mit konstanter Wahrscheinlichkeit auf Größe 4 schrumpfen. Die Wahrscheinlichkeit relevanter 2-Zeichen-Mutationen ist nach unten beschränkt durch

$$4 \cdot \frac{i}{2} \cdot \frac{2}{n} \left(1 - \frac{2}{n}\right)^{n/2-1} \cdot \binom{n/2}{1}^{-1} \geq \frac{8i}{en^2}.$$

Damit ergibt sich eine erwartete Zeit von  $\mathcal{O}(l^2(n^2/i)) = \mathcal{O}(n^4/i^3)$  bis zum Erreichen des Ziels.  $\square$

**Lemma 3.17.** *Für eine Partition  $x$  mit einem Block der Länge 4 gilt: Mit Wahrscheinlichkeit mindestens  $1/(3e)^4$  wird in  $\mathcal{O}(n^2/i)$  Schritten vom  $(1+1)$ -EA die Zahl der Blöcke um 2 reduziert.*

*Beweis.* Die Wahrscheinlichkeit, dass die nächsten vier relevanten Mutationen verkürzende 2-Zeichen-Mutationen sind, ist beschränkt durch  $1/(3e)^4$ : Solange der Block Länge  $j > 1$  hat, ist die Wahrscheinlichkeit einer solchen Mutation  $1/2e$ , für  $j = 1$  ist sie  $1/3e$ . Die Wahrscheinlichkeit einer relevanten Mutation ist nun nach unten beschränkt durch  $6i/(en^2)$ , wodurch sich die Schranke ergibt.  $\square$

Im schlechtesten Fall existieren  $n$  Blöcke in einer Partition, womit sich eine obere Schranke für die Optimierungszeit von

$$c \cdot \sum_{i=2}^{n/2} \frac{n^4}{(2i)^3} \leq c \cdot \sum_{i=4}^n \frac{n^4}{i^3} = \mathcal{O}(n^4)$$

für eine Konstante  $c$  ergibt.  $\square$

### 3.3 Graphen mit lokalen Optima

In diesem Abschnitt sollen Instanzen von Graphen betrachtet werden auf denen lokale Minima von  $\text{GBP}_G$  existieren. Um diese zu verlassen und das globale Optimum zu erreichen, benötigt der (1+1)-EA oder der SEMO im Erwartungswert exponentielle Laufzeit, da  $\mathcal{O}(n)$  Knoten gleichzeitig mutiert werden müssen. Ist die Wahrscheinlichkeit groß genug ein solches lokales Optimum zu erreichen, so ist auch die gesamte erwartete Laufzeit des jeweiligen Algorithmus exponentiell in der Anzahl der Knoten des Graphen.

Für alle folgenden Graphen existiert eine Bisektion mit Wert 0. Somit besteht die Pareto-Front  $P_0^*$  von  $\text{GBP}_G^0$  nur aus dem Element  $(0, 0)$  und die Pareto-Menge  $M_0^*$  aus den optimalen Bisektionen des jeweiligen Graphen. Daher fordern wir in der Analyse des global SEMO nur die Erzeugung eines optimalen Individuums  $x$  mit  $\text{zeros}(x) = 0$ . Dann ist auch die Pareto-Front durch den global SEMO abgedeckt.

#### 3.3.1 Der Vier-Cliquen-Graph

Als erste Analyse betrachten wir den Vier-Cliquen-Graphen (s. Abbildung 3.6), welcher sich sowohl für den (1+1)-EA als auch für den SEMO als schwierig erweist.

**Definition 3.18.** *Ein Graph wird als Vier-Cliquen-Graph bezeichnet, wenn dieser aus vier Cliques  $A, B, C, D$  der Größe  $n$  mit je einer Kante zwischen  $v_A \in A$  und  $v_B \in B$ , sowie  $v_C \in C$  und  $v_D \in D$  besteht, d.h.  $E = \{\{v, w\} \mid v, w \in M, M \in \{A, B, C, D\}\} \cup \{\{v_A, v_B\}, \{v_C, v_D\}\}$ . Die Knoten  $v_A, v_B, v_C, v_D$  bezeichnen wir als Verbindungsknoten, die Kanten zwischen diesen Knoten als Verbindungskanten.*

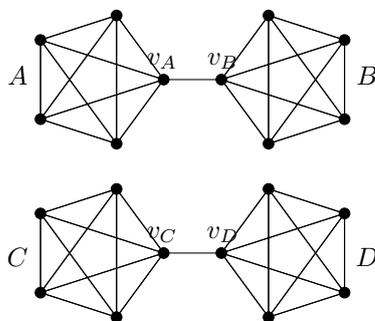


Abbildung 3.6: Ein Vier-Cliquen-Graph der Größe  $|V| = 20$

Offensichtlich ist eine Bisektion, die  $A$  und  $B$  von  $C$  und  $D$  trennt optimal, während eine Einteilung mit  $A$  und  $B$  auf verschiedenen Seiten - und damit  $C$  und  $D$  ebenfalls auf verschiedenen Seiten - ein lokales Optimum darstellt, aus dem nur eine Mutation von  $2n$  Knoten führt.

Wir analysieren zuerst das Verhalten des (1+1)-EA und der RLS. Dazu betrachten wir zunächst einen Vier-Cliquen-Graphen ohne seine Verbindungskanten. Es existieren dann sechs äquivalente Minima, die der (1+1)-EA bzw. die RLS mit genau gleicher

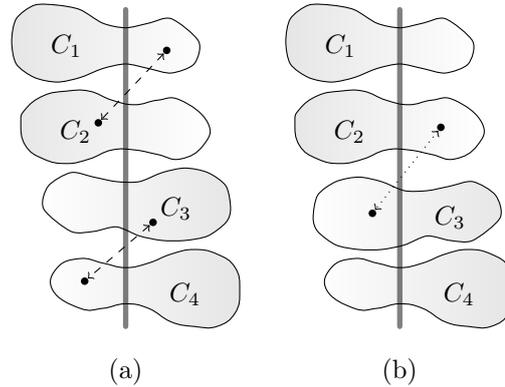


Abbildung 3.7: Schematische Verteilungen der Cliques eines initialen Individuums. Links: Je zwei Cliques haben die Mehrheit ihrer Knoten auf der gleichen Seite. Durch gestrichelte Linien sind akzeptierte Mutationen dargestellt, die eine Imbalance senken. Rechts: Die Mehrheit von drei Cliques befindet sich auf einer Seite. Die gepunktete Linie stellt eine akzeptierte Mutation dar, die  $|k_3|$  senkt.

Wahrscheinlichkeit erreichen kann. Das sind die folgenden drei Partitionen, sowie die jeweiligen Spiegelbilder:

$$\begin{array}{c|c} A & D \\ \hline B & C \end{array} \quad \begin{array}{c|c} A & B \\ \hline C & D \end{array} \quad \begin{array}{c|c} A & B \\ \hline D & C \end{array}$$

Falls der (1+1)-EA durch die Verbindungskanten nicht beeinflusst wird, so würde er mit Wahrscheinlichkeit  $2/3$  in einem der lokalen Optima landen. Wir beschreiben dafür zunächst das Verhalten der RLS auf dem Vier-Cliques-Graph *ohne* Verbindungskanten.

Hierzu ordnen wir die Cliques nach der Anzahl der Knoten, die sich nach der Initialisierung in  $L$  befinden. Seien diese Cliques absteigend geordnet  $C_1, C_2, C_3, C_4$  (s. Abbildung 3.7). Wir gehen im Folgenden zunächst davon aus, dass diese Ordnung nicht nur streng ist, sondern zusätzlich keine zwei Mengen mit  $|C_i \cap L| = |C_{i+1} \cap L| + 1$  existieren. Dafür bestimmen wir, welche Mutationen von der RLS akzeptiert werden. Die RLS und der (1+1)-EA können diese Ordnung jedoch zwischenzeitlich brechen. Dafür werden wir allerdings zeigen, dass dieser Fall zu keiner Veränderung des Verhaltens führt. Wir bestimmen zunächst die Wahrscheinlichkeit für die RLS in einem der lokalen Optima gefangen zu werden.

### Eine untere Schranke für die RLS

Basierend auf der Ordnung führen wir nun die *Imbalance*  $k_i$  jeder Clique ein, welche sich aus  $|C_i \cap L| = \frac{n}{2} + k_i$  für  $i = 1, 2$  bzw.  $|C_i \cap R| = \frac{n}{2} + k_i$  für  $i = 3, 4$  ergibt. Damit ergibt sich die Fitnessfunktion in Abhängigkeit der Imbalancen zu

$$\sum_{i=1}^4 \binom{\frac{n}{2} + k_i}{2} \binom{\frac{n}{2} - k_i}{2} = n^2 - k_1^2 - k_2^2 - k_3^2 - k_4^2.$$

Es lässt sich dann für eine Mutation der Knoten  $v \in L$  und  $w \in R$  Folgendes bestimmen: Eine Mutation wird genau dann akzeptiert, wenn

- $v \in C_1$  und  $w \in C_1$ ,
- $v \in C_2$  und  $w \in C_1 \cup C_2$ ,
- $v \in C_3$  und  $w \in C_1 \cup C_2 \cup C_3$ ,
- $v \in C_4$  und  $w \in C_1 \cup C_2 \cup C_3 \cup C_4$ .

Mutationen, welche die Imbalance zweier Cliques erhöhen sorgen für eine Abnahme der Fitness und somit zur Akzeptanz des erzeugten Individuums durch die RLS. Dies gilt auch dann, wenn z.B.  $k_3$  negativ ist, da aufgrund der strengen Ordnung dann  $|k_3| < k_2 < k_1$  ist (s. Abbildung 3.7 (b)). Entsprechend werden Mutationen, die die Imbalance beider Cliques senken, verworfen. Falls Knoten zweier Cliques so mutiert werden, dass die Imbalance einer Clique  $C_i$  steigt, während die der anderen Clique  $C_j$  sinkt, berechnet sich die Änderung der Fitness durch

$$\Delta \text{GBP}_G(x) = k_i^2 - (k_i + 1)^2 + k_j^2 - (k_j - 1)^2 = 2(k_j - k_i) - 2.$$

Eine solche Mutation wird also nur akzeptiert, wenn  $k_i \geq k_j - 1$ , womit sich basierend auf der Ordnung der Cliques die oben angegebenen Mutationen ergeben. Ein Tausch zweier Knoten einer Clique ist natürlich immer auch möglich. Da sich für die Clique dabei aber keine Änderung ergibt, können wir diesen Fall unbeachtet lassen.

Es scheint als könnten die Mengen  $C_1 \cap L$  und  $C_4 \cap R$  nur an Größe gewinnen und niemals abnehmen. Falls diese jedoch in ihrer Imbalance von einer der anderen Cliques „überholt“ werden, so ist auch ein Abnehmen möglich. In diesem Fall soll die Indizierung der Cliques insofern geändert werden, dass die Ordnung erhalten bleibt. Es gilt stets  $k_1 + k_2 = k_3 + k_4$ , damit  $|L| = |R| = 2n$  erhalten bleibt. Außerdem ist der Verlauf von  $k_1 + k_2$  während der RLS monoton steigend, da  $k_2$  nur durch Mutationen gesenkt wird bei denen  $k_1$  erhöht wird und  $k_1$  nicht sinken kann.

Durch das Einfügen der Verbindungskanten wird die Fitnessfunktion um maximal 2 geändert. Werden zwei Imbalancen erhöht, so wird diese Mutation weiterhin akzeptiert. Eine Senkung zweier Imbalancen erzeugt eine Änderung der Fitnessfunktion um

$$\Delta \text{GBP}_G(x) \geq 2(k_j + k_i) - 4,$$

was für  $k_i + k_j > 2$  in jedem Falle zu einer Verschlechterung führt. Fordern wir  $|C_i \cap L| > |C_{i+1} \cap L| + 2$  werden solche Mutationen vermieden. Für Mutationen, die ein  $k_i$  erhöhen und ein  $k_j$  senken ergibt sich durch diese Verschärfung der Ordnung ebenfalls keine Änderung.

**Lemma 3.19.** *Für eine Clique  $C_i$  mit  $k_i = n^\kappa$ ,  $0 < \kappa < 1$  und  $k_2 + k_3 = \Omega(n^\epsilon)$ ,  $\epsilon > 0$  gilt: Mit exponentiell großer Wahrscheinlichkeit erhöht die RLS  $k_i$  bis  $k_i = \Omega(n)$  ist.*

*Beweis.* Zu Beginn sei festzustellen, dass  $k_1$  und  $k_4$  niemals abnehmen werden, da sie durch die Umsortierung stets die Imbalance der Cliques repräsentieren, für die kein Tausch mit Knoten anderer Cliques akzeptiert wird. Für einen Seitenwechsel müssten  $k_i$  Knoten innerhalb der Clique getauscht werden, was mit Wahrscheinlichkeit  $\exp(-\Omega(n^\kappa))$  geschieht. Dagegen existiert immer eine Mutation mit Wahrscheinlichkeit mindestens  $1/4n^2$  welche die Imbalance erhöht, solange  $k_1$  bzw.  $k_4$  kleiner  $n/2$  ist.

Betrachten wir also Mutationen, die  $k_2$  oder  $k_3$  senken. Dies sind Mutationen, welche  $v \in C_2 \cap L$  und  $w \in C_1 \cap R$  bzw.  $v \in C_4 \cap L$  und  $w \in C_3 \cap R$  tauschen (s. Abbildung 3.7 (a)). Wir zeigen die Aussage nur für  $k_2 = n^\kappa$ . Für  $k_3$  verläuft der Beweis völlig analog mit vertauschten Seiten.

Eine Abnahme von  $k_2$  geschieht mit Wahrscheinlichkeit

$$P_{\text{RLS}}(k_2^-) = \frac{1}{4n^2} \binom{n}{2} + k_2 \binom{n}{2} - k_1 = \frac{1}{16} - \frac{k_1 - k_2}{8n} - \frac{k_1 k_2}{4n^2}, \quad (3.5)$$

die Wahrscheinlichkeit einer Erhöhung von  $k_2$  liegt dagegen bei

$$P_{\text{RLS}}(k_2^+) = \frac{1}{4n^2} \binom{n}{2} - k_3 + \frac{n}{2} - k_4 \binom{n}{2} - k_2 = \frac{1}{8} - \frac{k_2 + k_3 + k_4}{8n} + \frac{(k_3 + k_4)k_2}{4n^2}. \quad (3.6)$$

Sei nun  $\alpha := k_1/n$ ,  $\beta := k_2/n$ ,  $\gamma := k_3/n$ ,  $\delta := k_4/n$  und damit  $\alpha, \beta, \gamma, \delta \in [0, 1/2]$ . Da  $k_1 + k_2 = k_3 + k_4$  gilt, ist auch  $\alpha + \beta = \gamma + \delta$ . Zuletzt können wir annehmen, dass  $\alpha < 1/2$  ist, da sonst  $C_1$  vollständig auf der linken Seite ist und damit  $P(k_2^-) = 0$  gilt. Außerdem ist  $\beta = \mathcal{O}(n^{-(1-\kappa)})$ .

Insgesamt ergibt sich mit  $\alpha = \gamma + \delta - \beta$  nun

$$P_{\text{RLS}}(k_2^-) = \frac{1}{16} - \frac{\gamma + \delta - 2\beta}{8} - \frac{\alpha\beta}{4} = \frac{1}{16} - \frac{\gamma + \delta}{8} + \mathcal{O}(n^{-(1-\kappa)}) \quad (3.7)$$

und

$$P_{\text{RLS}}(k_2^+) = \frac{1}{8} - \frac{\gamma + \delta + \beta}{8} + \frac{\beta(\gamma + \delta)}{4} = \frac{1}{8} - \frac{\gamma + \delta}{8} - \mathcal{O}(n^{-(1-\kappa)}). \quad (3.8)$$

Nach Fellers Ungleichung für das Gamblers-Ruin-Problem (s. Kapitel 1.8) ergibt sich eine Wahrscheinlichkeit dafür, dass der Fall  $k_2 = n^\kappa/2$  vor  $k_2 = \epsilon n$  für ein  $\epsilon > 0$  eintritt von

$$\frac{r^{n^\kappa/2} - r^{\epsilon n - n^\kappa/2}}{1 - r^{\epsilon n - n^\kappa/2}} = r^{n^\kappa/2} \cdot \frac{1 - r^{\epsilon n - n^\kappa}}{1 - r^{\epsilon n - n^\kappa/2}},$$

wobei gilt

$$r = \frac{P_{\text{RLS}}(k_2^-)}{P_{\text{RLS}}(k_2^+)} = \frac{1 - 2(\gamma + \delta) + \mathcal{O}(n^{-(1-\kappa)})}{2 - 2(\gamma + \delta) - \mathcal{O}(n^{-(1-\kappa)})} \leq \frac{1}{2} + \mathcal{O}(n^{-(1-\kappa)}),$$

da  $\gamma + \delta = \alpha + \mathcal{O}(n^{-(1-\kappa)}) < 1/2 + \mathcal{O}(n^{-(1-\kappa)})$  und  $x/(1+x) \leq 1/2$  für  $x \leq 1$ . Die Wahrscheinlichkeit ist also höchstens

$$c \cdot \left( \frac{1 - 2(\gamma + \delta) + \mathcal{O}(n^{-(1-\kappa)})}{2 - 2(\gamma + \delta) - \mathcal{O}(n^{-(1-\kappa)})} \right)^{n^\kappa/2} = \exp(-\Omega(n^\kappa)),$$

mit  $c < 1$ , da  $r^{\epsilon n - n^\kappa} < r^{\epsilon n - n^\kappa/2}$  ist.

Zu Beginn gingen wir davon aus, dass für die Ordnung der Cliques gilt  $|C_i \cap L| > |C_{i+1} \cap L| + 2$ . Nun zeigen wir, dass auch der Fall  $|C_i \cap L| \geq |C_{i+1} \cap L|$  zu keiner dauerhaften Senkung einer Imbalance führen kann. Falls die strenge Ordnung für die Beziehung von  $C_1$  und  $C_2$  verletzt wird, so beeinträchtigt dies nur das Verhalten der RLS bei einem Tausch von zwei Knoten aus  $C_1$  und  $C_2$ , also dem Fall des Senkens

einer Imbalance und Steigerns der anderen. Solange dieser Zustand erhalten bleibt kann  $k_1$  sinken. Die Funktion  $k_1 + k_2$  ist aber weiterhin monoton steigend, da  $k_2$  im Gegenzug steigt. Die Aussage bleibt also erhalten. Dies gilt analog für  $k_3 \geq k_4 - 2$ . Nach Voraussetzung gilt  $k_2 = \Omega(n^\epsilon) - k_3$ , womit eine Verletzung der Ordnung in anderen Fällen nicht auftritt. Dieser Zustand ändert sich auch nicht, da  $k_2 + k_3$  mit exponentiell großer Wahrscheinlichkeit nur erhöht wird.  $\square$

**Lemma 3.20.** *Sobald für alle Cliques gilt  $k_i = \Omega(n^\epsilon)$  für ein beliebiges  $\epsilon > 0$ , wechselt in der RLS mit exponentiell großer Wahrscheinlichkeit für keine Clique die Majorität ihrer Knoten die Schnittseite.*

*Beweis.* Jede akzeptierte Mutation, welche nicht zwei Knoten der gleichen Clique tauscht, erhöht die Imbalance mindestens einer Clique um den Wert 1 und senkt den Wert maximal einer anderen Clique um 1. Die gesamte Imbalance  $k_1 + k_2 + k_3 + k_4$  nimmt also niemals ab. Die Wahrscheinlichkeit für eine Zunahme liegt bei mindestens  $1/(4n^2)$  (der Tausch zweier Knoten aus  $C_2 \cap L$  und  $C_3 \cap R$ ), so dass in polynomieller Zeit eines der lokalen Optima erreicht wird. Nach Voraussetzung gilt Lemma 3.19 und die Imbalance einer Clique mit  $k_i = o(n)$  und  $k_i = \Omega(n^\epsilon)$  wird mit exponentiell hoher Wahrscheinlichkeit niemals wesentlich sinken. Falls also eine Clique mit  $k_i = \Omega(n)$  an Imbalance verliert wird weiterhin gelten  $k_i = \Omega(n^{1-\epsilon})$ .  $\square$

**Theorem 3.21.** *Die RLS auf  $\text{GBP}_G$ , für einen Vier-Cliquen-Graphen  $G = (V, E)$  mit  $|V| = 4n$ , fällt mit Wahrscheinlichkeit  $2/3 - \mathcal{O}(1/n^{1-\epsilon})$  für ein beliebiges  $\epsilon > 0$  in eines der lokalen Optima mit Wert 2 und hat damit eine erwartete Laufzeit von unendlich.*

*Beweis.* Wir gehen davon aus, dass die Folge der Zufallsbits bekannt ist und vergleichen den Verlauf der RLS auf dem Graphen mit Verbindungskanten mit dem Verlauf ohne Verbindungskanten.

Sobald alle Cliques Imbalance  $\Omega(n^\epsilon)$  besitzen ist nach Lemma 3.20 mit exponentiell hoher Wahrscheinlichkeit bereits festgelegt in welchem Optimum die RLS landet. Dies ist unabhängig davon, ob sich die Verbindungskanten im Graphen befinden oder nicht. Ab diesem Punkt kann die RLS auf beiden Graphen vollkommen andere Mutationen durchführen, was mit exponentiell großer Wahrscheinlichkeit keinen Einfluss auf das Ergebnis haben wird.

Wir wollen also den Einfluss der Verbindungskanten auf die Anfangsphase der RLS bestimmen, bei der für mindestens eine Clique  $C_i$  gilt  $k_i = o(n^\epsilon)$ . Mit einer Fehlerwahrscheinlichkeit von höchstens  $2/n$  pro Mutationsschritt wird keiner der beiden Verbindungsknoten mutiert. Dann verhalten sich beide Läufe der RLS gleich.

Wenn  $k_2 + k_3 \geq 2$  gilt, gelten die Wahrscheinlichkeiten für eine Vergrößerung bzw. Verringerung von  $k_2$  und  $k_3$  aus den Gleichungen (3.5) und (3.6). Für einen beliebigen Anfangswert für  $k_2$  von  $z \geq 2$ ,  $z = o(n^\epsilon)$  ist die Wahrscheinlichkeit, den Zustand  $k_2 = z - 1$  zu erreichen, bevor  $k_2 = n^\epsilon$  gilt nach Lemma 1.8 beschränkt durch

$$\frac{2^{-1} - 2^{-(n^\epsilon - z)}}{1 - 2^{-(n^\epsilon - z)}} + o(n^{-(1-\epsilon)}) \leq \frac{1}{2} + o(n^{-(1-\epsilon)}).$$

Außerdem gilt nach Lemma 1.9, dass die erwartete Laufzeit bis zum Eintreten von  $k_2 = z - 1$  oder  $k_2 = n^\epsilon$  beschränkt ist durch

$$\frac{n^\epsilon - z}{d} \cdot \frac{1 - r}{1 - r^{n^\epsilon - b}} - \frac{1}{d} = \mathcal{O}(n^\epsilon),$$

mit  $d = \text{P}_{\text{RLS}}(k_2^+) - \text{P}_{\text{RLS}}(k_2^-) \geq 1/16 - o(n^{-(1-\epsilon)})$ . Das gleiche gilt für  $k_3$ .

Mit Wahrscheinlichkeit mindestens  $1/2 - o(n^\epsilon/n)$  gilt  $k_2, k_3 = \Omega(n^\epsilon)$  dann in Zeit  $\mathcal{O}(n^\epsilon)$ , falls zu Beginn  $k_2 + k_3 \geq 2$  galt. Durch höchstens zwei Mutationen von Knoten aus  $C_2 \cap L$  und  $C_3 \cap R$  kann dieser Fall erreicht werden, wenn gilt  $k_2 + k_3 < 2$ . Wir erinnern daran, dass  $k_2 + k_3 \geq 0$  gilt, da  $k_2 \geq -k_3$  ist. Eine solche Mutation hat Wahrscheinlichkeit

$$\frac{1}{4n^2} \binom{n}{2} + k_2 \binom{n}{2} + k_3 \geq \frac{1}{16}.$$

Mit konstanter Wahrscheinlichkeit werden zwei solche Mutationen nacheinander ausgeführt, bevor  $k_2$  und  $k_3$  wieder verändert werden, womit sich eine Laufzeit von  $\mathcal{O}(n^\epsilon)$  für die Anfangsphase ergibt.

Somit ist die Wahrscheinlichkeit eines Einflusses der Verbindungskanten nach oben beschränkt durch  $\mathcal{O}(n^\epsilon/n)$ . Die Wahrscheinlichkeit, dass die RLS auf dem Vier-Cliquen-Graphen mit Verbindungskanten das lokale Optimum erzeugt, ist also mindestens  $2/3 - \mathcal{O}(n^\epsilon/n)$ .  $\square$

### Eine untere Schranke für den (1+1)-EA

Während bei der RLS die Mutationswahrscheinlichkeiten noch explizit berechnet werden konnten, gestaltet sich dies beim (1+1)-EA wegen seiner Mutationen von mehr als zwei Zeichen wesentlich schwieriger. Zusätzlich zu einer parallelen Abhandlung der in der RLS seriell ausführbaren Schritten sind auch Mutationen möglich, die Zustände erreichen, welche von der RLS nicht unmittelbar erreicht werden können. Somit besteht nun die Möglichkeit auch  $k_1$  und  $k_4$  zu senken, sowie eine höhere Wahrscheinlichkeit für die Verringerung von  $k_2$  bzw.  $k_3$  (s. Abbildung 3.8). Es gilt zwar weiterhin  $k_1 + k_2 = k_3 + k_4$ , jedoch kann eine Erhöhung von  $k_1$ , welche zu einer starken Verbesserung der Fitness führt, auch bei einer Senkung von  $k_2$  um mehr als 1 akzeptiert werden. Damit gilt insbesondere nicht mehr die Monotonie von  $k_1 + k_2$  bzw.  $k_3 + k_4$  wie bei der RLS.

Da wir nicht wie in Abbildung 3.8 eine Menge von expliziten Tauschen von Knotenpaaren bei einer Mutation erhalten, sondern nur eine Menge von Knoten, welche die Schnittseite ändern, wollen wir zur Analyse der nun möglichen Mutationen die geänderten Knoten in Tupel einteilen. Wir bezeichnen die Menge der mutierten Knoten einer Mutation  $M$ , die sich vor der Mutation in  $L$  befinden als  $M_L$  und die aus  $R$  als  $M_R$ . Damit bilden wir eine bijektive Relation  $\mathcal{R}_M$  auf  $M_L \times M_R$  und betrachten die enthaltenen Tupel als Knotenpaare, die zu vertauschen sind. Zunächst erstellen wir die Relation  $\mathcal{R}'_M$  indem für jede Clique  $C_i$  so lange Tupel aus  $(M_L \cap C_i) \times (M_R \cap C_i)$  zu  $\mathcal{R}'_M$  hinzugefügt und die Knoten aus  $M_L$  bzw.  $M_R$  entfernt werden, bis entweder gilt  $M_L \cap C_i = \emptyset$  oder  $M_R \cap C_i = \emptyset$ . Nun ordnen wir die restlichen Knoten absteigend nach dem Index der zugehörigen Cliques zu  $M_L = \{v_{l_1}, \dots, v_{l_m}\}$ ,  $M_R = \{v_{r_1}, \dots, v_{r_m}\}$  und bilden darauf die Relation  $\mathcal{R}''_M = \{(v_{l_i}, v_{r_i}) \mid 1 \leq i \leq m\}$ . Dann erhalten wir  $\mathcal{R}_M = \mathcal{R}'_M \cup \mathcal{R}''_M$  (s. Abbildung 3.9). Hierfür definieren wir nun durch  $\Delta_{(v,w)}$   $\text{GBP}_G$

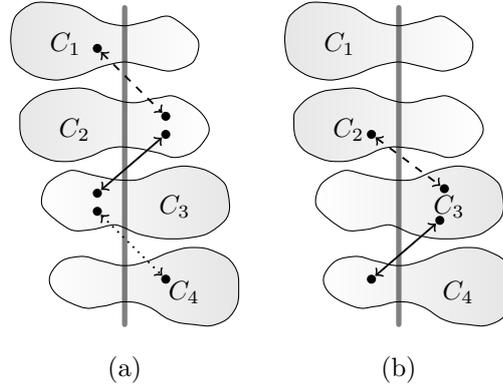


Abbildung 3.8: Links: Eine Senkung der Imbalance von  $C_1$  und  $C_4$  kann erfolgen wenn zwei Knoten entsprechend der gestrichelten bzw. gepunkteten Linie getauscht werden und zusätzlich Knoten mutiert werden, die die Imbalance einer Clique erhöhen, wie durch die durchgezogene Linie dargestellt. Rechts: Der durch die gestrichelte Linie dargestellte Tausch wird nur vom (1+1)-EA akzeptiert, wenn z.B. zusätzlich ein Tausch, wie durch die durchgezogenen Linie dargestellt, erfolgt, der eine genügend große Verbesserung mit sich bringt.

die Änderung der Fitness durch den Tausch  $(v, w) \in \mathcal{R}_M$ . In der Relation  $\mathcal{R}_M$  gibt es bei einer Mutation niemals zwei Tupel, von denen eines  $k_i$  verringert und das andere  $k_i$  erhöht, da Tupel aus  $\mathcal{R}'_M$  kein  $k_i$  verändern, während in  $\mathcal{R}''_M$  Knoten aus  $C_i$  nur auf einer Seite der Relation vorhanden sind.

Die Mutationen, die  $k_2$  verringern, sind dann solche, bei denen ein  $(v, w) \in \mathcal{R}''_M$  mit  $v \in C_2$  existiert. Tupel mit  $w \in C_1$  tragen dabei zu einer Verbesserung der Fitness bei (entspricht der  $k_2$  verringernden Mutation der RLS), Tupel mit  $w \in C_3 \cup C_4$  fordern weitere Tupel in  $\mathcal{R}''_M$ , damit  $M$  vom (1+1)-EA akzeptiert wird. Falls  $w \in C_4$  gilt, so ist  $C_4 \cap M_L = \emptyset$ , da diese bei der Konstruktion von  $\mathcal{R}'_M$  eliminiert wurden und es existieren keine Tupel aus  $C_4 \times C_3$  in  $\mathcal{R}''_M$ , welche die Fitness verbessern könnten. Nach Konstruktion gibt es keine  $(v', w') \in \mathcal{R}''_M$  mit  $w' \in C_2$ . Somit muss für  $w \in C_4$  ein Tupel  $(v', w') \in C_2 \times C_1$  mutiert werden, damit  $M$  akzeptiert wird (s. Abbildung 3.9 (c)). Falls hingegen gilt  $w \in C_3$  und keine Tupel aus  $C_2 \times C_4$  in  $\mathcal{R}''_M$  existieren, so kann ein  $(v', w') \in C_4 \times C_3$  die Fitness verbessern und zur Akzeptanz der Mutation führen (s. Abbildung 3.8 (b)). Dies ist der einzige Fall, bei dem keine Mutation zweier Knoten  $v' \in C_2 \cap L$  und  $w' \in C_1 \cap R$  nötig ist.

Da  $P_{\text{RLS}}(k_2^-)$  aus Gleichung 3.7 eine obere Schranke für die Wahrscheinlichkeit des (1+1)-EA für eine Mutation mit  $(C_2 \times C_1) \cap \mathcal{R}''_M \neq \emptyset$  angibt, ist die Wahrscheinlichkeit

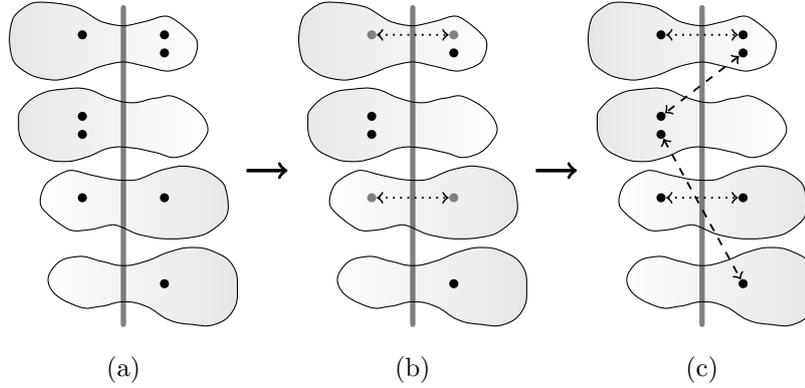


Abbildung 3.9: Die Relation  $\mathcal{R}_M$  auf den Knoten, die durch die Mutation  $M$  mutiert werden, wird in zwei Schritten gebildet. In (b) ist die Relation  $\mathcal{R}'_M$  abgebildet, welche nur Paare von Knoten gleicher Cliques beinhaltet. Die entfernten Knoten sind grau markiert. In (c) sind zusätzlich Knotenpaare aus  $\mathcal{R}''_M$  eingezeichnet.

$k_2 = \Theta(n^\kappa)$  um genau 1 zu verringern nach oben beschränkt durch

$$\begin{aligned}
P(k_2^{-1}) &\leq P_{\text{RLS}}(k_2^-) \\
&\quad + \frac{1}{(2n)^2} \binom{2n}{2}^{-1} \binom{n}{2} + k_2 \binom{n}{2} + k_3 \binom{n}{2} - k_4 \binom{n}{2} + k_3 \\
&= P_{\text{RLS}}(k_2^-) \\
&\quad + \frac{1}{(2n)^4} \binom{2n}{2}^{-1} \frac{n^4 + 4n^3k_3 + 4n^2k_3^2 - 2n^3k_4 - 8n^2k_3k_4 - 8nk_3^2k_4}{16} \\
&\quad + \mathcal{O}(n^{-(1-\kappa)}) \\
&\stackrel{(1)}{\leq} P_{\text{RLS}}(k_2^-) + \frac{1 + 4\gamma - 2\delta + 4\gamma^2 - 8\gamma\delta - 8\gamma^2\delta}{2^8} \frac{2n}{2n-1} + \mathcal{O}(n^{-(1-\kappa)}) \\
&\stackrel{(2)}{\leq} P_{\text{RLS}}(k_2^-) + \frac{1 + 2\gamma + 4\gamma^2}{2^8} \frac{2n}{2n-1} + \mathcal{O}(n^{-(1-\kappa)}) \\
&\stackrel{(3)}{\leq} \frac{1}{16} - \frac{\gamma + \delta}{8} + \frac{1}{128} + \mathcal{O}(n^{-(1-\kappa)}),
\end{aligned}$$

mit  $k_1 = \alpha n$ ,  $k_2 = \beta n$ ,  $k_3 = \gamma n$ ,  $k_4 = \delta n$  für  $\alpha, \beta, \gamma, \delta \in [0, 1/2]$  (1), sowie  $\gamma + \delta = \alpha + \beta \leq 1/2 + \mathcal{O}(n^{-(1-\kappa)})$  und  $\gamma \leq \delta$ , also  $\gamma \leq 1/4 + \mathcal{O}(n^{-(1-\kappa)})$  (2) und für  $n \geq 4$  (3).

Eine Mutation  $M$ , die  $k_2$  erhöht, enthält Tupel  $(v, w) \in \mathcal{R}''_M$ ,  $w \in C_2$  die sich in zwei Klassen einteilen lassen. Zum einen die schon in der RLS möglichen Tauschpaare mit  $v \in C_3 \cup C_4$ . Andererseits Tupel mit  $v \in C_1$ , die nur in Kombination mit genügend anderen Tupeln der ersten Klasse oder  $(v', w') \in \mathcal{R}''_M \cap (C_4 \times C_3)$  dazu führen, dass  $M$  akzeptiert wird. Für Vergrößerungen von  $k_2$  um den Wert 1 kann die Wahrscheinlichkeit einer Mutation, die nur durch den zweiten Typ herbei geführt wird beliebig klein werden. Wenn  $k_3$  und  $k_4$  sich nur um einen Faktor  $n^\epsilon$  unterscheiden, so ist die Verbesserung der

Fitness durch diesen Tausch nur  $\Delta_{(v',w')} \text{GBP}_G = \mathcal{O}(k_4 - k_3) = \mathcal{O}(n^\epsilon)$ . Gilt jedoch  $k_1 = \Omega(n)$  und  $k_2 = \mathcal{O}(n^\epsilon)$  sind  $\Omega(n^{1-\epsilon})$  zusätzliche Knotenpaare von Nöten damit die Mutation akzeptiert wird. Die Wahrscheinlichkeit einer solchen Mutation ist nach oben beschränkt durch  $\exp(-\Omega(n^{1-\epsilon}))$ .

Es lässt sich also nur schwer eine bessere allgemeine untere Schranke für die Wahrscheinlichkeit einer Vergrößerung von  $k_2 = \Theta(n^\kappa)$  um 1 zeigen als die Übertragung von Gleichung 3.8 aus dem Beweis für die RLS zu

$$\begin{aligned} \mathbb{P}(k_2^{+1}) &\geq \left(\frac{1}{8} - \frac{\gamma + \delta}{8}\right) \left(1 - \frac{1}{2n}\right)^{2n-1} - \mathcal{O}(n^{-(1-\kappa)}) \\ &\geq \left(\frac{1}{8} - \frac{\gamma + \delta}{8}\right) \frac{1}{e} - \mathcal{O}(n^{-(1-\kappa)}). \end{aligned}$$

Der zusätzliche Faktor  $1/e$ , der nun in die untere Schranke eingeht, führt zu einer erheblichen Erschwernis des Beweises für eine Vergrößerung von  $k_2$ , wie er bei der RLS geführt wurde. Für die verwendete Ungleichung von Feller aus Kapitel 1.8 ist nun

$$r = \frac{\mathbb{P}(k_2^{-1})}{\mathbb{P}(k_2^{+1})} < \frac{1,016 - 2(\gamma + \delta) + \mathcal{O}(n^{\kappa-1})}{2e - 2(\gamma + \delta)e - \mathcal{O}(n^{\kappa-1})}.$$

Damit ist  $r$  für  $(\gamma + \delta) \leq (e - 1)/(e - 2)$  nicht besser als durch  $r \leq 1$  abzuschätzen. Unter Verwendung dieser Schranken kann der Beweis der RLS nur auf den (1+1)-EA mit veränderter Mutationswahrscheinlichkeit übertragen werden.

**Theorem 3.22.** *Der (1+1)-EA mit Mutationswahrscheinlichkeit  $(2/|V|)^{1+\lambda}$  für  $\lambda \geq \max\{-\ln \ln(2-2\epsilon)/\ln(2n), 2\epsilon\}$  auf  $\text{GBP}_G$ , für einen Vier-Cliquen-Graphen  $G = (V, E)$  mit  $|V| = 4n$ , fällt in eines der lokalen Optima mit Wert 2 mit Wahrscheinlichkeit  $2/3 - \mathcal{O}(n^{-1+\epsilon+\lambda})$  mit  $0 < \epsilon < 1$  und hat damit eine exponentiell große erwartete Laufzeit.*

*Beweis.* Wir zeigen dazu wieder, dass die Imbalance jeder Clique mit großer Wahrscheinlichkeit steigt und die Mehrheit der Knoten auf der ihnen zugeordneten Schnittseite bleiben, sobald  $k_i = \Omega(n^{\epsilon+\lambda})$  für alle Cliquen  $C_i$  gilt. Dazu beweisen wir eine untere Schranke für Wahrscheinlichkeit  $k_2$  zu steigern. Der Beweis für die untere Schranke der Wahrscheinlichkeit einer Steigerung von  $k_3$  verläuft analog. Weil stets  $k_1 \geq k_2$  bzw.  $k_4 \geq k_3$  ist, gilt damit diese Schranke entsprechend auch für  $k_1$  und  $k_4$ .

**Lemma 3.23.** *Für eine Clique mit Imbalance  $k_i = n^\kappa$ ,  $0 < \kappa < 1$  und  $k_2 + k_3 = \Omega(n^\epsilon)$ ,  $\epsilon > 0$  gibt es ein  $\lambda \geq 2\epsilon$ , so dass der (1+1)-EA mit Mutationswahrscheinlichkeit  $(2/|V|)^{1+\lambda}$  die Imbalance  $k_i$  mit Wahrscheinlichkeit exponentiell nahe an 1 erhöht.*

*Beweis.* Wir betrachten zunächst nur Mutationen, die  $k_2$  um 1 ändern und zeigen dann, dass durch das Einführen von Mutationen, die mehr Knoten mutieren, das Verhalten des (1+1)-EA kaum beeinflusst wird. Für die Steigerung und Senkung von

$k_2$  gilt

$$\begin{aligned} \mathbb{P}(k_2^{+1}) &\geq \frac{1}{(2n)^{1+\lambda}} \binom{2n}{1}^{-1} \left(\frac{n}{2} - k_3 + \frac{n}{2} - k_4\right) \left(\frac{n}{2} - k_2\right) \left(1 - \frac{1}{(2n)^{1+\lambda}}\right)^{2n-1} \\ &\geq \left(\frac{1}{2^{3+\lambda} + n^\lambda} - \frac{\gamma + \delta}{2^{3+\lambda} n^\lambda}\right) \left(1 - \frac{1}{(2n)^{1+\lambda}}\right)^{2n-1} - \mathcal{O}(n^{-(1+\lambda-\kappa)}) \\ &\stackrel{(*)}{\geq} \left(\frac{1}{2^{3+\lambda} n^\lambda} - \frac{\gamma + \delta}{2^{3+\lambda} n^\lambda}\right) e^{-1/(2n)^\lambda} - \mathcal{O}(n^{-(1+\lambda-\kappa)}), \end{aligned}$$

wobei  $(*)$  nach Kapitel 1.5.1 gilt, sowie

$$\begin{aligned} s\mathbb{P}(k_2^{-1}) &\leq \frac{1}{(2n)^{2+\lambda}} \left(\frac{n}{2} + k_2\right) \left(\frac{n}{2} - k_1\right) + \mathcal{O}(n^{-\lambda}) \\ &\leq \frac{1}{2^{4+\lambda} n^\lambda} - \frac{\gamma + \delta}{2^{3+\lambda} n^\lambda} + \mathcal{O}(n^{-2\lambda}) + \mathcal{O}(n^{-(1+\lambda-\kappa)}) \end{aligned}$$

und es ist

$$r = \frac{\mathbb{P}(k_2^{-1})}{\mathbb{P}(k_2^{+1})} \leq \frac{e^{1/(2n)^\lambda}}{2} + \mathcal{O}(n^{-\min\{2\lambda, 1+\lambda-\kappa\}}) \leq 1 - \varphi$$

mit konstantem  $0 < \varphi < 1$  für ein geeignet großes  $\lambda$ . Somit ist wie in Theorem 3.21 die Wahrscheinlichkeit für  $k_2 = n^\kappa$  den Wert  $n^\kappa/2$  zu erreichen, bevor  $n^{\kappa+\epsilon}$  erreicht wurde, nach Lemma 1.8 nach oben beschränkt durch  $\exp(-\Omega(n^\kappa))$ . Dies betrachten wir nun genauer. In einem Abschnitt der Laufzeit von  $T = n^{\kappa+\lambda+\epsilon}$  Schritten ist die erwartete Anzahl von Mutationen, die  $k_2$  um 1 erhöhen, mindestens

$$\mathbb{E}(X_T^{+1}) = \mathbb{P}(k_2^{+1}) \cdot T = \Omega(n^{\kappa+\epsilon}),$$

die erwartete Anzahl von Mutationen, die  $k_2$  um genau 1 senken lässt sich ähnlich abschätzen durch

$$\mathbb{E}(X_T^{-1}) = \mathbb{P}(k_2^{-1}) \cdot T = \mathcal{O}(n^{\kappa+\epsilon}).$$

Mit den Chernoff-Ungleichungen aus Kapitel 1.5.2 ergeben sich dann die folgenden oberen und unteren Schranken für die Anzahl von Mutationen, die  $k_2$  um genau 1 verändern:

$$\begin{aligned} \mathbb{P}\left(X_T^{+1} \leq \left(1 - \frac{\varphi}{2(2-\varphi)}\right) \mathbb{E}(X_T^{+1})\right) &\leq \exp(-\Omega(n^{\kappa+\epsilon})) \\ \mathbb{P}\left(X_T^{-1} \geq \left(1 + \frac{\varphi}{2(2-\varphi)}\right) \theta \mathbb{E}(X_T^{-1})\right) &\leq \exp(-\Omega(n^{\kappa+\epsilon})), \end{aligned}$$

für ein geeignetes  $\theta$ , so dass  $\theta \mathbb{E}(X_T^{-1}) = \Omega(n^{\kappa+\epsilon})$  gilt.

Damit ist die Erhöhung von  $k_2$  mittels solcher Mutationen mit Wahrscheinlichkeit

exponentiell nahe an 1 nach unten beschränkt durch

$$\begin{aligned}
X_T^{+1} - X_T^{-1} &> \left(1 - \frac{\varphi}{2(2-\varphi)}\right) \mathbb{E}(X_T^{+1}) - \left(1 + \frac{\varphi}{2(2-\varphi)}\right) \mathbb{E}(X_T^{-1}) \\
&= \left(1 - \frac{\varphi}{2(2-\varphi)}\right) \mathbb{P}(k_2^{+1}) \cdot T - \left(1 + \frac{\varphi}{2(2-\varphi)}\right) \mathbb{P}(k_2^{-1}) \cdot T \\
&\stackrel{(*)}{\geq} \left(1 - \frac{\varphi}{2(2-\varphi)}\right) \mathbb{P}(k_2^{+1}) \cdot T - \left(1 + \frac{\varphi}{2(2-\varphi)}\right) (1-\varphi) \mathbb{P}(k_2^{+1}) \cdot T \\
&= \frac{\varphi}{2(2-\varphi)} (2(2-\varphi) - 2 + \varphi) \mathbb{P}(k_2^{+1}) \cdot T \\
&= \frac{\varphi}{2} \mathbb{P}(k_2^{+1}) \cdot T = \Omega(n^{\kappa+\epsilon}),
\end{aligned}$$

wobei bei  $(*)$   $r \leq 1 - \varphi$  verwendet wurde. Nach Zeit  $T$  ist also die Imbalance  $k_2$  auf  $\Omega(n^{\kappa+\epsilon})$  gestiegen.

Mutationen, die  $k_2$  um mehr als 1 senken ändern daran mit einer Wahrscheinlichkeit exponentiell nahe an 1 nichts. Die Wahrscheinlichkeit dieser Mutationen ist beschränkt durch die Wahrscheinlichkeit mindestens zwei Knoten aus  $C_2 \cap L$  zu mutieren. Es ist

$$\begin{aligned}
\mathbb{P}(k_2^{-\geq 2}) &\leq \frac{1}{(2n)^{2(1+\lambda)}} \left(\frac{n}{2} + k_2\right)^2 \\
&\leq \frac{1}{2^{3+2\lambda} n^{2\lambda}},
\end{aligned}$$

womit sich eine erwartete Anzahl solcher Mutationen in  $T$  Schritten ergibt durch

$$\mathbb{E}(X_T^{-\geq 2}) \leq \frac{n^{\kappa-\lambda+\epsilon}}{2^{2+2\lambda}}.$$

Nach der ersten Chernoff-Ungleichung (1.1) ist die Wahrscheinlichkeit für das Auftreten von  $\omega(n^{\kappa-\lambda+\epsilon})$  dieser Mutationen beschränkt durch

$$\mathbb{P}(X^{-\geq 2} \geq 2 \frac{n^{\kappa-\lambda+\epsilon}}{2^{2+2\lambda}}) \leq \exp(-\Omega(n^{\kappa-\lambda+\epsilon})).$$

In einer Mutation werden mit Wahrscheinlichkeit  $\exp(-\Omega(n^\epsilon))$  nur  $\Omega(n^\epsilon)$  Knoten mutiert. Somit sinkt  $k_2$  mit Wahrscheinlichkeit exponentiell nahe an 1 nur um höchstens  $o(n^{\kappa-\lambda+2\epsilon})$ . Da gilt  $\lambda \geq 2\epsilon$  ist somit der Fall  $k_2 \leq n^\kappa/2$  mit der angegebenen Wahrscheinlichkeit ausgeschlossen.

Wir benötigen jetzt noch eine passende untere Schranke für  $\lambda$ . Da  $\varphi$  eine Konstante ist können wir  $\lambda$  asymptotisch abschätzen durch

$$\begin{aligned}
\frac{e^{1/(2n)^\lambda}}{2} &\leq 1 - \varphi \\
\Leftrightarrow 1/(2n)^\lambda &\leq \ln(2 - 2\varphi) \\
\Leftrightarrow \lambda &\geq \frac{-\ln \ln(2 - 2\varphi)}{\ln(2n)}.
\end{aligned}$$

Für ein beliebiges  $\lambda$  welches die beiden unteren Schranken  $2\epsilon$  und die eben gezeigt nicht unterschreitet gilt also die Behauptung.  $\square$

Wir gehen wieder davon aus, dass uns die Zufallsbits bekannt sind und vergleichen den (1+1)-EA auf dem Vier-Cliquen-Graphen  $G = (V, E)$  mit dem (1+1)-EA auf  $G' = (V, E')$  mit  $E' = E \setminus \{\{v_A, v_B\}, \{v_C, v_D\}\}$ , also ohne die Verbindungskanten.

Falls alle Cliques Imbalance  $\Omega(n^\epsilon)$  aufweisen, so ist mit hoher Wahrscheinlichkeit wieder bereits „entschieden“ in welchem lokalen Optimum der (1+1)-EA landet. Nur mit exponentiell geringer Wahrscheinlichkeit wechselt die Majorität einer Clique die Seite des Schnittes. Da für einen Schnitt größer 2 immer mindestens eine Mutation die Fitness der Bisektion senkt ist in erwarteter Zeit  $\mathcal{O}(n^3)$  eines der Optima erreicht.

Betrachten wir also wieder das Verhalten des (1+1)-EA für  $k_i = o(n^\epsilon)$ . Auch hier behandeln wir zunächst den Prozess, der  $k_i$  nur um den Wert 1 ändert und fügen nachträglich mögliche Mehr-Zeichen-Mutationen ein. Wieder nach Feller ist für  $k_2 = z$  die Wahrscheinlichkeit des Falles  $k_2 = z - 1$  bevor  $k_2 = n^\epsilon$  gilt nach unten beschränkt durch

$$\frac{r^1 - r^{n^\epsilon - z + 1}}{1 - r^{n^\epsilon - z + 1}} \leq r \leq 1 - \varphi.$$

Die erwartete Zeit bis einer der Fälle eintritt ist nach Lemma 1.9 beschränkt durch

$$\frac{n^\epsilon - z + 1}{\varphi P(k_2^{+1})} \cdot \frac{1 - r}{1 - r^{n^\epsilon - z + 1}} - \frac{1}{\varphi P(k_2^{+1})} \leq \frac{n^\epsilon - z + 1}{\varphi P(k_2^{+1})} = \mathcal{O}(n^{\epsilon + \lambda})$$

und somit gilt nach erwarteter Zeit  $\mathcal{O}(n^{\epsilon + \lambda})$  für die Imbalance  $k_2 = n^\epsilon$ , wenn nur Mutationen zugelassen sind, die  $k_1$  um den Wert 1 ändern.

Nun soll der Einfluss der Mehr-Zeichen-Mutationen untersucht werden. In Zeit  $\mathcal{O}(n^{\epsilon + \lambda})$  treten nach Chernoff mit Wahrscheinlichkeit exponentiell nahe an 1 nicht mehr als  $\mathcal{O}(n^{\epsilon - \lambda}) = \mathcal{O}(n^{-\epsilon})$  Mutationen auf, die  $k_2$  um mindestens 2 senken, da  $\lambda \geq 2\epsilon$  ist: Sei  $X^-$  die Anzahl solcher Mutationen in  $\mathcal{O}(n^{\epsilon + \lambda})$  Schritten. Dann gilt

$$\mathbb{E}(X^-) \leq \frac{n^{\epsilon - \lambda}}{2^{2 + 2\lambda}}$$

und somit ist die Wahrscheinlichkeit einer Anzahl dieser Mutationen von mehr als  $n^{(\epsilon - \lambda)/2}$  beschränkt durch

$$\mathbb{P}(X^- \geq n^{(\epsilon - \lambda)/2}) \leq \exp(-\Omega(n^{\lambda - \epsilon})) = \exp(-\Omega(n^\epsilon)).$$

Mit Fehlerwahrscheinlichkeit  $\mathcal{O}(1/n^{1 - \epsilon})$  kommt also keine solche Mutation vor bis für alle Cliques  $C_i$  gilt  $k_i = \Omega(n^\epsilon)$ .

Während der ersten  $\mathcal{O}(n^{\epsilon + \lambda})$  Schritte des (1+1)-EA können Mutationen, in denen einer der Verbindungsknoten mutiert wird, das Ergebnis gegenüber eines Laufes des (1+1)-EA auf  $G'$  ändern. Mit der Wahrscheinlichkeit von höchstens  $4/n$  einen solchen Verbindungsknoten zu mutieren ergibt sich eine Schranke für die Wahrscheinlichkeit eines Einflusses von  $\mathcal{O}(n^{\epsilon + \lambda}/n)$  und damit eine untere Schranke für die Wahrscheinlichkeit in eines der lokalen Optima zu geraten von  $2/3 - \mathcal{O}(n^{-1 + \epsilon + \lambda})$ .  $\square$

Trotz der komplexeren Mutationen des (1+1)-EA mit der in 2.2.2 beschriebenen Mutationswahrscheinlichkeit und der damit verbundenen Schwierigkeiten in der Analyse lässt sich eine exponentielle erwartete Laufzeit für den Vier-Cliquen-Graphen zeigen. Relativ einfach lässt sich Folgendes zeigen:

**Theorem 3.24.** *Der (1+1)-EA mit Mutationswahrscheinlichkeit  $1/(2n)$  benötigt auf einem Vier-Cliquen-Graph  $G = (V, E)$  der Größe  $|V| = 4n$  eine erwartete Laufzeit von  $\Omega(n^n 2^{-3n})$ .*

*Beweis.* Mit Wahrscheinlichkeit mindestens  $1/2^{4n}$  wird eines der lokalen Optima mit Wert 2 als initiales Individuum des (1+1)-EA erzeugt. Ein optimales Individuum kann dann nur erzeugt werden, wenn zwei Cliques mit je  $n$  Knoten die Seite wechseln. Eine Mutation, die ein  $x^{\text{opt}}$  erzeugt, welche also genau  $n$  bestimmte Knoten in  $L$  mutiert, wird mit einer Wahrscheinlichkeit von höchstens

$$P(x^{\text{opt}} \mid x^{\text{subopt}}) \leq \frac{1}{(2n)^n} \left(1 - \frac{1}{2n}\right)^{2n-n} \geq \frac{1}{e(2n)^n}$$

durchgeführt, was in einer erwarteten Zeit von mindestens  $\Omega((2n)^n)$  geschieht. Somit ergibt sich eine untere Schranke von

$$\frac{n^n}{2^{3n}}$$

für die erwartete Laufzeit.  $\square$

Diese Schranke lässt sich allerdings noch um einen Faktor  $\Theta(2^{4n} n^{-1/2})$  verbessern. Dazu zeigen wir für eine größere Menge von initialen Individuen eine erwartete exponentielle Laufzeit und schätzen diese genauer ab.

**Theorem 3.25.** *Auf  $\text{GBP}_G$  für einen Vier-Cliquen-Graph  $G = (V, E)$  der Größe  $|V| = 4n$  benötigt der (1+1)-EA mit Mutationswahrscheinlichkeit  $2/|V|$  eine erwartete Laufzeit von  $\Omega((2n)^{n-1/2})$ .*

*Beweis.* Die Wahrscheinlichkeit ein initiales Individuum  $z$  für den (1+1)-EA zu erzeugen, bei dem zwei der vier Cliques schon vollständig einer Seite zugeordnet sind (s. Abbildung 3.10) liegt bei mindestens

$$P(x^{\text{init}} = z) = \binom{2n}{n} \cdot \binom{4n}{2n}^{-1}.$$

Mit den unteren und oberen Schranken für Binomialkoeffizienten aus 1.5.3 ergibt sich

$$\binom{2n}{n} \geq 2^{2n} n^{-1/2}$$

und

$$\binom{4n}{2n} \leq \frac{1}{2\sqrt{\pi}} n^{-1/2} 2^{4n+1/2}$$

und damit eine untere Schranke für die Wahrscheinlichkeit von

$$P(x^{\text{init}} = z) \geq 2^{-2n-1/2}.$$

**Lemma 3.26.** *Der (1+1)-EA auf  $\text{GBP}_G$  für einen Vier-Cliquen-Graphen  $G$  senkt mit Wahrscheinlichkeit exponentiell nahe an 1 die Imbalancen  $k_2$  und  $k_3$  nicht mehr, wenn zu Beginn gilt  $k_1 = k_4 = n/2$ .*

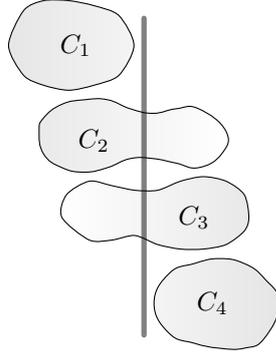


Abbildung 3.10: Schematische Darstellung eines Individuums mit zwei einseitig zugeordneten Cliques. Ein solches Individuum tritt als initiales Individuum des (1+1)-EA mit einer Wahrscheinlichkeit von mindestens  $1/2^{2n+1/2}$  auf.

*Beweis.* Solange  $k_1 = k_4 = n/2$  gilt, ist  $k_2 = k_3$ , da immer  $k_1 + k_2 = k_3 + k_4$  gelten muss. Wir gehen zunächst davon aus, dass  $k_2 = k_3 = o(n) = \mathcal{O}(n^{1-\epsilon})$  für ein  $\epsilon > 0$  ist. Existiert nun für eine Mutation  $M$  ein Tupel  $(v, w) \in \mathcal{R}'_M$  mit  $v \in C_1$  oder  $w \in C_4$ , so ist die Änderung der Fitness durch dieses Tupel  $\Delta_{(v,w)} \text{GBP}_G = \Omega(n)$ . Wir bezeichnen ein solches Tupel als vom Typ A. Außer dieser Art von Tupel existieren nur noch  $(v', w') \in \mathcal{R}''_M \cap (C_2 \times C_3)$  (Typ B) und  $(v'', w'') \in \mathcal{R}''_M \cap (C_3 \times C_2)$  (Typ C). Letztere sorgen ebenfalls für eine Verschlechterung der Fitness und schließen nach Konstruktion von  $\mathcal{R}_M$  die Existenz von einem  $(v', w')$  vom Typ B aus. Demnach wird eine Mutation  $M$  nur akzeptiert, wenn gilt  $\mathcal{R}''_M \cap (C_2 \times C_3) \neq \emptyset$ . Da diese eine Änderung der Fitness um  $\Delta_{(v',w')} \text{GBP}_G = -\mathcal{O}(n^{1-\epsilon})$  mit sich bringen, wird  $M$  mit  $(v, w) \in \mathcal{R}'_M$  vom Typ A nur akzeptiert, wenn  $\Omega(n^{1/(1-\epsilon)})$  Mutationen des Typs B vorliegen. Dies geschieht jedoch nur mit exponentiell kleiner Wahrscheinlichkeit. Also werden mit Wahrscheinlichkeit nahezu 1 von jeder akzeptierten Mutation  $k_2$  und  $k_3$  erhöht bis gilt  $k_2 = k_3 = \Omega(n)$ .

Sei nun  $k_2 = k_3 = \epsilon n$ . Ab diesem Zeitpunkt können auch  $k_1$  und  $k_4$  wieder sinken. Dann ist die Fitness der Partition  $x$  gegeben durch

$$\begin{aligned} \text{GBP}_G(x) &= n^2 - k_1^2 - k_2^2 - k_3^2 - k_4^2 \\ &= n^2 - 2 \left(\frac{n}{2}\right)^2 - 2(\epsilon n)^2. \end{aligned}$$

Jedes Individuum mit  $k_2 = o(n)$  oder  $k_3 = o(n)$  hat eine Fitness, die um  $\epsilon n^2 - o(n^2)$  kleiner ist. Eine Mutation die ein solches Individuum erzeugt wird also nicht akzeptiert. Da ein Tausch von  $\Omega(n)$  Knoten exponentiell geringe Wahrscheinlichkeit hat, wird mit einer exponentiell an 1 nahen Wahrscheinlichkeit  $k_2$  und  $k_3$  erhöht bis gilt  $k_1 = k_2 = k_3 = k_4 = n/2$ .  $\square$

Ist der (1+1)-EA in das lokale Optimum konvergiert, kann nur der Tausch von zwei der vier Cliques den Wert der Bisektion senken. Die Wahrscheinlichkeit einer

Mutation die genau  $n$  bestimmte Knoten jeder Seite vertauscht ist

$$P(x^{\text{opt}} | x^{\text{subopt}}) = \frac{1}{(2n)^n} \left(1 - \frac{1}{2n}\right)^n \binom{2n}{n}^{-1} \leq \frac{n^{1/2}}{(2n)^n 2^{2n}}$$

Somit ergibt sich eine untere Schranke für die erwartete Laufzeit des (1+1)-EA auf einem Vier-Cliquen-Graph von

$$P(x^{\text{init}} = z) \cdot \frac{(2n)^n 2^{2n}}{n^{1/2}} \geq (2n)^{n-1/2}.$$

□

### Eine untere Schranke für den global SEMO

Da schon für den (1+1)-EA eine umfassende Analyse nur mit geänderter Mutationswahrscheinlichkeit möglich war und eine Analyse des global SEMO wegen seiner Population stets schwieriger ist, zeigen wir für den global SEMO nur eine relativ einfache untere Schranke:

**Theorem 3.27.** *Der global SEMO benötigt auf  $\text{GBP}_G^0$  mit einem Vier-Cliquen-Graphen  $G = (V, E)$  der Größe  $|V| = 4n$  eine erwartete Laufzeit von  $\Omega(2^{-0,34n} n^{2n-2})$ .*

*Beweis.* Mit Wahrscheinlichkeit mindestens  $1/3^{4n}$  wird eines der lokalen Optima als initiale Population  $\mathbb{P} = \{x^{\text{subopt}}\}$  des global SEMO erzeugt. Es ist  $\text{GBP}_G^0(x^{\text{subopt}}) = (2, 0)$ . Alle erzeugten Individuen  $y$  mit  $f_0(y) \geq 2$  werden somit verworfen. Ein erzeugtes  $y$  kann nur akzeptiert werden, wenn entweder gilt  $f_0(y) \leq 1$  oder  $\text{zeros}(y) = 0 \wedge f_0(y) \leq 2$ . Somit ist die Größe der Population beschränkt durch 3.

Für den Fall  $\text{zeros}(y) = 0 \wedge f_0(y) \leq 2$  gibt es nur die Kandidaten  $y = x^{\text{opt}}$ , in dem eine Anordnung der Cliques ohne Verbindungskante im Schnitt vorliegt, oder eines der anderen lokalen Optima. Alle haben Hamming-Abstand mindestens  $2n$  von  $x^{\text{subopt}}$  und werden demnach mit exponentiell geringer Wahrscheinlichkeit erzeugt. Bevor wir die genaue Wahrscheinlichkeit bestimmen, betrachten wir deshalb den Fall, dass ein  $y$  mit  $f_0(y) \geq 2$  erzeugt wird, was mit deutlich höherer Wahrscheinlichkeit auftritt.

Solange  $x^{\text{subopt}} \in \mathbb{P}$  ist, können genau zwei weitere Individuen  $x'$  mit  $f_0(y') = 1$  und  $x''$  mit  $f_0(x'') = 0$  in der Population enthalten sein. Im Individuum  $x'$  muss mindestens einer der vier Verbindungsknoten von  $G$  ein freier Knoten sein, damit eine der Verbindungskanten nicht im Schnitt liegt. In  $x''$  müssen mindestens zwei Verbindungsknoten nicht eingeteilt sein. Mit Wahrscheinlichkeit

$$P(x' | x^{\text{subopt}}) \geq \frac{1}{3} \cdot \frac{4}{4n} \cdot \frac{1}{2} \left(1 - \frac{1}{4n}\right)^{4n-1} \geq \frac{1}{6en}$$

wird  $x'$  erzeugt, indem  $x^{\text{subopt}}$  gewählt und genau einer der Verbindungsknoten so mutiert wird, dass dieser ein freier Knoten ist. Dies geschieht in einer erwarteten Zeit von  $\mathcal{O}(n)$ .

Dann kann mit Wahrscheinlichkeit

$$P(x'' | x') \geq \frac{1}{3} \cdot \frac{2}{4n} \cdot \frac{1}{2} \left(1 - \frac{1}{4n}\right)^{4n-1} \geq \frac{1}{12en}$$

ein  $x''$  aus  $x'$  erzeugt werden, indem einer der beiden Verbindungsknoten der verbliebenen Verbindungskante zu einem freien Knoten mutiert wird. Auch dies passiert in erwarteter Zeit  $\mathcal{O}(n)$ .

In erwarteter Zeit  $\mathcal{O}(n)$  besteht also die Population aus den drei Individuen  $x^{\text{subopt}}, x', x''$  mit  $\text{GBP}_G^0(x') = (1, 1)$  und  $\text{GBP}_G^0(x'') = (0, 2)$ . Diese kann nur verändert werden, wenn in  $x'$  und  $x''$  ein anderer Verbindungsknoten zum freien Knoten wird oder ein Individuum  $x$  mit Hamming-Abstand  $H(x, x^{\text{subopt}}) \geq 2n$  erzeugt wird.

Um ausgehend von einer Population  $\mathbb{P}^{\text{subopt}} := \{x^{\text{subopt}}, x', x''\}$  das globale Optimum zu erreichen, müssen dann entweder  $2n - 2$  Knoten in  $x''$ ,  $2n - 1$  Knoten in  $x'$  oder  $2n$  Knoten in  $x^{\text{subopt}}$  mutiert und anders eingeteilt werden. Dies geschieht mit einer Wahrscheinlichkeit von höchstens

$$\begin{aligned} \mathbb{P}(x^{\text{opt}} \mid \mathbb{P}^{\text{subopt}}) &\leq \left( \frac{1}{3} \frac{1}{(4n)^{2n-2}} \cdot \frac{1}{2^{2n-2}} + \frac{1}{3} \frac{1}{(4n)^{2n-1}} \cdot \frac{1}{2^{2n-1}} + \frac{1}{3} \frac{1}{(4n)^{2n}} \cdot \frac{1}{2^{2n}} \right) \\ &\leq \frac{1}{(4n)^{2n-2}} \cdot \frac{1}{2^{2n-2}} \end{aligned}$$

und benötigt daher eine erwartete Laufzeit von mindestens  $2^{6n-6} n^{2n-2}$ .

Damit ergibt sich eine untere Schranke für die erwartete Laufzeit des global SEMO auf  $\text{GBP}_G^0$  mit einem Vier-Cliquen-Graphen  $G$  von

$$\frac{1}{3^{4n}} \cdot 2^{6n-6} n^{2n-2} = 2^{(6-4 \log 3)n-6} n^{2n-2} \geq 2^{-0,34n-6} n^{2n-2}.$$

□

### 3.3.2 Der 2n-n-Cliquen-Graph

Für die letzten beiden Graphen soll nur eine Analyse des (1+1)-EA und des global SEMO ausgehend von einem lokalen Optimum durchgeführt werden. Diese weisen auf einige interessante Vor- und Nachteile der beiden Algorithmen hin. So können wir für den ersten Graphen zeigen, dass der global SEMO in polynomieller erwarteter Zeit das lokale Optimum verlassen kann und das globale erreicht. Der (1+1)-EA muss dagegen in einer einzigen Mutation die Positionen von  $\mathcal{O}(n)$  Knoten ändern. Für den zweiten Graphen zeigt sich dagegen ein umgekehrtes Verhalten. So benötigt der global SEMO exponentielle erwartete Laufzeit um ein Optimum zu erreichen, wenn vorher eine bestimmte Population generiert wurde. Der (1+1)-EA schafft dies jedoch in erwarteter polynomieller Zeit.

Die Häufigkeit, mit der dabei ein solches lokales Optimum im Prozess der Optimierung des (1+1)-EA bzw. des global SEMO auftritt, werden wir in Kapitel 4 experimentell bestimmen.

Als ersten Graphen betrachten wir den 2n-n-Cliquen-Graphen:

**Definition 3.28.** *Ein Graph  $G = (V, E)$  mit  $|V| = 6n$  wird als 2n-n-Cliquen-Graph bezeichnet, wenn dieser aus zwei Cliquen  $A_1, A_2$  der Größe  $2n$  auf den Knoten  $v_1, \dots, v_{2n}$  und  $v_{3n+1}, \dots, v_{5n}$  und zwei Cliquen  $B_1, B_2$  der Größe  $n$  auf den Knoten  $v_{2n+1}, \dots, v_{3n}$  und  $v_{5n+1}, \dots, v_{6n}$  besteht. Dabei existieren je zwei Kanten von einem Knoten aus  $B_i$  zu zwei Knoten aus  $A_i$ , für  $i = 1, 2$ . Es ist dann*

$E = E_{A_1} \cup E_{A_2} \cup E_{B_1} \cup E_{B_2} \cup E_{V_B}$  mit  $E_M = \{\{v, w\} \mid v, w \in M\}$  und  $E_{V_B} = \{\{v_{2j+3in-1}, w_{j+(3i+2)n}\}, \{v_{2j+3in}, w_{j+(3i+2)n}\} \mid 1 \leq j \leq n, i = 0, 1\}$ .

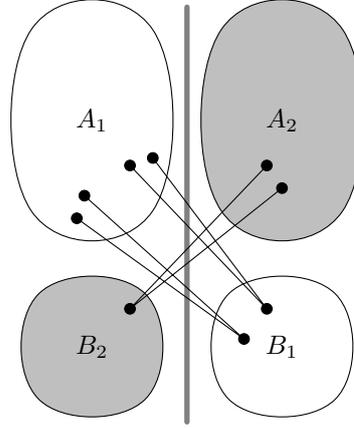


Abbildung 3.11: Schematische Darstellung eines Suboptimums des  $2n$ - $n$ -Cliquen-Graphen.

Offensichtlich ist die Bisektion  $L = A_1 \cup B_1$ ,  $R = A_2 \cup B_2$  ein globales Minimum mit Fitness 0, während  $L = A_1 \cup B_2$ ,  $R = A_2 \cup B_1$  ein lokales Minimum mit Wert  $4n$  darstellt (s. Abbildung 3.11). Wir werden zeigen, dass dieses vom  $(1+1)$ -EA nur durch eine Mutation von  $2n$  Knoten verlassen werden kann. Der SEMO hingegen ist in der Lage aus dem lokalen Optimum in erwarteter polynomieller Zeit zu entkommen.

### Der $(1+1)$ -EA auf dem $2n$ - $n$ -Cliquen-Graphen

**Theorem 3.29.** *Der  $(1+1)$ -EA auf  $GBP_G$ , für einen  $2n$ - $n$ -Cliquen-Graphen  $G = (V, E)$  mit  $|V| = 6n$ , benötigt erwartete exponentielle Laufzeit um das lokale Optimum  $L = A_1 \cup B_2$ ,  $R = A_2 \cup B_1$  zu verlassen.*

*Beweis.* In der Bisektion des betrachteten lokalen Optimums liegen genau  $4n$  Kanten im Schnitt. Wir zeigen nun, dass nahezu alle Mutationen des lokalen Optimums verworfen werden.

Sei  $M$  die Menge der Knoten, welche durch eine Mutation verändert werden. Betrachten wir einen beliebigen Knoten  $v \in M$ , welcher o.B.d.A. von der Menge  $L$  in die Menge  $R$  überführt wird und sei  $C$  die Clique dieses Knotens. Werden durch die Mutation  $k = |M \cap C|$  Knoten aus der Clique  $C$  von  $L$  nach  $R$  überführt, so schneiden nach der Mutation, je nach Größe der Clique  $C$ , mindestens  $n - k$  Kanten von  $v$  zu den verbliebenen Knoten der Clique  $C$  die Bisektion. Insgesamt sind dann also mindestens  $k(n - k)$  neue Kanten der Clique  $C$  nach der Mutation Schnittkanten. Es werden jedoch höchstens  $2k$  der Verbindungskanten zu diesen Knoten aus dem Schnitt entfernt. Weiterhin müssen ebenso  $k$  Knoten aus der Menge  $R$  in die Menge  $L$  überführt werden, um das Gleichgewicht der beiden Mengen zu erhalten. Auch für diese Knoten gilt das eben gezeigte.

Damit eine Mutation akzeptiert wird muss also gelten:  $2k(n-k) - 4k \leq 0$ . Dies ist äquivalent zu  $2k(n-k-2) \leq 0$  und es muss damit  $k = 0$  oder  $k \geq n-2$  gelten. Es müssen also mindestens  $n-2$  Knoten aus  $L$  mutiert werden, damit eine Veränderung der Population des (1+1)-EA erreicht wird. Die Wahrscheinlichkeit einer solchen Mutation kann mit Hilfe der Stirling-Formel aus Kapitel 1.5.3 abgeschätzt werden zu:

$$\frac{1}{(3n)^{n-2}} \binom{3n}{n-2} \leq \frac{1}{(n-2)!} = \exp(-\Omega(n \log n)).$$

Die erwartete Laufzeit, bis eine solche Mutation eintritt, ist also nach unten beschränkt durch  $\exp(\Omega(n \log n))$ .  $\square$

### Der global SEMO auf dem $2n$ - $n$ -Cliques-Graphen

**Theorem 3.30.** *Der global SEMO auf  $\text{GBP}_G^0$ , für einen  $2n$ - $n$ -Cliques-Graphen  $G = (V, E)$  mit  $|V| = 6n$ , benötigt eine erwartete Zeit von  $\mathcal{O}(n^4)$ , um aus dem Suboptimum  $L = A_1 \cup B_2$ ,  $R = A_2 \cup B_1$  zu entkommen und das Optimum zu erreichen.*

*Beweis.* Wir gehen davon aus, dass ein Individuum  $x^{\text{subopt}}$  mit  $\text{zeros}(x^{\text{subopt}}) = 0$  und  $L(x^{\text{subopt}}) = A_1 \cup B_2$ ,  $R(x^{\text{subopt}}) = A_2 \cup B_1$  Teil der Population ist. Der Beweis verläuft dann in zwei Phasen. Als erstes wollen wir die Zeit abschätzen, bis ein Individuum  $x^{2n}$  mit  $\text{zeros}(x^{2n}) \leq 2n$  und  $f_0(x^{2n}) = 0$  erzeugt wurde. In der zweiten Phase wird dann die Dauer bestimmt, bis ein optimales Individuum  $x^{\text{opt}}$  mit  $\text{GBP}_G^0(x^{\text{opt}}) = (0, 0)$  erzeugt wurde.

**Erste Phase** Sei  $\mathbb{X}_i^I \subseteq \mathcal{S}^*$ , für  $0 \leq i \leq 2n$  und  $x \in \mathbb{X}_i^I$  genau dann, wenn  $f_0(x) = 4n - 2i$  und  $\text{zeros}(x) \leq i$  gilt. Da die Funktion  $f_0(x)$  für alle Individuen einer Klasse gleich ist, sind alle Elemente aus  $\mathbb{X}_i^I$  miteinander vergleichbar. Somit kann höchstens ein Individuum aus jedem  $\mathbb{X}_i^I$  in der Population  $\mathbb{P}$  enthalten sein. Weiterhin ist die Größe der Population durch  $4n$  beschränkt, da nur Individuen mit Fitness kleiner oder gleich  $4n$  enthalten sind. Andere Individuen werden von  $x^{\text{subopt}}$  dominiert. Offensichtlich ist  $x^{\text{subopt}} \in \mathbb{X}_0^I$ .

Betrachten wir nun die anderen Individuen der Klassen  $\mathbb{X}_i^I$ . Wir gehen zuerst davon aus, dass ein  $x \in \mathbb{X}_i^I$  nicht der Partition  $x^{\text{subopt}}$  mit möglichen freien Knoten entspricht, also die Schnittkanten nicht aus  $A_1 \times B_1$  bzw.  $A_2 \times B_2$  sind. Dann müssen die Schnittkanten jedoch durch Kanten innerhalb der Cliques zustande kommen. Da aber  $f_0(x) \leq 4n$  sein muss, damit  $x$  akzeptiert wird, können sich dann höchstens vier Knoten auf der anderen Seite befinden, als der Rest der zugehörigen Clique. Ist ein solches Individuum in der Population enthalten, so wird mit Wahrscheinlichkeit

$$P(x' | x) \geq \frac{1}{4n} \cdot \frac{1}{6n} \left(1 - \frac{1}{6n}\right)^{6n-1}$$

dieses Individuum  $x$  aus der Population gewählt und einer dieser „falsch“ eingeordneten Knoten entfernt oder korrigiert. Dies geschieht in erwarteter Zeit  $\mathcal{O}(n^2)$ . Nach spätestens vier solcher Mutationen liegt dann ein Individuum  $x' \in \mathbb{X}_0^I$  vor. Es kann also nur eine konstante Anzahl dieser Mutationen vorkommen, bis  $\mathbb{X}_{2n}^I \cap \mathbb{P} \neq \emptyset$  gilt.

Wir gehen nun davon aus, dass ein  $x \in \mathbb{X}_i^I$  mit  $A_1$  und  $B_1$  bzw.  $A_2$  und  $B_2$  auf unterschiedlichen Seiten Teil der Population ist. Um ein Individuum  $x' \in \mathbb{X}_{i+1}^I$  zu

erzeugen, genügt das Entfernen eines Knotens der beiden  $n$ -Cliques  $B_1$  und  $B_2$ . Die Wahrscheinlichkeit ein Individuum  $x$  aus  $\mathbb{X}_i^I$  zu wählen und eine solche Mutation durchzuführen, ist

$$P(x' | x) \geq \frac{1}{4n} \cdot (2n - i) \cdot \frac{1}{6n} \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{6n}\right)^{6n-1} \geq \frac{2n - i}{48e n^2}.$$

Die erwartete Zeit  $T_{i+1}^I$  bis ein Individuum aus  $\mathbb{X}_{i+1}^I$  erzeugt wurde, wenn  $\mathbb{P} \cap \mathbb{X}_i^I \neq \emptyset$  gilt, ist damit  $E(T_{i+1}^I) \leq (48e n^2)/(2n - i)$ . Ein so erzeugtes Individuum kann nur (schwach) dominiert werden, wenn ein Individuum  $y$  mit besserer oder gleicher Fitness  $f_0(y)$  und höchstens genau so großem zeros( $y$ ) existiert. Dann gilt aber  $\bigcup_{j=i+1}^{2n} \mathbb{X}_j^I \cap \mathbb{P} \neq \emptyset$  und wir sind dem Ziel höchstens näher.

Die erwartete Zeit  $T^I$  bis ein Individuum aus  $x^{2n} \in \mathbb{X}_{2n}^I$  erzeugt wurde, ist damit beschränkt durch

$$E(T^I) \leq \sum_{i=0}^{2n-1} \frac{48e n^2}{2n - i} = 48e n^2 \sum_{j=1}^{2n} \frac{1}{j} = 48e n^2 H_{2n} = \mathcal{O}(n^2 \log n),$$

mit der  $2n$ -ten harmonischen Zahl  $H_{2n}$ .

**Zweite Phase** Für die zweite Phase definieren wir die Teilmengen des Suchraums  $\mathbb{X}_{2n}^{II}, \dots, \mathbb{X}_0^{II}$ , wobei  $x \in \mathbb{X}_i^{II}$  genau dann gilt, wenn  $\text{GBP}_G^0(x) = (0, i)$  ist. Alle Individuen einer Teilmenge  $\mathbb{X}_i^{II}$  haben also die gleiche Fitness. Für je zwei Individuen  $x^i \in \mathbb{X}_i^{II}$ ,  $x^j \in \mathbb{X}_j^{II}$  wird  $x_i$  (schwach) von  $x^j$  dominiert, wenn  $j \leq i$  ist. Ansonsten wird  $x^j$  von  $x^i$  dominiert. Daher kann aus allen Teilmengen zusammen nur genau ein Individuum in der Population enthalten sein. Es gilt also  $|\mathbb{P} \cap \bigcup_{i=0}^{2n} \mathbb{X}_i^{II}| = 1$ . Für das Individuum  $x^{2n}$ , welches zu Beginn dieser Phase in der Population enthalten ist, ist  $f_0(x^{2n}) = 0$  und  $\text{zeros}(x^{2n}) \leq 2n$ . Also ist  $x^{2n} \in \mathbb{X}_l^{II}$  für ein  $l \leq 2n$ . Weiterhin wird  $x^{2n}$  nur von Individuen  $y$  mit  $\text{zeros}(y) \leq 2n$  (schwach) dominiert. Somit gilt während der gesamten zweiten Phase  $\text{zeros}(x) \leq 2n$  für alle  $x \in \mathbb{P}$ .

Sobald ein Individuum  $x \in \mathbb{X}_i^{II}$  Teil der Population ist, muss für alle Individuen  $y \in \mathbb{P}$   $\text{zeros}(y) \leq i$  gelten, da  $y$  sonst von  $x$  dominiert werden würde. Daher ist die Größe der Population dann durch  $i$  beschränkt. Die Population verkleinert sich also in jedem Schritt, mit dem wir uns dem Ziel nähern.

Da  $f_0(x) = 0$  für  $x \in \mathbb{X}_i^{II}$  gilt, müssen in  $x$  alle eingeteilten Knoten einer Clique  $C$  auf der selben Seiten der Partition liegen. Ansonsten existieren Schnittkanten dieser Clique der Bisektion. Eine weitere Eigenschaft von  $x$  zeigen wir im folgenden Lemma.

**Lemma 3.31.** *Für  $x \in \mathbb{X}_i^{II}$ , mit mindestens je einem eingeteilten Knoten aus  $A_1$  und  $A_2$ , liegen die eingeteilten Knoten aus  $A_1$  und  $A_2$  auf verschiedenen Schnittseiten.*

*Beweis.* Wir gehen davon aus, dass ein  $x$  existiert, in dem sich die eingeteilten Knoten aus  $A_1$  und  $A_2$  auf der gleichen Schnittseite befinden und führen dies zum Widerspruch. Seien o.B.d.A. die eingeteilten Knoten aus  $A_1 \cup A_2$  in  $L(x)$ . Falls auch die eingeteilten Knoten einer weiteren Clique  $B_j$  in  $L(x)$  enthalten sind, so müssen  $|A_1| + |A_2| + |B_j| - |L(x)| \geq 2n$  Knoten dieser Cliques freie Knoten sein. Es ist aber  $\text{zeros}(x) \leq 2n$ , also gibt es genau  $2n$  freie Knoten aus  $A_1 \cup A_2 \cup B_j$ . Dann müssen aber alle Knoten aus  $B_{1-j}$  in  $R(x)$  eingeteilt sein und es existieren Verbindungskanten zu  $A_{1-j}$  im Schnitt.

Gehen wir also davon aus, dass die eingeteilten Knoten aus  $B_1$  und  $B_2$  in  $R(x)$  liegen. Sei dann  $k_1$  die Anzahl der eingeteilten Knoten in  $B_1$  und  $k_2$  die aus  $B_2$ . Dann gibt es  $2n - k_1 - k_2$  freie Knoten in  $B_1 \cup B_2$ . Da  $f_0(x) = 0$  gilt, darf es keine Verbindungskanten im Schnitt geben. Somit müssen mindestens  $2k_1$  Knoten aus  $A_1$  und  $2k_2$  Knoten aus  $A_2$  freie Knoten sein, nämlich genau die Knoten, die eine Verbindungskante mit einem eingeteilten Knoten aus  $B_1$  bzw.  $B_2$  teilen. Somit müssen mindestens  $2n + k_1 + k_2$  freie Knoten existieren. Da aber  $\text{zeros}(x) \leq 2n$  gilt, muss  $k_1 = k_2 = 0$  sein. Damit sind alle Knoten aus  $B_1$  und  $B_2$  freie Knoten und somit gibt es  $2n$  freie Knoten aus  $B_1 \cup B_2$ . Dann müssen aber alle Knoten aus  $A_1$  und  $A_2$  eingeteilt sein. Die Cliques  $A_1$  und  $A_2$  passen jedoch vollständig eingeteilt nicht auf eine Schnittseite alleine, denn es ist  $|A_1| + |A_2| = 4n$ , jedoch  $|L(x)|, |R(x)| \leq 3n$ . Somit müssen in jedem  $x \in \mathbb{X}_i^{II}$  eingeteilte Knoten aus  $A_1$  und  $A_2$  auf verschiedenen Schnittseiten liegen.  $\square$

Wir benötigen dieses Lemma um zu zeigen, dass das Einsortieren der freien Knoten auf die richtige Seite stets möglich ist und diese nicht bereits mit  $3n$  eingeteilten Knoten gefüllt ist.

Für  $\{x\} = \mathbb{X}_i^{II} \cap \mathbb{P}$  lassen sich nun vier Fälle unterscheiden:

1. Alle Knoten aus  $B_1$  und  $B_2$  sind freie Knoten. Dann ist  $i = 2n$  und es kann ein  $x' \in \mathbb{X}_{i-1}^{II}$  aus  $x$  erzeugt werden, indem einer der  $i$  freien Knoten aus  $B_1$  und  $B_2$  auf die jeweils korrekte Seite eingeordnet wird, so dass keine Verbindungskanten im Schnitt liegen. Solch eine Seite existiert für beide Cliques, da nach Lemma 3.31 nicht  $A_1$  und  $A_2$  auf einer Schnittseite liegen. Somit gilt für diesen Fall  $|L(x)| = |R(x)| = 2n$  und es ist möglich einen freien Knoten korrekt einzuordnen. Dies geschieht mit Wahrscheinlichkeit

$$P(x' | x) \geq \frac{1}{i} \cdot i \cdot \frac{1}{6n} \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{6n}\right)^{6n-1} \geq \frac{1}{12en},$$

womit sich eine erwartete Zeit für eine solche Mutation von  $E(T_i^{II,1}) \leq 12en$  ergibt.

2. Nur aus einer der Cliques  $B_1$  und  $B_2$  sind Knoten eingeordnet. Sei o.B.d.A.  $B_1$  die Clique der eingeordneten Knoten. Dann kann durch Einsortieren eines freien Knotens ein  $x \in \mathbb{X}_{i-1}^{II}$  erstellt werden.

Es müssen sich wieder alle eingeordneten Knoten einer Clique auf einer Seite befinden, sonst gäbe es Schnittkanten. Befinden sich dann  $k$  eingeordnete Knoten von  $B_1$  auf einer Seite, während sich eine unbestimmte Anzahl von eingeordneten Knoten aus  $A_1$  auf der gegenüberliegenden Seite befinden, so müssen mindestens  $2k$  Knoten der Clique  $A_1$  nicht eingeordnet sein, damit keine Verbindungskanten im Schnitt liegen. Damit ergibt sich eine Anzahl von mindestens  $(n - k) + 2k = n + k$  nicht eingeordneten Knoten aus  $A_1 \cup B_1$ , sowie zusätzlich den  $n$  nicht eingeordneten Knoten aus  $B_2$ . Es gilt aber  $\text{zeros}(x) \leq 2n$  für alle  $x \in \mathbb{X}_i^{II}$ . Somit muss  $k = 0$  gelten. Also befinden sich die eingeordneten Knoten aus  $B_1$  auf der Seite von  $A_1$ .

Aus  $x$  kann dann ein  $x' \in \mathbb{X}_{i-1}^{II}$  erzeugt werden, indem einer der freien Knoten korrekt einsortiert wird, d.h. Knoten aus  $A_1$ ,  $A_2$  und  $B_1$  auf die Seite, auf der die restlichen Knoten der Clique liegen, Knoten aus  $B_2$  auf die Seite von  $A_2$ . Nach Lemma 3.31 sind die Seiten von  $A_1$  und  $A_2$  verschieden. Somit ist ein

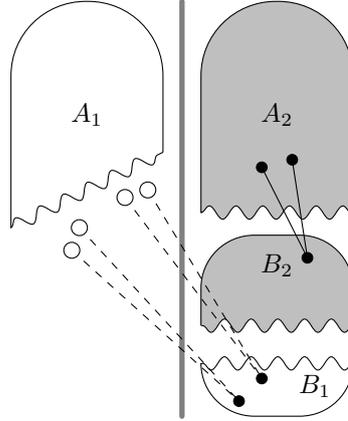


Abbildung 3.12: Schematische Darstellung einer Partition des  $2n$ - $n$ -Cliquen-Graphen gemäß des Falles 3.b) der Analyse für den global SEMO. Alle Knoten aus  $A_1$ , welche eine Kante mit den eingordneten Knoten aus  $B_1$  gemeinsam haben sind freie Knoten.

Einsortieren auf die jeweilige Seite möglich. Die Wahrscheinlichkeit dafür ist beschränkt durch

$$P(x' | x) \geq \frac{1}{i} \cdot i \cdot \frac{1}{6n} \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{6n}\right)^{6n-1} \leq \frac{1}{12en}$$

und hat eine erwartete Zeit von  $E(T_i^{II,2}) \leq 12en$  bis zum Eintreten dieses Falles.

3. Es gibt in  $B_1$  und  $B_2$  je mindestens einen Knoten der eingordnet ist, aber nicht alle Knoten aus  $B_1$  und  $B_2$  sind eingordnet. Dann ist die erwartete Zeit, bis ein  $x' \in \mathbb{X}_{i-1}^{II}$  erzeugt wird beschränkt durch  $E(T_i^{II,3}) \leq T := 144en^2 + 12en$ , wie wir nun zeigen werden. Der Beweis dazu lässt sich in die beiden folgenden Unterfälle aufspalten:
  - a) Die Mengen der eingordneten Knoten aus  $B_1$  und  $B_2$  befinden sich auf unterschiedlichen Schnittseiten. Nach gleicher Argumentation wie unter Punkt 2 müssen die Knoten aus  $B_1$  auf der Seite von  $A_1$  und die aus  $B_2$  auf der von  $A_2$  sein. Demnach kann wie unter Punkt 2 ein beliebiger freier Knoten in erwarteter Zeit  $12en \leq T$  korrekt eingordnet und damit ein  $x' \in \mathbb{X}_{i-1}^{II}$  erzeugt werden.
  - b) Die Mengen der eingordneten Knoten beider Cliques befinden sich auf der selben Schnittseite (s. Abbildung 3.12). Dieser Fall kann durch eine für uns ungünstige Mutation eintreten und erschwert den Prozess. Die falsch eingordneten Knoten können nun nicht einfach entfernt werden, da die Fitness  $f_0$  eines solchen Individuums weiterhin bei 0 bleibt, die Funktion  $\text{zeros}(x)$  nimmt jedoch zu. Somit würde eine solche Mutation verworfen werden.

Wie in Abbildung 3.12 seien nun o.B.d.A. die eingeteilten Knoten aus  $A_1$  in  $L(x)$ , alle anderen in  $R(x)$ . Falls  $j > 0$  freie Knoten existieren, welche so eingordnet werden können, dass keine Kante zum Schnitt hinzugefügt

wird, so kann ein  $x' \in \mathbb{X}_{i-1}^{II}$  mit Wahrscheinlichkeit

$$P(x' | x) \geq \frac{1}{i} \cdot j \cdot \frac{1}{6n} \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{6n}\right)^{6n-1} \geq \frac{j}{12e in} \geq \frac{1}{12e in}$$

erzeugt werden. Dies geschieht in einer erwarteten Zeit von höchstens  $12e in \leq T$ .

Gehen wir also nun davon aus, dass kein Knoten mehr eingefügt werden kann, ohne eine Kante zur Bisektion hinzuzufügen. In diesem Fall gilt  $|R(x)| = 3n$ , da sonst Knoten aus  $A_2$  oder  $B_2$  eingeteilt werden könnten. Weiterhin haben alle freien Knoten aus  $A_1$  eine Kante zu einem eingeteilten Knoten aus  $B_1$  und analog alle freien Knoten aus  $B_1$  eine Kante zu einem eingeteilten Knoten aus  $A_1$  (diese sind durch gestrichelte Linien in Abbildung 3.12 dargestellt). Ansonsten könnte einer der freien Knoten eingeteilt werden, ohne dass eine Kante dem Schnitt hinzugefügt wird.

Über ein Individuum  $y$  mit  $f_0(y) = 1$  kann dann wieder ein  $x' \in \mathbb{X}_{i-1}^{II}$  erzeugt werden. Durch Einfügen eines freien Knotens aus  $A_1$ , kann solch ein  $y$  erzeugt werden, für das  $\text{zeros}(y) = i - 1$  und  $f_0(y) = 1$  gilt. Dies geschieht mit Wahrscheinlichkeit

$$P(y | x) \geq \frac{1}{i} \cdot j \cdot \frac{1}{6n} \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{6n}\right)^{6n-1} \geq \frac{j}{12e in} \geq \frac{1}{12e in}$$

für  $j = |A_1 \setminus L(x)| > 0$  und daher in erwarteter Zeit höchstens  $12e in$ .

Wir beachten dabei den Fall, dass  $y$  dominiert wird und zeigen, dass dies zu keinem Nachteil führt.

**Lemma 3.32.** *Aus einem  $y \in \mathbb{P}$  mit  $f_0(y) = 1$  und  $\text{zeros}(y) \leq i - 1$  kann in einer erwarteten Zeit höchstens  $T' = 144e in^2$  ein Individuum  $x' \in \mathbb{X}_{i-1}^{II}$  erzeugt werden.*

*Beweis.* Für ein beliebiges  $y \in \mathbb{P}$  mit  $f_0 = 1$  und  $\text{zeros}(x) \leq i - 1$  existiert genau eine Schnittkante  $s \in E$ .

**Fall (i):** Falls  $s \in E_C$  für eine Clique  $C$  ist, so müssen aus  $C$  zwei Knoten eingeordnet sein, welche auf verschiedenen Schnittseiten liegen. Ein beliebiger Schnitt durch eine Clique der Größe mindestens 3 führt zu mindestens 2 Schnittkanten. Es ist aber  $f_0(y) = 1$  und somit müssen genau zwei Knoten aus  $C$  eingeordnet sein. Weiterhin ist  $\text{zeros}(y) \leq i - 1$  und  $i \leq 2n$  für die zweite Phase, also  $\text{zeros}(y) \leq 2n - 1$ .

Betrachten wir zunächst den Fall, dass  $C$  einer der Cliques  $A_1$  oder  $A_2$  entspricht. Sei dies o.B.d.A. die Clique  $A_1$ . Dann sind  $2n - 2$  Knoten aus  $A_1$  freie Knoten. Es verbleibt dann noch höchstens ein freier Knoten aus den restlichen Cliques. Die Cliques  $A_2$ ,  $B_1$  und  $B_2$  sind also nahezu vollständig eingeordnet. Da keine weitere Schnittkante existiert, müssen also die eingeteilten Knoten aus  $A_2$  und  $B_2$  auf der selben Schnittseite liegen. Sei dies o.B.d.A. die rechte Seite  $R(x)$ . Da ein Knoten aus  $A_1$  auf jeder Seite liegt, ist auch ein Knoten in  $R(x)$  enthalten. Somit muss ein Knoten aus  $A_2 \cup B_2$  ein freier Knoten sein, denn sonst ist  $|R(x)| \geq 3n + 1$ . Also sind  $2n - 2$  Knoten aus  $A_1$  und ein Knoten aus  $A_2 \cup B_2$  freie Knoten. Dies entspricht der maximalen Anzahl freier Knoten, die  $y$  haben kann. Daher

sind alle Knoten aus  $B_1$  eingeteilt und es existiert eine Verbindungskante von  $B_1$  zu dem Knoten in  $R(x) \cap A_2$ . Dies widerspricht jedoch der Forderung  $f_0(y) = 1$ . Somit kann  $C$  keine der Cliques  $A_1$  oder  $A_2$  sein.

Also ist  $C$  eine der beiden Cliques  $B_1$  und  $B_2$ . Sei dann o.B.d.A.  $C = B_1$  und seien die eingeteilten Knoten aus  $A_1$  in  $L(x)$ . Wir zeigen dann, dass  $L(x) < 3n$  ist und somit der Knoten  $\{v\} = B_1 \cap R(x)$  durch eine Mutation die Seite wechseln kann, um zu einer Knoteneinteilung  $x' \in \mathbb{X}_{i-1}^{II}$  zu führen. Eine solche Mutation hat Wahrscheinlichkeit

$$P(x' | y) \leq \frac{1}{i} \cdot \frac{1}{6n} \cdot 12 \left(1 - \frac{1}{6n}\right)^{6n-1} \leq \frac{1}{12e in}.$$

Die erwartete Dauer dafür ist dann  $12e in \leq T'$ .

Für  $C = B_1$  sind  $n - 2$  freie Knoten aus  $B_1$ . Es verbleiben daher noch höchstens  $n + 1$  freie Knoten für die restlichen Cliques. Nach der Argumentation von Fall 2 gilt wieder: die eingeteilten Knoten von  $A_2$  und  $B_2$  liegen auf der selben Schnittseite. Da es nur  $n + 1$  freie Knoten aus  $A_1 \cup A_2 \cup B_2$  gibt, können nicht alle Cliques auf der selben Schnittseite liegen und die eingeteilten Knoten aus  $A_2$  und  $B_2$  müssen in  $R(x)$  liegen. Somit gilt  $L(x) \leq 2n + 1 < 3n$  und der Knoten  $v$  mit  $\{v\} = B_1 \cap R(x)$  kann wie oben beschrieben mutiert werden.

**Fall (ii):** Falls von jeder Clique mehr als zwei Knoten eingeordnet sind, so kann es sich nicht um eine Kante aus  $E_{A_1} \cup E_{A_2} \cup E_{B_1} \cup E_{B_2}$  handeln. Also ist  $s \in E_{VB}$ . Sei im Folgenden o.B.d.A.  $s \in A_1 \times B_1$  und seien die eingeteilten Knoten aus  $A_1$  in  $L(x)$ , die aus  $B_1$  in  $R(x)$ . Die Schnittkante verläuft somit zwischen zwei Knoten  $v_k \in B_1 \cap R(x)$  und  $v_l \in A_1 \cap L(x)$ .

Durch eine Mutation von zwei Knoten kann dann ein  $x'$  mit  $\text{zeros}(x') = i - 1$  und  $f_0(x') = 0$  erstellt werden: Dazu wird der Knoten  $v_k \in B_1 \cap R(x)$  zu einem freien Knoten mutiert, während gleichzeitig der Knoten  $v_m \in A_1 \cap L(x)$  in  $L(x)$  eingeordnet wird, welcher die zweite Verbindungskante mit  $v_k$  gemein hat (es ist also  $m = l \pm 1$ , je nach Wert von  $l$ ). Dies geschieht mit einer Wahrscheinlichkeit von

$$P(x' | y) \geq \frac{1}{i} \cdot \frac{1}{(6n)^2} \cdot \frac{1}{2^2} \cdot \left(1 - \frac{1}{6n}\right)^{6n-2} \geq \frac{1}{144e in^2}.$$

Die erwartete Dauer, bis eine solche Mutation eintritt ist nach oben beschränkt durch  $144e in^2 = T'$ .  $\square$

Somit ist die erwartete Zeit für diesen Fall, bis aus  $x \in \mathbb{X}_i^{II}$  ein  $x' \in \mathbb{X}_{i-1}^{II}$  erzeugt wird, höchstens der Summe aus der erwarteten Zeit, bis ein  $y$  mit  $f_0(y) = 1$  und  $\text{zeros}(y) = i - 1$  generiert wird und der erwarteten Zeit daraus  $x'$  zu erzeugen:

$$E(T_i^{II,3}) \leq 144e in^2 + 12e in.$$

4. Alle Knoten aus  $B_1$  und  $B_2$  sind eingeordnet. Für ein solches  $x$  muss  $B_1$  auf der Schnittseite der eingeordneten Knoten aus  $A_1$  und  $B_2$  auf der Seite der eingeordneten Knoten aus  $A_2$  liegen. Sonst gibt es Verbindungskanten im Schnitt.

Sind Knoten aus  $A_1$  und  $A_2$  eingeordnet, so befinden sich diese nach Lemma 3.31 auf verschiedenen Seiten. Also befinden sich auch  $B_1$  und  $B_2$  auf verschiedenen Seiten. Dann kann ein  $x' \in \mathbb{X}_{i-1}^{II}$  erzeugt werden, indem ein freier Knoten aus  $A_1$  bzw.  $A_2$  korrekt eingeordnet wird.

Falls dagegen nicht aus beiden Cliques  $A_1$  und  $A_2$  ein Knoten eingeordnet ist, so muss also mindestens eine Clique nur aus freien Knoten bestehen. Da  $i \leq 2n$  ist, kann höchstens eine Clique  $A_j$  nur aus freien Knoten bestehen. Dann aber sind alle Knoten aus  $A_{1-j}$  eingeordnet und die Seite, auf der  $A_{1-j}$  und  $B_{1-j}$  liegen kann  $B_j$  nicht mehr beinhalten. Somit liegen auch dann die Cliques  $B_1$  und  $B_2$  auf verschiedenen Seiten und Knoten aus  $A_j$  können eingefügt werden, ohne dass eine Verbindungskante zum Schnitt hinzugefügt wird.

Es kann also mit Wahrscheinlichkeit

$$P(x' | x) \geq \frac{1}{i} \cdot i \cdot \frac{1}{6n} \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{6n}\right)^{6n-1} \geq \frac{1}{12en}$$

einer der  $i$  freien Knoten korrekt eingeordnet werden und somit ein  $x' \in \mathbb{X}_{i-1}^{II}$  erzeugt werden. Dies geschieht in erwarteter Zeit  $E(T_i^{II,4}) \leq 12en$ .

Insgesamt ergibt sich eine obere Schranke in der erwarteten Dauer den Fall  $\mathbb{X}_{i-1}^{II} \cap \mathbb{P} \neq \emptyset$  zu erreichen, wenn  $\mathbb{X}_i^{II} \cap \mathbb{P} \neq \emptyset$  gegeben ist, von

$$E(T_i^{II}) \leq \max_{1 \leq j \leq 4} \{E(T_i^{II,j})\} \leq 144en^2 + 12en.$$

Die erwartete Zeit bis zum Abschluss der zweiten Phase ist damit

$$\begin{aligned} E(T^{II}) &\leq \sum_{i=1}^{2n} (144en^2 + 12en) = (144en^2 + 12en) \sum_{i=1}^{2n} i \\ &= (144en^2 + 12en) \cdot n(2n-1) = \mathcal{O}(n^4). \end{aligned}$$

Bis zum Erzeugen eines Individuums  $x^{\text{opt}} \in \mathbb{X}_0^{II}$  sind also im Erwartungswert insgesamt  $\mathcal{O}(n^4)$  Schritte nötig.  $\square$

### 3.3.3 Der Clique-Ring-Graph

Als letzte Klasse von Graphen mit lokalen Optima analysieren wir den Clique-Ring-Graphen. Dieser besitzt ein lokales Optimum für den Schnitt, aus welchem der (1+1)-EA in erwarteter polynomieller Zeit entkommt, während der global SEMO im Erwartungswert exponentielle Rechenzeit benötigt um von dem lokalen Optimum ausgehend das globale Optimum zu erreichen. Das Verhalten ist also genau umgekehrt wie bei dem vorherigen Graphen.

**Definition 3.33.** *Ein Graph  $G = (V, E)$  ist ein Clique-Ring-Graph, wenn  $G$  für  $V = \{v_1, \dots, v_{6n}\}$  zwei Cliques  $A_1$  und  $A_2$  der Größe  $2n$  auf den Knoten  $v_1, \dots, v_{2n}$  und  $v_{3n+1}, \dots, v_{5n}$  beinhaltet, sowie zwei Ringe  $B_1$  und  $B_2$  der Größe  $n$  auf  $v_{2n+1}, \dots, v_{3n}$  und  $v_{5n+1}, \dots, v_{6n}$ . Weiterhin existiert eine Verbindungskante zwischen einem ausgezeichneten Verbindungsknoten aus  $A_i$  und zwei aufeinander folgenden Verbindungsknoten aus  $B_i$  für  $i = 1, 2$  (s. Abbildung 3.13). Es gilt  $E = E_{A_1} \cup E_{A_2} \cup E_{B_1} \cup$*

$E_{B_2} \cup \{\{v_1, v_{2n+1}\}, \{v_1, v_{2n+2}\}, \{v_{3n+1}, v_{5n+1}\}, \{v_{3n+1}, v_{5n+2}\}\}$  mit  $E_{A_k} = \{\{v_i, v_j\} \mid 3(k-1)n+1 \leq i, j \leq (3(k-1)+2)n\}$  und  $E_{B_k} = \{\{v_i, v_{i+1}\} \mid (3(k-1)+2)n+1 \leq i \leq (3(k-1)+3)n\}$ .

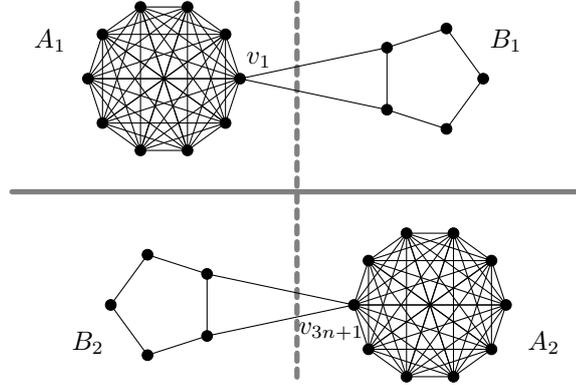


Abbildung 3.13: Ein Clique-Ring-Graph der Größe  $|V| = 30$ . Die durchgezogene Linie stellt den optimalen Schnitt dar, die gestrichelte ein lokales Optimum.

Der Clique-Ring-Graph  $G = (V, E)$  hat ein lokales Minimum mit einer Bisektion von 4 Kanten, wenn  $A_1 \cup B_2$  und  $A_2 \cup B_1$  auf verschiedenen Schnittseiten liegen. Dagegen hat  $G$  eine optimale Bisektion mit einem Wert von 0, wenn gilt  $L = A_1 \cup B_1$  und  $R = A_2 \cup B_2$  (s. Abbildung 3.13).

### Der (1+1)-EA auf dem Clique-Ring-Graph

**Theorem 3.34.** *Der (1+1)-EA benötigt erwartete Zeit  $\mathcal{O}(n^4)$  für die Optimierung von  $\text{GBP}_G$  für einen Clique-Ring-Graphen  $G = (V, E)$  mit  $6n$  Knoten, falls das Suboptimum  $z$  mit  $L(z) = A_1 \cup B_2$ ,  $R(z) = A_2 \cup B_1$  erzeugt und akzeptiert wurde.*

*Beweis.* Die Fitness von  $z$  ist  $\text{GBP}_G(z) = 4$  (s. Abbildung 3.13). Durch eine Mutation, welche die Knoten  $v_{2n+1}, v_{2n+2}$  und  $v_{5n+1}, v_{5n+2}$  vertauscht, kann ein Individuum  $y$  erzeugt werden, für das ebenfalls gilt  $\text{GBP}_G(y) = 4$  (s. Abbildung 3.14). Diese Mutation hat eine Wahrscheinlichkeit von

$$P(y \mid z) \geq \frac{1}{(3n)^2} \cdot \left(1 - \frac{1}{3n}\right)^{3n-2} \cdot \binom{3n}{2}^{-1} \geq \frac{2}{e(6n)^4},$$

da genau zwei Knoten aus  $L(x)$  und zwei aus  $R(x)$  ausgewählt werden müssen. Somit tritt eine solche Mutation in erwarteter Zeit  $\mathcal{O}(n^4)$  auf, falls keine andere Mutation durchgeführt wird, die akzeptiert wird. Den Fall einer anderen Mutation betrachten wir als nächstes.

Mit nur exponentiell kleiner Wahrscheinlichkeit werden in einer Mutation  $\Omega(n)$  Knoten mutiert. Gehen wir also davon aus, dass die Anzahl der mutierten Knoten in  $o(n)$  liegt. Eine Mutation, bei der Knoten aus den Cliquen  $A_1$  und  $A_2$  verändert werden, führt dann zu  $\Omega(n)$  neuen Schnittkanten. Eine solche Mutation wird daher

nicht akzeptiert. Sei  $M$  die Menge der Knoten, die von einer Mutation mutiert werden. Dann kann eine Mutation nur akzeptiert werden, wenn  $M \subset (B_1 \cup B_2)$  ist. Es dürfen also nur Knoten der Ringe mutiert werden.

Ist  $\{v_{2n+1}, v_{2n+2}, v_{5n+1}, v_{5n+2}\} \not\subset M$ , wird also mindestens einer dieser Knoten nicht mutiert, so führt dies zu einer Bisektion mit mehr als vier Kanten: Da nicht beide Ringe komplett mutiert werden, existieren für jeden Ring mindestens zwei Pfade zwischen  $B_i \cap M$  und  $B_i \setminus M$ . Wie eben gezeigt, können wir davon ausgehen, dass für eine akzeptierte Mutation keine Knoten aus  $A_1$  und  $A_2$  mutiert werden. Daher müssen aus  $B_1$  und  $B_2$  gleich viele Knoten mutiert werden. Ist also  $M \neq \emptyset$ , so schneiden von jedem Ring bereits zwei Kanten die Bisektion. Zusätzlich schneidet aber die zu  $v \in \{v_{2n+1}, v_{2n+2}, v_{5n+1}, v_{5n+2}\} \setminus M$  adjazente Verbindungskante den Schnitt. Damit also für ein erzeugtes  $x \neq z$  mit Hamming-Abstand  $H(z, x) = o(n)$  gilt  $\text{GBP}_G(x) \leq 4$  müssen alle Knoten  $v_{2n+1}, v_{2n+2}, v_{5n+1}, v_{5n+2}$  mutiert werden.

Wurde ein Individuum  $y$  erzeugt und akzeptiert, in welchem mindestens zwei Knoten der Ringe  $B_1$  und  $B_2$  auf der anderen Seite des Schnittes als der Rest des Ringes liegen, so erfolgt die weitere Analyse analog zur Analyse des Rings in Kapitel 3.2. Falls währenddessen irgendwann wieder das Individuum  $z$  erzeugt wird, so kann in Zeit  $\mathcal{O}(n^4)$  wieder ein  $y \neq z$  mit  $\text{GBP}_G(y) = 4$  erzeugt werden, das die oben beschriebenen Eigenschaften hat.

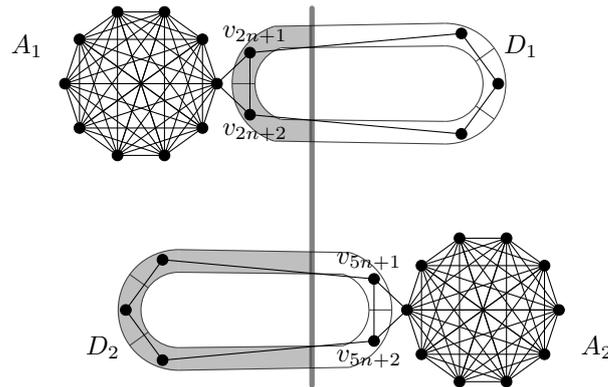


Abbildung 3.14: Eine solche Bisektion kann aus dem lokalen Optimum vom (1+1)-EA in Zeit  $\mathcal{O}(n^4)$  erreicht werden. Dann ist ein Random Walk möglich, das das Optimum erreichen kann.

Die Knotenmengen  $D_1 := B_1 \cap R(y)$  und  $D_2 := B_2 \cap L(y)$ , sowie der jeweilige Rest des Ringes können nun als Blöcke betrachtet werden (s. Abbildung 3.14). Somit existieren genau vier Blöcke, solange kein Individuum  $x$  mit  $\text{GBP}_G(x) < 4$  oder  $z$  erzeugt wurde. Aufgrund der Balance der Mengen  $L$  und  $R$  und der Feststellung, dass keine Knoten aus  $A_1$  und  $A_2$  mutiert werden, gilt stets  $|D_1| = |D_2|$ . Wir können also nur die Größe beider Blöcke um den gleichen Wert verändern. Die Mutationen, die Knoten an den Rändern der Blöcke mutieren und somit dessen Größe ändern bezeichnen wir wieder als *relevante* Mutationen.

Eine relevante Mutation, welche eine Änderung der Größe der Blöcke von mehr als zwei Knoten erreicht, muss an den Rändern von  $D_1$  mindestens drei Knoten mutieren und an den Rändern von  $D_2$  ebenso, da stets  $|D_1| = |D_2|$  gilt. In Zeit  $\mathcal{O}(n^4)$  wird somit

mit einer Fehlerwahrscheinlichkeit von  $\mathcal{O}(1/n^2)$  keine relevante Mutation durchgeführt, die die Größe der Blöcke  $D_1$  und  $D_2$  um mehr als 2 Knoten verändert. Wir gehen im Folgenden davon aus, dass keine solche Mutation erfolgt.

Solange die Blöcke  $D_1$  und  $D_2$  aus mehr als zwei Knoten bestehen ist analog zu Kapitel 3.2 die Wahrscheinlichkeit einer Abnahme der Blockgröße höchstens gleich der Wahrscheinlichkeit einer Zunahme (höchstens, da für das Individuum  $y$  nach Abbildung 3.14 die Wahrscheinlichkeit einer Zunahme von  $D_1$  bzw.  $D_2$  deutlich geringer ist). Mit Wahrscheinlichkeit mindestens  $1/2$  verkleinert somit die Hälfte der tatsächlich aufgetretenen relevanten Mutationen, die mehr als einen Knoten pro Block mutieren, die Blöcke  $D_1$  und  $D_2$ . Wir wollen nun davon ausgehen, dass dies der Fall ist.

Nach Lemma 1.5.4 ist die Wahrscheinlichkeit nach  $\mathcal{O}(n^2)$  relevanten 2-Zeichen-Mutationen die beiden Blöcke  $D_1$  und  $D_2$  auf Größe 3 zu schrumpfen durch eine Konstante nach unten beschränkt. Somit tritt das Ereignis nach einer erwarteten Anzahl von  $\mathcal{O}(n^2)$  dieser Mutationen auf. Die Wahrscheinlichkeit einer relevanten 2-Zeichen-Mutation ist nach unten beschränkt durch

$$12 \cdot \frac{1}{3n} \cdot \left(1 - \frac{1}{3n}\right)^{3n-1} \cdot \binom{3n}{1}^{-1} \geq \frac{12}{9en^2},$$

da aus  $L$  einer der vier zu den Schnittkanten adjazenten Knoten gewählt werden kann und mit einem der drei, nicht zur selben Schnittkante adjazenten Knoten aus  $R$  vertauscht werden kann. Somit gilt  $|D_1| = |D_2| = 3$  in erwarteter Zeit  $\mathcal{O}(n^4)$ .

Mit einer konstanten Wahrscheinlichkeit von mindestens  $1/(3e)^3$  sind dann die nächsten drei relevanten Mutationen die Blöcke  $D_1$  und  $D_2$  verkürzende 2-Zeichen-Mutationen. Somit wird in erwarteter Zeit  $\mathcal{O}(n^4)$  das Optimum von  $\text{GBP}_G$  erreicht.  $\square$

### Der global SEMO auf dem Clique-Ring-Graph

**Theorem 3.35.** *Der global SEMO benötigt eine erwartete exponentielle Laufzeit für  $\text{GBP}_G^0$  mit dem Clique-Ring-Graphen  $G = (V, E)$  der Größe  $|V| = 6n$ , falls zu einem Zeitpunkt  $\tau$  für die Population gilt  $\{z, z', z''\} \subseteq \mathbb{P}$  für*

- $z$  mit  $L(z) = A_1 \cup B_2$ ,  $R(z) = A_2 \cup B_1$ ,
- $z'$  mit  $z'_i = 0$  für ein  $i \in \{1, 3n+1\}$  und  $z'_j = z_j$  für alle  $j \neq i$ ; der Verbindungsknoten von  $A_1$  oder  $A_2$  ist also ein freier Knoten,
- $z''$  mit  $z''_1 = z''_{3n+1} = 0$  und  $z''_j = z_j$  für alle  $j \notin \{1, 3n+1\}$ ; es sind also beide Verbindungsknoten aus  $A_1$  und  $A_2$  freie Knoten.

*Beweis.* Der Beweis gliedert sich in drei Teile, die wir in den nachfolgenden Lemmata beweisen werden. Wir zeigen zunächst, dass die Individuen  $z'$  und  $z''$  fortlaufend in der Population enthalten sein werden, wobei  $z'$  von dem äquivalenten Individuum mit anderem  $i \in \{1, 3n+1\}$  für  $z'_i = 0$  ersetzt werden kann. Wir bezeichnen das Individuum  $z'$  als  $z'^1$ , falls  $i = 1$  ist und als  $z'^2$ , falls  $i = 3n+1$  ist. Ist von  $z'$  die Rede, so ist ein beliebiges dieser beiden Individuen gemeint. Dann kann einzig das Individuum  $z$  bzw. ein anderes Individuum  $x$  mit  $\text{zeros}(x) = 0$  wesentlich verändert werden. Mutationen bewegen ein solches  $x$  wie beim Beweis des (1+1)-EA auf dem Clique-Ring-Graph entsprechend einem Random Walk. Mit hoher Wahrscheinlichkeit wird allerdings aus  $z'$  regelmäßig wieder  $z$  mit  $L(z) = A_1 \cup B_2$ ,  $R(z) = A_2 \cup B_1$

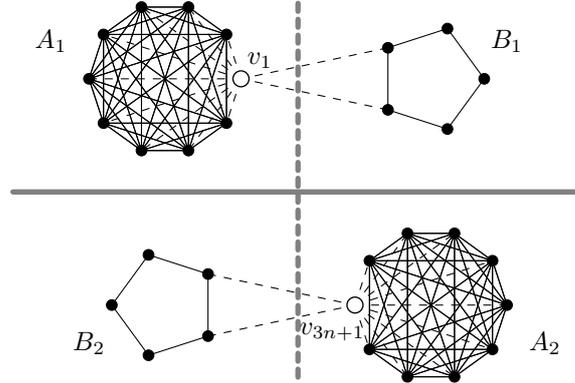


Abbildung 3.15: Durch Entfernen von einem der Knoten  $v_1, v_{3n+1}$  kann ein Individuum  $z'$  mit  $\text{GBP}_G^0(z') = (2, 1)$  erzeugt werden. Sind beide Knoten freie Knoten so ergibt sich das Individuum  $z''$  mit  $\text{GBP}_G^0(z'') = (0, 2)$ .

generiert, bevor ein Optimum erreicht werden kann. Somit wird eine mögliche erreichte Distanz im Suchraum wieder zunichte gemacht.

Es ist  $\text{GBP}_G^0(z) = (4, 0)$ ,  $\text{GBP}_G^0(z') = (2, 1)$  und  $\text{GBP}_G^0(z'') = (0, 2)$  (s. Abbildung 3.15), die Individuen dominieren sich demnach nicht gegenseitig. Das Individuum  $z''$  kann nur durch ein Individuum  $x$  mit gleichem Funktionswert  $f_0(x) = 0$  und  $\text{zeros}(x) \leq 2$  ersetzt werden. Jedes durch eine Mutation erzeugte Individuum  $y$  mit  $\text{zeros}(y) > 2$  wird in jedem Falle verworfen. Zum Zeitpunkt  $\tau$  gilt daher  $\{z, z', z''\} = \mathbb{P}$  und für alle Zeitpunkte  $t \geq \tau$  ist  $|\mathbb{P}| \leq 3$ .

Für das globale Optimum  $x^{\text{opt}}$  ist  $L(x^{\text{opt}}) = A_1 \cup B_1$  und  $R(x^{\text{opt}}) = A_2 \cup B_2$ . Die beiden Ringe  $B_1$  und  $B_2$  (oder die beiden Cliques  $A_1$  und  $A_2$ ) müssen also im Prozess des global SEMO, ausgehend von der Population  $\{z, z', z''\}$ , komplett die Schnittseite wechseln, damit das globale Optimum erreicht wird.

Mit dem folgenden Lemma zeigen wir, dass mit exponentiell kleiner Wahrscheinlichkeit ein Individuum erstellt wird, welches  $z''$  dominiert. Somit bleibt  $z''$  mit hoher Wahrscheinlichkeit durchgehend Bestandteil der Population.

**Lemma 3.36.** *Solange die Population  $\mathbb{P}$  aus dem Individuum  $z''$  und höchstens zwei beliebigen  $y', y$  besteht, wobei für den Hamming-Abstand der Individuen  $y'$  und  $y$  gilt  $H(y', z'') = o(n)$ ,  $H(y, z'') = o(n)$ , wird mit Wahrscheinlichkeit  $\exp(-\Omega(n))$  ein Individuum  $x \neq z''$  erzeugt, für das gilt  $x \preceq_{\text{GBP}_G^0} z''$ .*

*Beweis.* Mit Wahrscheinlichkeit  $\exp(-\Omega(n))$  werden  $\Omega(n)$  Zeichen mutiert. Somit ist mit Wahrscheinlichkeit  $1 - \exp(-\Omega(n))$  der Hamming-Abstand eines erzeugten Individuums  $H(x, z'') = o(n)$ , denn  $x$  kann nur durch Mutation aus  $z'', y'$  oder  $y$  erstellt werden. Weiterhin muss  $\text{zeros}(x) \leq 2$  gelten, damit  $x \preceq_{\text{GBP}_G^0} z''$  gelten kann. Für  $A_1$  und  $A_2$  gilt daher  $|A_1 \cap L(x)| = \Omega(n)$ ,  $|A_2 \cap R(x)| = \Omega(n)$  und für die Mengen  $B_1$  und  $B_2$  gilt  $B_1 \cap R(x) \neq \emptyset$  und  $B_2 \cap L(x) \neq \emptyset$ .

Falls sich Knoten der Cliques  $A_1$  oder  $A_2$  auf der anderen Schnittseite als der Rest der Cliques befindet, also  $A_1 \cap R(x) \neq \emptyset$  oder  $A_2 \cap L(x) \neq \emptyset$  gilt, so schneiden  $\Omega(n)$  Kanten der Cliques den Schnitt, da  $\text{zeros}(x) \leq 2$  ist. Für  $x \preceq_{\text{GBP}_G^0} z''$  muss daher

mit exponentiell nahe an 1 liegender Wahrscheinlichkeit  $|A_1 \cap L(x)| \geq 2n - 1$  und  $|A_2 \cap R(x)| \geq 2n - 1$  gelten.

Für Knoten aus  $B_1$  unterscheiden wir die folgenden drei Fälle:

1.  $|B_1 \cap L(x)| = 0$ : In diesem Fall existieren zwei Kanten von  $v_1$  zu den beiden Knoten  $v_{2n+1}, v_{2n+2} \in B_1$ , die den Schnitt kreuzen, falls alle drei Knoten eingeordnet sind. Für  $x \preceq_{\text{GBP}_G^0} z''$  muss gelten  $f_0(x) = 0$ . Somit darf entweder  $v_1$  keiner Menge zugeordnet werden, dann muss nur ein Knoten aus  $A_1 \cup B_1$  ein freier Knoten sein. Andernfalls muss für die beiden Knoten  $v_{2n+1}, v_{2n+2}$  gelten  $x_{2n+1} = x_{2n+2} = 0$ . Dann liegt die Anzahl der freien Knoten aus  $A_1 \cup B_1$  bei mindestens 2.
2.  $B_1 \cap L(x) = \{v\}$ : Falls sich genau ein Knoten  $v$  aus  $B_1$  auf der linken Seite der Bisektion befindet, so existieren zum einen zwei Kanten von  $v$  zu Knoten aus  $B_1$ . Weiterhin existiert mindestens eine Verbindungskante von  $v_1$  zu  $w \in B_1 \setminus \{v\}$ . Gilt  $f_0(x) = 0$ , so muss einerseits entweder  $v_1$  oder  $w$  ein freier Knoten sein, zusätzlich muss  $v$  oder einer der zu  $v$  adjazenten Knoten ein freier Knoten sein. Somit ist die Zahl der freien Knoten aus  $A_1 \cup B_1$  mindestens 2.
3.  $|B_1 \cap L(x)| \geq 2$ : Wenn sich mindestens zwei Knoten aus  $B_1$  auf der linken Seite des Schnittes befindet, so existieren in  $G$  mindestens zwei Pfade aus der Menge  $B_1 \cap L(x)$  in die Menge  $B_1 \cap R(x)$ . Da der Hamming-Abstand nur  $o(n)$  ist, ist in  $B_1 \cap R(x)$  mehr als ein Knoten enthalten. Demnach sind auch hier mindestens zwei freie Knoten aus  $A_1 \cup B_1$  notwendig, um diese Pfade zu unterbrechen, damit  $f_0(x) = 0$  gelten kann.

Die entsprechenden Fälle für  $B_2$  können analog behandelt werden. In jeder Menge  $A_i \cup B_i$ ,  $i = 1, 2$  ist mindestens ein freier Knoten notwendig, damit  $f_0(x) = 0$  gilt. In einem  $x$  mit  $\text{zeros}(x) \leq 2$  muss daher für beide Mengen der erste Fall vorliegen, wobei  $v_1 \in A_1$  und  $v_{3n+1} \in A_2$  die freien Knoten sind. Dann ist aber  $x = z''$ . In jedem anderen Fall kann  $z''$  von  $x$  weder schwach noch stark dominiert werden.  $\square$

Das folgende Lemma zeigt nun, dass auch das Individuum  $z'$  mit hoher Wahrscheinlichkeit ständig in der Population enthalten ist. Es wird dabei nur mit exponentiell geringer Wahrscheinlichkeit ein  $x$  mit  $x \preceq_{\text{GBP}_G^0} z'$  erzeugt, welches nicht  $z'^1$  oder  $z'^2$  entspricht.

**Lemma 3.37.** *Solange die Population  $\mathbb{P}$  aus dem Individuum  $z''$  und  $z'$ , sowie höchstens einem beliebigen  $y$  mit  $H(y, z) = o(n)$  besteht, so wird mit Wahrscheinlichkeit  $\exp(-\Omega(n))$  ein Individuum  $x \notin \{z'^1, z'^2\}$  erzeugt, für das gilt  $x \preceq_{\text{GBP}_G^0} z'^1$  bzw.  $x \preceq_{\text{GBP}_G^0} z'^2$ .*

*Beweis.* Mit Wahrscheinlichkeit  $1 - \exp(-\Omega(n))$  werden auch hier nur  $o(n)$  Zeichen mutiert. Es gilt  $H(z'', z) = 2$  und  $H(z', z) = 1$ . Somit ist der Hamming-Abstand gegeben durch  $H(x, z) = o(n)$  und es gilt  $B_1 \cap R(x) \neq \emptyset$  und  $B_2 \cap L(x) \neq \emptyset$ , mit der nur exponentiell geringen Fehlerwahrscheinlichkeit. Es ist außerdem  $|A_1 \cap L(x)| = \Omega(n)$  und  $|A_1 \cap R(x)| = \Omega(n)$ . Für ein Individuum  $x$  mit  $x \preceq_{\text{GBP}_G^0} z'$  ist  $f_0(x) \leq 2$ , also muss  $A_1 \cap R(x) = A_2 \cap L(x) = \emptyset$  gelten. Andernfalls liegen  $\Omega(n)$  Kanten der Cliques im Schnitt.

Für die Ringe  $B_1, B_2$  können wieder die drei Fälle aus Lemma 3.36 unterschieden werden. Ein Individuum  $x$  mit  $x \preceq_{\text{GBP}_G^0} z'$  kann diesmal nur höchstens einen freien

Knoten beinhalten, es dürfen jedoch bis zu zwei Kanten im Schnitt liegen. Wie bei der Unterscheidung der Fälle in Lemma 3.36 gezeigt, gibt es für jede der Mengen  $A_1 \cup B_1$  und  $A_2 \cup B_2$  ohne freie Knoten mindestens je zwei Schnittkanten. Damit  $f_0(x) \leq 2$  ist, muss also  $\text{zeros}(x) \geq 1$  sein und somit  $\text{zeros}(x) = 1$  für  $x \preceq_{\text{GBP}_G^0} z'$ . Durch das Entfernen eines Knotens aus  $L(x)$  oder  $R(x)$  müssen dann mindestens zwei Kanten aus der Bisektion genommen werden. Betrachten wir also, für welche der drei Fälle aus Lemma 3.36 dies möglich ist. Im Fall 2 kann durch Entfernen von  $v \in B_i$  der Wert der Bisektion um 2 gesenkt werden. Es bleibt jedoch mindestens eine Kante in der Menge  $A_i \cup B_i$  übrig und für ein solches Individuum ist  $f_0(x) \geq 3$ . Im dritten Fall kann mit einem freien Knoten nur eine Kante entfernt werden. Somit bleibt nur der erste Fall, wobei entweder  $v_1$  oder  $v_{3n+1}$  ein freier Knoten sein muss.

Für die Menge  $A_j \cup B_j, j \neq i$ , in der kein freier Knoten enthalten ist, dürfen dann nicht mehr als zwei Kanten den Schnitt kreuzen. Dies ist gegeben für die Fälle 1 und 3. Da wir jedoch davon ausgehen, dass  $A_1 \cap R(x) = A_2 \cap L(x) = \emptyset$  gilt, muss aufgrund der Balance von  $L(x)$  und  $R(x)$  gelten:  $|B_1 \cap L(x)| = |B_2 \cap R(x)|$ . Somit kann nicht gleichzeitig Fall 1 für  $B_i$  und Fall 3 für  $B_j, j \neq i$  eintreten und es muss auch für  $B_j$  Fall 1 vorliegen. Dann aber ist  $x \in \{z'^1, z'^2\}$ .  $\square$

Wir wissen nun, dass  $z''$  und  $z'$  mit Wahrscheinlichkeit  $1 - \exp(-\Omega(n))$  in der Population bleiben, bis ein Individuum  $x$  mit  $\text{zeros}(x) = 0$  und  $H(x, z) = \Omega(n)$  erzeugt und akzeptiert wurde. Nun widmen wir uns dem Random Walk der Individuen  $y$  mit  $\text{zeros}(y) = 0$  und zeigen, dass dieser mit großer Wahrscheinlichkeit eine so große Dauer benötigt, dass in der Zwischenzeit das Individuum  $z$  wieder erzeugt und akzeptiert wird. Dazu betrachten wir ein Zeitintervall der Länge  $T = n^{3/2}$  und zeigen, dass in dieser Zeit mit hoher Wahrscheinlichkeit kein  $x$  mit Hamming-Abstand  $H(z, x) = \Omega(n)$  erzeugt wird.

**Lemma 3.38.** *Gelte zu Beginn einer Phase von  $T = n^{3/2}$  Schritten, dass  $\{z, z', z''\} = \mathbb{P}$  ist. Dann ist mit Wahrscheinlichkeit  $1 - \exp(-\Omega(n^{1/2-\epsilon}))$  nach Zeit  $T$  für das Individuum  $y \in \mathbb{P}$  mit  $\text{zeros}(y) = 0$  der Hamming-Abstand zu  $z$  beschränkt durch  $H(y, z) = o(n)$ .*

*Beweis.* Wie in den vorherigen Lemmata gezeigt, bleiben  $z'$  und  $z''$  mit großer Wahrscheinlichkeit stets in der Population. Daher betrachten wir in diesem Beweis nur Individuen  $x$  mit  $\text{zeros}(x) = 0$ . Es existieren also keinerlei freie Knoten.

Die Individuen  $z, z'$  und  $z''$  haben paarweise Hamming-Abstände von höchstens 2. Für ein erstes, durch Mutation eines Individuums aus der Population  $\{z, z', z''\}$ , erzeugtes  $x \neq z$  ist damit  $H(x, z) = o(n)$  mit Wahrscheinlichkeit  $1 - \exp(-\Omega(n))$ . Dies gilt weiterhin, wenn statt  $z$  ein  $y \in \mathbb{P}$  mit  $\text{zeros}(y) = 0$  existiert, für das  $H(y, z) = o(n)$  gilt.

Ist dann ein  $v \in A_1$  in  $R(x)$ , so existieren  $\Omega(n)$  Kanten in die Menge  $L(x)$ , denn durch den beschränkten Hamming-Abstand ist  $|A_1 \cap L(x)| = \Omega(n)$ . Selbiges gilt für Knoten in  $A_2 \cap L(x)$ . In diesen Fällen ist also  $f_0(x) = \Omega(n)$  und  $x$  wird nicht akzeptiert. Also ist  $A_1 \subseteq L(x)$  und  $A_2 \subseteq R(x)$  für alle Individuen  $x \preceq_{\text{GBP}_G^0} z$  mit  $H(x, z) = o(n)$ . Der Hamming-Abstand kann demnach nicht durch Knoten aus  $A_1$  oder  $A_2$  erhöht werden und es bleiben nur die Knoten der Ringe  $B_1$  und  $B_2$ , welche in  $x$  eine andere Wertigkeit haben können als in  $z$ .

Für  $x$  mit  $H(x, z) = o(n)$  ist weiterhin  $B_1 \cap R(x) \neq \emptyset$  und  $B_2 \cap L(x) \neq \emptyset$  und es gibt mindestens zwei Pfade von Knoten aus  $(B_1 \cap L(x)) \cup \{v_1\}$  zu Knoten aus  $B_1 \cap R(x)$  und zwei Pfade von Knoten aus  $(B_2 \cap R(x)) \cup \{v_{3n+1}\}$  zu Knoten aus  $B_2 \cap L(x)$ . Damit  $x$  akzeptiert wird, dürfen nicht mehr als vier Kanten die Bisektion schneiden. Außerdem gilt für  $x \preceq_{\text{GBP}_G^0} z$  mit dem begrenzten Hamming-Abstand  $v_1 \in L(x)$  und  $v_{3n+1} \in R(x)$ . Somit muss es für jeden Ring genau zwei Pfade zwischen  $L(x)$  und  $R(x)$  geben, damit  $x$  akzeptiert wird.

Wir betrachten  $k := |B_1 \cap R(x)| = |B_2 \cap L(x)|$  als Maß für den Fortschritt. Die Gleichheit der Kardinalität der beiden Mengen ist durch die Balance von  $L(x)$  und  $R(x)$  gegeben und der Feststellung, dass wir  $A_1 \subseteq L(x)$  und  $A_2 \subseteq R(x)$  fordern können. Da die Menge  $B_1 \cap R(x)$  zusammenhängend ist, fügt jede Mutation von  $o(n)$  Knoten dieser Menge mindestens eine Kante zum Schnitt hinzu, die darin noch nicht enthalten war. Damit ein erzeugtes Individuum mit höherem  $k$  akzeptiert wird, muss deshalb mindestens eine der aktuellen Schnittkanten entfernt werden. Selbiges gilt für  $B_2$ . Aufgrund der Balance  $|B_1 \cap R(x)| = |B_2 \cap L(x)|$  müssen stets auf beiden Seiten Knoten mutiert werden, damit  $k$  erhöht wird. Daraus folgt, dass für eine akzeptierte Mutation, welche  $k$  erhöht, mindestens zwei der aktuellen Schnittkanten entfernt werden müssen. Dies ist nur möglich, wenn je einer der zu den Schnittkanten adjazenten Knoten mutiert wird.

Die Wahrscheinlichkeit  $k$  um mindestens 1 zu erhöhen ist damit beschränkt durch die Wahrscheinlichkeit  $x$  mit  $\text{zeros}(x) = 0$  aus der Population zu wählen und zwei dieser vier Knoten unterschiedlicher Seiten genau so zu mutieren, dass sie die Seite wechseln, also

$$\mathbb{P}(k^+) \leq \frac{1}{3} \cdot \frac{4}{(6n)^2} \cdot \frac{1}{2^2} = \frac{1}{108n^2}.$$

In Zeit  $T = n^{3/2}$  werden im Erwartungswert  $\mathbb{E}(X_T^+) = 1/(108n^{1/2})$  Mutation durchgeführt, welche  $k$  um mindestens 1 erhöhen. Nach der ersten Chernoff-Ungleichung aus Kapitel 1.5.2 werden mehr als  $(1 + \delta) \mathbb{E}(X_T^+)$  dieser Mutationen für ein  $\delta \geq 1$  mit Wahrscheinlichkeit höchstens  $\exp(-\mathbb{E}(X_T^+) \delta / 3)$  durchgeführt. Mit einem passenden  $\delta = \Theta(n^{1-\epsilon})$  für ein beliebig kleines  $\epsilon > 0$  gilt damit

$$\mathbb{P}(X_T^+ \geq n^{1/2-\epsilon}) \leq \exp(-\Omega(n^{1/2-\epsilon})).$$

Mit Wahrscheinlichkeit  $1 - \exp(-\Omega(n^{1/2-\epsilon}))$  werden also weniger als  $n^{1/2-\epsilon}$  solcher Mutation im Zeitintervall  $T$  durchgeführt.

Eine Mutation von  $\Omega(n^{1/2})$  Knoten tritt in polynomieller Zeit mit Wahrscheinlichkeit  $\exp(-\Omega(n^{1/2}))$  auf. Daher kann mit Wahrscheinlichkeit  $1 - \exp(n^{1/2-\epsilon})$  in Zeit  $T$  der Wert  $k$  um höchstens  $n^{1/2-\epsilon} \cdot n^{1/2} = n^{1-\epsilon}$  verändert werden. Es gilt somit  $k = o(n)$  nach Zeit  $T$  für das Individuum  $y \in \mathbb{P}$  mit  $\text{zeros}(y) = 0$  mit der angegebenen Wahrscheinlichkeit. Da damit nicht alle Knoten der Ringe die Seite gewechselt haben können, existieren weiterhin vier Kanten zwischen  $L(y)$  und  $R(y)$  und es gilt  $z \preceq_{\text{GBP}_G^0} y$ .  $\square$

Wir schließen den Beweis ab, indem wir zeigen, dass in dieser Zeit mit großer Wahrscheinlichkeit  $z$  erzeugt und akzeptiert wird. Aus  $z' \in \mathbb{P}$  mit  $z' \in \{z'^1, z'^2\}$  wird mit Wahrscheinlichkeit mindestens

$$\mathbb{P}(z \mid z') \geq \frac{1}{3} \cdot \frac{1}{6n} \cdot \frac{1}{2} \cdot \left(1 - \frac{1}{6n}\right)^{6n-1} \geq \frac{1}{36en}$$

das Individuum  $z$  mit  $\text{GBP}_G^0(z) = (4, 0)$  erzeugt, indem der freie Knoten von  $z'$  auf die einzig mögliche Seite einsortiert wird. In Zeit  $T = n^{3/2}$  werden im Erwartungswert mindestens  $\mathbb{E}(X'_T) = n^{1/2}/(36e)$  solcher Mutationen durchgeführt. Mit der zweiten Chernoff-Ungleichung aus Kapitel 1.5.2 erhalten wir eine obere Schranke für die Wahrscheinlichkeit, weniger als  $\mathbb{E}(X'_T)/2$  solcher Mutationen durchzuführen. Es ist

$$\mathbb{P}(X'_T \leq \frac{1}{2} \mathbb{E}(X'_T)) \leq \exp(-\frac{1}{8} \mathbb{E}(X'_T)) = -\exp(-\Omega(n^{1/2})),$$

falls  $z'$  nicht dominiert wird. Da in Zeit  $T$ , mit Wahrscheinlichkeit  $1 - \exp(-\Omega(n^{1/2-\epsilon}))$  kein  $x$  mit  $H(z, x) = \Omega(n)$  erzeugt und in  $\mathbb{P}$  aufgenommen wird, ist nach Lemma 3.37  $\{z'^1, z'^2\} \cap \mathbb{P} \neq \emptyset$ . Mit Wahrscheinlichkeit  $1 - \exp(-\Omega(n^{1/2}))$  wird in diesem Zeitintervall irgendwann das Individuum  $z$  erzeugt und akzeptiert und es sind wieder die Anfangsbedingungen gegeben.

Die erwartete Zeit bis eine der geforderten Aussagen nicht zutrifft ist nach unten beschränkt durch  $\exp(\Omega(n^{1/2-\epsilon}))$  für ein beliebig kleines  $\epsilon > 0$ .  $\square$



## 4 Experimente

Für die fünf Klassen von Graphen aus Kapitel 3 haben wir entweder nur untere oder nur obere Schranken für erwartete Laufzeiten vorgestellt. Für 2n-n-Cliquen-Graphen und Clique-Ring-Graphen haben wir sogar nur eine erwartete Laufzeit für den Fall gezeigt, dass der (1+1)-EA oder der global SEMO in einem bestimmten lokalen Optimum startet. Außerdem wurde für Ringe nur eine obere Schranke für den (1+1)-EA gezeigt, jedoch keine für den global SEMO.

Deshalb werden in diesem Kapitel Experimente mit dem (1+1)-EA bzw. dem global SEMO für jeden der fünf Graphen vorgestellt und somit das in Kapitel 3 gezeigte mit dem tatsächlichen Verhalten verglichen. Wie in den theoretischen Analyse betrachten wir auch hier nur den global SEMO auf der Funktion  $\text{GBP}_G^0$ .

### 4.1 Der Zwei-Cliquen-Graph

Die Resultate der Untersuchung des (1+1)-EA auf dem Zwei-Cliquen-Graph legen ein quadratisches Wachstum der Laufzeit in der Anzahl der Knoten des Graphen nahe. Abbildung 4.1 zeigt für Zwei-Cliquen-Graphen mit  $2 \leq |V| \leq 200$ , in Schritten von 2 Knoten, für 500 Läufe pro Graph-Instanz die durchschnittliche Laufzeit sowie die Standardabweichung. Zusätzlich ist die Funktion  $f(x) = 1,05x^2$  eingezeichnet, welche eine gute Übereinstimmung mit den experimentell bestimmten Werten aufweist. Die in Kapitel 3.1.2 bestimmte obere Schranke scheint also für die ausgewerteten Graphen scharf zu sein. Es ist jedoch eine große Streuung der Werte zu beobachten, so dass nur eine beschränkte Signifikanz der Auswertung vorhanden ist.

Der global SEMO hingegen kommt anscheinend mit einer geringeren Laufzeit aus als durch die obere Schranke in Kapitel 3.1.1 abgeschätzt. Zudem fällt die Standardabweichung wesentlich geringer aus als beim (1+1)-EA. Ebenso für Zwei-Cliquen-Graphen der Größe  $2 \leq |V| \leq 200$ , in Schritten von 2 Knoten, für 500 Läufe pro Graphinstanz, sind die Mittelwerte und die Standardabweichung in Abbildung 4.2 zu finden. Als eine gute Ausgleichskurve erweist sich die Funktion  $g(x) = ax^{1,8}$  mit  $a \approx 3,13$ . In den Experimenten zeigt der global SEMO somit interessanterweise eine deutlich bessere Laufzeit, als die in der Analyse in Kapitel 3.1.1 gezeigte obere Schranke von  $\mathcal{O}(n^2 \log n)$ . Diese starke Abweichung von der oberen Schranke mag daran liegen, dass wir die Populationsgröße in Teilen der Analyse durchgehend nur durch die maximale Populationsgröße von  $n$  abschätzen konnten. Diese deckt sich zumindest mit der durchschnittlichen maximalen Populationsgröße der Läufe in Abbildung 4.3. Die Werte der durchschnittlichen maximalen Populationsgröße eines Laufes scheinen sich für wachsendes  $n$  der Geraden  $x/4 + 4,8$  anzunähern.

Bei einem Vergleich des (1+1)-EA und des global SEMO zeigt sich, dass die Funktion  $g(x)$  sogar asymptotisch besser als  $f(x)$  ist. Damit könnten wir die Hoffnung verbinden, dass der global SEMO für sehr große Zwei-Cliquen-Graphen im Durchschnitt weniger

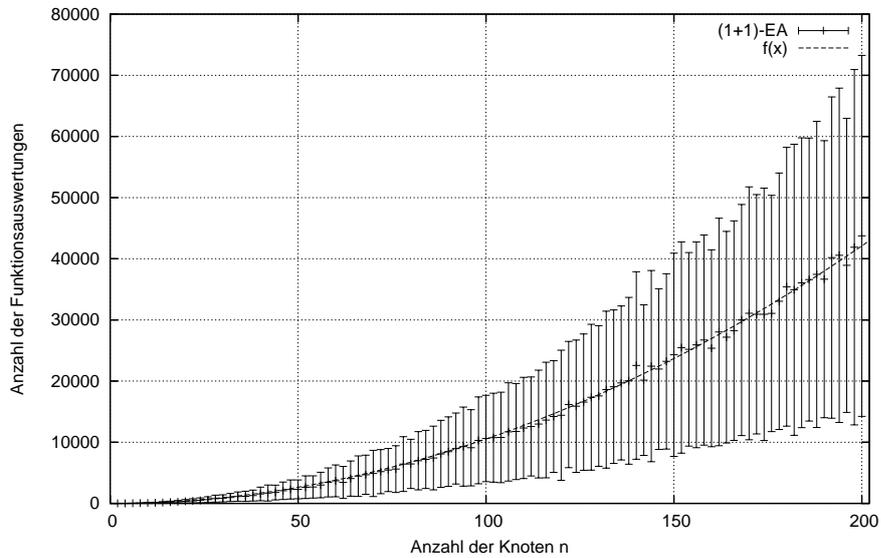


Abbildung 4.1: Durchschnittliche Laufzeit sowie Standardabweichung des (1+1)-EA auf Zwei-Cliquen-Graphen der Größe  $|V| = 2$  bis  $|V| = 200$ , in Schritten von 2 Knoten, gemittelt über 500 Läufe. Zusätzlich eingezeichnet:  $f(x) = 1,05x^2$ .

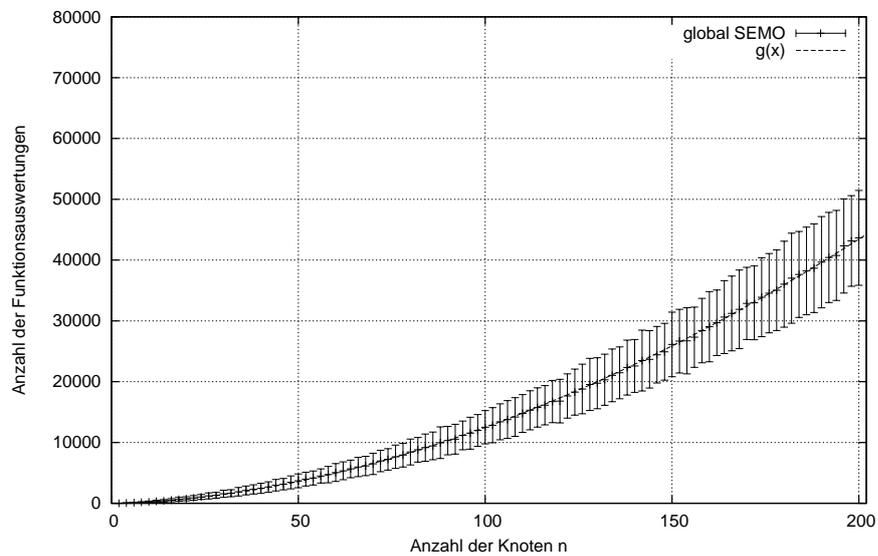


Abbildung 4.2: Durchschnittliche Laufzeit sowie Standardabweichung des global SEMO auf Zwei-Cliquen-Graphen der Größe  $|V| = 2$  bis  $|V| = 200$ , in Schritten von 2 Knoten, gemittelt über 500 Läufe. Zusätzlich eingezeichnet:  $g(x) = ax^{1,8}$ , mit  $a \approx 3,13$ .

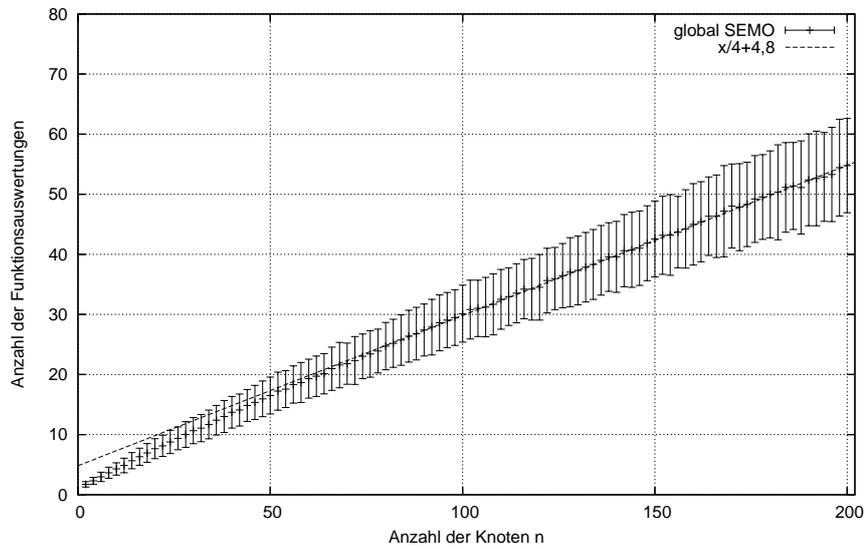


Abbildung 4.3: Durchschnittliche maximale Populationsgröße sowie Standardabweichung des global SEMO auf dem Zwei-Cliquen-Graph der Größe  $2 \leq |V| \leq 200$  gemittelt über 500 Läufe pro Instanz. Außerdem wurde die Funktion  $y = x/4 + 4,8$  eingezeichnet.

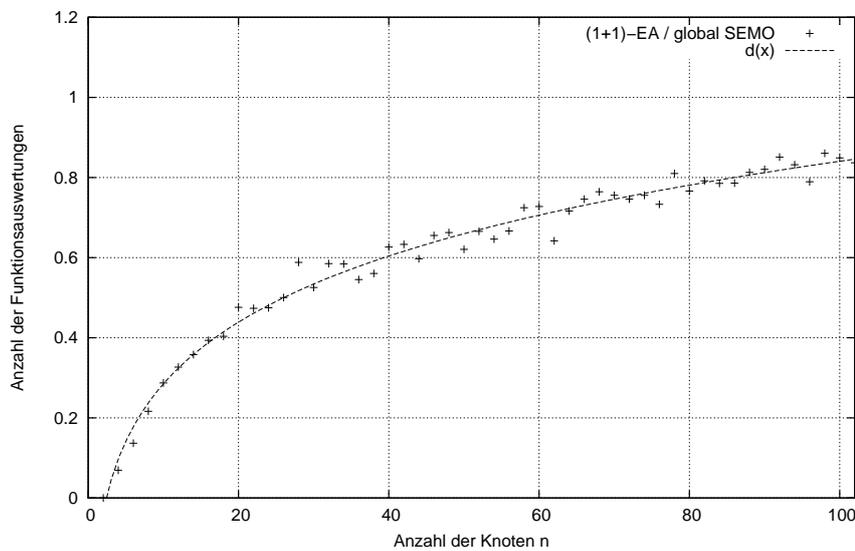


Abbildung 4.4: Das Verhältnis  $r(n) = (1+1)\text{-EA}(n)/\text{global SEMO}(n)$ , wobei  $(1+1)\text{-EA}(n)$  die durchschnittliche Laufzeit des  $(1+1)\text{-EA}$  und  $\text{global SEMO}(n)$  die des global SEMO für einen Zwei-Cliquen-Graphen der Größe  $|V| = n$  über 500 Läufe bezeichnen soll. Zusätzlich eingezeichnet:  $d(x) = ax^{0,2} + b$  mit  $a \approx 1,71$  und  $b \approx -1,86$ .

Funktionsauswertungen als der (1+1)-EA benötigt, um ein Optimum zu erzeugen. Für einen Vergleich der Laufzeiten des (1+1)-EA und des global SEMO wurde deshalb für jedes  $n = |V|$  das Verhältnis  $r(n)$  aus der durchschnittlichen Laufzeit des (1+1)-EA und der durchschnittlichen Laufzeit des global SEMO gebildet und in Abbildung 4.4 eingezeichnet. Die gemessenen Werte von  $r(n)$  sind kleiner als 1, da der (1+1)-EA in den Experimenten im Durchschnitt eine geringere Laufzeit als der global SEMO aufweist. Es ist allerdings ein Anstieg von  $r(n)$  zu beobachten, der für steigende  $n$  abflacht. Die Funktion  $d(x) = ax^{0.2} + b$  mit  $a \approx 1,71$  und  $b \approx -1,86$  zeigte eine gute Übereinstimmung mit dem Verlauf der Werte von  $r(n)$ . Dies entspricht auch asymptotisch dem Verhältnis der für die Messungen der beiden Algorithmen verwendeten Ausgleichskurven  $f(x)$  und  $g(x)$ . Diese Beobachtung liefert einen Hinweis darauf, dass der global SEMO für den Zwei-Cliquen-Graph eine asymptotisch bessere Laufzeit haben könnte, als der (1+1)-EA. Da die Messreihe den Wert 1 jedoch nicht überschreitet, kann nur spekuliert werden, ob der global SEMO für sehr große  $n$  tatsächlich eine bessere erwartete Laufzeit hat, als der (1+1)-EA.

## 4.2 Ringe

Für Ringe der Größe  $n$  zeigt sich, dass die obere Schranke der Analyse aus Kapitel 3.2 recht gut mit den gemessenen Durchschnittswerten übereinstimmt. Es gibt also Grund zur Annahme, dass die gezeigte obere Schranke von  $\mathcal{O}(n^4)$  scharf ist. Trotz der durch den Random-Walk zu erwartenden starken Varianz liegen die Mittelwerte

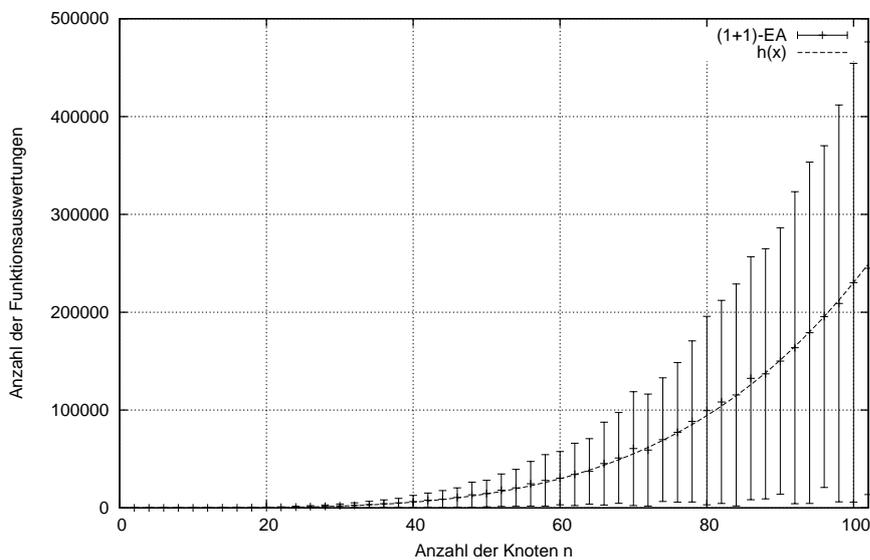


Abbildung 4.5: Durchschnittliche Laufzeit sowie Standardabweichung des (1+1)-EA auf Ringen der Größe  $|V| = 2$  bis  $|V| = 200$ , in Schritten von 2 Knoten, gemittelt über 500 Läufe. Zusätzlich eingezeichnet:  $h(x) = ax^4$  mit  $a \approx 0,0023$ .

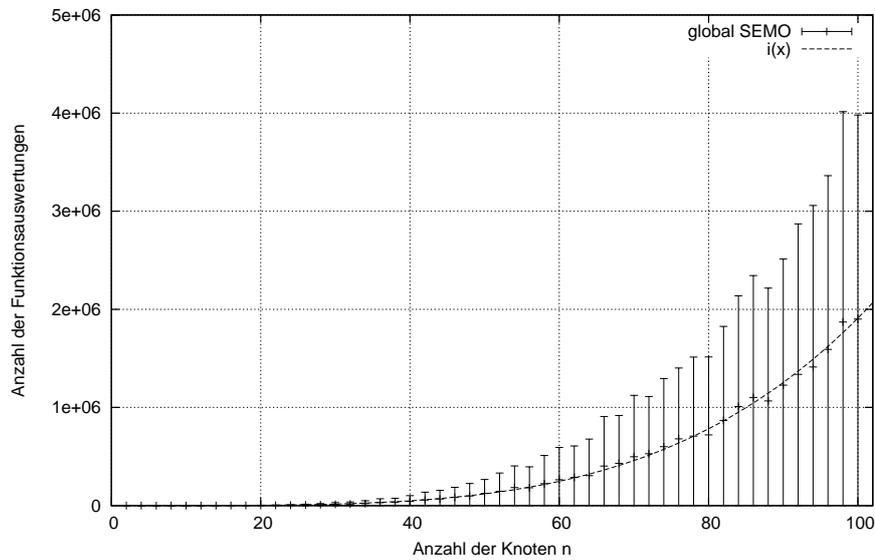


Abbildung 4.6: Durchschnittliche Laufzeit sowie Standardabweichung des global SEMO auf Ringen der Größe  $|V| = 2$  bis  $|V| = 200$ , in Schritten von 2 Knoten, gemittelt über 500 Läufe. Zusätzlich eingezeichnet:  $i(x) = ax^4$  mit  $a \approx 0,019$ .

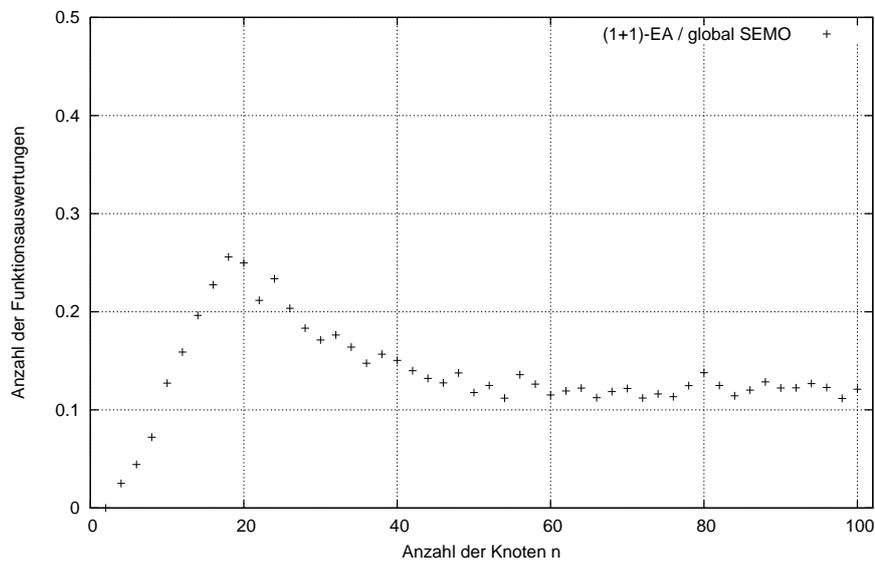


Abbildung 4.7: Das Verhältnis  $r(n) = (1+1)\text{-EA}(n) / \text{global SEMO}(n)$ , wobei  $(1+1)\text{-EA}(n)$  die durchschnittlichen Laufzeit des  $(1+1)\text{-EA}$  und  $\text{global SEMO}(n)$  die des global SEMO für einen Ring der Größe  $|V| = n$  über 500 Läufe bezeichnen soll.

der Messungen in guter Näherung auf der Funktion  $h(x) = ax^4$  für  $a \approx 0,0023$ . Die Datenerhebung erfolgte dabei über Ringe der Größe  $|V| = 2$  bis  $|V| = 100$  in Schritten von 2 Knoten. Wegen der großen Standardabweichung wurden für jede Graphinstanz 1000 Läufe durchgeführt. Die Mittelwerte der Laufzeiten sowie deren Standardabweichung ist in Abbildung 4.5 aufgetragen.

Ähnlich sieht der Graph für den global SEMO in Abbildung 4.6 aus. Für die gleichen Graphen, jedoch aufgrund der hohen Laufzeit nur für 500 Läufe pro Graphinstanz, wurde hier die Anzahl der Funktionsauswertungen gemessen. Die Resultate des global SEMO unterscheiden sich von denen des (1+1)-EA anscheinend nur in einem konstanten Faktor. Dieser Faktor liegt allerdings im Durchschnitt bei fast 10 und sorgt somit bei den Experimenten für eine merkbar längere Laufzeit des global SEMO. Die Funktion  $i(x) = ax^4$  mit  $a \approx 0,019$  zeigt eine gute Übereinstimmung mit den gemessenen Werten.

Wie vermutet, gestaltet sich ein Random Walk für den SEMO als schwerer als für den (1+1)-EA, da durch die vorhandene Population einzelne Änderungen leicht wieder „überschrieben“ werden können. Dennoch scheint die Komplexität beider Algorithmen gleich zu sein. Für einen Vergleich der durchschnittlichen Laufzeiten des (1+1)-EA mit denen des global SEMO wurde wieder das Verhältnis  $r(n)$  gebildet. Im Zähler von  $r(n)$  stehen dabei die durchschnittlichen Laufzeiten des (1+1)-EA für einen Ring mit  $n$  Knoten, im Nenner die des global SEMO. Das Ergebnis ist in Abbildung 4.7 zu finden. Die Werte von  $r(n)$  scheinen für größere  $n$  nahezu unverändert im Bereich um 0,12 zu liegen. Dies kann als weiteres Argument dafür betrachtet werden, dass die Komplexität des (1+1)-EA und die des global SEMO gleich sind.

### 4.3 Der Vier-Cliquen-Graph

Für den Vier-Cliquen-Graph wurden 500 Läufe pro Instanz der Graphen von  $|V| = 12$  bis  $|V| = 200$  durchgeführt. Da der Graph aus vier gleich großen Cliquen besteht, wurden die Graphinstanzen in Schritten von 4 Knoten erzeugt. Als Abbruchkriterium des (1+1)-EA wurde eine Fitness der Bisektion von weniger oder gleich 2 gewählt. Für  $|V| \geq 12$  hat ein lokales Optimum dann einen Wert von 2, während das globale Optimum eine Fitness von 0 aufweist. Dies gilt, da bei  $|V| \geq 12$  jede Clique mindestens drei Knoten enthält. Für ein Individuum, in dem mindestens ein Knoten einer Clique auf der anderen Seite des Schnittes als die Mehrheit der Knoten der Clique liegt, fallen mindestens Kosten  $|V|/4 - 1$  an. Gleichzeitig muss für einen Knoten einer anderen Clique das Gleiche gelten, damit die Balance von  $L$  und  $R$  gegeben ist. Somit hat ein solches Individuum eine Fitness von mindestens  $|V|/2 - 2 \geq 4$ .

Gemessen wurde dann die Anzahl der erfolgreichen Läufe, bei denen eine Fitness von 0 erreicht wurde, sowie die Anzahl der Läufe, die in einem lokalen Optimum mit Wert 2 endeten. In Abbildung 4.8 ist der Anteil der suboptimalen Läufe eingezeichnet. Man sieht eine recht deutliche Entwicklung, bei der für wachsendes  $|V|$  die Wahrscheinlichkeit in ein lokales Optimum zu fallen gegen etwa 2/3 geht. Dies deckt sich mit der theoretischen Analyse aus Kapitel 3.3.1 für die RLS und den (1+1)-EA mit veränderter Mutationswahrscheinlichkeit  $p_m < 2/|V|$ . Wie allerdings zu erwarten ist, ist auch beim (1+1)-EA für wachsendes  $n = |V|$  kein großer Einfluss der Verbindungskante mehr zu erkennen und jede Anordnung der Cliquen scheint mit nahezu gleicher Wahrscheinlichkeit erreicht zu werden. Dass für kleinere  $n$  die Wahrscheinlichkeit, das

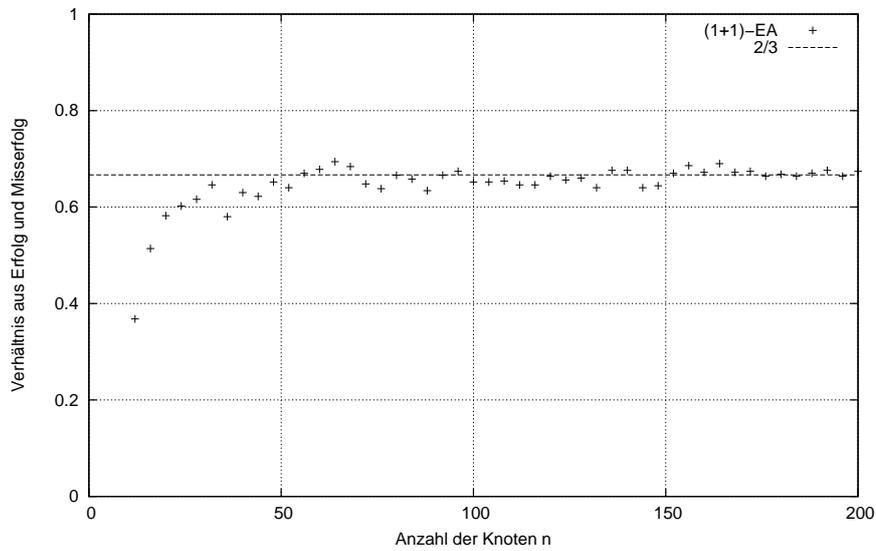


Abbildung 4.8: Der relative Anteil von 500 Läufen pro Graph, welcher in dem lokalen Optimum des Vier-Cliquen-Graph landet, für  $12 \leq n \leq 200$  in Schritten von 4 Knoten. Abgebrochen wurde, sobald der (1+1)-EA eine Bisektion mit Wert  $\leq 2$  gefunden hat. Für  $n \geq 12$  entspricht dies dem lokalen Optimum.

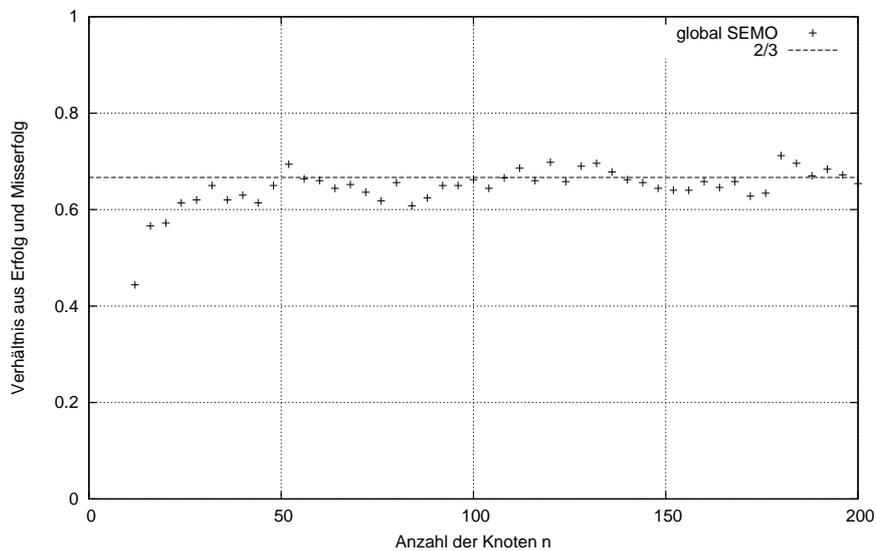


Abbildung 4.9: Der relative Anteil von 500 Läufen pro Graph, welcher in dem lokalen Optimum des Vier-Cliquen-Graph landet, für  $12 \leq n \leq 200$  in Schritten von 4 Knoten.

lokale Optimum mit Fitness 2 zu erreichen, geringer ist, ist auch zu erwarten, da dafür der Einfluss der beiden Verbindungskanten größer ist.

Ähnliche Resultate konnten auch für den global SEMO gezeigt werden. Ebenfalls für Vier-Cliquen-Graphen der Größen  $|V| = 12$  bis  $|V| = 200$ , in Schritten von 4 Knoten, wurden 500 Experimente pro Graphinstanz durchgeführt. Die erhaltenen Verhältnisse  $r(n)$  aus der gemessenen Anzahl der Läufe, die in dem lokalen Optimum mit Fitness 2 landeten und derer, die das globale Optimum erreichten, ist in Abbildung 4.9 zu finden. Auch hier sammeln sich die Werte von  $r(n)$  für  $n > 50$  dicht im Bereich um den Wert  $2/3$ .

#### 4.4 Der $2n$ - $n$ -Cliquen-Graph

Für den  $2n$ - $n$ -Cliquen-Graphen haben wir in Kapitel 3.3.2 gezeigt, dass der  $(1+1)$ -EA aus einem bestimmten lokalen Optimum nur mit erwarteter exponentieller Laufzeit entkommen kann. Im Experiment mit  $2n$ - $n$ -Cliquen-Graphen der Größen  $|V| = 6$  bis  $|V| = 198$  und 500 Läufen pro Instanz wurde gemessen, wie oft ein solches lokales Optimum vom  $(1+1)$ -EA erreicht wird. Aufgrund der Struktur des Graphen, welcher aus zwei Cliques  $A_1$  und  $A_2$  der Größe  $2n$  und zwei Cliques  $B_1$  und  $B_2$  der Größe  $n$  besteht, wurde die Größe der  $2n$ - $n$ -Cliquen-Graphen in Schritten von 6 Knoten erhöht. Als Abbruchkriterium wurde zunächst ein Schnitt mit Wert 0 gefordert oder die Aufteilung,  $L = A_1 \cup B_2$ ,  $R = A_2 \cup B_1$  bzw. deren Spiegelbild.

Überraschenderweise zeigte sich in diesem Experiment, dass noch ein weiteres suboptimales lokales Optimum existiert, für das der  $(1+1)$ -EA sehr lange Rechenzeit benötigt. Bei diesem befinden sich die beiden Cliques  $B_1$  und  $B_2$  auf der selben Seite, während eine der Cliques  $A_1$  oder  $A_2$  vollständig der anderen Seite zugeteilt ist. Die verbleibende Clique hat dann Knoten sowohl in  $L$ , als auch in  $R$  (s. Abbildung 4.10). Man kann sich leicht überlegen, dass der  $(1+1)$ -EA auch für dieses lokale Optimum

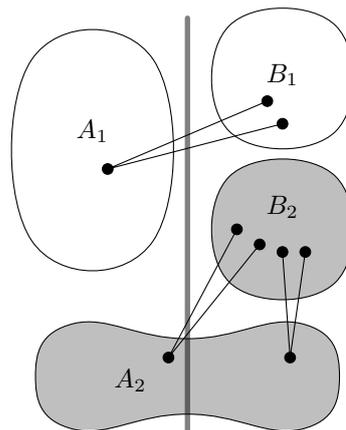


Abbildung 4.10: Schematische Darstellung des unerwartet eingetretenen Suboptimums des  $2n$ - $n$ -Cliquen-Graphen.

exponentielle Rechenzeit benötigt. Dies wollen wir nur kurz an einem Beispiel erörtern. Dabei verwenden wir die Einteilung der Cliques aus Abbildung 4.10. Werden  $k$  Knoten aus der Clique  $B_1$  auf die andere Schnittseite verschoben, schneiden damit  $k(n-k)$  Kanten der Clique die Bisektion. Gleichzeitig liegen  $2k$  Verbindungskanten weniger im Schnitt. Wenn dann  $k$  Knoten aus der über beide Seiten verteilten Clique  $A_2$  die Seite wechseln, so schneiden aus dieser Clique  $(n-k)(n+k)$  Kanten, anstatt der vorher  $n^2$  Kanten. Auch hier werden  $2k$  Kanten aus dem Schnitt entfernt. Somit ergibt sich eine Änderung des Schnittes von

$$k(n-k) - 2k + (n-k)(n+k) - n^2 - 2k = k(n-k) - 4k - k^2 = k(n-2k-4).$$

Diese Änderung wird nur akzeptiert, wenn  $k = 0$  oder  $k \geq n/2 - 2$  gilt.

Daher wurde das Abbruchkriterium erweitert, so dass auch beim Erzeugen eines solchen lokalen Optimums der Lauf beendet wurde. Für die beiden auftretenden lokalen Optima mit Wert größer 0 wurde dann für jede Instanz des 2n-n-Cliquen-Graph der Anteil der Läufe, die mit einem solchen beendet wurden in Abbildung 4.11 eingetragen. Für wachsendes  $|V|$  nimmt dabei der Anteil der neu entdeckten lokalen Optima ab, während der Anteil der Läufe, die in dem analysierten lokalen Optimum endeten, zunimmt und sich bei etwa 0,3 anhäuft. Dies spricht für eine konstante Wahrscheinlichkeit des Auftretens dieses Falles.

In den Experimenten des global SEMO auf dem 2n-n-Cliquen-Graph wurde dagegen

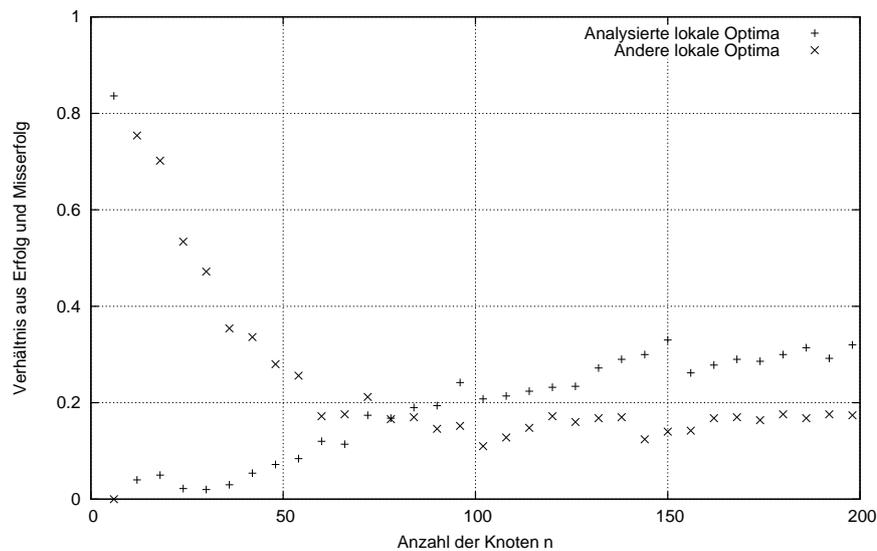


Abbildung 4.11: Relative Häufigkeit des Auftretens eines lokalen Optimums bei der Optimierung der Bisektion eines 2n-n-Cliquen-Graphen durch den (1+1)-EA. Es wurde dabei unterschieden, ob das untersuchte lokale Optimum aus Kapitel 3.3.2 (analysierte lokale Optima) oder das hier beschriebene (andere lokale Optima) erzeugt wurde. Die Erhebung erfolgt für 2n-n-Cliquen-Graphen der Größe  $|V| = 6$  bis  $|V| = 198$  in Schritten von 6 Knoten und 500 Läufen pro Graphinstanz.

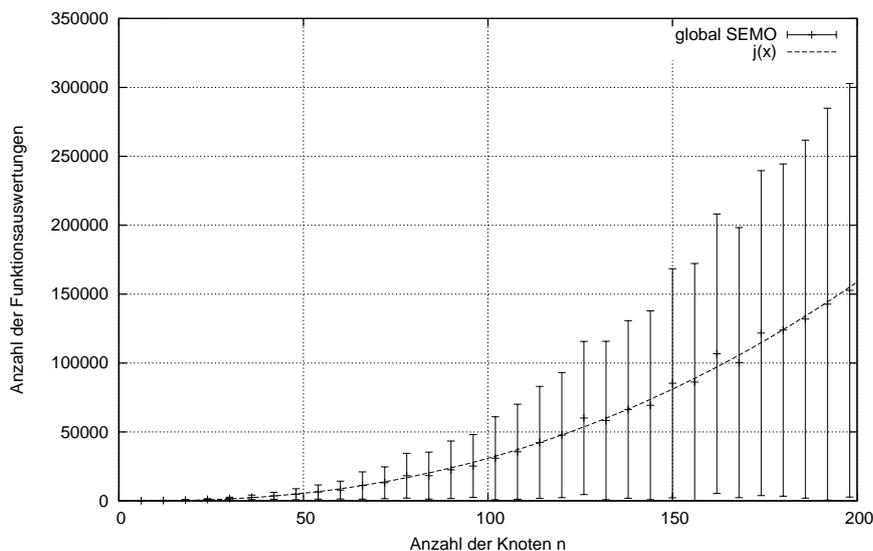


Abbildung 4.12: Durchschnittliche Laufzeit sowie Standardabweichung des global SEMO auf  $2n$ - $n$ -Cliquen-Graphen der Größe  $|V| = 6$  bis  $|V| = 198$ , in Schritten von 6 Knoten, gemittelt über 500 Läufe. Zusätzlich eingezeichnet:  $j(x) = ax^2 \log(x) - bx^2$  mit  $a \approx 1,27$  und  $b \approx 2,75$

die Laufzeit gemessen, bis ein optimales Individuum  $x^{\text{opt}}$  mit  $\text{GBP}_G^0(x^{\text{opt}}) = (0, 0)$  erzeugt wurde. Mit ähnlichen Überlegungen wie zur Analyse des global SEMO in Kapitel 3.3.2, kann man auch für das neu entdeckte lokale Optimum zu dem Schluss kommen, dass der global SEMO aus diesem in polynomieller Zeit ein  $x^{\text{opt}}$  erzeugen kann. Dies spiegelt sich auch in den Experimenten wieder. Für  $|V| = 6$  bis  $|V| = 198$  in Schritten von 6 Knoten wurde in 500 Läufen die benötigte Zeit gemessen, bis eine optimale Bisektion erzeugt wurde. Das Ergebnis findet sich in Abbildung 4.12. Als Ausgleichsfunktion liefert die Funktion  $j(x) = ax^2 \log(x) - bx^2$  mit  $a \approx 1,27$  und  $b \approx 2,75$  eine gute Näherung an die durchschnittlichen gemessenen Laufzeiten. Es scheint demnach, als könnte der global SEMO die lokalen Optima des  $2n$ - $n$ -Cliquen-Graph gut überwinden und das globale Optimum erreichen. Weiterhin ist die aufgrund der Experimente vermutete Komplexität von  $\Theta(x^2 \log(x))$  deutlich geringer als die gezeigte Schranke von  $\mathcal{O}(n^4)$ . Eine Erklärung könnte sein, dass der Fall 3 b) aus der Analyse in Kapitel 3.3.2 in den durchgeführten Experimenten entweder nicht hinreichend oft vorkommt oder dieser Fall bei einem Auftreten nicht besonders lange anhält.

## 4.5 Der Clique-Ring-Graph

Ähnlich den Experimenten des  $2n$ - $n$ -Cliquen-Graphen wurden auch für den Clique-Ring-Graphen für den einen Algorithmus die Laufzeiten ermittelt, während für den anderen die Häufigkeiten eines lokalen Optimums bestimmt wurden. Da das Verhalten für diesen Graphen umgekehrt ist und der (1+1)-EA in der Lage ist dem lokalen

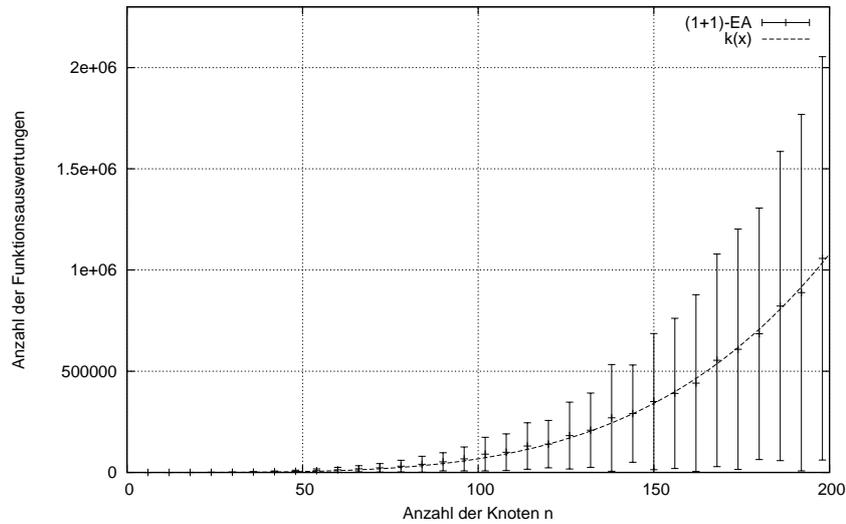


Abbildung 4.13: Durchschnittliche Laufzeit sowie Standardabweichung des (1+1)-EA auf Clique-Ring-Graphen der Größe  $|V| = 12$  bis  $|V| = 198$ , in Schritten von 6 Knoten, gemittelt über 500 Läufe. Zusätzlich eingezeichnet:  $k(x) = ax^4$  mit  $a \approx 6,7 \cdot 10^{-4}$

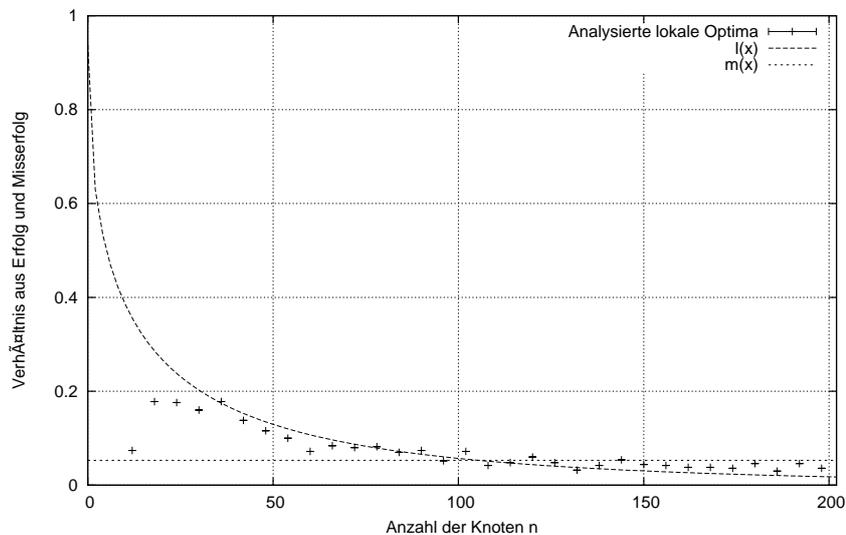


Abbildung 4.14: Relative Häufigkeit mit der die in Theorem 3.35 geforderte Population bei der Optimierung von  $\text{GBP}_G^0$  eines  $2n$ - $n$ -Cliques-Graphen  $G$  durch den global SEMO. Die Erhebung erfolgt für Clique-Ring-Graphen der Größe  $|V| = 6$  bis  $|V| = 198$  in Schritten von 6 Knoten und 500 Läufen pro Graphinstanz. Zusätzlich eingezeichnet:  $l(x) = a \exp(-0,28 \cdot \sqrt{x})$  für  $a \approx 0,94$  und  $m(x) = b$  für  $b \approx 0,053$

Optimum, in dem die Verbindungskanten im Schnitt liegen, zu entfliehen, wurden für den (1+1)-EA die Laufzeiten bis zum Erzeugen eines Optimums gemessen. Als Abbruchkriterium wurde entsprechend eine Bisektion mit Wert 0 gefordert. Die gemessenen Durchschnittslaufzeiten mit deren Standardabweichung für Clique-Ring-Graphen der Größen  $|V| = 12$  bis  $|V| = 198$  von je 500 Läufen sind in Abbildung 4.13 zu sehen. Da der Graph aus zwei Cliques der Größe  $2n$  und zwei Ringen der Größe  $n$  besteht, wurde die Größe der Graphen in Schritten von 6 Knoten erhöht. Außerdem gibt es zwei Verbindungskanten zu jedem Ring. Daher müssen die Ringe aus mindestens zwei Knoten bestehen. Also hat der kleinste mögliche  $2n$ - $n$ -Clique-Graph  $|V| = 12$  Knoten. Als Ausgleichskurve in Abbildung 4.13 wurde  $k(x) = ax^4$  mit  $a \approx 6,7 \cdot 10^{-4}$  verwendet. Dies entspricht der gezeigten oberen Schranke für das lokale Optimum aus Kapitel 3.3.3. Außerdem stimmt diese durch die Experimente induzierte Vermutung einer Laufzeit von  $\Theta(n^4)$  mit den Experimenten zu Ringen überein. In diese Graphen müssen neben den beiden Cliques zwei Ringe vollständig einer Seite zugeordnet werden. Dabei können die Knoten der Ringe zu Beginn völlig zufällig auf beide Schnittseiten verteilt sein. Sind die Cliques also zu einem Zeitpunkt während der Optimierung vollständig einer Seite zugeordnet, so ergibt sich ein Random Walk wie in der Analyse des Ringes in Kapitel 3.2 und der des lokalen Optimum für diese Graphen in Kapitel 3.3.3.

Für den global SEMO wurde in den je 500 Läufen für Instanzen des Clique-Ring-Graphen der Größe  $|V| = 12$  bis  $|V| = 198$  gemessen, wie oft eine Population entsprechend Theorem 3.35 aus Kapitel 3.3.3 erstellt wurde. Als Abbruchkriterium dient also die Existenz einer solchen Population oder das Erzeugen einer Bisektion mit Wert 0. Der relative Anteil der Läufe, welche in der Population des lokalen Optimums mit 4 Schnittkanten landete, ist in Abbildung 4.14 eingezeichnet.

Leider kann aus der Grafik keine zufriedenstellende Aussage über die Wahrscheinlichkeit einer solchen suboptimalen Population gewonnen werden. Die eingezeichneten Kurven  $l(x) = a \exp(-0,28 \cdot \sqrt{x})$  mit  $a \approx 0,94$  und  $m(x) = b$  für  $b \approx 0,053$  liefern beiden keine zufriedenstellende Näherung. Die Parameter  $a$  und  $b$  wurden dabei mittels Ausgleichsrechnung durch die Methode kleinster Quadrate für die Werte mit  $n \geq 60$  bestimmt. Die Funktion  $l(x)$  scheint schneller gegen 0 zu konvergieren als der Verlauf der Punkte. Falls dem so ist hätte der global SEMO eine erwartete exponentielle Laufzeit, da die Wahrscheinlichkeit eines solchen lokalen Optimums geringer ist als die erwartete Laufzeit ausgehend von dem lokalen Optimum. Dies sind jedoch nur Spekulationen, da die Menge und Genauigkeit der erhobenen Daten nicht ausreichend ist, um einen genauen Verlauf zu bestimmen. Die Quadratsumme der Residuen für die Werte mit  $n \geq 60$  ergibt für  $l(x)$  den Wert  $5,3 \cdot 10^{-3}$ , für  $m(x)$  den Wert  $6,6 \cdot 10^{-3}$ . Die durch Ausgleichsrechnung gefundene Konstante  $b$  scheint somit keine bessere Abschätzung zu bieten.

## 5 Zusammenfassung

Wir haben in dieser Arbeit drei unterschiedliche Verfahren zur Optimierung der Bisektion eines Graphen untersucht. Während die RLS und der (1+1)-EA mit einer binären Kodierung arbeiten, welche eine beliebige vollständige Bisektion der Knotenmenge  $V$  des Graphen beschreibt, wird für den global SEMO ein erweiterter Suchraum verwendet. In dieser bezeichnet eine ternär kodierte Zeichenkette eine partielle Einteilung der Knoten, wobei Knoten existieren dürfen, die noch keiner der Mengen der Bisektion zugeteilt sind. In einer Mutation des (1+1)-EA werden im Erwartungswert zwei Zeichen der Repräsentation verändert, bei dem global SEMO dagegen nur eines. Für die Fitness eines Individuums des global SEMO stehen uns mehrere untere Schranken aus Kapitel 2.1 zur Verfügung.

Für einen Zwei-Cliquen-Graphen  $G$ , der aus zwei Cliques besteht, welche nur durch eine Kante verbunden sind, wurde für den global SEMO gezeigt, dass die Pareto-Front für  $\text{GBP}_G^3$  nur in einer erwarteten exponentiellen Zeit in der Anzahl der Knoten erreicht wird. Dagegen konnte für die Funktion  $\text{GBP}_G^0$  eine obere Schranke für die Zeit, bis die Pareto-Front erreicht wurde, von  $\mathcal{O}(n^2 \log(n))$  mit der Anzahl der Knoten  $n$  gezeigt werden. Weiterhin wurde für den (1+1)-EA auf einem Zwei-Cliquen-Graphen eine obere Schranke von  $\mathcal{O}(n^2)$  für die Dauer gezeigt, bis eine optimale Bisektion erzeugt wird.

Für Ringe wurde nur für den (1+1)-EA eine obere Schranke der Laufzeit von  $\mathcal{O}(n^4)$  gezeigt. Diese ergibt sich durch Random Walks, welche das Individuum des (1+1)-EA auf den Plateaus der Fitnessfunktion vollführen muss.

Für Vier-Cliquen-Graphen, die aus genau zwei nicht zusammenhängenden Zwei-Cliquen-Graphen bestehen, wurde wieder für beide Algorithmen eine Schranke für die erwartete Laufzeit bewiesen. Diesmal handelt es sich allerdings um eine untere Schranke, da für diesen Graph ein lokales Optimum existiert, von welchem das globale Optimum nur durch Mutation von linear vielen Zeichen erreicht werden kann. Eine genaue Analyse wurde dabei für die RLS geführt. Diese Analyse wurde auf den (1+1)-EA mit veränderter Mutationswahrscheinlichkeit  $p < 2/|V|$  übertragen. Für den (1+1)-EA mit Mutationswahrscheinlichkeit  $p_m = 2/|V|$  wurde dann eine erwartete exponentielle Laufzeit für eine bestimmte Menge von initialen Individuen gezeigt. Auf den global SEMO konnte dies nur für eine Menge von zwei initialen Individuen übertragen werden.

Zum Schluss der theoretischen Analysen wurden zwei Graphen vorgestellt, auf denen entweder der (1+1)-EA oder der global SEMO Schwierigkeiten mit lokalen Optima haben, während der jeweils andere Algorithmus in erwarteter polynomieller Zeit, von einem bestimmten lokalen Optimum ausgehend, das globale Optimum erreicht. So wurde für den  $2n$ - $n$ -Cliquen-Graphen gezeigt, dass der (1+1)-EA eine erwartete exponentielle Laufzeit hat, wenn während der Optimierung ein bestimmtes, lokal optimales Individuum erzeugt wurde. Der global SEMO kann jedoch, wenn ein

solches Individuum in der Population enthalten ist, das globale Optimum in erwarteter Laufzeit von  $\mathcal{O}(n^4)$  erreichen. Für den Clique-Ring-Graphen zeigte sich dagegen ein umgekehrtes Verhalten. Erzeugt der (1+1)-EA ein bestimmtes Individuum, so wird in erwarteter Zeit  $\mathcal{O}(n^4)$  eine optimale Bisektion erzeugt. Dagegen benötigt der global SEMO eine erwartete exponentielle Laufzeit um eine optimale Bisektion zu erzeugen, wenn zu einem Zeitpunkt der Optimierung eine bestimmte Population vorliegt.

Abschließend wurden Experimente zu jeder theoretischen Analyse durchgeführt, um die gezeigten Aussagen in der Praxis auf deren Relevanz und Genauigkeit hin zu überprüfen. Dabei zeigte sich, dass die Laufzeit des global SEMO im Durchschnitt eine Abhängigkeit von der Größe des Graphen hat, die geringer zu sein scheint als die gezeigte obere Schranke. Die gemessenen Laufzeiten des (1+1)-EA zeigen dagegen eine gute Übereinstimmung mit der oberen Schranke aus dem Theoriekapitel. Ebenso bestätigen auch Experimente des (1+1)-EA auf Ringen die dafür gezeigte obere Schranke. Interessanterweise schafft auch der global SEMO mit seiner Population einen Random Walk auf dem Ring mit scheinbar gleicher Komplexität. Für den Vier-Cliquen-Graphen konnte gezeigt werden, dass beide Algorithmen mit hoher Wahrscheinlichkeit ein lokales Optimum mit zwei Schnittkanten erreichen. Aus diesem ist ein Entkommen nur in erwarteter exponentieller Zeit möglich. Weiterhin wurden für den  $2n$ - $n$ -Cliquen-Graphen experimentell Belege dafür gefunden, dass der global SEMO eine erwartete Laufzeit von  $\Theta(n^2 \log(n))$  hat um das globale Optimum zu erzeugen, während der (1+1)-EA mit vermuteter konstanter Wahrscheinlichkeit eine exponentielle Rechenzeit benötigt. Dabei konnte außerdem ein in der theoretischen Analyse unbeachtetes lokales Optimum entdeckt werden. Dies zeigt, dass auch für eine im wesentlichen theoretische Arbeit eine praktische Auswertung der Ergebnisse durchaus sinnvoll sein kann. Zuletzt wurden Werte für den (1+1)-EA für den Clique-Ring-Graphen gemessen, welche eine erwartete Laufzeit von  $\Theta(n^4)$  nahelegen. Für den global SEMO auf diesem Graphen konnten jedoch keine überzeugenden Hinweise auf die Wahrscheinlichkeit gefunden werden, mit der eine Population erzeugt wird, für die der global SEMO erwartete exponentielle Laufzeit benötigt.

## 5.1 Ausblick

Für die gezeigten oberen Schranken ist es sicherlich auch interessant, untere Schranken zu finden, um einen festen Rahmen für die erwarteten Laufzeiten zu bestimmen. Dazu wäre vor allem eine untere Schranke für die Größe der Population des global SEMO für viele Analysen von Interesse.

Weiterhin existieren einige Resultate zu anderen randomisierten Suchheuristiken für Zufallsgraphen, wie in Kapitel 1.1.1 erwähnt. Leider ließen sich die Ergebnisse dieser Arbeiten nicht mit einfachen Mitteln auf evolutionäre Algorithmen übertragen. Eine Analyse in dieser Richtung ist aber sicherlich wünschenswert.

# Literaturverzeichnis

- [1] David Aldous and Umesh V. Vazirani. “Go with the winners” algorithms. In *IEEE Symposium on Foundations of Computer Science*, pages 492–501, 1994.
- [2] T. Bui, C. Heigham, C. Jones, and T. Leighton. Improving the performance of the Kernighan-Lin and simulated annealing graph bisection algorithms. In *DAC '89: Proceedings of the 26th ACM/IEEE Conference on Design automation*, pages 775–778, New York, NY, USA, 1989. ACM.
- [3] T. N. Bui. On bisecting random graphs. Technical Report MIT/LCS/TR-287, Massachusetts Institute of Technology, Cambridge, MA, USA, 1983.
- [4] Thang Nguyen Bui and Curt Jones. Finding good approximate vertex and edge partitions is NP-hard. *Inf. Process. Lett.*, 42(3):153–159, 1992.
- [5] Thang Nguyen Bui, F. Thomson Leighton, Soma Chaudhuri, and Michael Sipser. Graph bisection algorithms with good average case behavior. *Combinatorica*, 7(2):171–191, 1987.
- [6] Thang Nguyen Bui and Byung Ro Moon. Genetic algorithm and graph partitioning. *IEEE Transactions on Computers*, 45(7):841–855, 1996.
- [7] Ted Carson and Russell Impagliazzo. Hill-climbing finds random planted bisections. In *SODA '01: Proceedings of the Twelfth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 903–909, Philadelphia, PA, USA, 2001. Society for Industrial and Applied Mathematics.
- [8] Tassos Dimitriou and Russell Impagliazzo. Go with the winners for graph bisection. In *SODA '98: Proceedings of the Ninth Annual ACM-SIAM symposium on Discrete Algorithms*, pages 510–520, Philadelphia, PA, USA, 1998. Society for Industrial and Applied Mathematics.
- [9] William Feller. *An Introduction to Probability Theory and Its Applications*, volume 1. Wiley, January 1968.
- [10] Ariel Felner. Finding optimal solutions to the graph partitioning problem with heuristic search. *Annals of Mathematics and Artificial Intelligence*, 45(3-4):293–322, 2005.
- [11] M. R. Garey, D. S. Johnson, and L. Stockmeyer. Some simplified NP-complete problems. In *STOC '74: Proceedings of the Sixth Annual ACM Symposium on Theory of Computing*, pages 47–63, New York, NY, USA, 1974. ACM.
- [12] Oliver Giel. Runtime analysis of a simple multi-objective evolutionary algorithm. In Jürgen Branke, Kalyanmoy Deb, Kaisa Miettinen, and Ralph E. Steuer, editors, *Practical Approaches to Multi-Objective Optimization*, number 04461 in Dagstuhl Seminar Proceedings. Internationales Begegnungs- und Forschungszentrum fuer Informatik (IBFI), Schloss Dagstuhl, Germany, 2005.

- 
- [13] Torben Hagerup and C. Rüb. A guided tour of Chernoff bounds. *Inf. Process. Lett.*, 33(6):305–308, 1990.
  - [14] Mark Jerrum and Gregory B. Sorkin. Simulated annealing for graph bisection. In *IEEE - Symposium on Foundations of Computer Science*, pages 94–103, 1993.
  - [15] David S. Johnson, Cecilia R. Aragon, Lyle A. McGeoch, and Catherine Schevon. Optimization by simulated annealing: an experimental evaluation. Part i, graph partitioning. *Oper. Res.*, 37(6):865–892, 1989.
  - [16] B. W. Kernighan and S. Lin. An efficient heuristic procedure for partitioning graphs. *The Bell system technical journal*, 49(1):291–307, 1970.
  - [17] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science, Number 4598, 13 May 1983*, 220, 4598:671–680, 1983.
  - [18] Marco Laumanns, Lothar Thiele, Eckart Zitzler, Emo Welzl, and Kalyanmoy Deb. Running time analysis of multi-objective evolutionary algorithms on a simple discrete optimization problem. In *PPSN VII: Proceedings of the 7th International Conference on Parallel Problem Solving from Nature*, pages 44–53, London, UK, 2002. Springer-Verlag.
  - [19] Harpal Maini, Kishan Mehrotra, Chilukuri Mohan, and Sanjay Ranka. Genetic algorithms for graph partitioning and incremental graph partitioning. In *Supercomputing '94: Proceedings of the 1994 Conference on Supercomputing*, pages 449–457, Los Alamitos, CA, USA, 1994. IEEE Computer Society Press.
  - [20] A. M. Patel and L. C. Cote. Partitioning for VLSI placement problems. In *DAC '81: Proceedings of the 18th Conference on Design Automation*, pages 411–418, Piscataway, NJ, USA, 1981. IEEE Press.
  - [21] Pantelimon Stanica. Good lower and upper bounds on binomial coefficients. *Journal of Inequalities in Pure and Applied Mathematics*, 2(1):171–191, 2001.

# Symbolverzeichnis

$\delta_{ij}$	Kronecker-Delta
$\epsilon$	eine kleine Konstante echt größer 0
$\text{GBP}_G^i(x)$	Tupel aus $f_i(x)$ und $\text{zeros}(x)$ für eine partielle Bisektion $x \in \mathcal{S}^*$
$\text{GBP}_G(x)$	Wert einer vollständigen Bisektion $x \in \mathcal{S}$
$\mathbb{P}$	Population des SEMO
$\mathbb{X}_i^I$	Teilmenge des Suchraums $\mathcal{S}^*$
$\mathcal{S}$	Menge der vollständigen Bisektionen
$\mathcal{S}^*$	Menge der partiellen Bisektionen
$\mathbb{N}$	Menge der natürlichen Zahlen
$\mathbb{R}$	Menge der reellen Zahlen
$\mathbb{Z}$	Menge der ganzen Zahlen
$\preceq$	Halbordnung auf den Tupeln der Bildmenge von $\text{GBP}_G^i$
$\preceq_{\text{GBP}_G^i}$	Quasiordnung auf $\mathcal{S}^*$ nach $\text{GBP}_G^i$
$\text{wrongs}(x)$	Anzahl falsch eingeordneter Knoten des Zwei-Cliquen-Graphen
$\text{zeros}(x)$	Anzahl der 0en in einer ternären Zeichenkette $x \in \{0, 1, 2\}^n$
$F^*$	Pareto-Front
$f_0, \dots, f_3$	Untere Schranken für partielle Bisektionen $x \in \mathcal{S}^*$
$G = (V, E)$	Graph mit Knotenmenge $V$ und Kantenmenge $E$
$H(x, y)$	Hamming-Abstand der Zeichenketten $x$ und $y$
$i, j$	Indizes verschiedener Variablen
$L(x)$	Menge der Knoten einer Bisektion $x \in \mathcal{S}^*$ auf der linken Schnittseite
$P^*$	Pareto-Menge
$R(x)$	Menge der Knoten einer Bisektion $x \in \mathcal{S}^*$ auf der rechten Schnittseite
$T$	eine Zeitspanne
$t$	ein Zeitpunkt

$u, v, w$	Knoten eines Graphen $G$
$X$	eine Zufallsvariable; meist die Anzahl von Mutationen
$x, y, z$	Individuen des (1+1)-EA oder des global SEMO