



Estimating the Size of a Chord-based P2P System at Run Time

Andreas Binzenhöfer, Dirk Staehle, and Robert Henjes

University of Würzburg

binzenhoefer@informatik.uni-wuerzburg.de

Inhalt

- ▷ Kurze Erläuterung des Chord-Algorithmus
- ▷ Wozu benötigt man einen Schätzer?
- ▷ Analytisches Modell des Schätzers
- ▷ Ausgewählte Ergebnisse
- ▷ Zusammenfassung

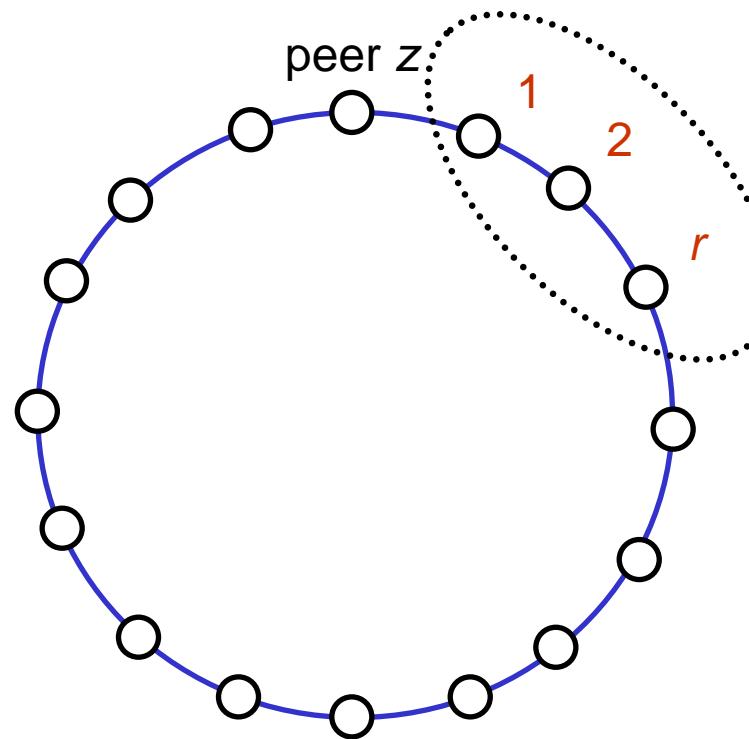


Chord Architektur

Chord: Strukturiertes P2P Netz mit Ringtopologie

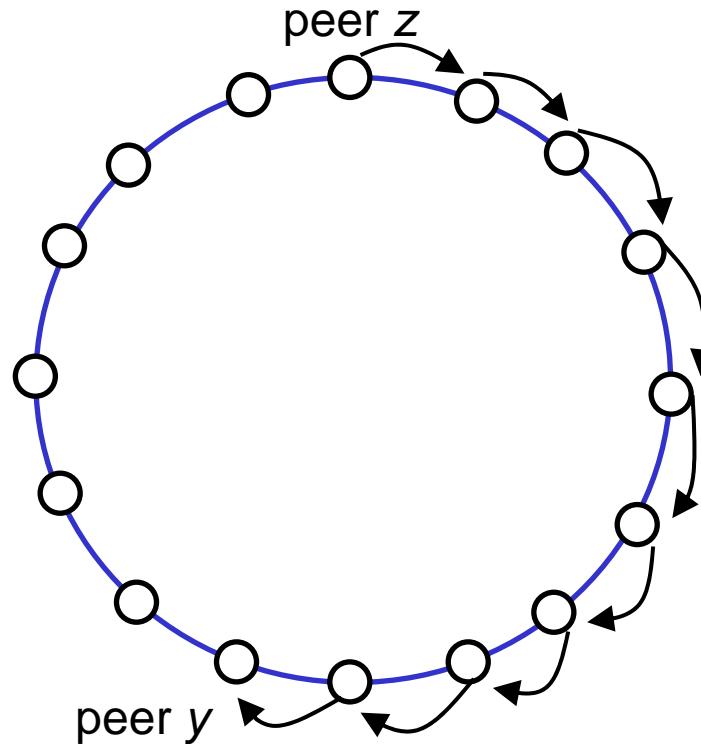
2^m mögliche IDs

n Peers

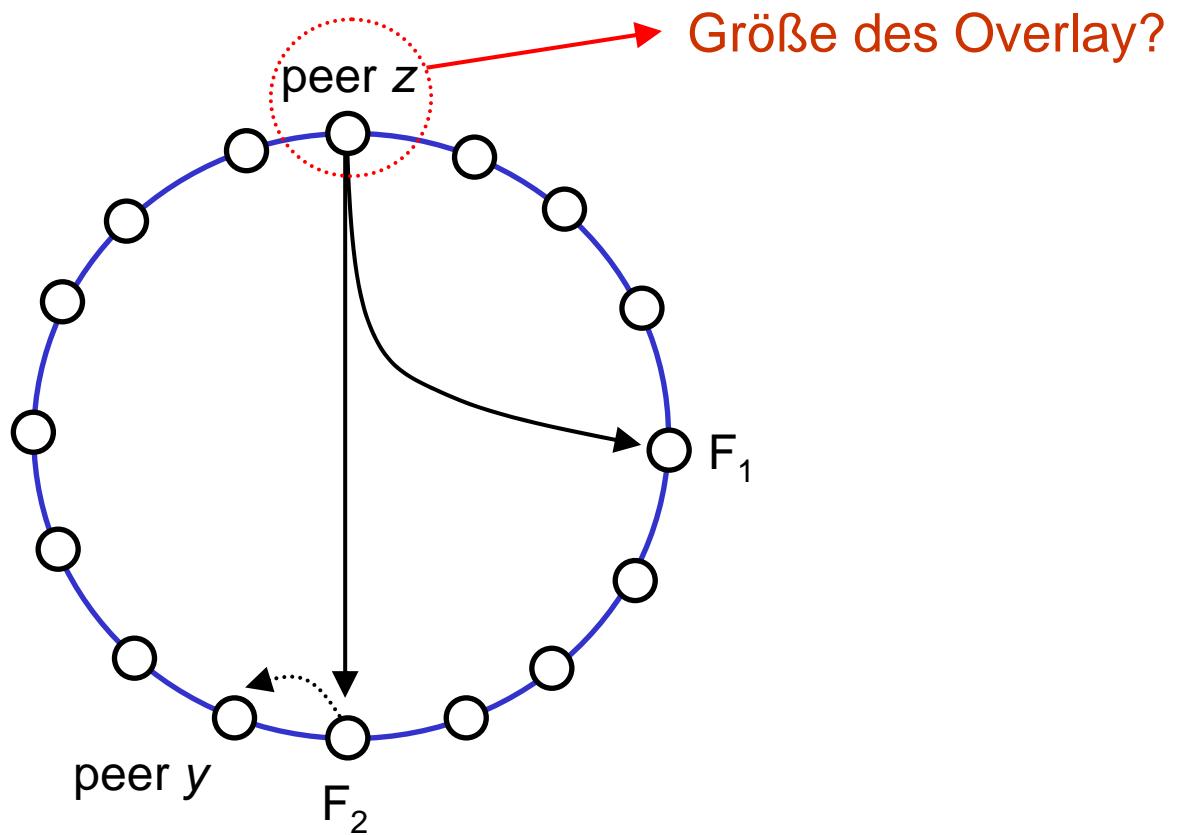


Jedes Peer speichert Zeiger auf seine $r = \log_2(n)$ direkten Nachbarn

Chord: Suche



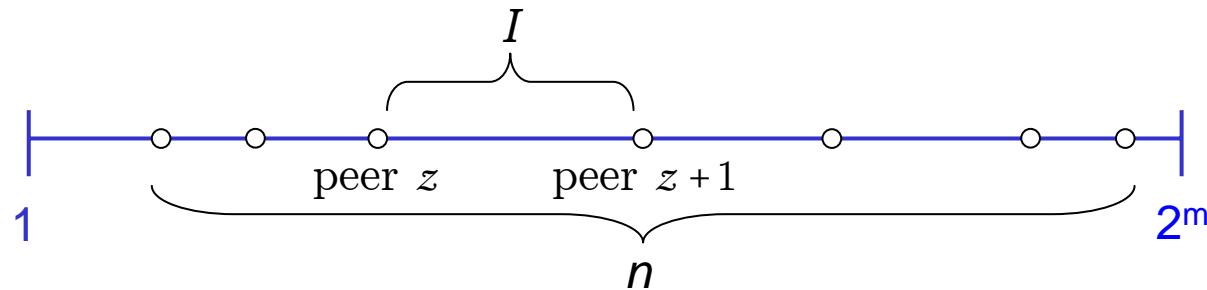
Chord: Fingerliste



Ein einzelnes Peer kennt die Größe n des Overlay nicht

Schätzen der Overlay-Größe

- ▷ Die Größe n des Overlay wird benötigt, um
 - die Größe des Systems als Betreiber zu kennen
 - Timeouts in rekursiven Suchen zu setzen
 - die Stabilität des Overlay-Netzes zu garantieren
 - $\log_2(n)$ berechnen zu können

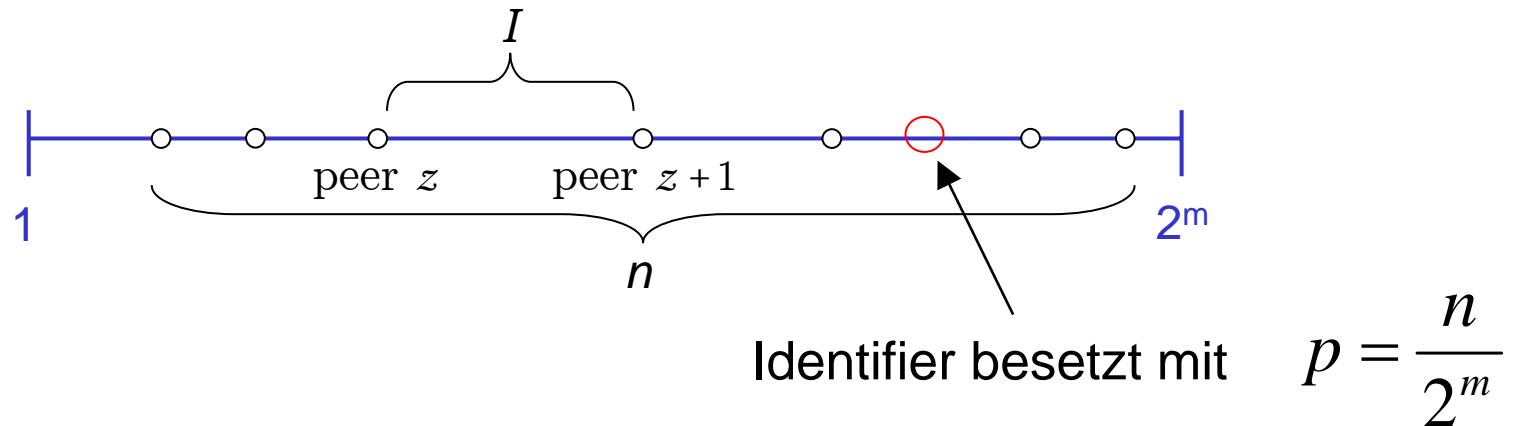


Schätzen der Overlay-Größe

$$2^m \gg n \gg 1$$

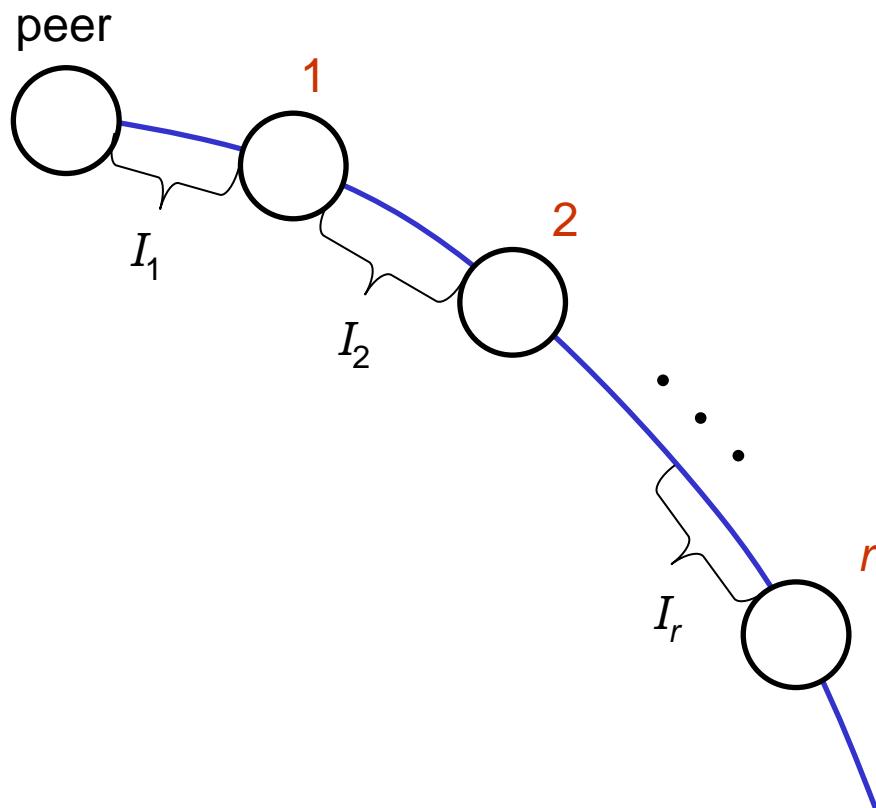
$$P(I = i) \approx (1 - p)^i \cdot p \quad \Rightarrow \quad I \sim \text{geom}(p) \text{ with } p = \frac{n}{2^m}$$

\Rightarrow Die Zufallsvariable I ist näherungsweise geometrisch verteilt

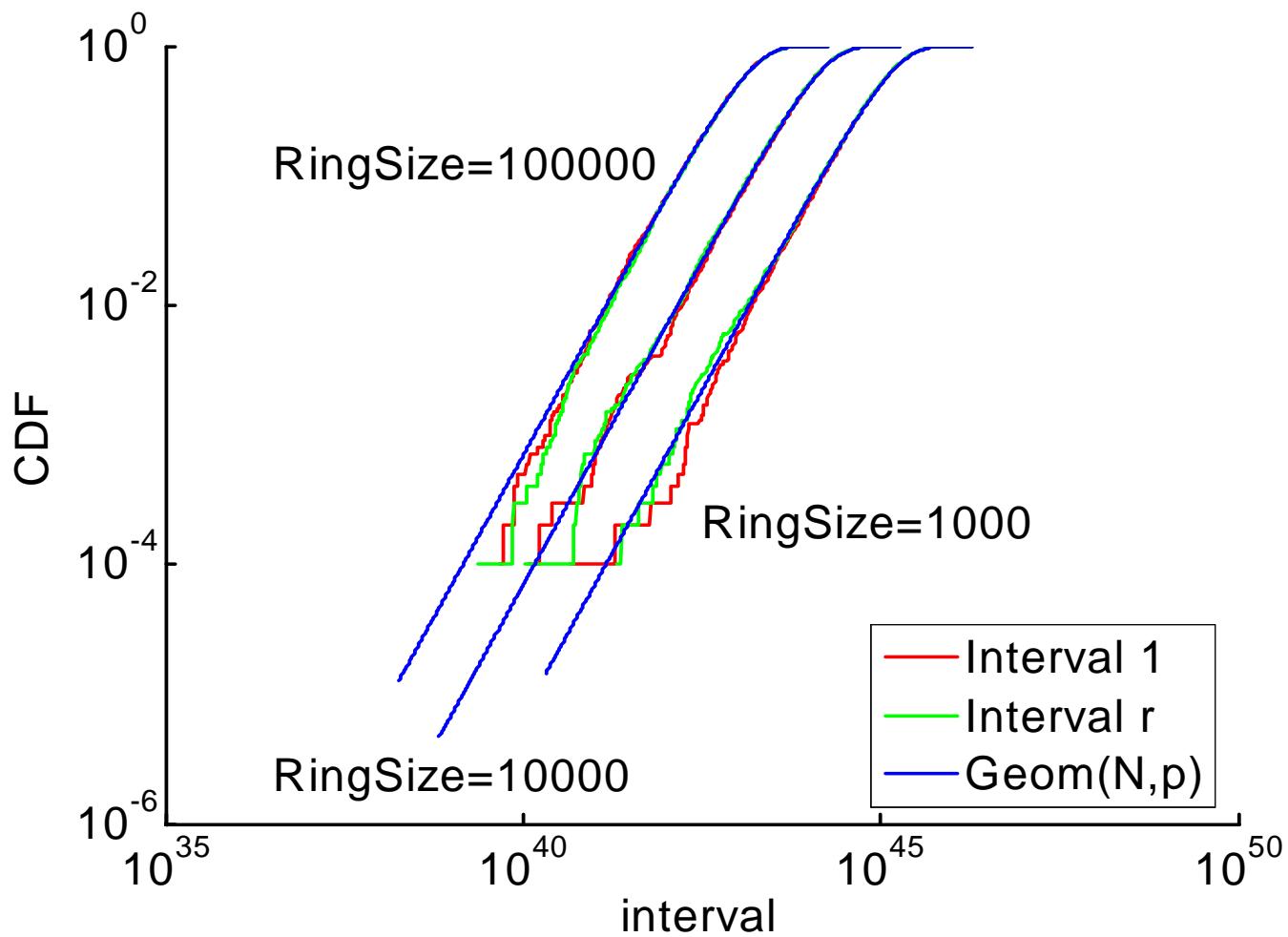


Ein Blick auf die Nachfolgerliste

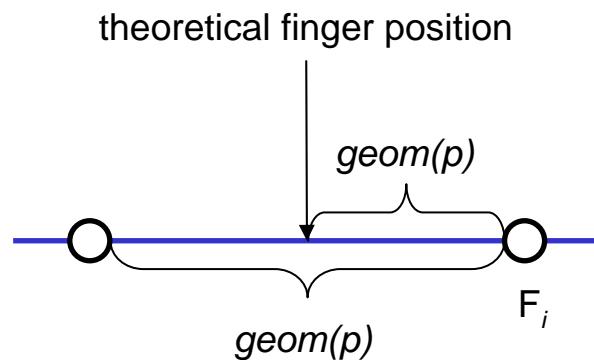
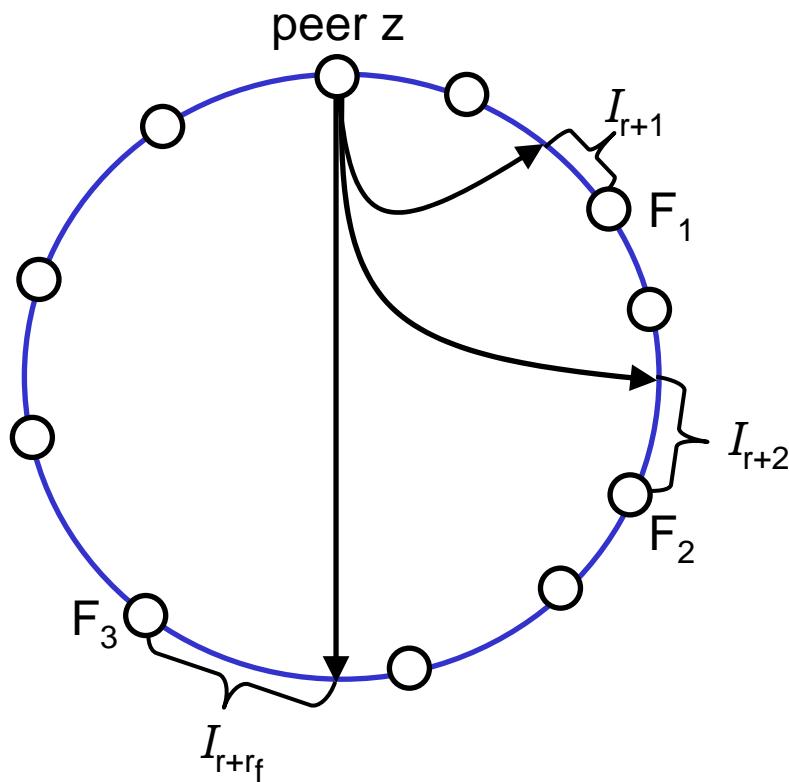
Jedes Peer speichert eine Liste mit r Nachfolgern $\Rightarrow r$ Realisierungen von I



Monte Carlo Simulation



Ein Blick auf die Fingerliste



Jedes Peer hat $r_f = \log_2(n)$ Fingers \Leftrightarrow weitere r_f Realisierungen von I

Schätzung der Größe n des Overlay-Netzes

- ▷ Maximum Likelihood Estimator für p :

$$\hat{p} = \frac{1}{\bar{I}(r + r_f) + 1}$$

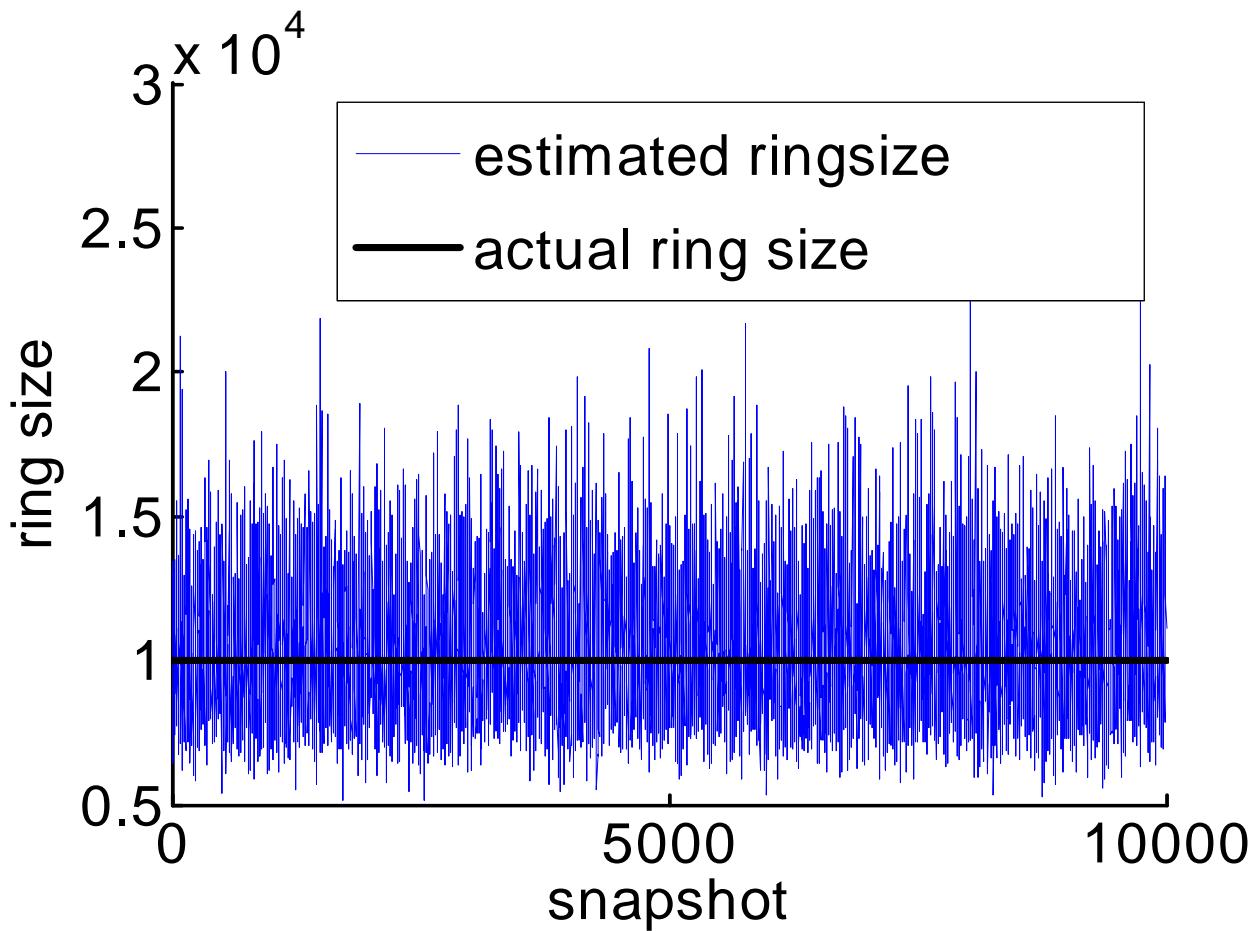
- ▷ und die zugehörigen Konfidenzintervalle:

$$\hat{p} \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}^2(1-\hat{p})}{r + r_f}}$$

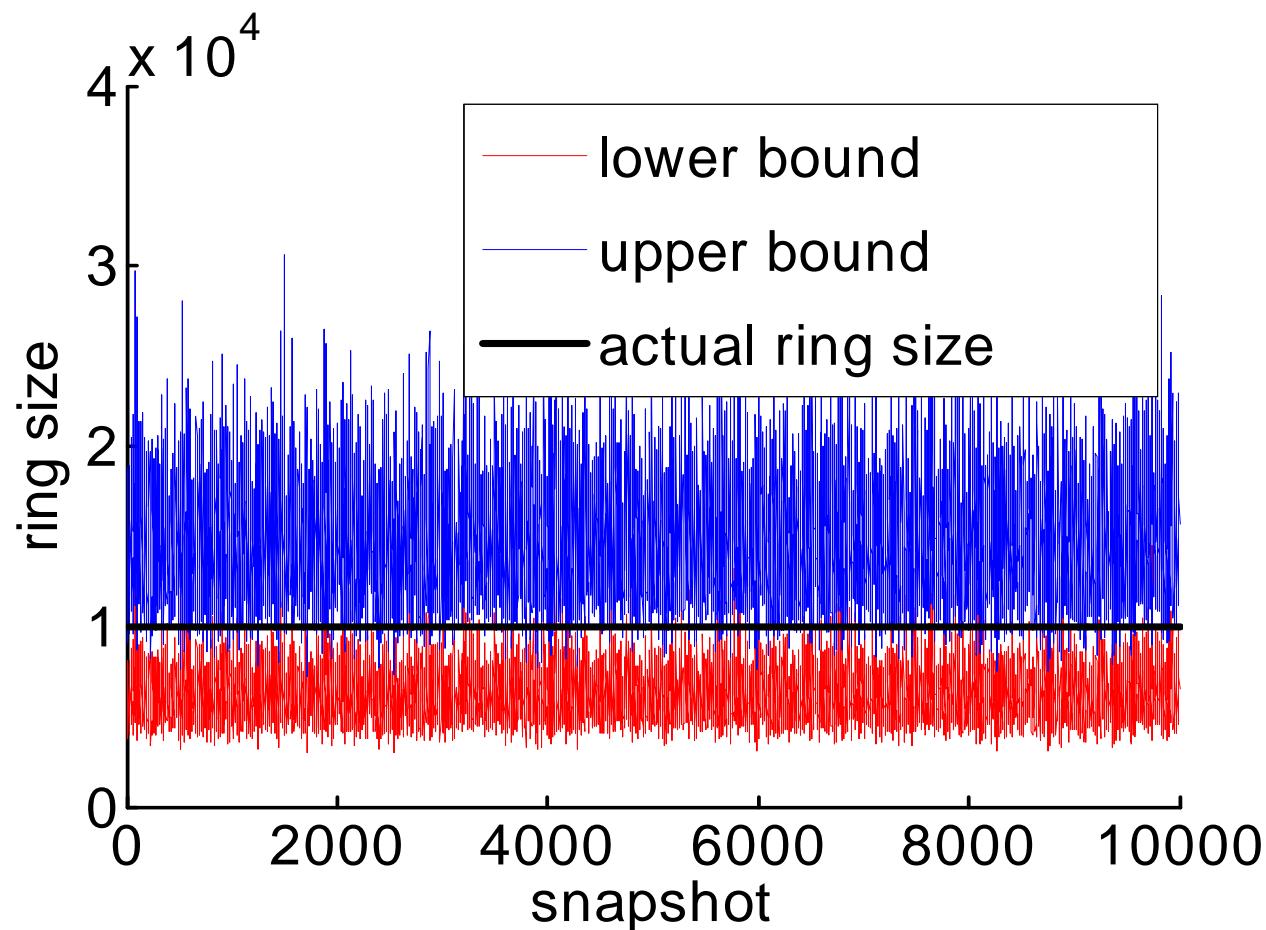
- ▷ Damit: $\hat{n} = \hat{p} \cdot 2^m$



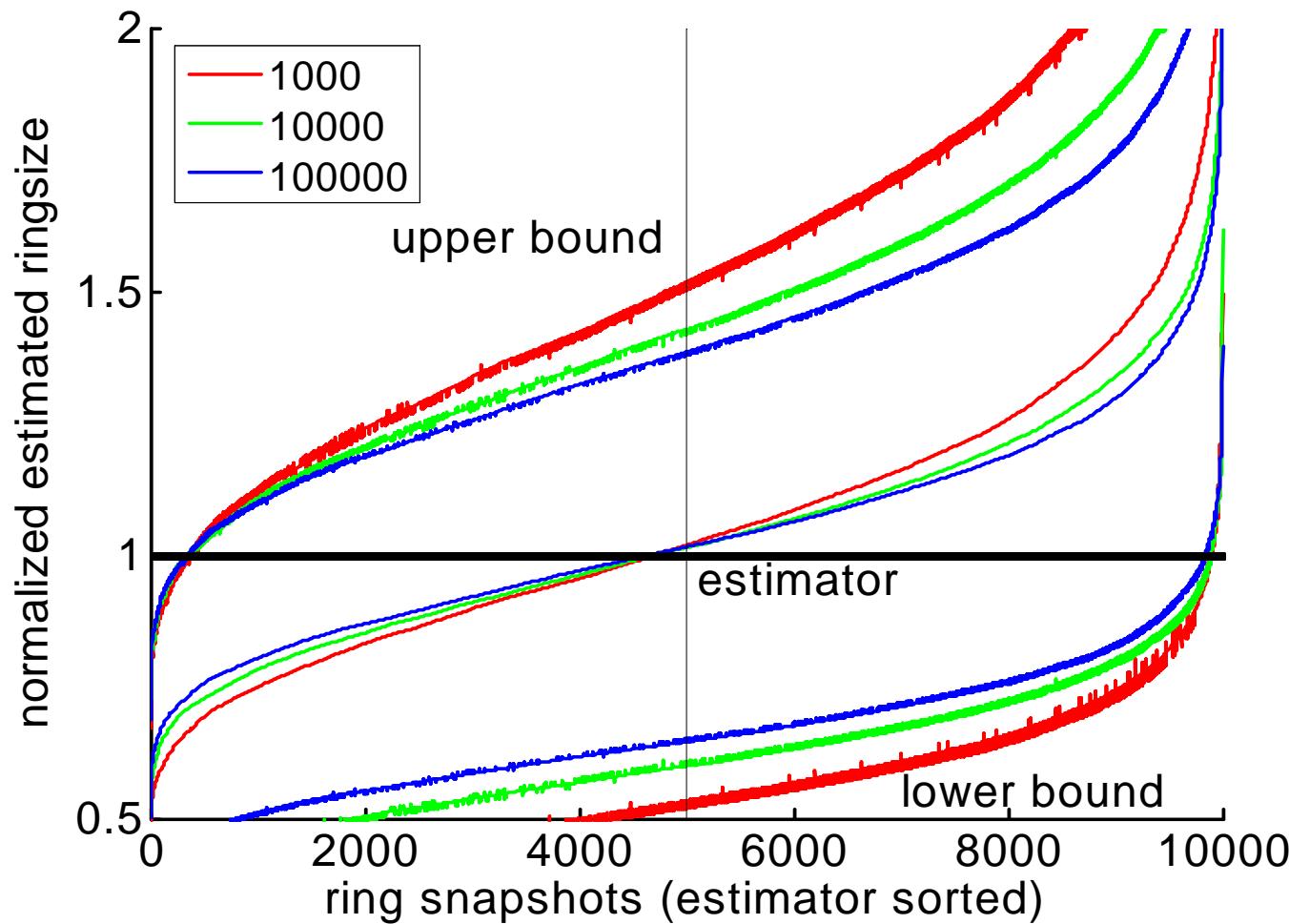
Ergebnisse: Normaler Schätzer



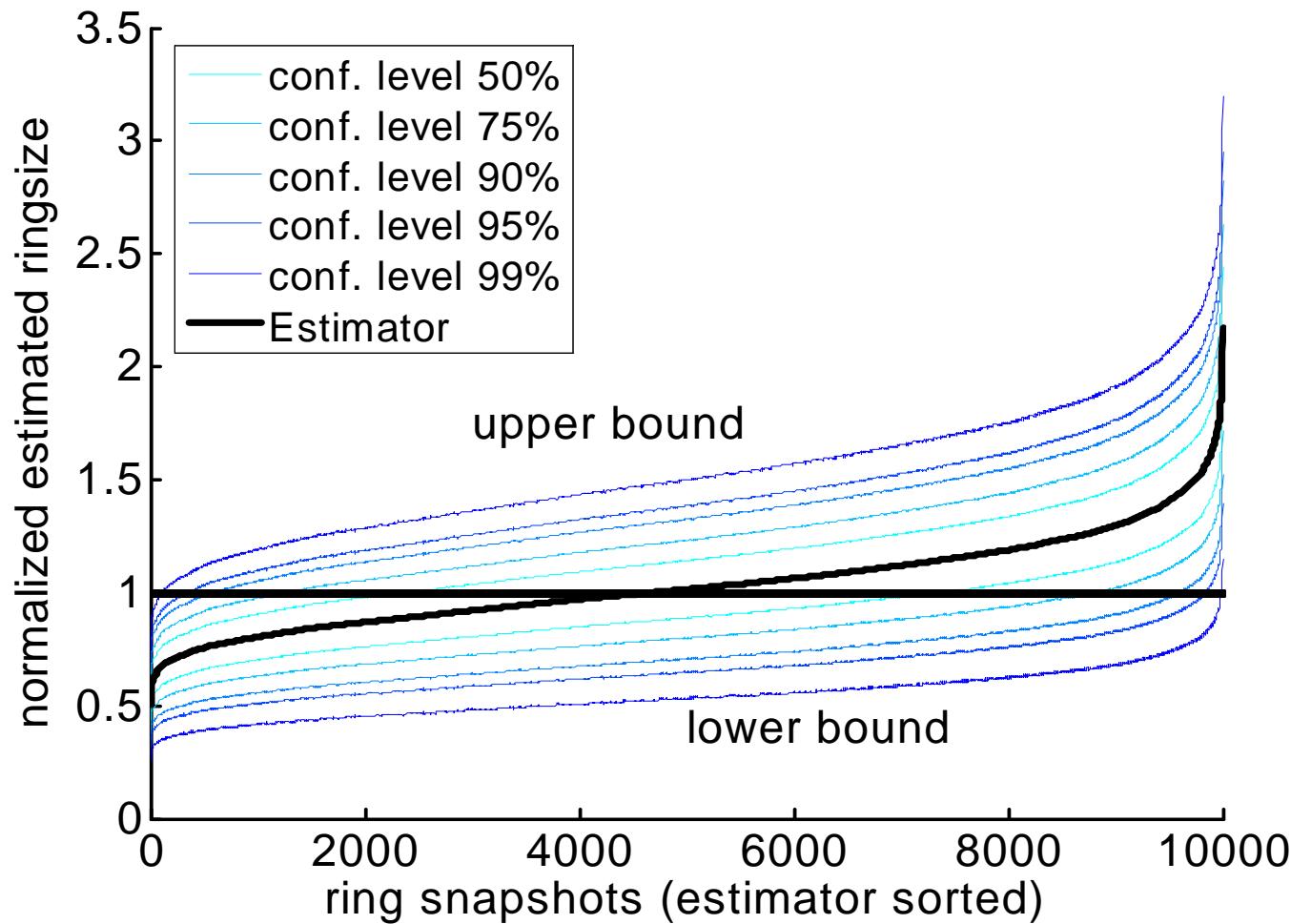
Ergebnisse: Obere und untere Grenze



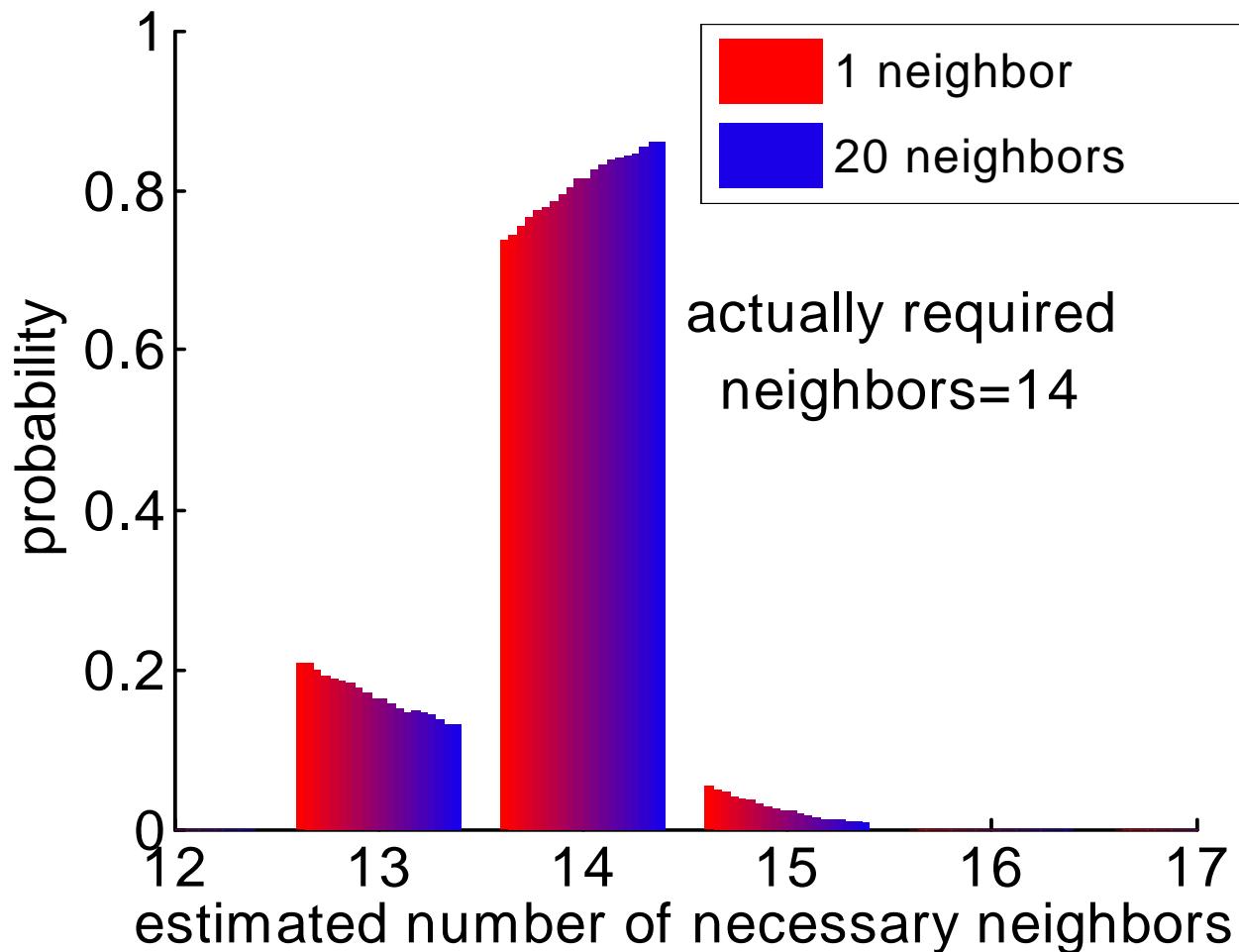
Ergebnisse: Obere und untere Grenze



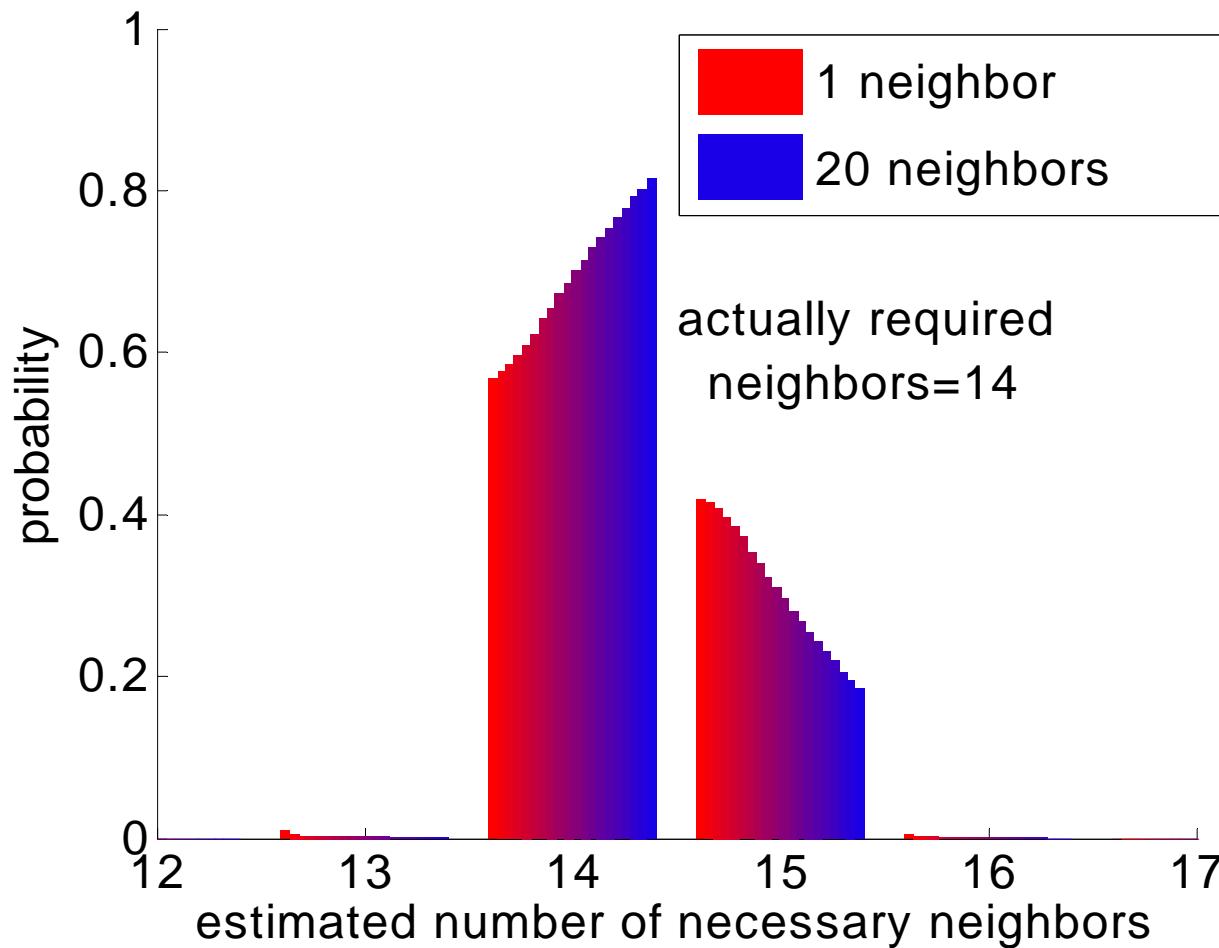
Ergebnisse: Obere und untere Grenze



Ergebnisse: Benötigte Nachbarn (Schätzer)



Ergebnisse: Benötigte Nachbarn (Obere Grenze)



Zusammenfassung

- ▷ Die Größe n ist ein häufig benötigter Parameter
- ▷ Schätzer nutzt die speziellen Eigenschaften von Chord aus
- ▷ Geschätzte Größe liegt grob zwischen $0.5n$ und $2n$
- ▷ Obere und untere Grenze können als Schätzer benutzt werden
- ▷ Besonders geeignet um $\log_2(n)$ zu berechnen (Overlay-Stabilität)



Q&A

