

Algorithmische Geometrie

Geometrische Graphen – euklidisch und hyperbolisch



Wiederholung: Poincaré Disk

Punkte

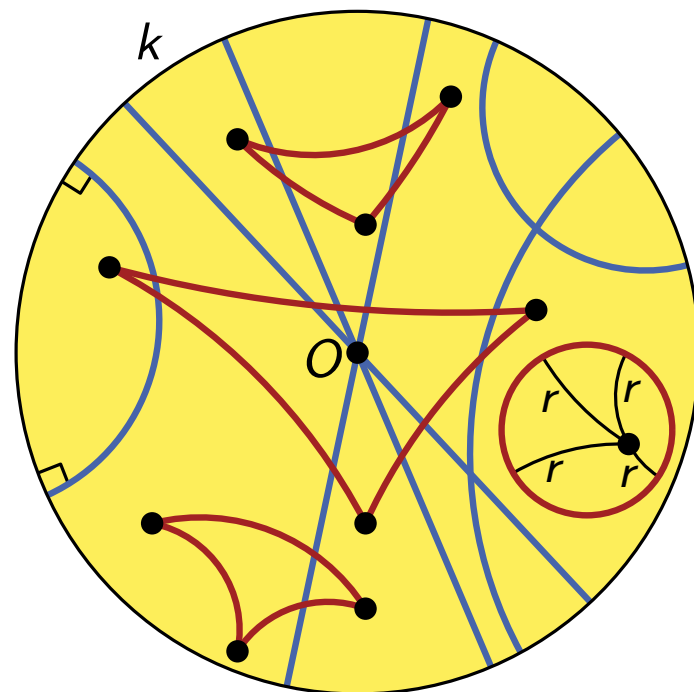
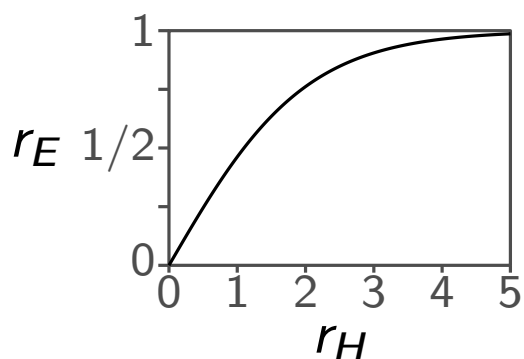
- betrachte einen (euklidischen) Kreis k mit Radius 1 um den Punkt O
- sei $\mathcal{P}$ die Menge der Punkte im Inneren des Kreises

Geraden

- sei $\mathcal{G}$ die Vereinigung von:
 - Menge der offenen Strecken durch O mit Endpunkten auf k
 - Menge der offenen Kreisbögen in k , die senkrecht auf k stehen

Beobachtung

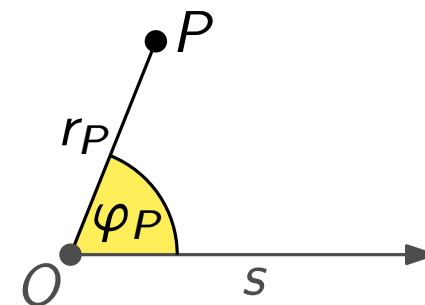
- nah an O : sehr ähnlich zum Euklidischen
- um die Besonderheiten der hyperbolischen Ebene zu nutzen müssen wir weit von O weg
- Problem: wir sind dann sehr schnell sehr nah am Rand von k
- unterschiedliche Radien optisch dann nicht mehr unterscheidbar



(Native) Polarkoordinaten

Polarkoordinaten

- Referenz: Ursprung O , Strahl s mit Anfangspunkt O
- **Radius** r_P von P : $d(O, P)$
- **Winkel** φ_P von P : Winkel zwischen s und OP
- (r_P, φ_P) sind die **Polarkoordinaten** von P



Native Polarkoordinaten

- d ist die hyperbolische Distanz
- insbesondere ist d **nicht** die euklidische Distanz in der Poincaré Disk

Distanz zwischen Punkten in Polarkoordinaten

- betrachte zwei Punkte (in nativen Polarkoordinaten)

$$A = (r_A, \varphi_A) \text{ und } B = (r_B, \varphi_B)$$

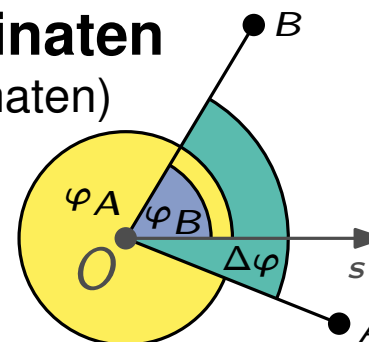
- sei $\Delta\varphi$ ihre Winkeldistanz:

$$\Delta\varphi = \min\{|\varphi_A - \varphi_B|, 2\pi - |\varphi_A - \varphi_B|\}$$

- dann: $d(A, B) = \operatorname{arcosh} [\cosh(r_A) \cosh(r_B) - \sinh(r_A) \sinh(r_B) \cos(\Delta\varphi)]$

$$\approx \log [2 \cdot (e^{r_A}/2 \cdot e^{r_B}/2 - e^{r_A}/2 \cdot e^{r_B}/2 \cdot \cos(\Delta\varphi))]$$

$$= \log [e^{r_A+r_B} \cdot (1 - \cos(\Delta\varphi))/2] = r_A + r_B - \log \left[\frac{2}{1 - \cos(\Delta\varphi)} \right]$$



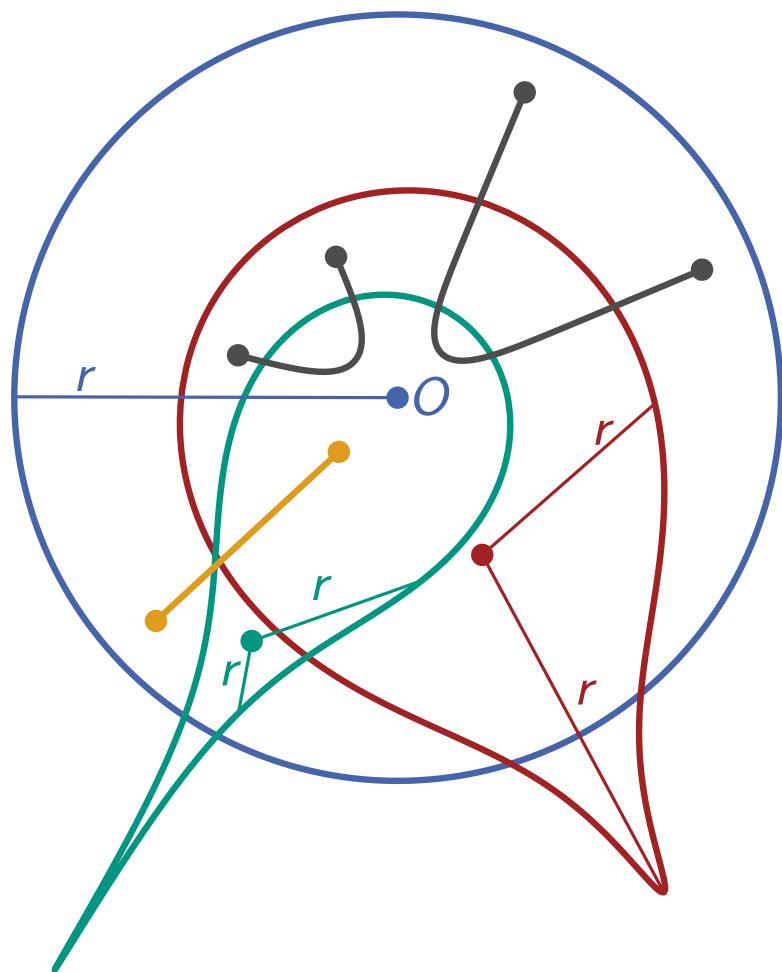
$$\sinh(x) = \frac{e^x - e^{-x}}{2}$$

$$\cosh(x) = \frac{e^x + e^{-x}}{2}$$

Das native Modell

Natives Modell

- nutze nativen Polarkoordinaten für jeden Punkt der hyperbolischen Ebene
- tu so, als wären das ganz normale euklidische Polarkoordinaten



- Kreise um den Ursprung sind Kreise
- Kreise mit anderen Mittelpunkten sehen Tränenförmig aus
- Strecken auf Ursprungsger. sind Strecken
- ihre zu sehende Länge ist die korrekte hyperbolische Länge
- andere Strecken sind Richtung Ursprung gekrümmt
- Darstellung ist **nicht** Winkelerhaltend

$$d(A, B) \approx r_A + r_B - \log \left[\frac{2}{1 - \cos(\Delta\varphi)} \right]$$

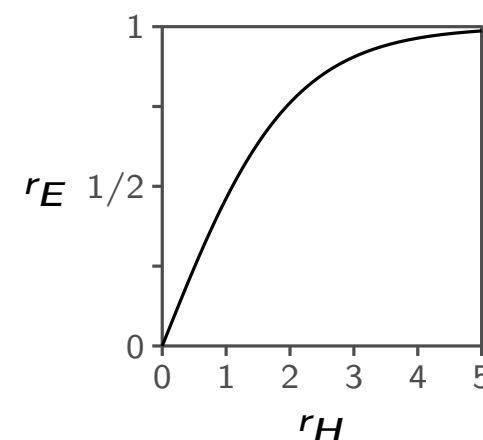
Poincaré vs. nativ

Vorteile der Poincaré Disk

- Darstellung ist Winkelerhaltend
- Geraden sind euklidische Kreisbögen, Kreise sind euklidische Kreise
→ das sind uns vertraute Objekten
- wir können unsere euklidische Intuition bemühen, um Einsichten über die hyperbolische Ebene zu zeigen

Nachteile der Poincaré Disk

- die (euklidische) Distanz vom Ursprung geht sehr schnell gegen 1 für wachsende hyperbolische Distanz
- möchte man etwas hyperbolisches anschauen, das nicht sehr nah am Ursprung liegt, so liegt alles auf dem Rand der Disk und man sieht nichts mehr
- anfälliger für numerische Probleme



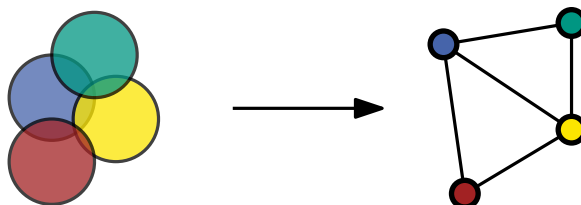
Heuristik zur Wahl des Modells

- visuelle Darstellung von hyperbolischen Daten → natives Modell
- Berechnungen auf Koordinaten → natives Modell (oder auch: Hyperboloid)
- Verständnisfragen/Beweise → Poincaré Disk

Unit-Disk Graphen

Definition

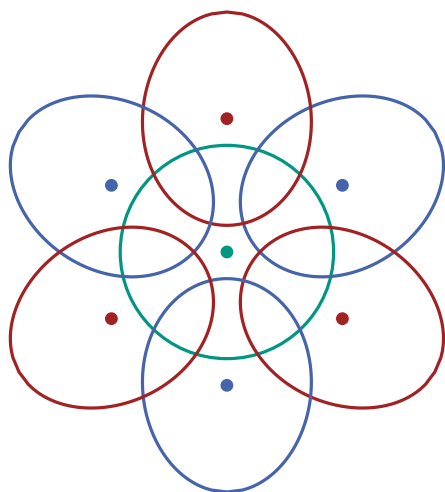
Ein Graph ist ein **Unit-Disk Graph**, wenn er der geometrische Schnittgraph einer Menge von Disks mit Radius 1 ist.



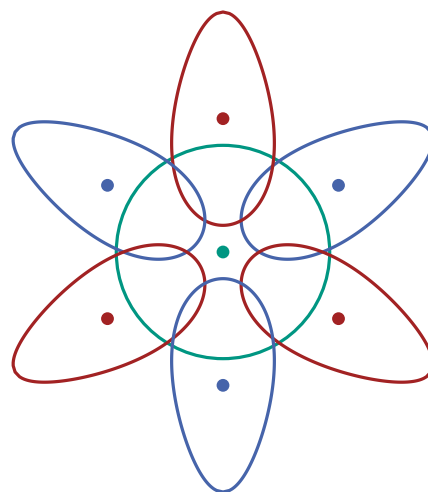
Jetzt hyperbolisch

- kann man im Prinzip erstmal analog definieren
- Achtung: der Radius macht auf einmal einen Unterschied

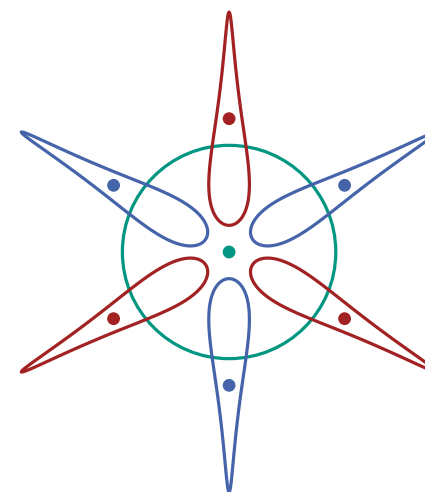
Kreisradius 1



Kreisradius 2



Kreisradius 4



Hyperbolische Unit-Disk Graphen

Definition

Beachte: die Disks der Schnittrepräsentation haben also Radius $R/2$

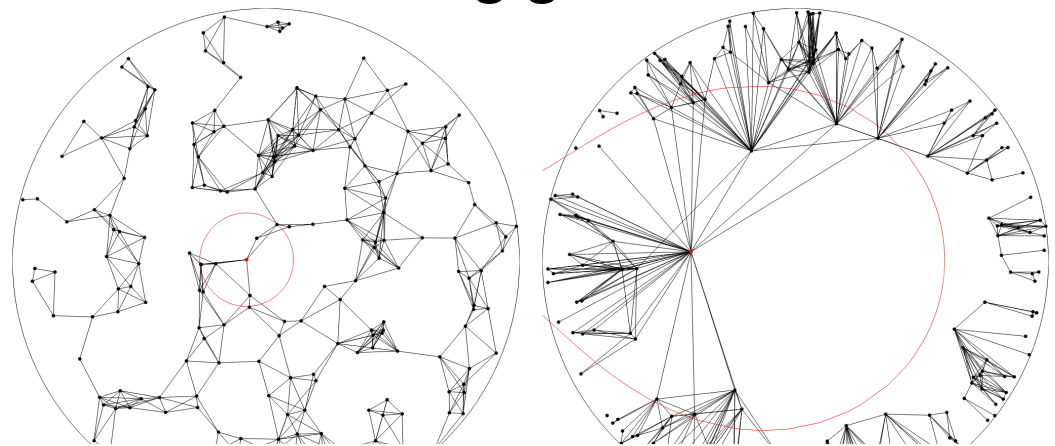
$G = (V, E)$ ist ein **hyperbolischer Unit-Disk Graph** wenn er eine hyperbolische Unit-Disk Repräsentation hat: eine Abbildung $p: V \rightarrow \mathbb{H}^2$ zusammen mit einem Radius R , sodass $\{u, v\} \in E \Leftrightarrow d(p(u), p(v)) \leq R$.

Warum der Name *Unit-Disk Graph*?

- der Begriff *Unit* ist hier als „einheitlich“ statt als „Länge = 1“ zu verstehen
- man erhält das gleiche, wenn man $R = 2$ fordert, aber die Krümmung der hyperbolischen Ebene nicht auf -1 festhält

Bekomme ich unterschiedliche Strukturen abhängig von R ?

- kleines R
 - wie im Euklidischen
 - regelmäßig / homogen
 - gitterartig
- großes R
 - unregelmäßig / inhomogen
 - hierarchisch



<https://thobl.github.io/hyperbolic-unit-disk-graph/>

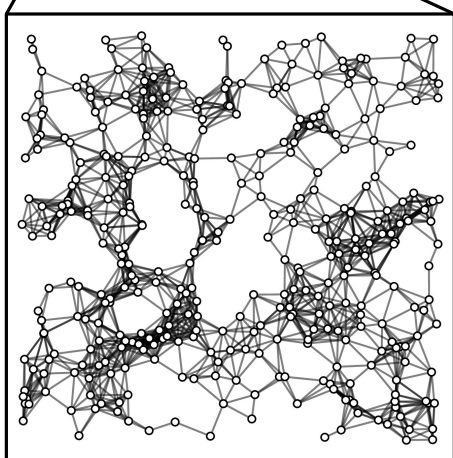
Wie hyperbolisch darfs sein?

kleiner Radius R $\longrightarrow$ großer Radius R

hyperbolische Unit-Disk Graphen

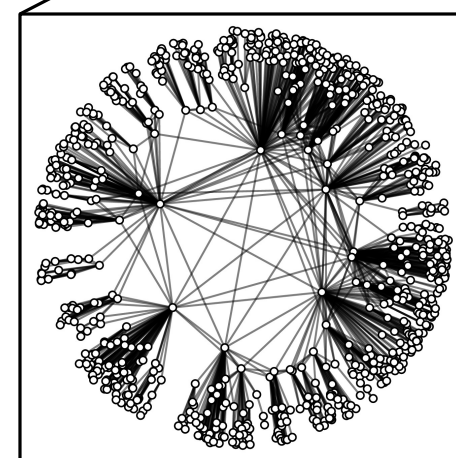
stark hyperbolische UD-Graphen

(euklidische) UD-Graphen



Situation

- euklidische UD-Graphen sind eine Teilklasse von hyperbolischen
- viele hyperbolische UD-Graphen sind nicht sehr hyperbolisch



Stark hyperbolische Unit-Disk Graphen

- Ziel: Gegenstück zu euklidischen UD-Graphen
- mit hierarchischer / heterogener Struktur
- Wie formalisieren wir das? Wie groß ist groß genug für R ?

Stark hyperbolisch

Definition

$G = (V, E)$ ist ein **hyperbolischer Unit-Disk Graph** wenn er eine hyperbolische Unit-Disk Repräsentation hat: eine Abbildung $p: V \rightarrow \mathbb{H}^2$ zusammen mit einem Radius R , sodass $\{u, v\} \in E \Leftrightarrow d(p(u), p(v)) \leq R$.

Definition

G ist ein **stark hyperbolischer Unit-Disk Graph**, wenn p alle Knoten in eine Disk mit Radius R abbildet.

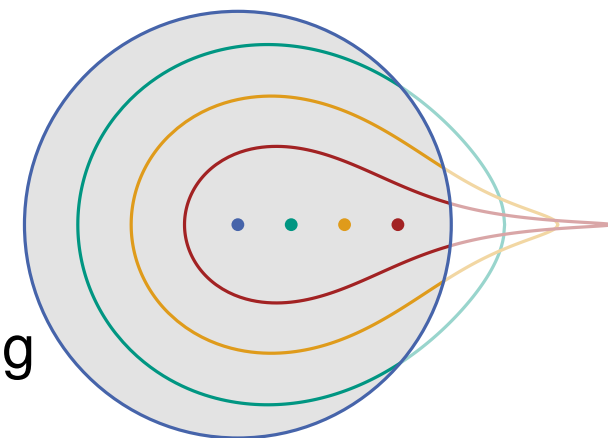
<https://thobl.github.io/hyperbolic-unit-disk-graph/>

Darstellung

- wähle Zentrum dieser Disk als Ursprung für die nativen Polarkoordinaten
- **Achtung:** nicht mit der Poincaré Disk verwechseln

Beobachtungen

- Knoten im Ursprung: mit allen anderen benachbart
- je weiter außen ein Knoten, desto kleiner sein Einflussbereich
- maximale Heterogenität: jede Distanz vom Ursprung liefert unterschiedlich großen Einflussbereich



Erzwungene Kanten

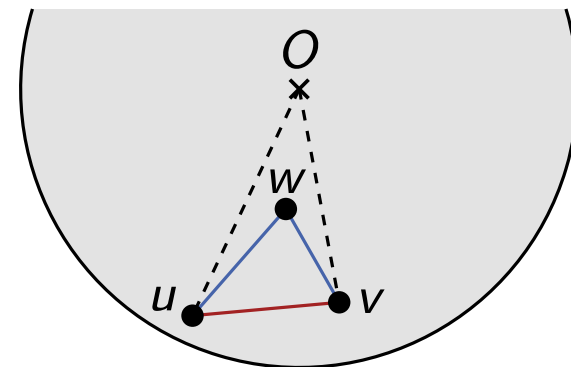
Situation

- stark hyperbolische UD-Repräsentation
- w liegt bezüglich Winkel zwischen u und v
- w liegt weiter innen als u und v

Theorem

(ohne Beweis)

Wenn $\{u, v\} \in E$, dann auch $\{u, w\}, \{v, w\} \in E$.



Anmerkung

- man kann nicht unbemerkt außen an einem inneren Knoten vorbeilaufen
- hierarchische Strukturen: je weiter innen desto höher in der Hierarchie
- Beweis: für Interessierte im Anhang

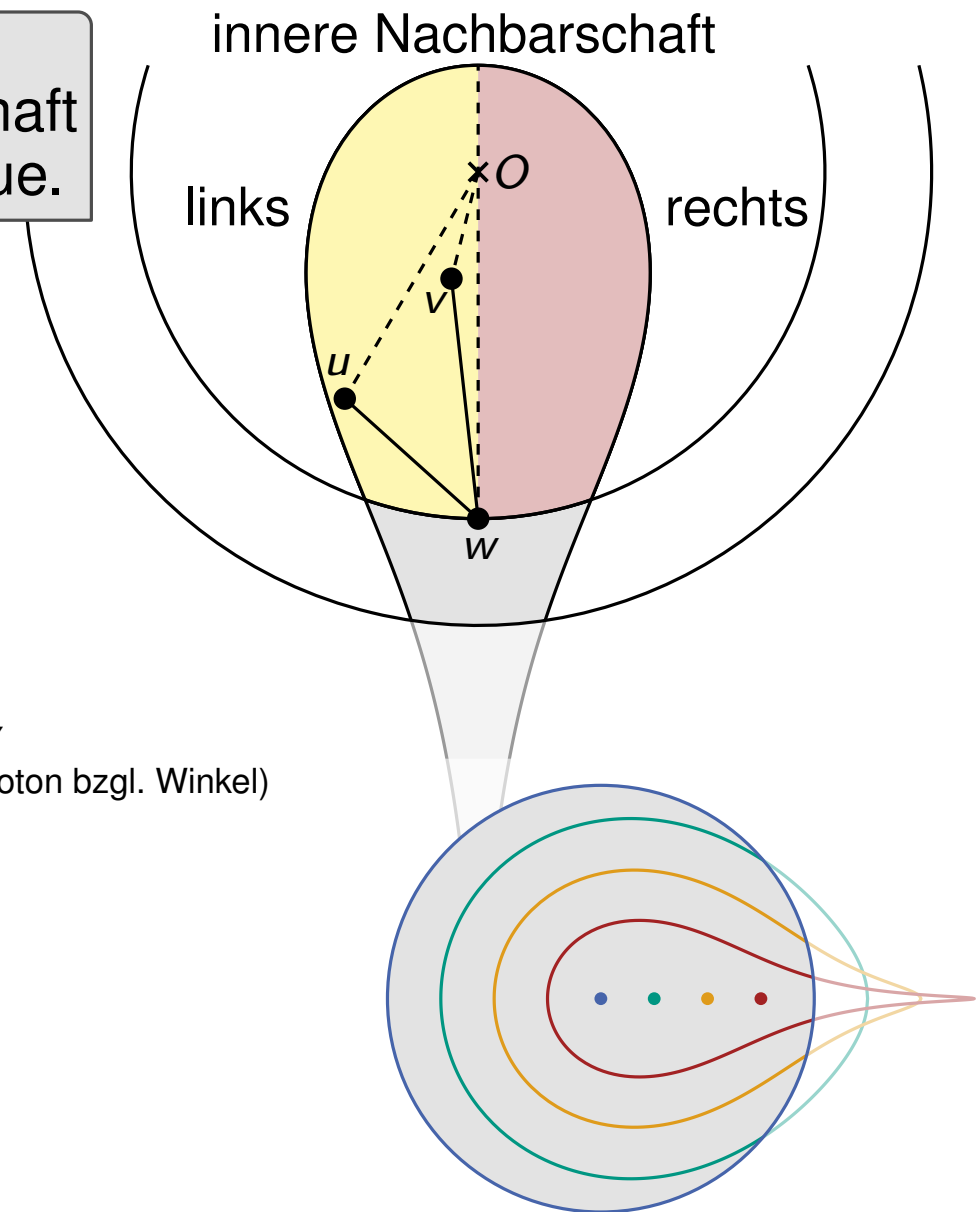
Innere Nachbarschaft (stark hyperbolisch)

Theorem

Die linke und rechte innere Nachbarschaft eines Knotens bilden jeweils eine Clique.

Beweis

- zeige: u und v sind verbunden
- w Richtung O verschieben verkleinert seine Nachbarschaft nicht
- verschiebe w , sodass $r_w = r_v$
 $\Rightarrow w$ bleibt mit u verbunden
(beachte: $d(u, w)$ wird ggf. größer, aber nicht größer als R)
- verschiebe w weiter, sodass $\varphi_w = \varphi_v$
 $\Rightarrow w$ bleibt mit u verbunden (Distanz monoton bzgl. Winkel)
- also: $\{u, v\} \in E$



Innere Nachbarschaft (stark hyperbolisch)

Theorem

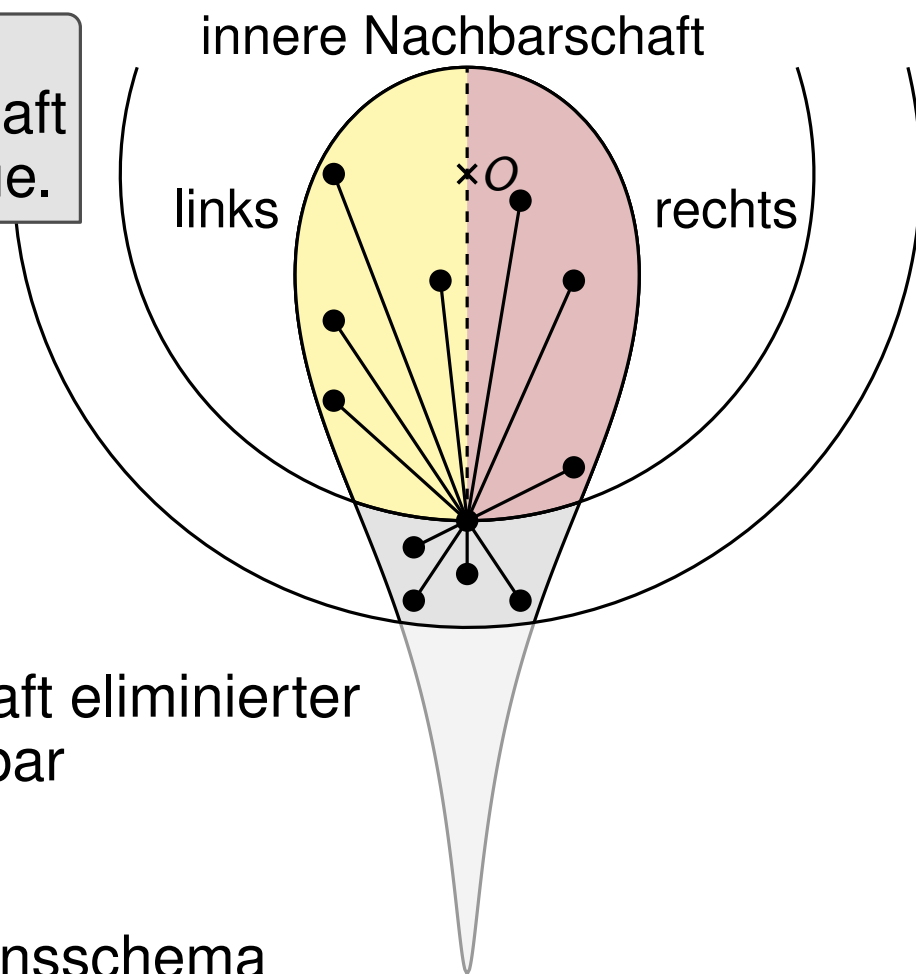
Die linke und rechte innere Nachbarschaft eines Knotens bilden jeweils eine Clique.

Folgerung

- sortiere Knoten $v_1, \dots, v_n$ von Außen nach Innen
- lösche Knoten nach und nach:
 $G_i = G[\{v_i, \dots, v_n\}]$
- v_i hat in G_i nur innere Nachbarn
- Eliminationsschema mit: Nachbarschaft eliminierten Knoten durch zwei Cliques überdeckbar

Vergleich: Chordale Graphen

- Graph chordal $\Leftrightarrow$ perfektes Eliminationsschema
- perfektes Eliminationsschema: Nachbarschaft eliminierten Knoten bilden Clique



Was wissen wir sonst noch so?

Erkennung von hyperbolischen UD-Graphen

- NP-schwer
- liegt in $\exists\mathbb{R}$
- offen: $\exists\mathbb{R}$ -Schwere
- offen: Komplexität für stark hyperbolische Graphen

Größte Clique finden

- für stark hyperbolische UD-Graphen in polynomieller Zeit

Baumweite für stark hyperbolische UD-Graphen

- balancierte Separatoren der Größe $O(\Delta)$ (max deg)
- Baumzerlegung entsprechender Weite
- ganz anders als bei Gittern

Routing auf stark hyperbolischen UD-Graphen

- lokales Routing (ähnlich zu greedy Routing) mit kleinem Stretch
- schlägt untere Schranke für allgemeine Graphen

Masterarbeit
Emil Dohse

Was soll das überhaupt?

Warum machen wir eigentlich Theorie?

- weil es Spaß macht
- um Beobachtungen zu erklären
- um Vorhersagen treffen zu können

Algorithmische Beispiele

■ sortieren	Bubblesort	Mergesort	Speedup
	$\Theta(n^2)$	$\Theta(n \log n)$	$\Theta\left(\frac{n}{\log n}\right)$
■ orthogonale Bereichsanfragen	k -d Tree	Range-Tree	Speedup
	$\Theta(\sqrt{n})$	$\Theta(\log^2 n)$	$\Theta\left(\frac{\sqrt{n}}{\log^2 n}\right)$

Vorhersage

- polynomieller asymptotischer Speedup
- beliebig großer Speedup, wenn man die Instanz groß genug wählt

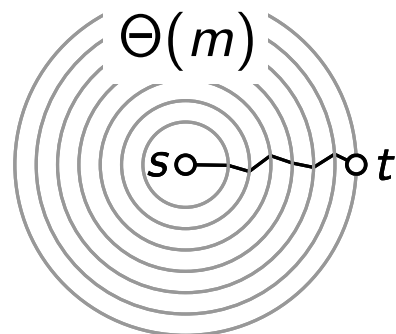
Schlussfolgerungen

- den langsameren Algorithmus braucht man nicht in Erwägung zu ziehen

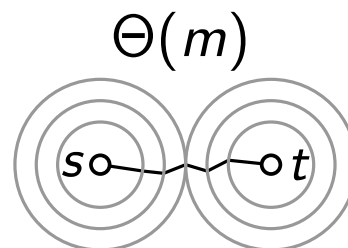
Beispiel: Bidirektionale Breitensuche

Kürzeste Wege in ungewichteten Graphen

Breitensuche (BFS)



bidirektionale BFS



Speedup

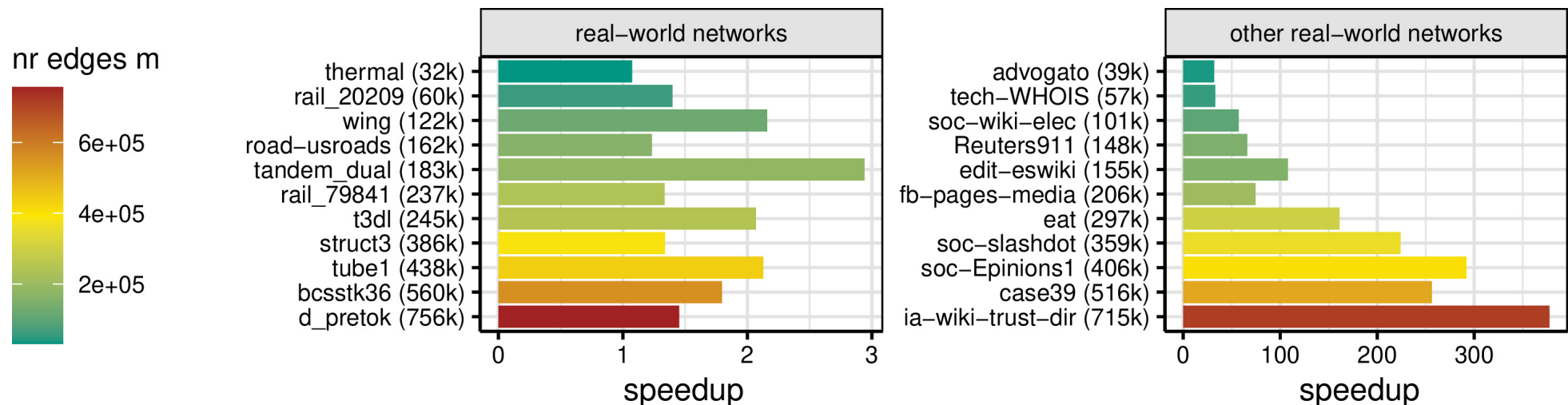
$\Theta(1)$

Vorhersage: Speedup
unabhängig von der Graphgröße

Schlussfolgerung

- Option 1: beides implementieren und experimentell vergleichen
 - wenn man fleißig ist bzw. die nötige Zeit hat
 - wenn man sich für konstante Faktoren interessiert
- Option 2: einfach BFS nutzen, da bidirektional den Aufwand nicht wert ist
 - wenn man wenig Zeit hat
 - wenn konstante Faktoren gerade nicht sehr relevant sind

Bidirektionale BFS: Experimenteller Vergleich



Beobachtung: sehr falsche Vorhersage für manche Instanzen

Was ist daran jetzt schlimm?

- Algorithmus schneller als erwartet → super
- außer: ggf. nicht implementiert, da ich dachte, es lohnt nicht
- außerdem: schon ein bisschen peinlich für die Theorie
 - Analyse von $\Theta(m)$ für bidirektionale BFS hat wenig Spaß gemacht
 - Theorie erklärt nicht das beobachtete Verhalten
 - Theorie sagt etwas falsches voraus

Make Theory Great Again

Wo liegt das Problem?

- theoretische Analysen betrachten häufig den Worst-Case
- typische Instanzen sind oft gutartig

Ziele

- formalisiere was „typische Instanze“ und „gutartig“ bedeutet
- analysiere Algorithmen bezüglich dieser Formalisierung
- keine Überraschungen: Erklärung für bzw. korrekte Vorhersage von beobachtetem Verhalten
- Laufzeit in Abhängigkeit struktureller Eigenschaften

Disclaimer

- verschiedene Methodiken sind für diese Ziele verschieden gut geeignet
- Methodik hierfür ungeeignet $\nRightarrow$ Methodik grundsätzlich nutzlos
- Beispiel: Worst-Case Analyse ist weiterhin ein wichtiges und in den meisten Fällen das richtige Tool

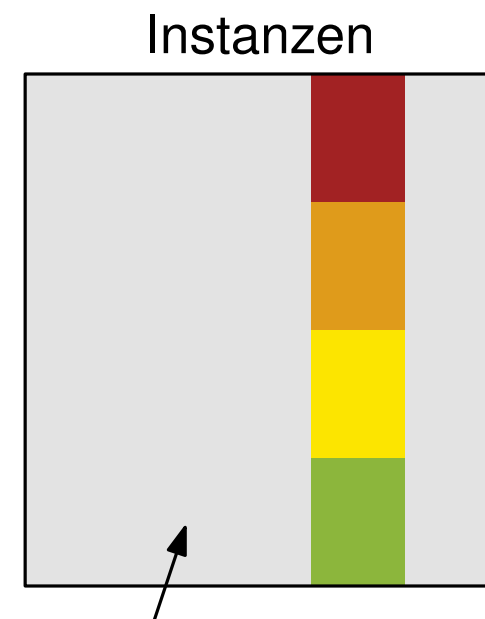
Was ist ein akzeptierbarer Fix?

Der Heilige Gral

- zeige: Instanzen mit einer gewissen Eigenschaft sind leicht
- verifiziere: praktische Instanzen haben die Eigenschaft
- die Eigenschaft sollte „natürlich“ sein

Warum ist der heilige Gral schwer zu finden?

- Worst-Case unter den Instanzen mit der Eigenschaft
 $\Rightarrow$ Eigenschaft muss ausreichend restriktiv sein
- sollte viele praktische Instanzen einschließen
 $\Rightarrow$ Eigenschaft darf nicht zu restriktiv sein



einfache praktische Instanzen

Schwächerer aber akzeptabler Fix

- Einschränkung der Analyse auf prototypische Instanzen
- Hoffnung: echte Instanzen verhalten sich ähnlich $\Rightarrow$ keine Überraschung

Was heißt das jetzt konkret?

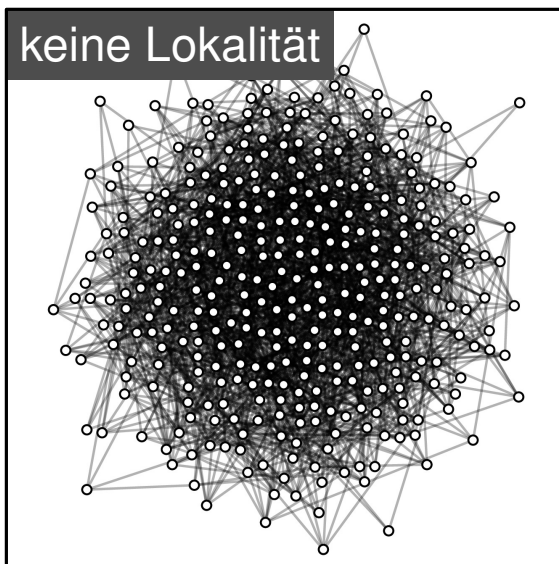
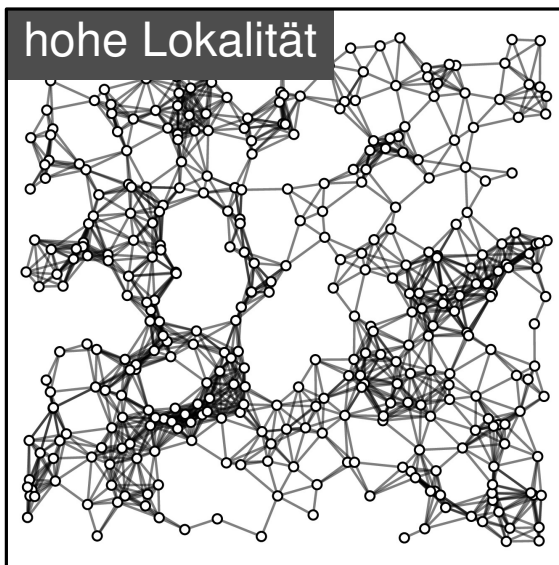
Methodik

- betrachte zufällig generierte Instanzen
- analysiere erwartete Laufzeit
- vergleiche Ergebnisse für verschiedene Wahrscheinlichkeitsverteilungen, die sich in nur einer Eigenschaft unterscheiden

Plan im Folgenden

- zwei wichtige Eigenschaften: Lokalität und Heterogenität
- Was bekommen wir damit für die bidirektionale Breitensuche?
- Generalisiert das von den prototypischen generierten Instanzen zu praktischen Instanzen?

Was ist Lokalität?



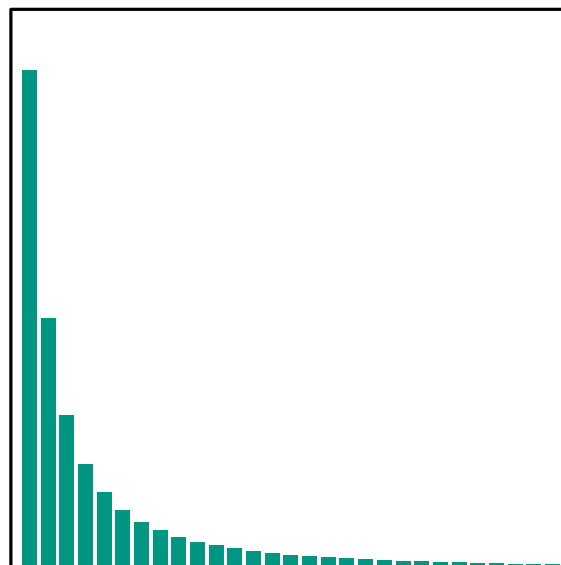
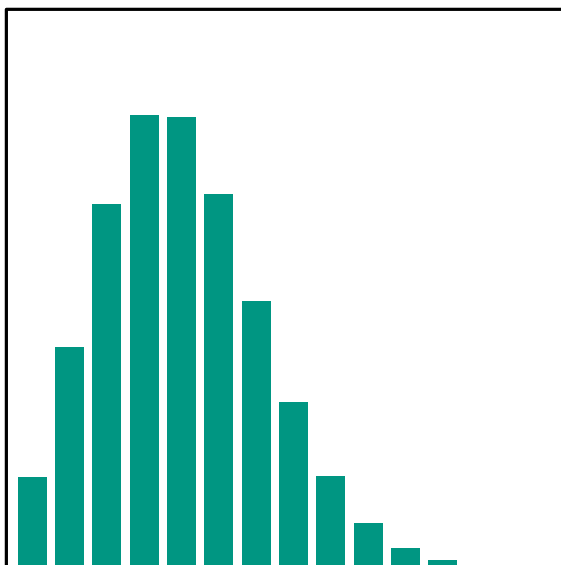
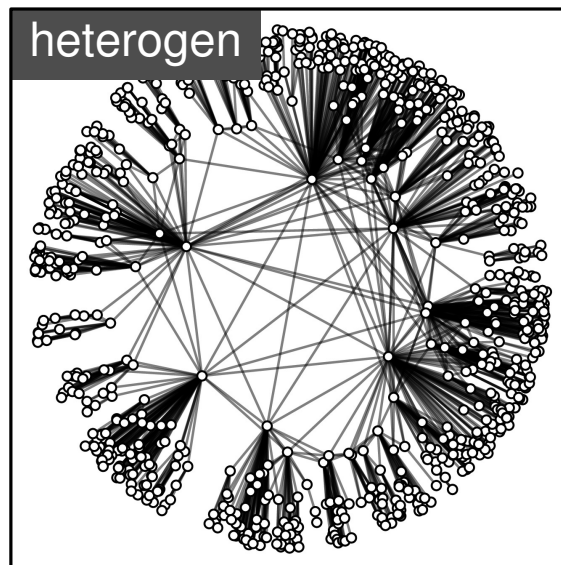
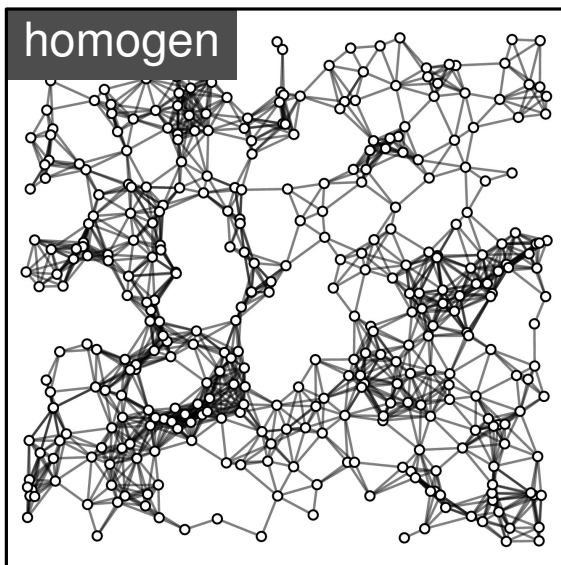
Charakteristische Eigenschaften

- kurze Kanten
- viele Dreiecke

Alternative Namen für das selbe Konzept

- Homophilie
- starkes Clustering
- Modularität
- Redundanz
- zugrundeliegende Geometrie
- tiefe Temperatur

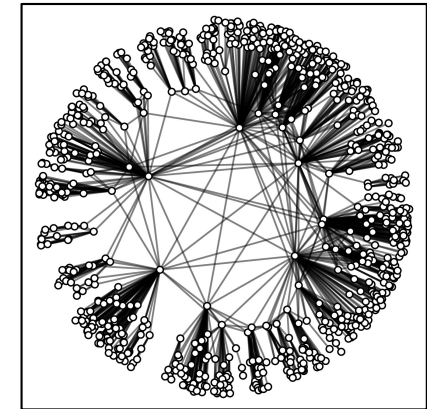
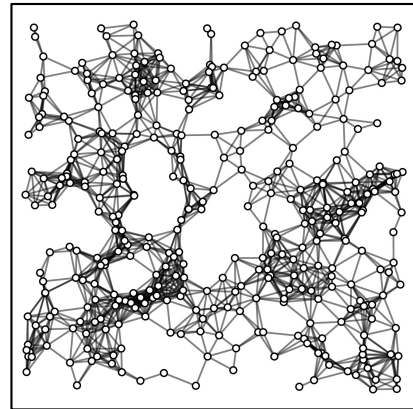
Was ist Heterogenität?



Gradverteilung

- kleine/große Varianz
- kleines/großes Maximum
- beide: kleiner Durchschnitt

Formalisierung mittels Zufallsgraphen



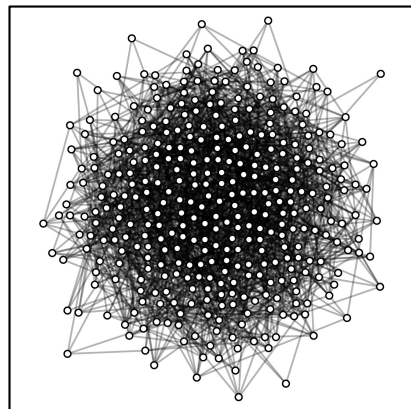
Heterogenität



geometrisch

hyperbolisch
GIRG

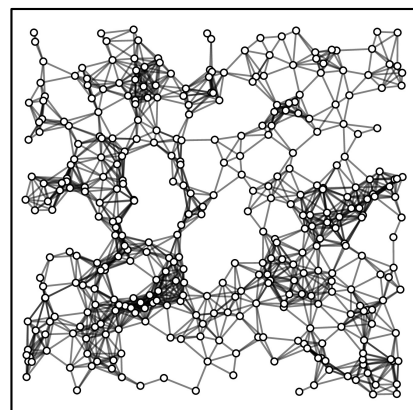
Lokalität



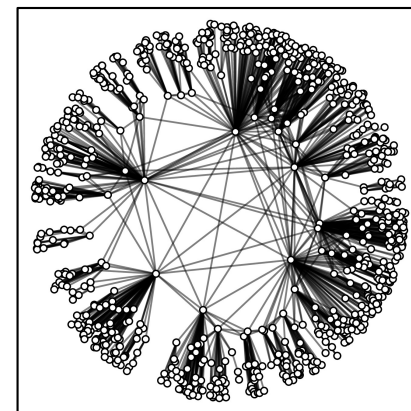
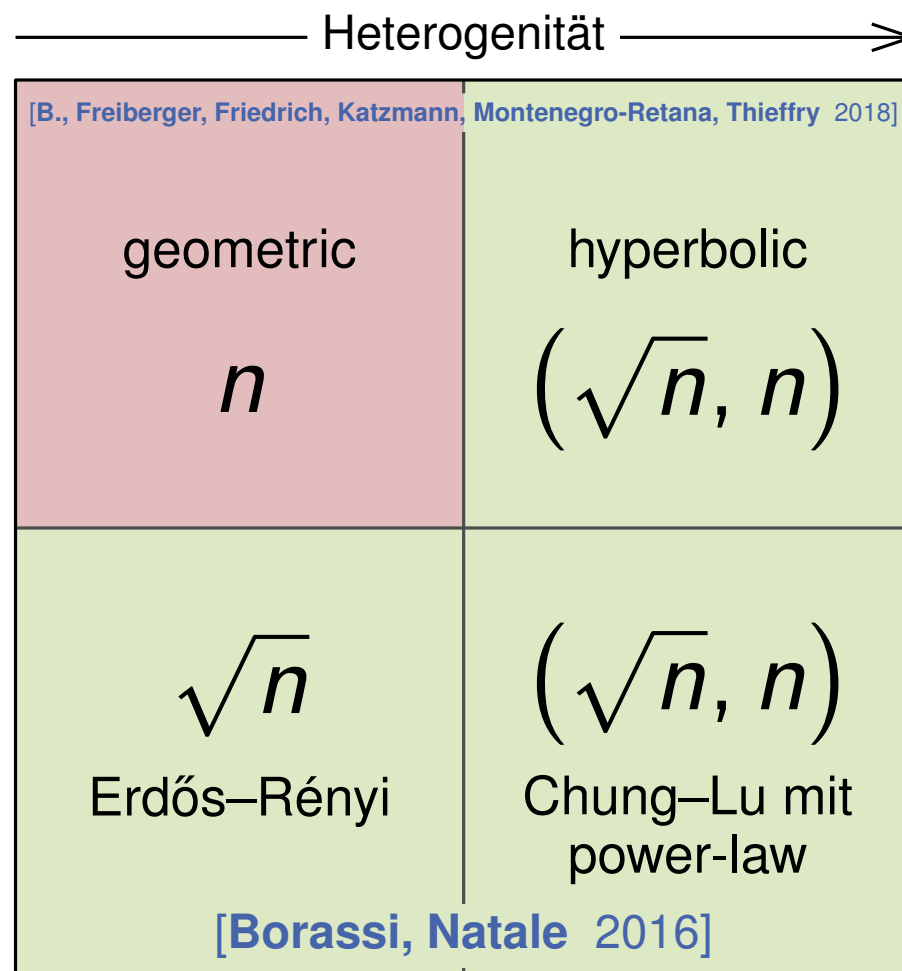
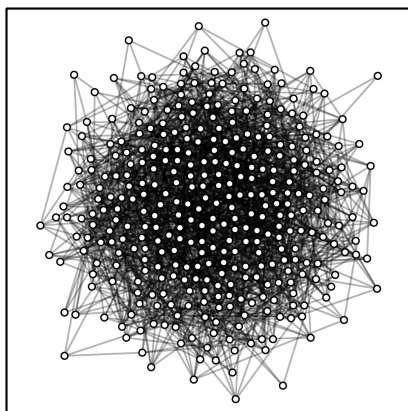
Erdős–Rényi
Gilbert

Chung–Lu
(soft) configuration

Vorhersagen für die bidirektionale BFS



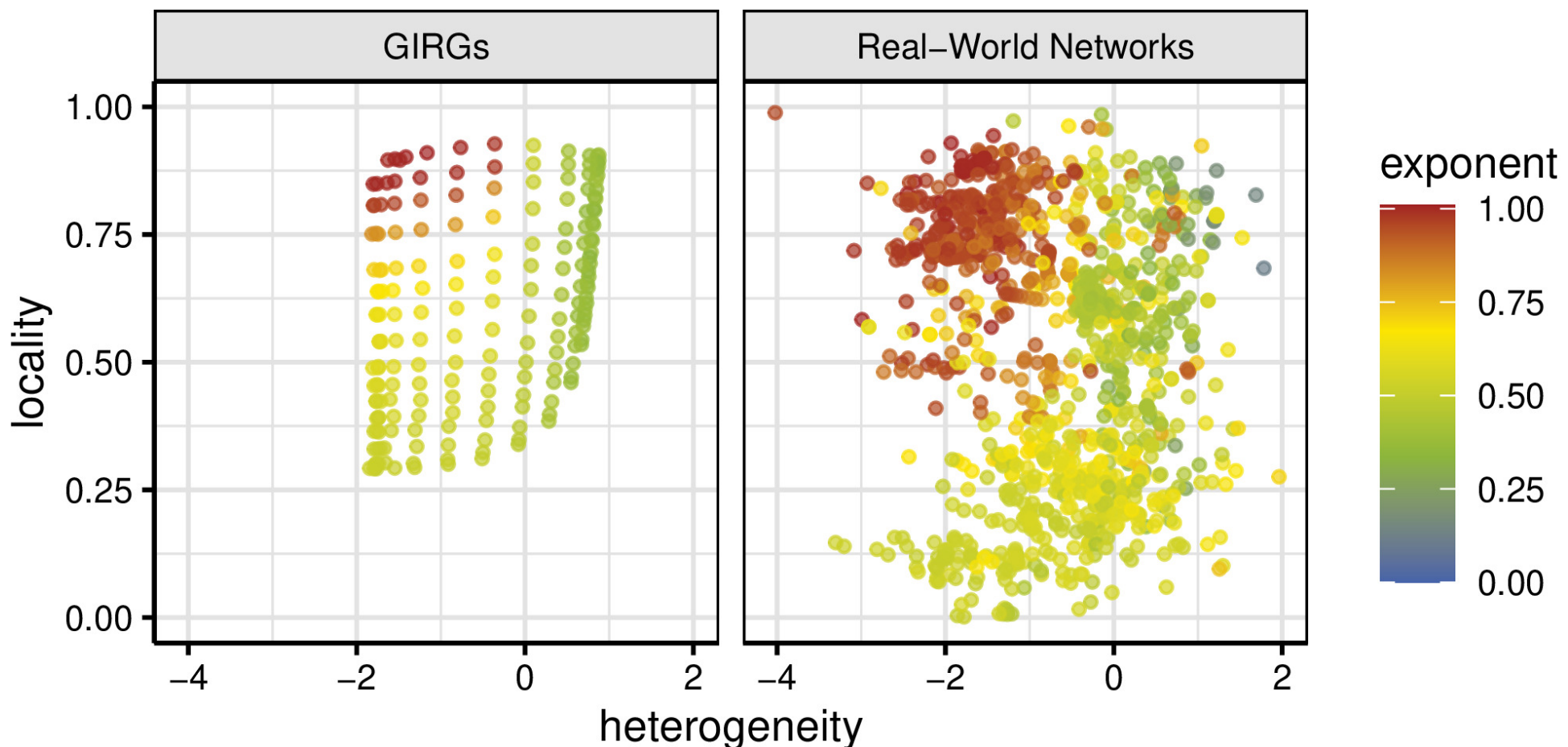
Lokalität



Generalisiert das für praktische Instanzen?

Vorhersage

- konstanter Speedup, wenn homogen und lokal
- sonst: asymptotischer polynomieller Speedup



Zusammenfassung

Hyperbolische Unit-Disk Graphen

- Generalisierung von euklidischen Unit-Disk Graphen
- interessanter Spezialfall: stark hyperbolische UD-Graphen
- viele offene Fragen

Algorithmen auf Zufallsgraphen

- Lokalität und Heterogenität als wichtige Eigenschaften
- Auswirkung dieser Eigenschaften auf Algorithmen
- Geometrie für Lokalität
- euklidisch $\rightarrow$ homogen; hyperbolisch $\rightarrow$ heterogen

Was gibt es sonst noch?

- andere Probleme auf hyperbolischen Zufallsgraphen: Vertex Cover (optimal + Approximation), SAT, Clique
- strukturelle Eigenschaften: Durchmesser, Cliques, Baumweite, ...
- effiziente Generatoren

Bonusfrage: Was meint Drost, wenn er von *Netzwerkeffekten* spricht?

Literaturhinweise

(Stark) hyperbolische Unit-Disk Graphen

- **Routing in Strongly Hyperbolic Unit Disk Graphs** (2021)
Thomas Bläsius, Tobias Friedrich, Maximilian Katzmann, Daniel Stephan
<https://arxiv.org/abs/2107.05518>

Hyperbolische Zufallsgraphen

- **Hyperbolic geometry of complex networks** (2010)
Krioukov, Papadopoulos, Kitsak, Vahdat, Boguñá
<https://doi.org/10.1103/PhysRevE.82.036106>
- **Random Hyperbolic Graphs: Degree Sequence and Clustering** (2012)
Luca Gugelmann, Konstantinos Panagiotou, Ueli Peter
https://doi.org/10.1007/978-3-642-31585-5_51

Bidirektionale Suche in hyperbolischen Zufallsgraphen

- **Efficient Shortest Paths in Scale-Free Networks with Underlying Hyperbolic Geometry** (2018)
Bläsius, Freiberger, Friedrich, Katzmann, Montenegro-Retana, Thieffry
<https://doi.org/10.4230/LIPIcs.ICALP.2018.20>

Ausführlicheres Video zur zweiten Hälfte von heute

- **Theoretical Algorithm Analysis meets Practical Data** (2021)
Thomas Bläsius, PODC-DARE Workshop
<https://www.youtube.com/watch?v=Do2FC3k0JMg>

Effiziente Generierung

- **Sampling Geometric Inhomogeneous Random Graphs in Linear Time** (2017)
Karl Bringmann, Ralph Keusch, Johannes Lengler
<https://doi.org/10.4230/LIPIcs.ESA.2017.20>
- **Efficiently Generating Geometric Inhomogeneous and Hyperbolic Random Graphs** (2019)
Bläsius, Friedrich, Katzmann, Meyer, Penschuck, Weyand
<https://doi.org/10.4230/LIPIcs.ESA.2019.21>

Bonus: Erzwungene Kanten

Situation

- stark hyperbolische UD-Repräsentation
- w liegt bezüglich Winkel zwischen u und v
- w liegt weiter innen als u und v

Theorem

Wenn $\{u, v\} \in E$, dann auch $\{u, w\}, \{v, w\} \in E$.

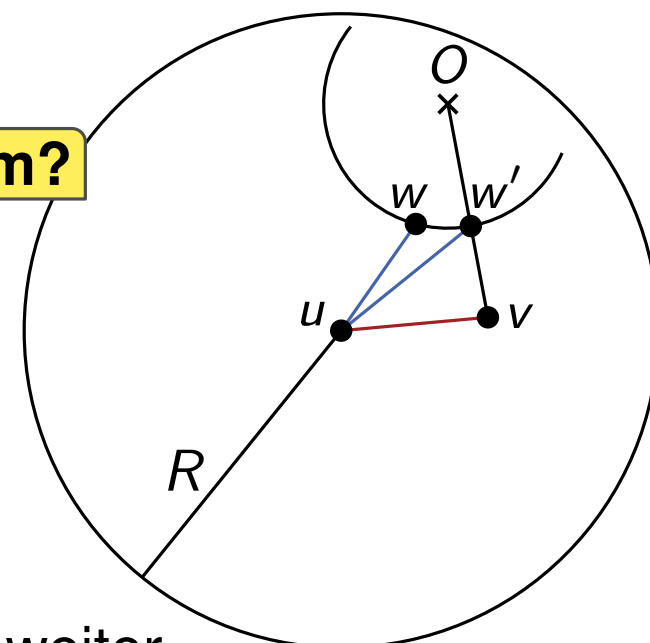
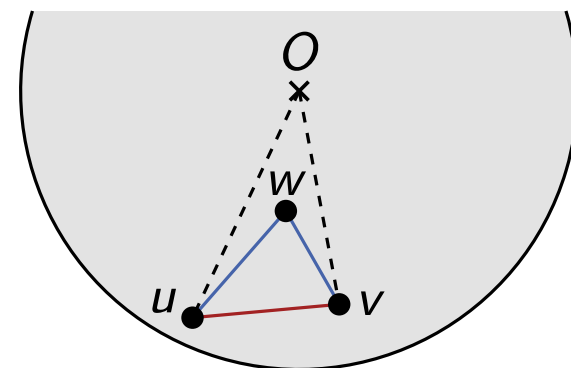
Beweis

- Kreis um u mit Radius R enthält v und O
- Kreis sind konvex $\Rightarrow$ Strecke $\overline{vO}$ liegt im Kreis
- w' liegt in dem Kreis $\Rightarrow d(u, w') \leq R$
- Distanz ist monoton bzgl. Winkeldifferenz
 $\Rightarrow d(u, w) \leq R$

Warum?

Anmerkung

- man kann also nicht unbemerkt außen an einem weiter innen liegenden Knoten vorbeilaufen
- hierarchische Strukturen: je weiter innen desto höher in der Hierarchie



(schematische Darstellung)