Karlsruhe Institute of Technology

# Probability and Computing – Random Graphs

Stefan Walzer | WS 2025/2026

# Organisation: Oral Exam

Karlsruhe Institute of Technology

## Dates and Time Slots

| | | |
|---|---|---|
| | 19.2. | *date of the last lecture* |
| We | 11.03. | Oral Exams |
| Th | 12.03. | Oral Exams |
| Fr | 13.03. | Oral Exams |
| We | 25.3. | Oral Exams |
| Th | 26.3. | Oral Exams |
| Fr | 27.3. | Oral Exams |

*Other dates may be possible on request.*

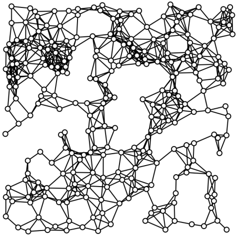
Available time slots:

- 10:00, 10:25, 10:50, 11:15
- 14:00, 14:25, 14:50, 15:15

- Registration via our secretary:
  - Anja Blancani (blancani@kit.edu)
  - cc to me (stefan.walzer@kit.edu)
  - Please specify:
    - your full name
    - matriculation number
    - subject of study (Studienfach)
    - version of the exam regulation (Version der Prüfungsordnung)
- Cancellation also via our secretary
- Location: Stefan's Office (50.34, Room 209).
- duration: 20 minutes
- scope: content of lectures *and exercises*

# Content

# Motivation 1: Average Case Analysis



## Theory-Practice Gap

Minimum Vertex Cover is APX-hard $\overset{???}{\longleftrightarrow}$ small vertex covers can often be computed efficiently in practice

$\rightsquigarrow$ relevant graph classes (e.g. social networks) are not worst-case.

## Bridging the Gap

1. Define a distribution $\mathcal{G}$ on graphs.
   - $\mathcal{G}$ should be realistic, i.e. model real world instances
   - $\mathcal{G}$ should have simple mathematical structure
2. Consider randomised complexity of handling $G \sim \mathcal{G}$.

## Goals

- model real world instances
- identify useful properties of these instances
- build algorithms exploiting these properties

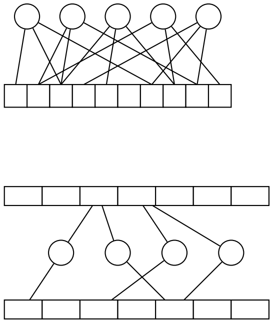
# Motivation 2: Data Structure Design

Karlsruhe Institute of Technology

**Stay tuned**

Random graphs occur naturally in

- cuckoo hash tables
- retrieval data structures
- perfect hash functions



# Motivation 3: Probabilistic Method

**Probabilistic Method for Graph Theory**

Show that graphs with a property $P$ exist by showing that a random graph $G$ satisfies $\Pr[G \text{ has } P] > 0$.

(we're not doing this)

# Content

# The Erdős-Renyi Model and Related Distributions

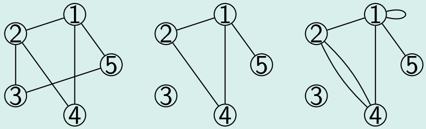
KIT
Karlsruhe Institute of Technology

## Original Erdős-Renyi Model $G(n, m)$: "Uniformly random graph with $n$ nodes and $m$ edges"

### Definition

Let $n \in \mathbb{N}$, $0 \leq m \leq \binom{n}{2}$. We use $G(n, m)$ to refer to a graph sampled uniformly from the set of all graphs with vertex set $[n]$ and $m$ edges.

### Example: $n = 5, m = 6$




probability $1/\binom{\binom{n}{2}}{m}$ 0 0

# The Erdős-Renyi Model and Related Distributions

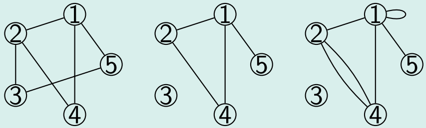
Original Erdős-Renyi Model $G(n, m)$: "Uniformly random graph with $n$ nodes and $m$ edges"

Gilbert Model $G(n, p)$: "Every edge with probability $p$"

## Definition

Let $n \in \mathbb{N}$ and $p \in (0, 1)$. We use $G(n, p)$ to refer to a graph with vertex set $[n]$ that contains each of the $\binom{n}{2}$ possible edges with probability $p$, independently from other edges.

## Example: $n = 5$




| probability | $p^6(1-p)^4$ | $p^4(1-p)^6$ | 0 |
|---|---|---|---|

# The Erdős-Renyi Model and Related Distributions

Original Erdős-Renyi Model $G(n, m)$: "Uniformly random graph with $n$ nodes and $m$ edges"

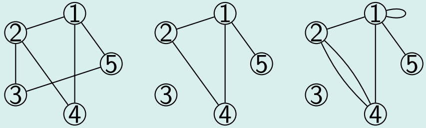
Gilbert Model $G(n, p)$: "Every edge with probability $p$"

Uniform Endpoint Model $G^{\mathrm{UE}}(n, m)$: "randomly attach the $2m$ endpoints of edges"

**Definition**

Let $n, m \in \mathbb{N}$ and $v_1, \ldots, v_{2m} \sim \mathcal{U}([n])$. We use $G^{\mathrm{UE}}(n, m)$ to refer to a multi-graph with vertex set $[n]$ and a multiset of edges that contains a copy of $\{v_{2i-1}, v_{2i}\}$ for each $i \in [m]$.

**Example:** $n = 5, m = 6$




probability $6! \cdot 2^6 \cdot 5^{-12}$   $0$   $6! \cdot 2^4 \cdot 5^{-12}$

# The Erdős-Renyi Model and Related Distributions

Original Erdős-Renyi Model $G(n, m)$: "Uniformly random graph with $n$ nodes and $m$ edges"

Gilbert Model $G(n, p)$: "Every edge with probability $p$"

Uniform Endpoint Model $G^{\mathrm{UE}}(n, m)$: "randomly attach the $2m$ endpoints of edges"

## Remarks

- for $p = m/\binom{n}{2}$ the three distributions are similar in many ways
- the original Erdős-Renyi model is often inconvenient to work with
- the uniform endpoint model is non-standard (we'll need it in later chapters)

# Plan for the Next Few Slides: Sparse Graphs

## Focus on Expected Degree $\lambda \in \mathcal{O}(1)$

- for $G(n, m)$ choose $m = \frac{\lambda n}{2}$ $\Rightarrow$ average vertex degree $\frac{2m}{n} = \lambda$
- for $G(n, p)$ choose $p = \frac{\lambda}{n-1}$ $\Rightarrow$ expected vertex degree $(n-1) \cdot p = \lambda$
- for $G^{\mathrm{UE}}(n, m)$ choose $m = \frac{\lambda n}{2}$ $\Rightarrow$ average vertex degree $\frac{2m}{n} = \lambda$ // loops contribute 2 to a vertex degree

# Plan for the Next Few Slides: Sparse Graphs

Karlsruhe Institute of Technology

## Focus on Expected Degree $\lambda \in \mathcal{O}(1)$

- for $G(n, m)$ choose $m = \frac{\lambda n}{2}$ $\Rightarrow$ average vertex degree $\frac{2m}{n} = \lambda$
- for $G(n, p)$ choose $p = \frac{\lambda}{n-1}$ $\Rightarrow$ expected vertex degree $(n-1) \cdot p = \lambda$
- for $G^{\text{UE}}(n, m)$ choose $m = \frac{\lambda n}{2}$ $\Rightarrow$ average vertex degree $\frac{2m}{n} = \lambda$ // loops contribute 2 to a vertex degree

## Goals

- Build intuition for properties of Erdős-Renyi graphs.
- Get a feeling for how to work with them.
- For simplicity: Focus on the Gilbert model only.

# Selected Properties of Erdős-Renyi Graphs

On the next few slides we consider:

**Vertex Degrees**

For large $n$, the degree of a given vertex is approximately Poisson distributed.

**Local Structure**

The neighbourhood around a vertex resembles a Galton-Watson tree.

**Degree Statistics**

The number of vertices of each degree is highly concentrated around its expectation.

**Largest Connected Component**

Size of the largest component is highly predictable.

# Content

# Degree Distribution of Erdős-Reyni Graphs

**Exercise: Degrees are approximately Poisson distributed**

For each $n \in \mathbb{N}$ consider $G(n, \lambda/n)$ and the degree $X_n \sim \text{Bin}(n-1, \lambda/n)$ of vertex 1. Moreover, let $X \sim \text{Pois}(\lambda)$. Then

$$X_n \xrightarrow{d} X \text{ for } n \to \infty.$$

The same holds for $G(n, \lfloor \lambda n/2 \rfloor)$ and $G^{\text{UE}}(n, \lambda n/2)$.

# Content

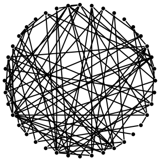
# The Number $N_d$ of Vertices of Degree $N_d$

## Notation

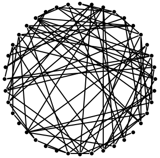
- Let $d \in \mathbb{N}$, $\lambda > 0$. We consider $G(n, \lambda/n)$. // Gilbert model
- Let $N_d := |\{v \in [n] \mid \deg(v) = d\}|$

## Is $N_d$ highly concentrated?

- Note: $(\deg(v))_{v \in [n]}$ are correlated.
- Otherwise $N_d$ would have a binomial distribution and we could use Chernoff bounds.



| $d$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| $N_d$ | 0 | 2 | 8 | 6 | 7 | 7 | 3 | 2 | 3 | 1 |



| $d$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| $N_d$ | 1 | 2 | 5 | 8 | 9 | 11 | 3 | 1 | 0 | 0 |

# Concentration of $N_d$

**Lemma (Near Independence of Degrees)**

Let $u \neq v$ be two vertices of $G(n, \lambda/n)$. Then $\Pr[\deg(u) = d, \deg(v) = d] = \Pr[\deg(u) = d]\Pr[\deg(v) = d] \pm \Theta(1/n)$.

# Concentration of $N_d$

## Lemma (Near Independence of Degrees)

Let $u \neq v$ be two vertices of $G(n, \lambda/n)$. Then $\Pr[\deg(u) = d, \deg(v) = d] = \Pr[\deg(u) = d] \Pr[\deg(v) = d] \pm \Theta(1/n)$.

## Proof

Let $\deg'(u) = \deg(u) - [\{u, v\} \in E]$ be the degree of $u$ when ignoring $\{u, v\}$ if present. Then

$$\Pr[\deg(u) \neq \deg'(u)] = \Pr[\{u, v\} \in E] = \lambda/n = \Theta(1/n).$$

The same holds for $\deg'(v) = \deg(v) - [\{u, v\} \in E]$. We conclude:

$$\begin{aligned}\Pr[\deg(v_1) = d, \deg(v_2) = d] &= \Pr[\deg'(v_1) = d, \deg'(v_2) = d] \pm \Theta(1/n) \\ &= \Pr[\deg'(v_1) = d] \Pr[\deg'(v_2) = d] \pm \Theta(1/n) = \Pr[\deg(v_1) = d] \Pr[\deg(v_2) = d] \pm \Theta(1/n).\end{aligned}$$

# Concentration of $N_d$

## Lemma (Near Independence of Degrees)

Let $u \neq v$ be two vertices of $G(n, \lambda/n)$. Then $\Pr[\deg(u) = d, \deg(v) = d] = \Pr[\deg(u) = d] \Pr[\deg(v) = d] \pm \Theta(1/n)$.

## Theorem

$\Pr[|N_d - np_d| \geq n^{2/3}] = \mathcal{O}(n^{-1/3})$ where $p_d = \Pr[\deg(1) = d] \approx e^{-\lambda} \frac{\lambda^d}{d!}$.

# Concentration of $N_d$

**Lemma (Near Independence of Degrees)**

Let $u \neq v$ be two vertices of $G(n, \lambda/n)$. Then $\Pr[\deg(u) = d, \deg(v) = d] = \Pr[\deg(u) = d] \Pr[\deg(v) = d] \pm \Theta(1/n)$.

**Theorem**

$\Pr[|N_d - np_d| \geq n^{2/3}] = \mathcal{O}(n^{-1/3})$ where $p_d = \Pr[\deg(1) = d] \approx e^{-\lambda} \frac{\lambda^d}{d!}$.

**Proof**

$\mathbb{E}[N_d] = ?$ $\mathbb{E}[N_d^2] = ?$ $\mathrm{Var}(N_d) = ?.$

# Concentration of $N_d$

## Lemma (Near Independence of Degrees)

Let $u \neq v$ be two vertices of $G(n, \lambda/n)$. Then $\Pr[\deg(u) = d, \deg(v) = d] = \Pr[\deg(u) = d] \Pr[\deg(v) = d] \pm \Theta(1/n)$.

## Theorem

$\Pr[|N_d - np_d| \geq n^{2/3}] = \mathcal{O}(n^{-1/3})$ where $p_d = \Pr[\deg(1) = d] \approx e^{-\lambda} \frac{\lambda^d}{d!}$.

## Proof

$\mathbb{E}[N_d] = ?$ $\quad \mathbb{E}[N_d^2] = ?$ $\quad \mathrm{Var}(N_d) = ?$.

$$\mathbb{E}[N_d] = \mathbb{E}[\sum_{v \in [n]} [\deg(v) = d]] = n \cdot \Pr[\deg(1) = d] = n \cdot p_d.$$

# Concentration of $N_d$

## Lemma (Near Independence of Degrees)

Let $u \neq v$ be two vertices of $G(n, \lambda/n)$. Then $\Pr[\deg(u) = d, \deg(v) = d] = \Pr[\deg(u) = d] \Pr[\deg(v) = d] \pm \Theta(1/n)$.

## Theorem

$\Pr[|N_d - np_d| \geq n^{2/3}] = \mathcal{O}(n^{-1/3})$ where $p_d = \Pr[\deg(1) = d] \approx e^{-\lambda}\frac{\lambda^d}{d!}$.

## Proof

$\mathbb{E}[N_d] = np_d$ $\qquad$ $\mathbb{E}[N_d^2] = ?$ $\qquad$ $\mathrm{Var}(N_d) = ?$.

$$\begin{aligned}
\mathbb{E}[N_d^2] &= \mathbb{E}\big[(\sum_{v\in[n]} [\deg(v) = d])^2\big] = \mathbb{E}\big[\sum_{u\in[n]}\sum_{v\in[n]} [\deg(u) = d, \deg(v) = d]\big] \\
&= \sum_{u\in[n]}\sum_{v\in[n]} \Pr[\deg(u) = d, \deg(v) = d] = \sum_{u\in[n]} \Pr[\deg(u) = d] + \sum_{u\in[n]}\sum_{v\neq u} \Pr[\deg(u) = d, \deg(v) = d] \\
&= n \cdot p_d + n \cdot (n-1) \cdot (p_d^2 \pm \mathcal{O}(1/n)) = n^2 p_d^2 \pm \mathcal{O}(n).
\end{aligned}$$

# Concentration of $N_d$

Karlsruhe Institute of Technology

## Lemma (Near Independence of Degrees)

Let $u \neq v$ be two vertices of $G(n, \lambda/n)$. Then $\Pr[\deg(u) = d, \deg(v) = d] = \Pr[\deg(u) = d] \Pr[\deg(v) = d] \pm \Theta(1/n)$.

## Theorem

$\Pr[|N_d - np_d| \geq n^{2/3}] = \mathcal{O}(n^{-1/3})$ where $p_d = \Pr[\deg(1) = d] \approx e^{-\lambda} \frac{\lambda^d}{d!}$.

## Proof

$\mathbb{E}[N_d] = np_d$ $\qquad$ $\mathbb{E}[N_d^2] = n^2 p_d^2 \pm \mathcal{O}(n)$ $\qquad$ $\mathrm{Var}(N_d) = ?.$

$$\mathrm{Var}(N_d) = \mathbb{E}[N_d^2] - \mathbb{E}[N_d]^2 \leq n^2 p_d^2 + \mathcal{O}(n) - (np_d)^2 = \mathcal{O}(n).$$

# Concentration of $N_d$

Karlsruhe Institute of Technology

**Lemma (Near Independence of Degrees)**

Let $u \neq v$ be two vertices of $G(n, \lambda/n)$. Then $\Pr[\deg(u) = d, \deg(v) = d] = \Pr[\deg(u) = d] \Pr[\deg(v) = d] \pm \Theta(1/n)$.

**Theorem**

$\Pr[|N_d - np_d| \geq n^{2/3}] = \mathcal{O}(n^{-1/3})$ where $p_d = \Pr[\deg(1) = d] \approx e^{-\lambda} \frac{\lambda^d}{d!}$.

**Proof**

$\mathbb{E}[N_d] = np_d$ $\qquad$ $\mathbb{E}[N_d^2] = n^2 p_d^2 \pm \mathcal{O}(n)$ $\qquad$ $\mathrm{Var}(N_d) = \mathcal{O}(n)$.

$$\text{Hence: } \Pr[|N_d - np_d| \geq n^{2/3}] = \Pr[|N_d - \mathbb{E}[N_d]| \geq n^{2/3}] \overset{\text{Cheb.}}{\leq} \frac{\mathrm{Var}(N_d)}{n^{4/3}} = \mathcal{O}(n^{-1/3}).$$

# Content

# Erdős-Renyi Graphs have Few Cycles

**Theorem: There are few short cycles in Erdős-Renyi graphs**

Let $C_k$ be the number of cycles of length $k$ in $G(n, \lambda/n)$ where $k, \lambda = \Theta(1)$. Then $\mathbb{E}[C_k] \leq \frac{\lambda^k}{2k} = \Theta(1)$.

# Erdős-Renyi Graphs have Few Cycles

Karlsruhe Institute of Technology

## Theorem: There are few short cycles in Erdős-Renyi graphs

Let $C_k$ be the number of cycles of length $k$ in $G(n, \lambda/n)$ where $k, \lambda = \Theta(1)$. Then $\mathbb{E}[C_k] \leq \frac{\lambda^k}{2k} = \Theta(1)$.

## Proof.

The number of potential cycles is $\underbrace{n(n-1)\cdot \ldots \cdot (n-k+1)}_{\text{sequences } (v_1, \ldots, v_k)} \cdot \underbrace{\frac{1}{k} \cdot \frac{1}{2}}_{\text{startpoint and direction irrelevant}}$

The probability that $(v_1, \ldots, v_k, v_1)$ is a cycle is $(\lambda/n)^k$. Hence:

$$\mathbb{E}[C_k] \leq \frac{n^k}{2k}\left(\frac{\lambda}{n}\right)^k = \frac{\lambda^k}{2k}. \qquad \square$$

# The Galton-Watson Branching Process

## Definition

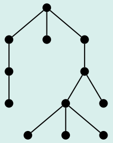
Let $\mathcal{D}$ be a distribution on $\mathbb{N}_0$ and $X_{i,j} \sim \mathcal{D}$ for $i, j \in \mathbb{N}$. Define $Z_0 = 1$ and $Z_i = \sum_{j=1}^{Z_{i-1}} X_{i,j}$ for $i \geq 1$.

## Intuition

- Start with a population of size $Z_1 = 1$.
- Each individual has a random number of decendents.
- Key question: What is the probability of extinction, i.e. for $\lim_{i \to \infty} Z_i = 0$?

## Galton-Watson *Tree*

| $Z_i$ |
|---|
| 1 |
| 3 |
| 2 |
| 3 |
| 3 |
| 0 |



| $X_{i,j}$ | 1 | 2 | 3 | 4 | ... |
|---|---|---|---|---|---|
| 1 | 3 | 1 | 0 | 2 | ... |
| 2 | 1 | 0 | 1 | 3 | ... |
| 3 | 1 | 2 | 2 | 0 | ... |
| 4 | 0 | 3 | 0 | 0 | ... |
| 5 | 0 | 0 | 0 | 2 | ... |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | |

# The Galton-Watson Branching Process

## Definition

Let $\mathcal{D}$ be a distribution on $\mathbb{N}_0$ and $X_{i,j} \sim \mathcal{D}$ for $i, j \in \mathbb{N}$. Define $Z_0 = 1$ and $Z_i = \sum_{j=1}^{Z_{i-1}} X_{i,j}$ for $i \geq 1$.

## Intuition

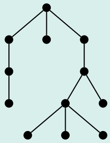
- Start with a population of size $Z_1 = 1$.
- Each individual has a random number of decendents.
- Key question: What is the probability of extinction, i.e. for $\lim_{i\to\infty} Z_i = 0$?

## Exercise: Galton-Watson Process with $\mathcal{D} = \text{Pois}(\lambda)$

If $\lambda \leq 1$ then the process goes extinct with probability 1.
If $\lambda > 1$ then the process survives with probability $s_\lambda > 0$.

## Galton-Watson *Tree*

| $Z_i$ |
|---|
| 1 |
| 3 |
| 2 |
| 3 |
| 3 |
| 0 |



| $X_{i,j}$ | 1 | 2 | 3 | 4 | ... |
|---|---|---|---|---|---|
| 1 | 3 | 1 | 0 | 2 | ... |
| 2 | 1 | 0 | 1 | 3 | ... |
| 3 | 1 | 2 | 2 | 0 | ... |
| 4 | 0 | 3 | 0 | 0 | ... |
| 5 | 0 | 0 | 0 | 2 | ... |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | |

# Local Structure of Erdős-Renyi Graphs

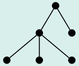
Karlsruhe Institute of Technology

## Theorem: The Neighbourhood of $v$ looks like a Galton Watson Tree

Let $R = \mathcal{O}(1)$. Let $H$ be an (ordered) tree of depth $R$ given by a sequence $c_1, \ldots, c_k$ specifiying the number of children of nodes in all layers except the last, in level order.
Let $\mathrm{GWT}(\lambda)|_R$ be the first $R$ layers of a $\mathrm{Pois}(\lambda)$-Galton-Watson tree.

$$\Pr[\mathrm{GWT}(\lambda)|_R = H] \overset{\text{(i)}}{=} \prod_{i=1}^{k} \Pr_{X\sim\mathrm{Pois}(\lambda)}[X = c_i] = \prod_{i=1}^{k} e^{-\lambda}\frac{\lambda^{c_i}}{c_i!}$$

## Example for $H$



$(c_1, c_2, c_3) = (2, 3, 0)$

# Local Structure of Erdős-Renyi Graphs

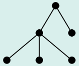
Karlsruhe Institute of Technology

## Theorem: The Neighbourhood of $v$ looks like a Galton Watson Tree

Let $R = \mathcal{O}(1)$. Let $H$ be an (ordered) tree of depth $R$ given by a sequence $c_1, \dots, c_k$ specifiying the number of children of nodes in all layers except the last, in level order.
Let $\mathrm{GWT}(\lambda)|_R$ be the first $R$ layers of a $\mathrm{Pois}(\lambda)$-Galton-Watson tree.
Let $G(n, \lambda/n)|_{v,R}$ be the (ordered) subgraph of $G(n, \lambda/n)$ induced by vertices with distance $\leq R$ from $v$.

$$\Pr[\mathrm{GWT}(\lambda)|_R = H] \overset{\text{(i)}}{=} \prod_{i=1}^{k} \Pr_{X \sim \mathrm{Pois}(\lambda)}[X = c_i] = \prod_{i=1}^{k} e^{-\lambda} \frac{\lambda^{c_i}}{c_i!} \overset{\text{(ii)}}{\approx} \Pr[G(n, \lambda/n)|_{v,R} = H].$$

## Example for $H$



$(c_1, c_2, c_3) = (2, 3, 0)$

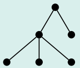
# Local Structure of Erdős-Renyi Graphs

## Theorem: The Neighbourhood of $v$ looks like a Galton Watson Tree

Let $R = \mathcal{O}(1)$. Let $H$ be an (ordered) tree of depth $R$ given by a sequence $c_1, \dots, c_k$ specifiying the number of children of nodes in all layers except the last, in level order.
Let $\mathrm{GWT}(\lambda)|_R$ be the first $R$ layers of a $\mathrm{Pois}(\lambda)$-Galton-Watson tree.
Let $G(n, \lambda/n)|_{v,R}$ be the (ordered) subgraph of $G(n, \lambda/n)$ induced by vertices with distance $\leq R$ from $v$.

$$\Pr[\mathrm{GWT}(\lambda)|_R = H] \overset{\text{(i)}}{=} \prod_{i=1}^{k} \Pr_{X \sim \mathrm{Pois}(\lambda)}[X = c_i] = \prod_{i=1}^{k} e^{-\lambda} \frac{\lambda^{c_i}}{c_i!} \overset{\text{(ii)}}{\approx} \Pr[G(n, \lambda/n)|_{v,R} = H].$$

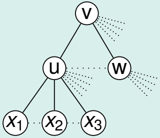
## Example for $H$



$(c_1, c_2, c_3) = (2, 3, 0)$

## Proof of (ii) by Example: The following has to "go right" for $G(n, \lambda/n)|_{v,R} = H$

| random variable | desired outcome | probability |
|---|---|---|
| $\deg(v) \sim \mathrm{Bin}(n-1, \lambda/n) \approx \mathrm{Pois}(\lambda)$ | 2 | $\approx e^{-\lambda}\frac{\lambda^2}{2!}$ |
| $[\{u, w\} \in E]$ | 0 | $1 - \frac{\lambda}{n} \approx 1$ |
| $\deg(u) - 1 \sim \mathrm{Bin}(n-3, \lambda/n) \approx \mathrm{Pois}(\lambda)$ | 3 | $\approx e^{-\lambda}\frac{\lambda^3}{3!}$ |
| $\deg(w) - 1 \sim \mathrm{Bin}(n-3, \lambda/n) \approx \mathrm{Pois}(\lambda)$ | 0 | $\approx e^{-\lambda}\frac{\lambda^0}{0!}$ |
| $[\{x_1, x_2\} \in E \vee \{x_2, x_3\} \in E \vee \{x_1, x_3\} \in E]$ | 0 | $(1 - \frac{\lambda}{n})^3 \approx 1$ |



# Local Structure of Erdős-Renyi Graphs

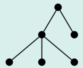
Karlsruhe Institute of Technology

## Theorem: The Neighbourhood of $v$ looks like a Galton Watson Tree

Let $R = \mathcal{O}(1)$. Let $H$ be an (ordered) tree of depth $R$ given by a sequence $c_1, \dots, c_k$ specifiying the number of children of nodes in all layers except the last, in level order.
Let $\mathrm{GWT}(\lambda)|_R$ be the first $R$ layers of a $\mathrm{Pois}(\lambda)$-Galton-Watson tree.
Let $G(n, \lambda/n)|_{v,R}$ be the (ordered) subgraph of $G(n, \lambda/n)$ induced by vertices with distance $\leq R$ from $v$.

$$\Pr[\mathrm{GWT}(\lambda)|_R = H] \overset{\text{(i)}}{=} \prod_{i=1}^{k} \Pr_{X\sim\mathrm{Pois}(\lambda)}[X = c_i] = \prod_{i=1}^{k} e^{-\lambda}\frac{\lambda^{c_i}}{c_i!} \overset{\text{(ii)}}{\approx} \Pr[G(n, \lambda/n)|_{v,R} = H].$$

## Example for $H$



$(c_1, c_2, c_3) = (2, 3, 0)$

## Corollaries

- $G(n, \lambda/n)|_{v,R} \xrightarrow{d} \mathrm{GWT}(\lambda)|_R$ // convergence in distribution for $n \to \infty$
- The number $N_H$ of "copies" of $H$ in $G(n, \lambda/n)$ satisfies $\mathbb{E}[N_H] \approx n \cdot \prod_{i=1}^{k} e^{-\lambda}\frac{\lambda^{c_i}}{c_i!}$.
  Concentration of $N_H$ can be proved much like we proved concentration of $N_d$ earlier.

# Content

# How does $G(n, \lambda/n)$ look like for different $\lambda$?

Karlsruhe Institute of Technology

## Theorem: Sudden Emergence of the Giant Component (Erdős, Renyi 1960)

Consider $G(n, \lambda/n)$. The following holds with probability approaching 1 for $n \to \infty$.

i. If $\lambda < 1$ then $G(n, \lambda/n)$ only has components of size $\mathcal{O}(\log n)$.
   Each component is a tree or pseudotree. // pseudotree: connected and # edges = # vertices
   $\hookrightarrow$ Intuition: GWT($\lambda$) dies out with probability 1.

ii. If $\lambda > 1$ then $G(n, \lambda/n)$ has one "giant" component of size $\approx s(\lambda) \cdot n$.
   $\hookrightarrow$ Intuition: $s(\lambda) > 0$ is the probability that GWT($\lambda$) is infinite
   $\approx$ probability that fixed vertex is in giant component

iii. If $\lambda = 1$ then the largest component of $G(n, \lambda/n)$ has size $\Theta(n^{2/3})$.
   $\hookrightarrow$ Intuition: ?

wild pseudotree
https://crowspath.org/love-trees/

# Content

# Locality: A Property of Networks in Practice

Karlsruhe Institute of Technology

## Observation: Locality in Practice

Take social networks. A friend of my friend is more likely to be my friend than a random person.

## Definition: Locality[1]

$L = \Pr[\{u, w\} \in E \mid \{v, u\} \in E \wedge \{v, w\} \in E]$ where $v, u, w$ are distinct (random) vertices.

[1] Similar numbers are sometimes called *clustering coefficient*.

# Locality: A Property of Networks in Practice

## Observation: Locality in Practice

Take social networks. A friend of my friend is more likely to be my friend than a random person.

## Definition: Locality[1]

$L = \Pr[\{u, w\} \in E \mid \{v, u\} \in E \wedge \{v, w\} \in E]$ where $v, u, w$ are distinct (random) vertices.

[1] Similar numbers are sometimes called *clustering coefficient*.

## Observation: No Locality in Erdős-Renyi Random Graphs

In $G(n, \lambda/n)$ we have $L = \frac{\lambda}{n} = \mathcal{O}(\frac{1}{n})$.

# Locality: A Property of Networks in Practice

Karlsruhe Institute of Technology

## Observation: Locality in Practice

Take social networks. A friend of my friend is more likely to be my friend than a random person.

## Definition: Locality[1]

$L = \Pr[\{u, w\} \in E \mid \{v, u\} \in E \wedge \{v, w\} \in E]$ where $v, u, w$ are distinct (random) vertices.

[1] Similar numbers are sometimes called *clustering coefficient*.

## Observation: No Locality in Erdős-Renyi Random Graphs

In $G(n, \lambda/n)$ we have $L = \frac{\lambda}{n} = \mathcal{O}(\frac{1}{n})$.

## Next: Random *Geometric* Graphs with $L = \Omega(1)$.

# Random Geometric Graphs

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

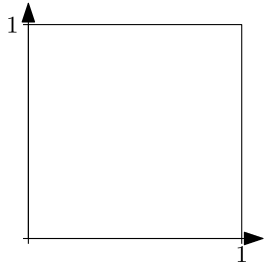
# Random Geometric Graphs

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple



# Random Geometric Graphs

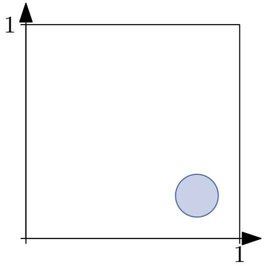
Karlsruhe Institute of Technology

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple



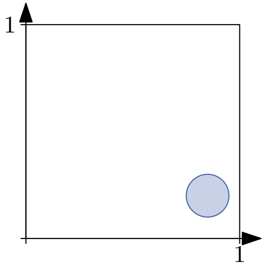
# Random Geometric Graphs

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple



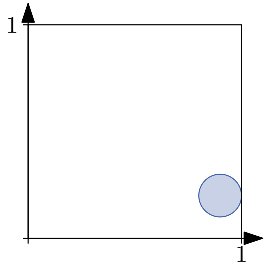
# Random Geometric Graphs

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple



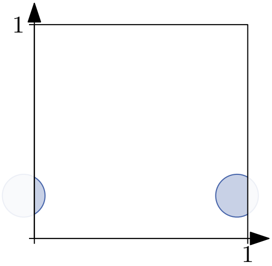
# Random Geometric Graphs

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple



# Random Geometric Graphs

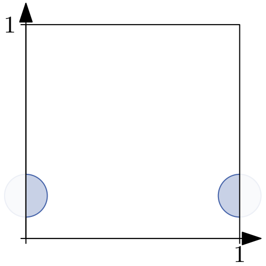
Karlsruhe Institute of Technology

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple



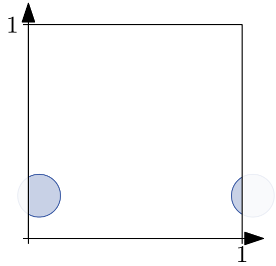
# Random Geometric Graphs

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple



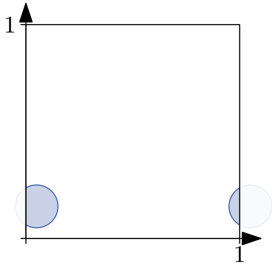
# Random Geometric Graphs

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple



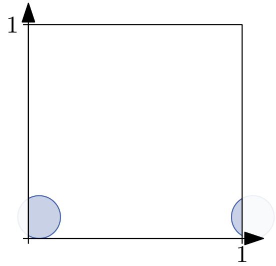
# Random Geometric Graphs

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple



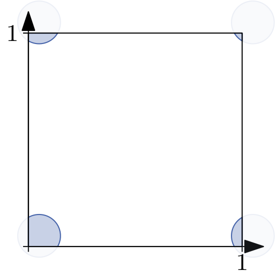
# Random Geometric Graphs

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple



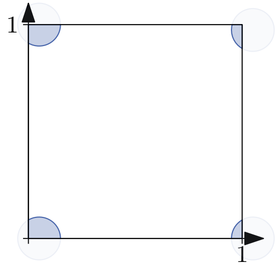
# Random Geometric Graphs

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple



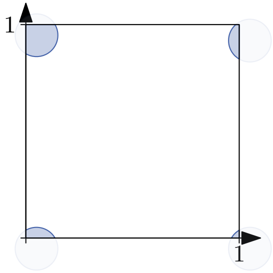
# Random Geometric Graphs

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple



# Random Geometric Graphs

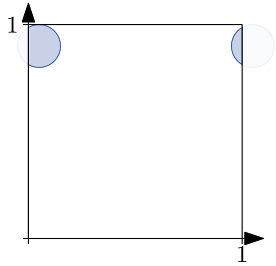
Karlsruhe Institute of Technology

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple



# Random Geometric Graphs

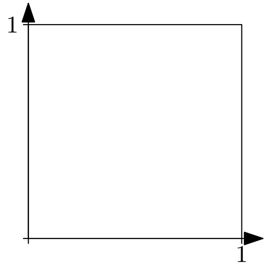
Karlsruhe Institute of Technology

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple



# Random Geometric Graphs

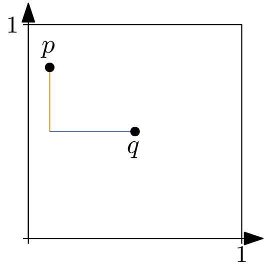
Karlsruhe Institute of Technology

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple
- metric: $L_\infty$ // $L_2$ is more common but less simple
  $\hookrightarrow \text{dist}((x_1, y_1), (x_2, y_2)) = \max(\text{dist}(x_1, x_2), \text{dist}(y_1, y_2))$.



# Random Geometric Graphs

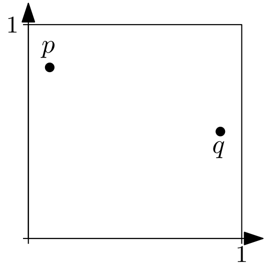
Karlsruhe Institute of Technology

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple
- metric: $L_\infty$ // $L_2$ is more common but less simple
  $\hookrightarrow \mathrm{dist}((x_1, y_1), (x_2, y_2)) = \max(\mathrm{dist}(x_1, x_2), \mathrm{dist}(y_1, y_2))$.



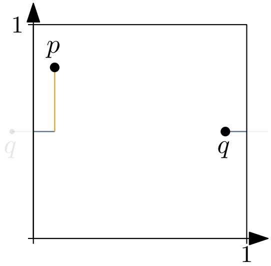
# Random Geometric Graphs

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple
- metric: $L_\infty$ // $L_2$ is more common but less simple
  $\hookrightarrow$ $\text{dist}((x_1, y_1), (x_2, y_2)) = \max(\text{dist}(x_1, x_2), \text{dist}(y_1, y_2))$.



# Random Geometric Graphs

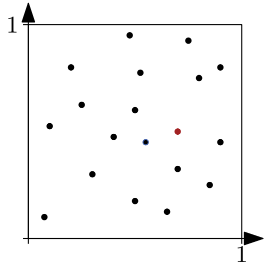
Karlsruhe Institute of Technology

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple
- metric: $L_\infty$ // $L_2$ is more common but less simple
  $\hookrightarrow \mathrm{dist}((x_1, y_1), (x_2, y_2)) = \max(\mathrm{dist}(x_1, x_2), \mathrm{dist}(y_1, y_2))$.
- vertex distribution: for $v \in [n]$: $P_v \sim \mathcal{U}(\mathbb{T}^2)$



# Random Geometric Graphs

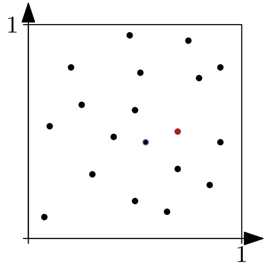
Karlsruhe Institute of Technology

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple
- metric: $L_\infty$ // $L_2$ is more common but less simple
  $\hookrightarrow \text{dist}((x_1, y_1), (x_2, y_2)) = \max(\text{dist}(x_1, x_2), \text{dist}(y_1, y_2))$.
- vertex distribution: for $v \in [n]$: $P_v \sim \mathcal{U}(\mathbb{T}^2)$
- edge “probability” is 0 or 1: $\{u, v\} \in E \Leftrightarrow \text{dist}(P_u, P_v) \le r$
  // not random when $P_u$ and $P_v$ are given



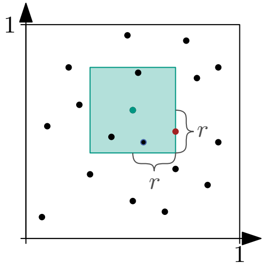
# Random Geometric Graphs

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple
- metric: $L_\infty$ // $L_2$ is more common but less simple
  $\hookrightarrow \text{dist}((x_1, y_1), (x_2, y_2)) = \max(\text{dist}(x_1, x_2), \text{dist}(y_1, y_2))$.
- vertex distribution: for $v \in [n]$: $P_v \sim \mathcal{U}(\mathbb{T}^2)$
- edge "probability" is 0 or 1: $\{u, v\} \in E \Leftrightarrow \text{dist}(P_u, P_v) \leq r$
  // not random when $P_u$ and $P_v$ are given



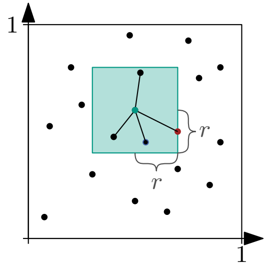
# Random Geometric Graphs

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple
- metric: $L_\infty$ // $L_2$ is more common but less simple
  $\hookrightarrow \text{dist}((x_1, y_1), (x_2, y_2)) = \max(\text{dist}(x_1, x_2), \text{dist}(y_1, y_2))$.
- vertex distribution: for $v \in [n]$: $P_v \sim \mathcal{U}(\mathbb{T}^2)$
- edge "probability" is 0 or 1: $\{u, v\} \in E \Leftrightarrow \text{dist}(P_u, P_v) \leq r$
  // not random when $P_u$ and $P_v$ are given



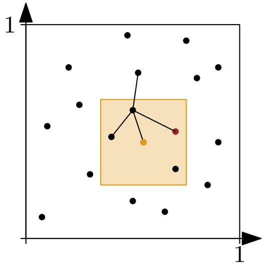
# Random Geometric Graphs

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple
- metric: $L_\infty$ // $L_2$ is more common but less simple
  $\hookrightarrow \text{dist}((x_1, y_1), (x_2, y_2)) = \max(\text{dist}(x_1, x_2), \text{dist}(y_1, y_2))$.
- vertex distribution: for $v \in [n]$: $P_v \sim \mathcal{U}(\mathbb{T}^2)$
- edge “probability” is 0 or 1: $\{u, v\} \in E \Leftrightarrow \text{dist}(P_u, P_v) \leq r$
  // not random when $P_u$ and $P_v$ are given



# Random Geometric Graphs

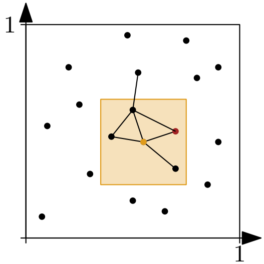
Karlsruhe Institute of Technology

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple
- metric: $L_\infty$ // $L_2$ is more common but less simple
  $\hookrightarrow$ $\text{dist}((x_1, y_1), (x_2, y_2)) = \max(\text{dist}(x_1, x_2), \text{dist}(y_1, y_2))$.
- vertex distribution: for $v \in [n]$: $P_v \sim \mathcal{U}(\mathbb{T}^2)$
- edge "probability" is 0 or 1: $\{u, v\} \in E \Leftrightarrow \text{dist}(P_u, P_v) \leq r$
  // not random when $P_u$ and $P_v$ are given



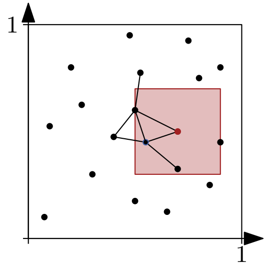
# Random Geometric Graphs

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple
- metric: $L_\infty$ // $L_2$ is more common but less simple
  $\hookrightarrow$ $\text{dist}((x_1, y_1), (x_2, y_2)) = \max(\text{dist}(x_1, x_2), \text{dist}(y_1, y_2))$.
- vertex distribution: for $v \in [n]$: $P_v \sim \mathcal{U}(\mathbb{T}^2)$
- edge "probability" is 0 or 1: $\{u, v\} \in E \Leftrightarrow \text{dist}(P_u, P_v) \leq r$
  // not random when $P_u$ and $P_v$ are given



# Random Geometric Graphs

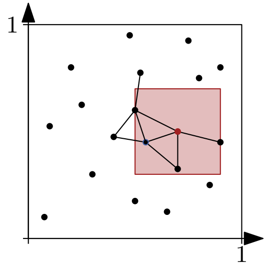
Karlsruhe Institute of Technology

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple
- metric: $L_\infty$ // $L_2$ is more common but less simple
  $\hookrightarrow \text{dist}((x_1, y_1), (x_2, y_2)) = \max(\text{dist}(x_1, x_2), \text{dist}(y_1, y_2))$.
- vertex distribution: for $v \in [n]$: $P_v \sim \mathcal{U}(\mathbb{T}^2)$
- edge "probability" is 0 or 1: $\{u, v\} \in E \Leftrightarrow \text{dist}(P_u, P_v) \leq r$
  // not random when $P_u$ and $P_v$ are given



# Random Geometric Graphs

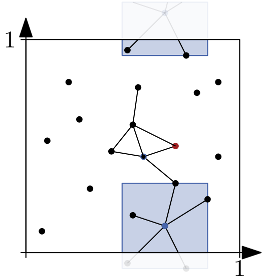
Karlsruhe Institute of Technology

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple
- metric: $L_\infty$ // $L_2$ is more common but less simple
  $\hookrightarrow \text{dist}((x_1, y_1), (x_2, y_2)) = \max(\text{dist}(x_1, x_2), \text{dist}(y_1, y_2))$.
- vertex distribution: for $v \in [n]$: $P_v \sim \mathcal{U}(\mathbb{T}^2)$
- edge "probability" is 0 or 1: $\{u, v\} \in E \Leftrightarrow \text{dist}(P_u, P_v) \leq r$
  // not random when $P_u$ and $P_v$ are given



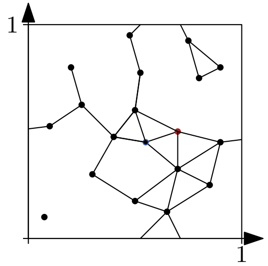
# Random Geometric Graphs

## Definition: Random Geometric Graph (RGG)

An RGG is obtained by distributing vertices in a metric space and connecting any two vertices with a probability depending on their distance.

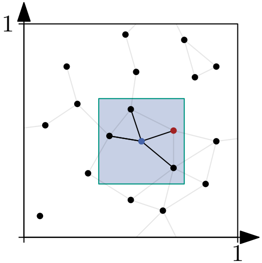
## Simple Example: $G^{\mathbb{T}^2}(n, r)$

- number of vertices: $n$
- space: 2-dimensional torus $\mathbb{T}^2 = [0, 1)^2$
  // standard unit square is more common but less simple
- metric: $L_\infty$ // $L_2$ is more common but less simple
  $\hookrightarrow$ $\text{dist}((x_1, y_1), (x_2, y_2)) = \max(\text{dist}(x_1, x_2), \text{dist}(y_1, y_2))$.
- vertex distribution: for $v \in [n]$: $P_v \sim \mathcal{U}(\mathbb{T}^2)$
- edge “probability” is 0 or 1: $\{u, v\} \in E \Leftrightarrow \text{dist}(P_u, P_v) \leq r$
  // not random when $P_u$ and $P_v$ are given



# Degree Distribution of $G^{\mathbb{T}^2}(n, r)$

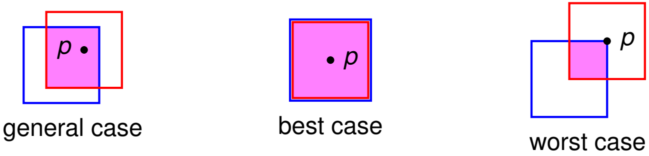
- Consider arbitrary $v \in [n]$.
- By symmetry of $\mathbb{T}^2$ each outcome of $P_v$ behaves the same.
- $\Pr[\{u, v\} \in E] = \Pr[P_u$ is in the $2r \times 2r$ square centered at $P_v] = 4r^2$.
- Hence $\deg(v) \sim \text{Bin}(n-1, 4r^2)$ and $\mathbb{E}[\deg(v)] = 4r^2(n-1)$.



# Locality in $G^{\mathbb{T}^2}(n, r)$

## Observation

Let $p, q \sim \mathcal{U}([-0.5, 0.5]^2)$ and $S_p$ the unit square around $p$.
Then $\Pr[q \in S_p] = \int_{-0.5}^{0.5} \int_{-0.5}^{0.5} (1 - |x|)(1 - |y|) dx\, dy = \frac{9}{16}$.

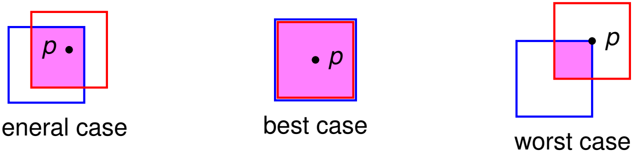


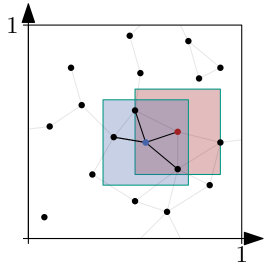
# Locality in $G^{\mathbb{T}^2}(n, r)$



**Observation**

Let $p, q \sim \mathcal{U}([-0.5, 0.5]^2)$ and $S_p$ the unit square around $p$.
Then $\Pr[q \in S_p] = \int_{-0.5}^{0.5} \int_{-0.5}^{0.5} (1 - |x|)(1 - |y|) dx\, dy = \frac{9}{16}$.




**Corollary**

By "rescaling" the observation we get $L = \Pr[\underbrace{\{u, w\} \in E}_{P_w \text{ in square around } P_u} \mid \underbrace{\{v, u\} \in E \wedge \{v, w\} \in E}_{P_u, P_w \text{ in square around } P_v}] = \frac{9}{16} = \Omega(1)$.

# Poissonisation



## Poissonised Variant $G^{\mathbb{T}^2}_{\text{Pois}}(n, r)$ of $G^{\mathbb{T}^2}(n, r)$

Replace the point set with a Poisson point process on $\mathbb{T}^2$ with intensity $n$.

$\hookrightarrow$ i.e. region of size $\lambda$ contains $\text{Pois}(\lambda n)$-many points, independent for disjoint regions

**Note:** $G^{\mathbb{T}^2}_{\text{Pois}}(n, r) \overset{\text{d}}{=} G^{\mathbb{T}^2}(N, r)$ where $N \sim \text{Pois}(n)$.

# Poissonisation

## Poissonised Variant $G^{\mathbb{T}^2}_{\text{Pois}}(n, r)$ of $G^{\mathbb{T}^2}(n, r)$

Replace the point set with a Poisson point process on $\mathbb{T}^2$ with intensity $n$.

$\hookrightarrow$ i.e. region of size $\lambda$ contains $\text{Pois}(\lambda n)$-many points, independent for disjoint regions

**Note:** $G^{\mathbb{T}^2}_{\text{Pois}}(n, r) \overset{d}{=} G^{\mathbb{T}^2}(N, r)$ where $N \sim \text{Pois}(n)$.

## Advantages

- No long-distance correlations. ✓
- $\text{Pois}(4r^2)$-distributed degrees. ✓

## Disadvantages

- Less natural in practice. ✗
- Number of vertices $N \sim \text{Pois}(n)$ not fixed. ✗

# Poissonisation

Karlsruhe Institute of Technology

## Poissonised Variant $G^{\mathbb{T}^2}_{\text{Pois}}(n,r)$ of $G^{\mathbb{T}^2}(n,r)$

Replace the point set with a Poisson point process on $\mathbb{T}^2$ with intensity $n$.

$\hookrightarrow$ i.e. region of size $\lambda$ contains $\text{Pois}(\lambda n)$-many points, independent for disjoint regions

**Note:** $G^{\mathbb{T}^2}_{\text{Pois}}(n,r) \stackrel{d}{=} G^{\mathbb{T}^2}(N,r)$ where $N \sim \text{Pois}(n)$.

## Advantages

- No long-distance correlations. ✓
- $\text{Pois}(4r^2)$-distributed degrees. ✓

## Disadvantages

- Less natural in practice. ✗
- Number of vertices $N \sim \text{Pois}(n)$ not fixed. ✗

## De-Poissonisation (an analogous result holds for de-Poissonising balls-into-bins)

Let $P$ be a graph property. If $P$ is very unlikely for $G^{\mathbb{T}^2}_{\text{Pois}}(n,r)$ then $P$ is unlikely for $G^{\mathbb{T}^2}(n,r)$:

$$\Pr[G^{\mathbb{T}^2}(n,r) \in P] = \Pr[G^{\mathbb{T}^2}_{\text{Pois}}(n,r) \in P \mid N = n] \le \frac{\Pr[G^{\mathbb{T}^2}_{\text{Pois}}(n,r) \in P]}{\Pr[N=n]} = \Theta(n^{1/2}) \Pr[G^{\mathbb{T}^2}_{\text{Pois}}(n,r) \in P].$$

# Content

# Scale-Free Networks

Karlsruhe Institute of Technology

## Semi-Formal Definition

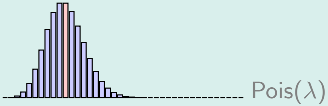
A scale-free network is a graph with a degree distribution that follows a power law (in an asymptotic sense)

## Practical Consequence

There are vertices of very high degree (*hubs*).

## Contrast: Erdős-Renyi

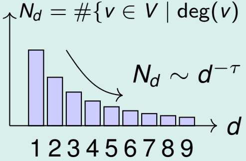
*exponentially* decreasing tail.




for $\lambda = 10$.

## Power Laws

$N_d = \#\{v \in V \mid \deg(v) = d\}$




$\tau \leq 1$: not a distribution

$1 < \tau \leq 2$: distribution, but $\mathbb{E}[\deg(v)] = \infty$

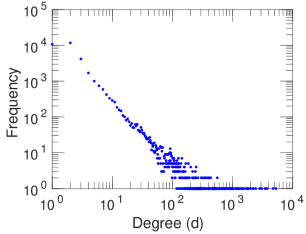
$2 < \tau \leq 3$: $\mathbb{E}[\deg(v)] < \infty$, but $\text{Var}(\deg(v)) = \infty$

$3 < \tau \leq 4$: variance $< \infty$, but higher moments are $\infty$

$\tau \in (2, 3]$ is especially popular

"Internet"

konect.cc/plot/degree.a.topology.full.png



# Scale-Free Networks

## Semi-Formal Definition

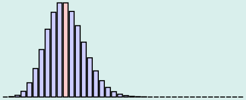
A scale-free network is a graph with a degree distribution that follows a power law (in an asymptotic sense)

## Practical Consequence

There are vertices of very high degree (*hubs*).

## Contrast: Erdős-Renyi

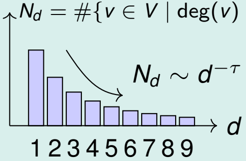
*exponentially* decreasing tail.




for $\lambda = 10$.

## Power Laws

$N_d = \#\{v \in V \mid \deg(v) = d\}$




$\tau \leq 1$: not a distribution

$1 < \tau \leq 2$: distribution, but $\mathbb{E}[\deg(v)] = \infty$

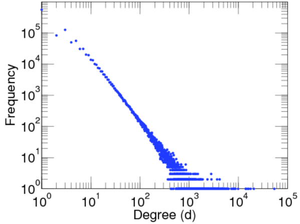
$2 < \tau \leq 3$: $\mathbb{E}[\deg(v)] < \infty$, but $\mathrm{Var}(\deg(v)) = \infty$

$3 < \tau \leq 4$: variance $< \infty$, but higher moments are $\infty$

$\tau \in (2, 3]$ is especially popular

"Youtube"

konect.cc/plot/degree.a.youtube-links.full.png



# Scale-Free Networks

Karlsruhe Institute of Technology

## Semi-Formal Definition

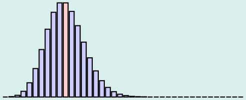
A scale-free network is a graph with a degree distribution that follows a power law (in an asymptotic sense)

## Practical Consequence

There are vertices of very high degree (*hubs*).

## Contrast: Erdős-Renyi

*exponentially* decreasing tail.



Pois($\lambda$)

for $\lambda = 10$.

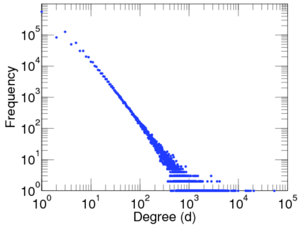
## The Name "Scale-Free"

From *Barabási: "Linked: The New Science of Networks", 2002*.

*In a random network [...] the vast majority of nodes have the same number of links [...]. Therefore, a random network has a characteristic scale in its node connectivity [...]. In contrast, the absence of a peak in a power-law degree distribution implies that [...] we see a continuous hierarchy of nodes, spanning from rare hubs to the numerous tiny nodes. There is no intrinsic scale in these networks. This is the reason my research group started to describe networks with power-law degree distribution as scale-free.*

"Youtube"

konect.cc/plot/degree.a.youtube-links.full.png



# A Scale-Free Random Geometric Graph

## Reminder: Random Geometric Graph (RGG)

Distribute vertices in a metric space and connect any two vertices with a probability depending on their distance.

## Definition: Geometric Inhomogeneous Random Graph (GIRG), Special Case

- number of vertices: $n$
- desired average degree $\lambda > 0$
- metric space $\mathbb{T}$ // more generally: $\mathbb{T}^d$ for $d \in \mathbb{N}$

# A Scale-Free Random Geometric Graph

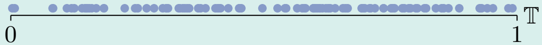
Karlsruhe Institute of Technology

**Reminder: Random Geometric Graph (RGG)**

Distribute vertices in a metric space and connect any two vertices with a probability depending on their distance.

**Definition: Geometric Inhomogeneous Random Graph (GIRG), Special Case**

- number of vertices: $n$
- desired average degree $\lambda > 0$
- metric space $\mathbb{T}$ // more generally: $\mathbb{T}^d$ for $d \in \mathbb{N}$
- for each $v$: position $x_v \sim \mathcal{U}(\mathbb{T})$



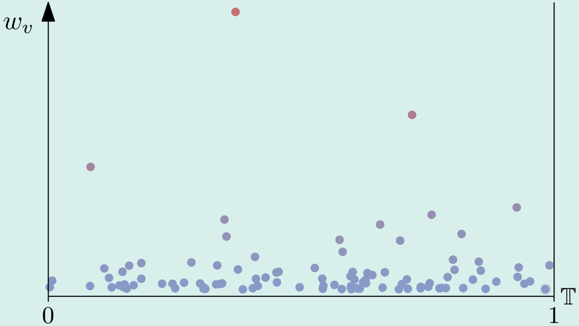
# A Scale-Free Random Geometric Graph

## Reminder: Random Geometric Graph (RGG)

Distribute vertices in a metric space and connect any two vertices with a probability depending on their distance.

## Definition: Geometric Inhomogeneous Random Graph (GIRG), Special Case

- number of vertices: $n$
- desired average degree $\lambda > 0$
- metric space $\mathbb{T}$ // more generally: $\mathbb{T}^d$ for $d \in \mathbb{N}$
- for each $v$: position $x_v \sim \mathcal{U}(\mathbb{T})$
- for each $v$: weight $w_v \sim \text{Par}(\tau - 1, 1)$
  the Pareto distribution is a power law distribution with exponent $\tau$



# A Scale-Free Random Geometric Graph

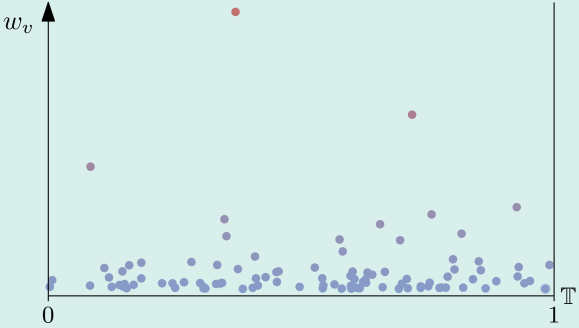
Karlsruhe Institute of Technology

## Reminder: Random Geometric Graph (RGG)

Distribute vertices in a metric space and connect any two vertices with a probability depending on their distance.

## Definition: Geometric Inhomogeneous Random Graph (GIRG), Special Case

- number of vertices: $n$
- desired average degree $\lambda > 0$
- metric space $\mathbb{T}$ // more generally: $\mathbb{T}^d$ for $d \in \mathbb{N}$
- for each $v$: position $x_v \sim \mathcal{U}(\mathbb{T})$
- for each $v$: weight $w_v \sim \text{Par}(\tau - 1, 1)$
  the Pareto distribution is a power law distribution with exponent $\tau$
- $\{u, v\} \in E \Leftrightarrow \text{dist}(x_u, x_v) \leq \frac{\lambda}{n} w_u w_v$
  $\Leftrightarrow \frac{n}{\lambda w_v} \leq \frac{w_u}{\text{dist}(x_u, x_v)}$.



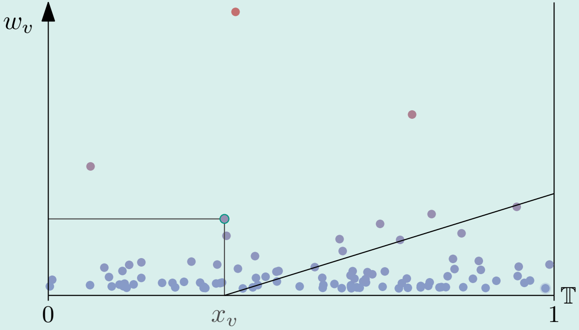
# A Scale-Free Random Geometric Graph

## Reminder: Random Geometric Graph (RGG)

Distribute vertices in a metric space and connect any two vertices with a probability depending on their distance.

## Definition: Geometric Inhomogeneous Random Graph (GIRG), Special Case

- number of vertices: $n$
- desired average degree $\lambda > 0$
- metric space $\mathbb{T}$ // more generally: $\mathbb{T}^d$ for $d \in \mathbb{N}$
- for each $v$: position $x_v \sim \mathcal{U}(\mathbb{T})$
- for each $v$: weight $w_v \sim \mathrm{Par}(\tau - 1, 1)$
  the Pareto distribution is a power law distribution with exponent $\tau$
- $\{u, v\} \in E \Leftrightarrow \mathrm{dist}(x_u, x_v) \le \frac{\lambda}{n} w_u w_v$
  $\Leftrightarrow \frac{n}{\lambda w_v} \le \frac{w_u}{\mathrm{dist}(x_u, x_v)}$.



# A Scale-Free Random Geometric Graph

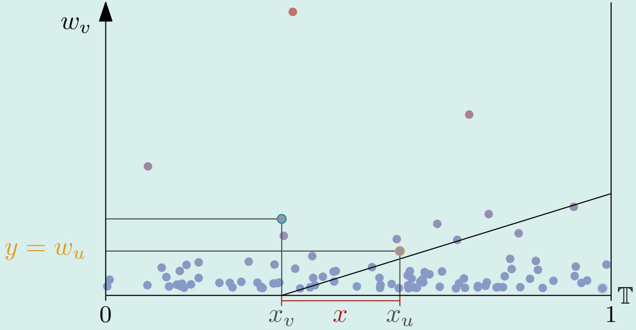
Karlsruhe Institute of Technology

## Reminder: Random Geometric Graph (RGG)

Distribute vertices in a metric space and connect any two vertices with a probability depending on their distance.

## Definition: Geometric Inhomogeneous Random Graph (GIRG), Special Case

- number of vertices: $n$
- desired average degree $\lambda > 0$
- metric space $\mathbb{T}$ // more generally: $\mathbb{T}^d$ for $d \in \mathbb{N}$
- for each $v$: position $x_v \sim \mathcal{U}(\mathbb{T})$
- for each $v$: weight $w_v \sim \text{Par}(\tau - 1, 1)$
  the Pareto distribution is a power law distribution with exponent $\tau$
- $\{u, v\} \in E \Leftrightarrow \text{dist}(x_u, x_v) \leq \frac{\lambda}{n} w_u w_v$
  $\Leftrightarrow \frac{n}{\lambda w_v} \leq \frac{w_u}{\text{dist}(x_u, x_v)}$.



# A Scale-Free Random Geometric Graph

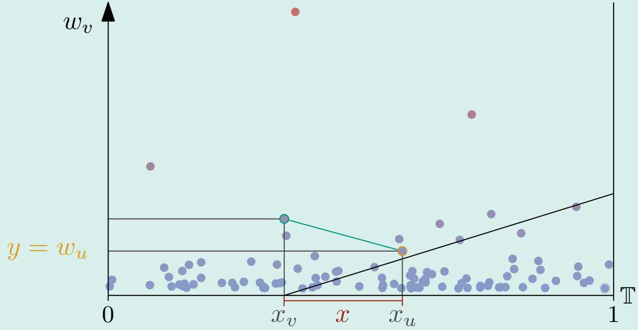
Karlsruhe Institute of Technology

## Reminder: Random Geometric Graph (RGG)

Distribute vertices in a metric space and connect any two vertices with a probability depending on their distance.

## Definition: Geometric Inhomogeneous Random Graph (GIRG), Special Case

- number of vertices: $n$
- desired average degree $\lambda > 0$
- metric space $\mathbb{T}$ // more generally: $\mathbb{T}^d$ for $d \in \mathbb{N}$
- for each $v$: position $x_v \sim \mathcal{U}(\mathbb{T})$
- for each $v$: weight $w_v \sim \mathrm{Par}(\tau - 1, 1)$
  the Pareto distribution is a power law distribution with exponent $\tau$
- $\{u, v\} \in E \Leftrightarrow \mathrm{dist}(x_u, x_v) \leq \frac{\lambda}{n} w_u w_v$
  $\Leftrightarrow \frac{n}{\lambda w_v} \leq \frac{w_u}{\mathrm{dist}(x_u, x_v)}$.



# A Scale-Free Random Geometric Graph

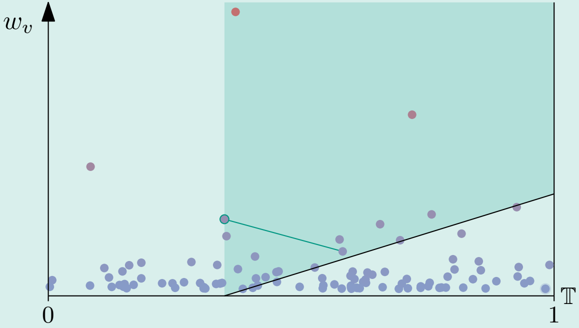
Karlsruhe Institute of Technology

## Reminder: Random Geometric Graph (RGG)

Distribute vertices in a metric space and connect any two vertices with a probability depending on their distance.

## Definition: Geometric Inhomogeneous Random Graph (GIRG), Special Case

- number of vertices: $n$
- desired average degree $\lambda > 0$
- metric space $\mathbb{T}$ // more generally: $\mathbb{T}^d$ for $d \in \mathbb{N}$
- for each $v$: position $x_v \sim \mathcal{U}(\mathbb{T})$
- for each $v$: weight $w_v \sim \text{Par}(\tau - 1, 1)$
  the Pareto distribution is a power law distribution with exponent $\tau$
- $\{u, v\} \in E \Leftrightarrow \text{dist}(x_u, x_v) \leq \frac{\lambda}{n} w_u w_v$
  $\Leftrightarrow \frac{n}{\lambda w_v} \leq \frac{w_u}{\text{dist}(x_u, x_v)}$.



# A Scale-Free Random Geometric Graph

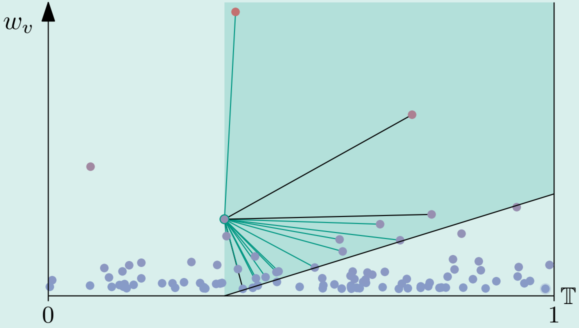
Karlsruhe Institute of Technology

## Reminder: Random Geometric Graph (RGG)

Distribute vertices in a metric space and connect any two vertices with a probability depending on their distance.

## Definition: Geometric Inhomogeneous Random Graph (GIRG), Special Case

- number of vertices: $n$
- desired average degree $\lambda > 0$
- metric space $\mathbb{T}$ // more generally: $\mathbb{T}^d$ for $d \in \mathbb{N}$
- for each $v$: position $x_v \sim \mathcal{U}(\mathbb{T})$
- for each $v$: weight $w_v \sim \mathrm{Par}(\tau - 1, 1)$
  the Pareto distribution is a power law distribution with exponent $\tau$
- $\{u, v\} \in E \Leftrightarrow \mathrm{dist}(x_u, x_v) \leq \frac{\lambda}{n} w_u w_v$
  $\Leftrightarrow \frac{n}{\lambda w_v} \leq \frac{w_u}{\mathrm{dist}(x_u, x_v)}$.



# A Scale-Free Random Geometric Graph

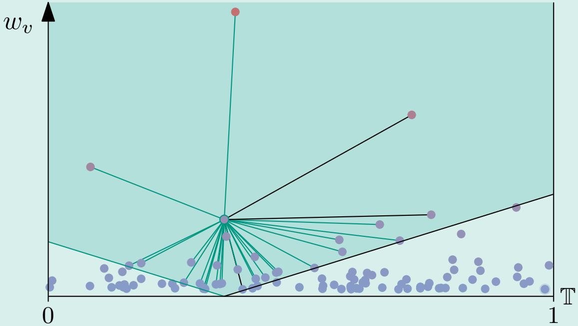
Karlsruhe Institute of Technology

## Reminder: Random Geometric Graph (RGG)

Distribute vertices in a metric space and connect any two vertices with a probability depending on their distance.

## Definition: Geometric Inhomogeneous Random Graph (GIRG), Special Case

- number of vertices: $n$
- desired average degree $\lambda > 0$
- metric space $\mathbb{T}$ // more generally: $\mathbb{T}^d$ for $d \in \mathbb{N}$
- for each $v$: position $x_v \sim \mathcal{U}(\mathbb{T})$
- for each $v$: weight $w_v \sim \text{Par}(\tau - 1, 1)$
  the Pareto distribution is a power law distribution with exponent $\tau$
- $\{u, v\} \in E \Leftrightarrow \text{dist}(x_u, x_v) \leq \frac{\lambda}{n} w_u w_v$
  $\Leftrightarrow \frac{n}{\lambda w_v} \leq \frac{w_u}{\text{dist}(x_u, x_v)}$.



# A Scale-Free Random Geometric Graph

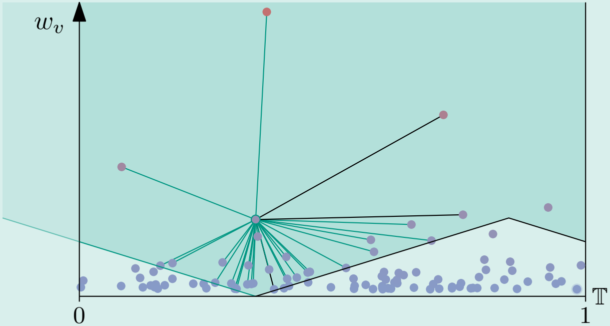
Karlsruhe Institute of Technology

## Reminder: Random Geometric Graph (RGG)

Distribute vertices in a metric space and connect any two vertices with a probability depending on their distance.

## Definition: Geometric Inhomogeneous Random Graph (GIRG), Special Case

- number of vertices: $n$
- desired average degree $\lambda > 0$
- metric space $\mathbb{T}$ // more generally: $\mathbb{T}^d$ for $d \in \mathbb{N}$
- for each $v$: position $x_v \sim \mathcal{U}(\mathbb{T})$
- for each $v$: weight $w_v \sim \mathrm{Par}(\tau - 1, 1)$
  the Pareto distribution is a power law distribution with exponent $\tau$
- $\{u, v\} \in E \Leftrightarrow \mathrm{dist}(x_u, x_v) \leq \frac{\lambda}{n} w_u w_v$
  $\Leftrightarrow \frac{n}{\lambda w_v} \leq \frac{w_u}{\mathrm{dist}(x_u, x_v)}$.



# A Scale-Free Random Geometric Graph

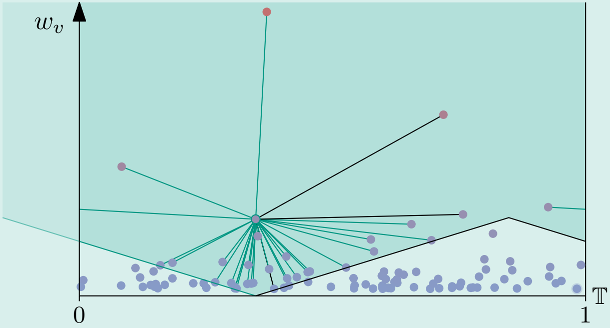
Karlsruhe Institute of Technology

## Reminder: Random Geometric Graph (RGG)

Distribute vertices in a metric space and connect any two vertices with a probability depending on their distance.

## Definition: Geometric Inhomogeneous Random Graph (GIRG), Special Case

- number of vertices: $n$
- desired average degree $\lambda > 0$
- metric space $\mathbb{T}$ // more generally: $\mathbb{T}^d$ for $d \in \mathbb{N}$
- for each $v$: position $x_v \sim \mathcal{U}(\mathbb{T})$
- for each $v$: weight $w_v \sim \text{Par}(\tau - 1, 1)$
  the Pareto distribution is a power law distribution with exponent $\tau$
- $\{u, v\} \in E \Leftrightarrow \text{dist}(x_u, x_v) \leq \frac{\lambda}{n} w_u w_v$
  $\Leftrightarrow \frac{n}{\lambda w_v} \leq \frac{w_u}{\text{dist}(x_u, x_v)}$.



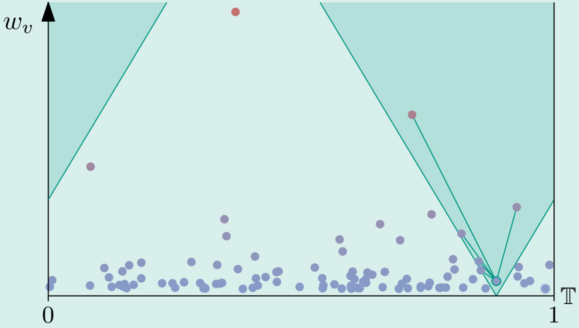
# A Scale-Free Random Geometric Graph

## Reminder: Random Geometric Graph (RGG)

Distribute vertices in a metric space and connect any two vertices with a probability depending on their distance.

## Definition: Geometric Inhomogeneous Random Graph (GIRG), Special Case

- number of vertices: $n$
- desired average degree $\lambda > 0$
- metric space $\mathbb{T}$ // more generally: $\mathbb{T}^d$ for $d \in \mathbb{N}$
- for each $v$: position $x_v \sim \mathcal{U}(\mathbb{T})$
- for each $v$: weight $w_v \sim \text{Par}(\tau - 1, 1)$
  the Pareto distribution is a power law distribution with exponent $\tau$
- $\{u, v\} \in E \Leftrightarrow \text{dist}(x_u, x_v) \leq \frac{\lambda}{n} w_u w_v$
  $\Leftrightarrow \frac{n}{\lambda w_v} \leq \frac{w_u}{\text{dist}(x_u, x_v)}$.



# A Scale-Free Random Geometric Graph

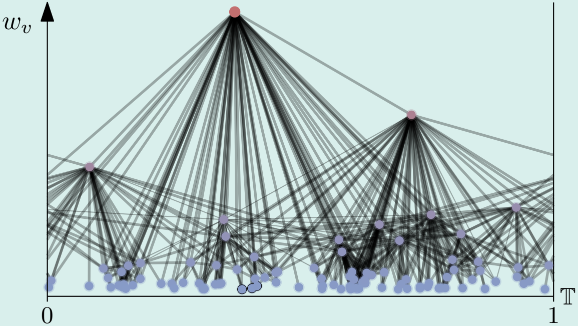
Karlsruhe Institute of Technology

## Reminder: Random Geometric Graph (RGG)

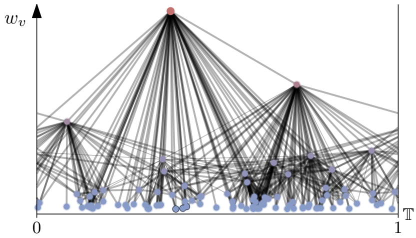
Distribute vertices in a metric space and connect any two vertices with a probability depending on their distance.

## Definition: Geometric Inhomogeneous Random Graph (GIRG), Special Case

- number of vertices: $n$
- desired average degree $\lambda > 0$
- metric space $\mathbb{T}$ // more generally: $\mathbb{T}^d$ for $d \in \mathbb{N}$
- for each $v$: position $x_v \sim \mathcal{U}(\mathbb{T})$
- for each $v$: weight $w_v \sim \text{Par}(\tau - 1, 1)$
  the Pareto distribution is a power law distribution with exponent $\tau$
- $\{u, v\} \in E \Leftrightarrow \text{dist}(x_u, x_v) \leq \frac{\lambda}{n} w_u w_v$
  $\Leftrightarrow \frac{n}{\lambda w_v} \leq \frac{w_u}{\text{dist}(x_u, x_v)}$.



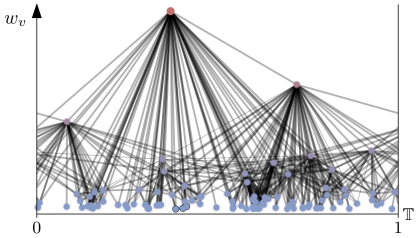
# How GIRGs are Useful




## GIRGs are Scale-Free

$\mathbb{E}[\deg(v) \mid w_v] = \Theta(w_v)$ and $\deg(v)$ follows a power law if $w_v$ does.

# How GIRGs are Useful




## GIRGs are Scale-Free

$\mathbb{E}[\deg(v) \mid w_v] = \Theta(w_v)$ and $\deg(v)$ follows a power law if $w_v$ does.

## GIRGs are a Good Model for Real World Networks (Bläsius, Fischbeck, 2022)

- consider two graph parameters: locality and heterogeneity ($\approx \log \mathrm{Var}(\deg(v))$).
- in many contexts, a real network behaves like a GIRG with the same parameters

---

*On the External Validity of Average-Case Analyses of Graph Algorithms, ESA 2022.*

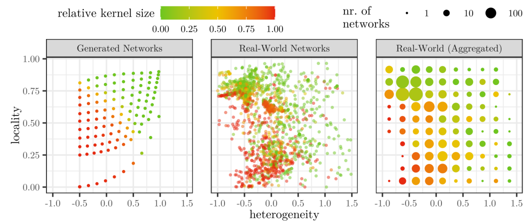



**Figure 7** The relative kernel size of the vertex cover domination rule.

# Hyperbolic Geometric Graphs

## Poincaré Model of Hyperbolic Geometry

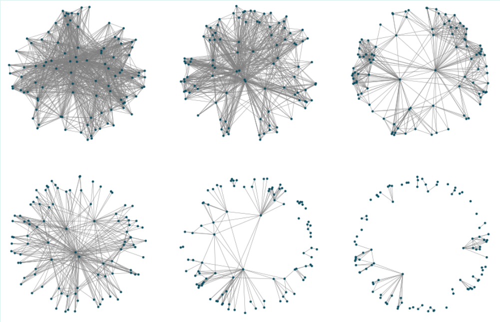
Illustration by M.C. Escher, Circle Limit III, 1959.

*(All creatures are congruent in hyperbolic space.)*

## Hyperbolic Random Graph (HGGs)

Sample points with bias towards the centre.
Connect points if distance is beneath a threshold.

https://commons.wikimedia.org/wiki/File:Hyperbolic_graphs.png



Can yield power law distribution for node degrees.

# Hyperbolic Geometric Graphs

## Poincaré Model of Hyperbolic Geometry

Illustration by M.C. Escher, Circle Limit III, 1959.

*(All creatures are congruent in hyperbolic space.)*

## Result (Bläsius, Friedrich, Katzmann, 2021)

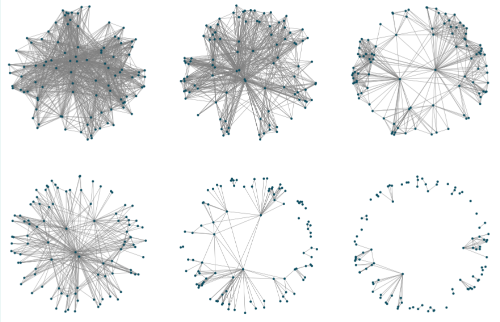
Vertex Cover can be Approximated on HGGs.

*Efficiently Approximating Vertex Cover on Scale-Free Networks with Underlying Hyperbolic Geometry, ESA 2021.*

## Hyperbolic Random Graph (HGGs)

Sample points with bias towards the centre.
Connect points if distance is beneath a threshold.

https://commons.wikimedia.org/wiki/File:Hyperbolic_graphs.png



Can yield power law distribution for node degrees.

# Preferential Attachment and the Barabasi-Albert Model

## How a Graph is Grown Over Time

- There is a parameter $m \in \mathbb{N}$.
- start with any graph on $\geq m$ nodes.
- add new nodes one by one
  - new node is connected to $m$ existing nodes
  - existing nodes are selected with probability proportional to their degree

# Preferential Attachment and the Barabasi-Albert Model

Karlsruhe Institute of Technology

## How a Graph is Grown Over Time

- There is a parameter $m \in \mathbb{N}$.
- start with any graph on $\geq m$ nodes.
- add new nodes one by one
  - new node is connected to $m$ existing nodes
  - existing nodes are selected with probability proportional to their degree

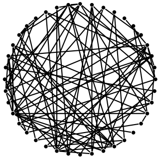
## Why the Model is Interesting

- node degrees approach a power law distribution with exponent 3
- model may explain *why* scale-free networks emerge in practice

# Summary

## Erdős-Renyi Random Graphs

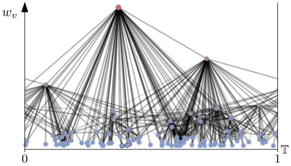
- simplest type of random graphs
- "Erdős-Renyi" refers to various related models
- arise in certain data structures (stay tuned)
- look locally like Poisson Galton-Watson Trees
- no locality or high-degree vertices



## Random Graphs for Average Case Analsis

Mimic properties of real world networks:

- locality // a friend of my friend is often my friend
  - arises naturally in random geometric graphs
- "scale-freeness" $\approx$ existence of hubs
  - assign weights to vertices (in GIRGs)
  - use hyperbolic geometry
  - use preferential attachment



# Appendix: Possible Exam Questions I

Karlsruhe Institute of Technology

- What is meant by the theory–practice gap in the context of graph algorithms?
- What might a theoretician try to overcome the gap?
- What is the classical model of Erdős and Rényi?
  - Which variants of the Erdős–Rényi model did we consider?
  - What can be said about the distribution of $\deg(v)$ when we set $\mathbb{E}[\deg(v)] = \lambda$?
  - What can be said about $N_d = |\{v \in [n] \mid \deg(v) = d\}|$?
  - We studied the $R$-neighborhood $G(n, \lambda/n)|_{v,R}$ of a vertex $v$.
    - What holds for the distribution of $G(n, \lambda/n)|_{v,R}$, and why?
    - What is a Galton–Watson tree?
    - What can be said about the extinction probability of a Poisson Galton–Watson tree?
  - What is meant by the “sudden emergence of the giant component”? State the result formally.
  - We considered a quantity $L$, called locality. How is it defined?
  - What locality do Erdős–Rényi graphs have?
- Name properties that distinguish real-world networks from Erdős–Rényi graphs.
  - Give an example of a geometric random graph. What is the locality in this model?

# Appendix: Possible Exam Questions II

- In what way might a poissonized model be more convenient?
- When is a network “scale-free”?
  - Give an example of a real-world network to which this property is attributed.
  - Describe at least two ways in which networks of this type can be generated.