The Maximum Block Size of Critical Random Graphs

VLADY RAVELOMANANA

IRIF – UMR CNRS 8243
University of Paris 7
vlad@univ-paris-diderot.fr

- in collaboration with V. Rasendrahasina & A. Rasoanaivo -
- supported by PEPS Hydrata and ANR BLANC Magnum -

Outline of the talk

- Introduction, motivations & previous works.
  - 2-path statistics.
  - Largest block.
  - Conclusion & Perspectives.

Erdős-Rényi random graphs

- Introduced simultaneously by
  - ERDŐS - RÉNYI (1959) :
    $G(n, M) \to$ a graph is chosen uniformly at random from the collection of all graphs which have n nodes and M edges.
  - GILBERT (1959) :
    $G(n, p) \to n$ vertices and each of the $\binom{n}{2}$ edges is included in the graph with probability p independent from every other edge.
- The *double-jump* phenomenon.
  - As $p < \frac{1}{n}$ largest connected component is of order $\mathcal{O}(\log n)$.
  - As $p \sim \frac{1}{n}$ many connected components of order $\mathcal{O}(n^{\frac{2}{3}})$.
  - As $p > \frac{1}{n}$ largest connected component is unique and of order $\mathcal{O}(n)$: the **giant component**.
- Many books BOLLOBÁS (1985, 2001), KOLCHIN (1999), JANSON, ŁUCZAK, RUCIŃSKI (2000), FRIEZE, KAROŃSKI (2016), REMCO VAN DER HOFSTAD (2016).

Why these models?

- **Advantages:** Simple and mathematically well-understood (allow rigorous results).
- **Drawbacks:** too homogeneous, unrealistic in many situations, etc ...

Erdős-Rényi random graphs are at the core of

- Probabilistic Methods NOGA, SPENCER (2000).
- Algorithmic Methods (cf. "**Algorithmic Theory of Random Graphs**" FRIEZE, MCDIARMID (2000).
- Statistical Physics Methods ("**Cavity method**", "**Replica Symmetry Breaking**").
- Combinatorics and AofA.

Some distances in random graphs

Metrics of interest include

- Longest path: maximum length of self avoiding path.
- Diameter: maximum hop distance between 2 nodes in a connected component.
- Circumference: number of vertices belonging to the largest cycle.
- Maximum block sizes: number of vertices in the largest 2-connected component.

Previous works on diameter

Theorem CHUNG, LU (2001).

Let $p = \frac{c}{n}$ and let C_1 be the largest component of $G(n, p)$. Then

- If $c < 1$ then $diam(C_1) = \mathcal{O}(\sqrt{\log n})$ a.a.s
- If $c > 1$ then $diam(C_1) = \Theta(\log n)$ a.a.s

Remark. There are more precise asymptotics in the case $c > 1$ by FERNOLTZ, RAMACHANDRAN (2007), BOLLOBÁS, JANSON, RIORDAN (2007) and in RIORDAN, WORMALD (2009).

BENJAMINI, KOZMA AND WORMALD : What about $c = 1$?

Previous works on diameter

Theorem CHUNG, LU (2001).

Let $p = \frac{c}{n}$ and let C_1 be the largest component of $G(n, p)$. Then

- If $c < 1$ then $diam(C_1) = \mathcal{O}(\sqrt{\log n})$ a.a.s
- If $c > 1$ then $diam(C_1) = \Theta(\log n)$ a.a.s

Remark. There are more precise asymptotics in the case $c > 1$ by FERNOLTZ, RAMACHANDRAN (2007), BOLLOBÁS, JANSON, RIORDAN (2007) and in RIORDAN, WORMALD (2009).

BENJAMINI, KOZMA AND WORMALD : What about $c = 1$?

Theorem NACHMIAS, PERES (2007).

If $p = \frac{1}{n}$, for any ε there exists $A = A(\varepsilon) < \infty$ such that for all large n

$$\mathbb{P}\Big[\mathrm{diam}(C_1) \notin \Big[A^{-1} n^{\frac{1}{3}}, A\, n^{\frac{1}{3}}\Big]\Big] < \varepsilon\,.$$

The result using random walks and continuum random trees

"*The continuum limit of critical random graphs*" (arXiv:0903.4730)

ADDAGIO-BERRY, BROUTIN, GOLDSMITH (2010)

Abstract

We consider the Erdős–Rényi random graph $G(n, p)$ inside the critical window, that is when $p = 1/n + \lambda n^{-4/3}$, for some fixed $\lambda \in \mathbb{R}$. We prove that the sequence of connected components of $G(n, p)$, considered as metric spaces using the graph distance rescaled by $n^{-1/3}$, converges towards a sequence of continuous compact metric spaces. The result relies on a bijection between graphs and certain marked random walks, and the theory of continuum random trees. Our result gives access to the answers to a great many questions about distances in critical random graphs. In particular, we deduce that the diameter of $G(n, p)$ rescaled by $n^{-1/3}$ converges in distribution to an absolutely continuous random variable with finite mean.

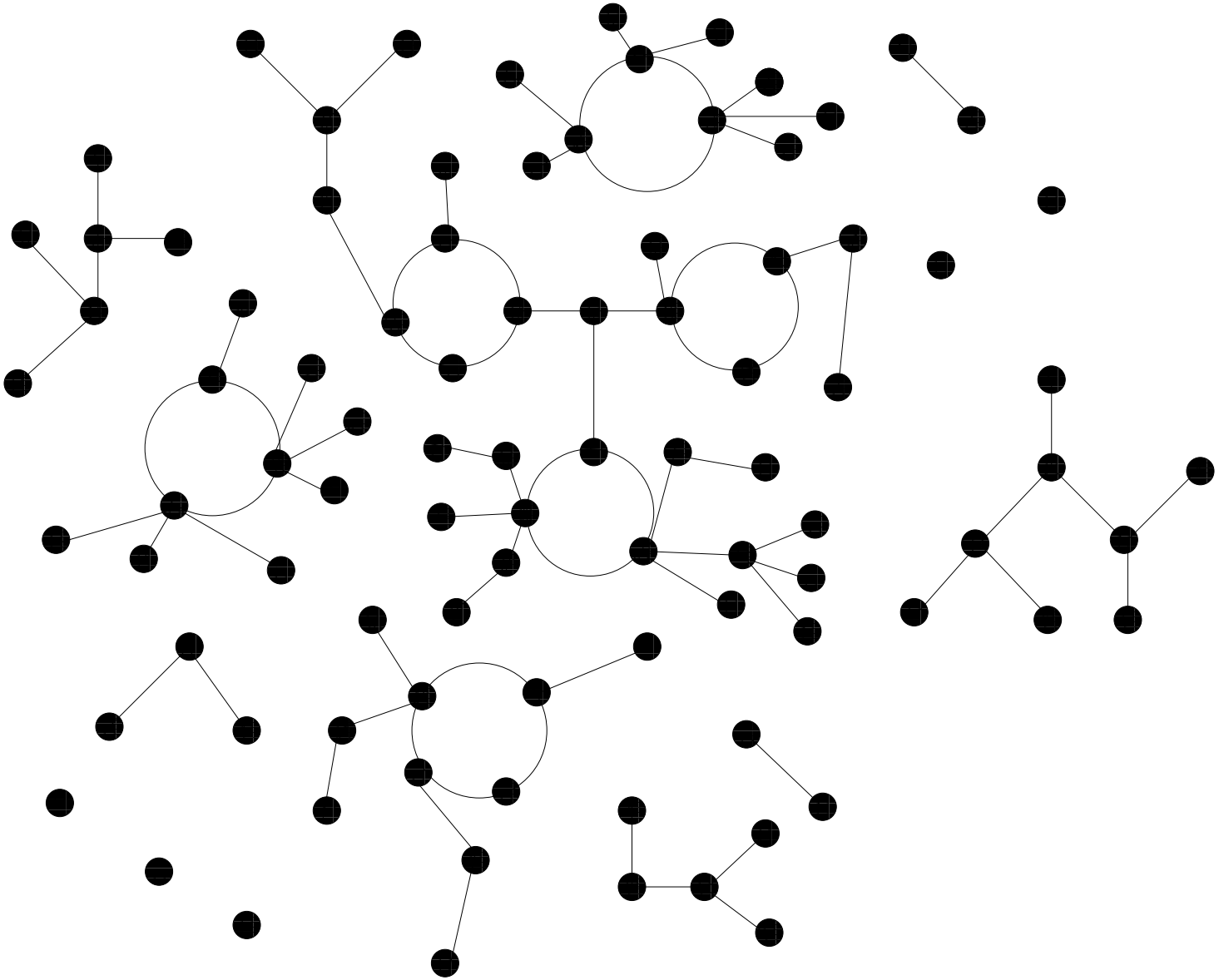
Keywords

Random graphs – Gromov–Hausdorff distance – Scaling limits – Continuum random tree – Diameter

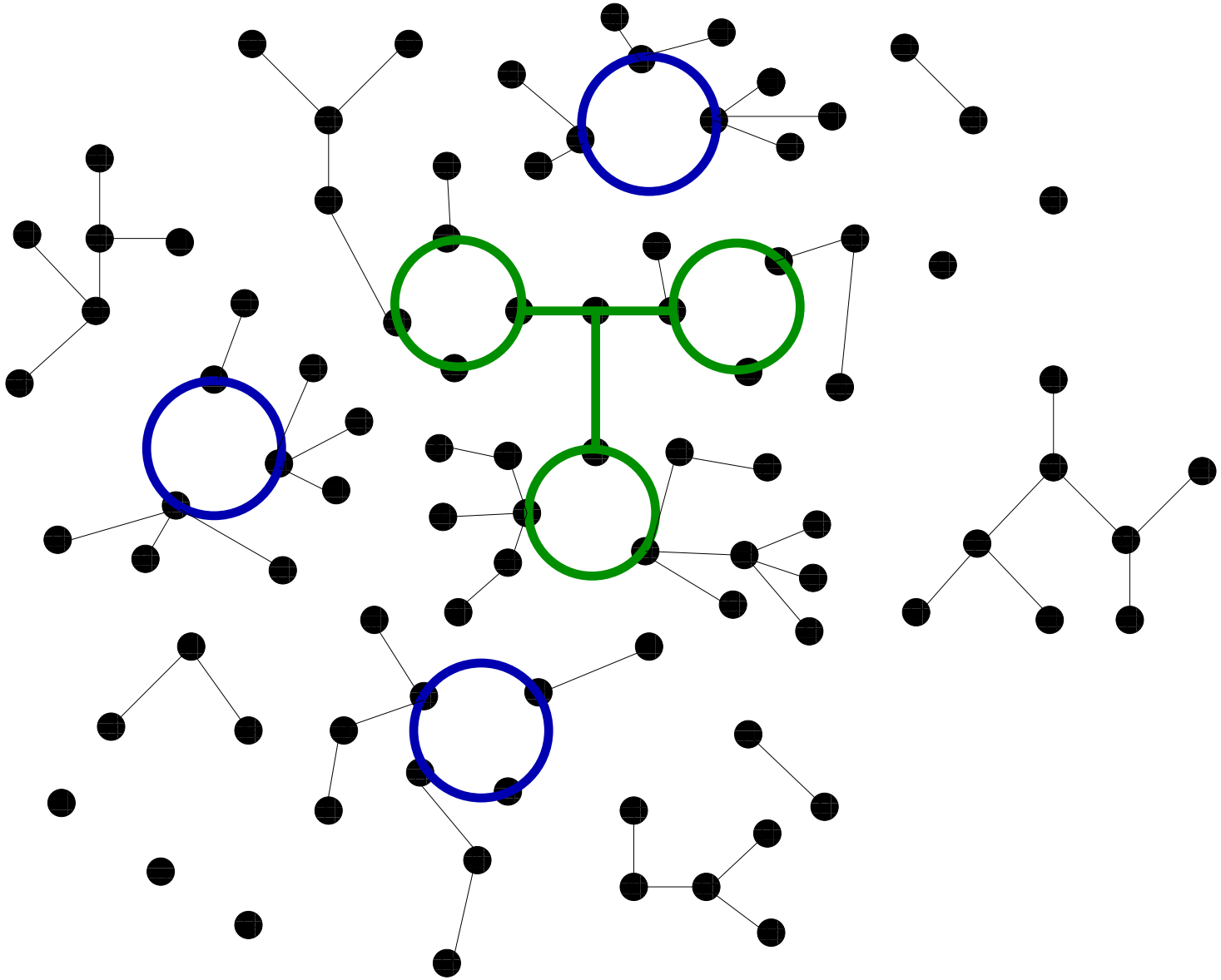
Outline of the talk

- Introduction, motivations & previous works.
- 2-path statistics.
- Largest block.
- Conclusion & Perspectives.

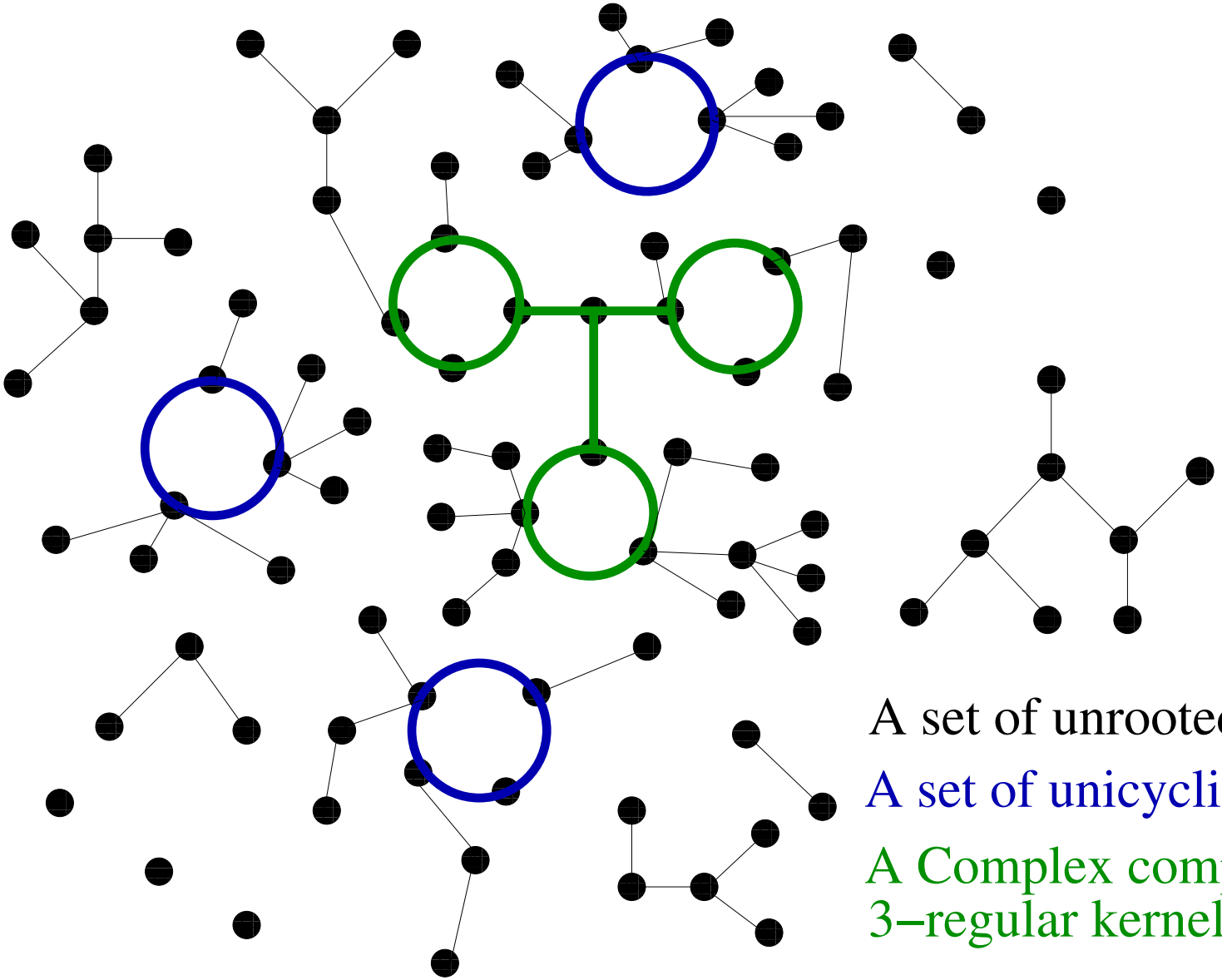
Typical anatomy of critical random graphs



Typical anatomy of critical random graphs



Typical anatomy of critical random graphs



A set of unrooted Trees

A set of unicyclic Components

A Complex component with 3–regular kernel

Generatingfunctionology of critical random graphs

Cayley's rooted trees

$$T(z) = ze^{T(z)} = \sum_{n\geq 1} n^{n-1} \frac{z^n}{n!} .$$

Unrooted trees (Cayley 1889)

$$U(z) = T(z) - \frac{T(z)^2}{2} = \sum_{n\geq 1} n^{n-2} \frac{z^n}{n!} .$$

Unicyclic components

$$V(z) = \frac{1}{2} \log \left(\frac{1}{1 - T(z)} \right) - \frac{T(z)}{2} - \frac{T(z)^2}{4} .$$

Complex component of excess R (R edges more than vertices – Wright 1977)

$$E_R(z) \asymp \frac{e_R}{(1 - T(z))^{3R}} \qquad \text{with} \qquad e_R = \frac{(6R)!}{2^{5R} 3^{2R} (3R)! (2R)!} .$$

Analytic combinatorics of critical random graphs

Using symbolic method of GF, the probability that a random graph $G(n, M)$ has a complex component of excess R is approximately

$$\begin{aligned} p_R(n, M) &\sim \frac{n!}{\binom{\binom{n}{2}}{M}} \left[z^n\right] \frac{U(z)^{n-M+R}}{(n-M+R)!} e^{V(z)} E_R(z) \\ &= \underbrace{\frac{n!}{\binom{\binom{n}{2}}{M}(n-M+R)!}}_{\text{STIRLING}} \times \underbrace{\frac{1}{2\pi i} \oint g_R(z) e^{nh(z)} dz}_{\text{CAUCHY}} . \end{aligned}$$

Theorem JANSON, KNUTH, ŁUZCAK, PITTEL (1993).

As $n \to \infty$, $M = \frac{n}{2} + \mu n^{2/3}$, $|\mu| = O(n^{1/12})$ and R is fixed

$$p_R(n, M) \sim \sqrt{2\pi}\, e_R\, A\left(3R + \frac{1}{2}, \mu\right)$$

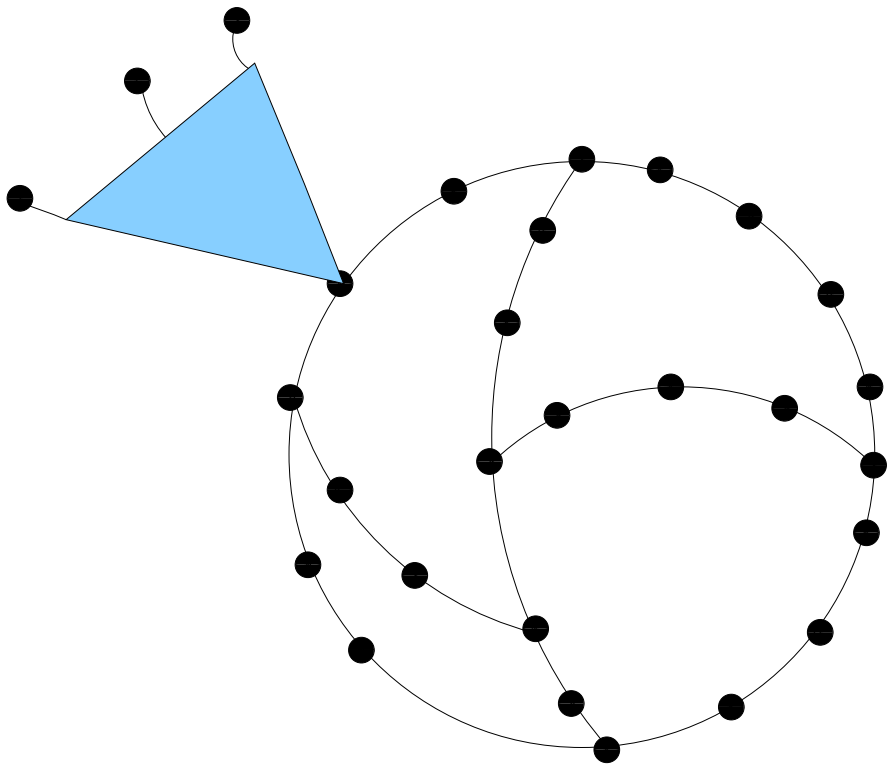
with

$$A(y, \mu) = \frac{e^{-\mu^3/6}}{3^{(y+1)/3}} \sum_{k \geq 0} \frac{\left(\frac{1}{2} 3^{2/3} \mu\right)^k}{k!\, \Gamma\left((y+1-2k)/3\right)} .$$

2-paths

Def.

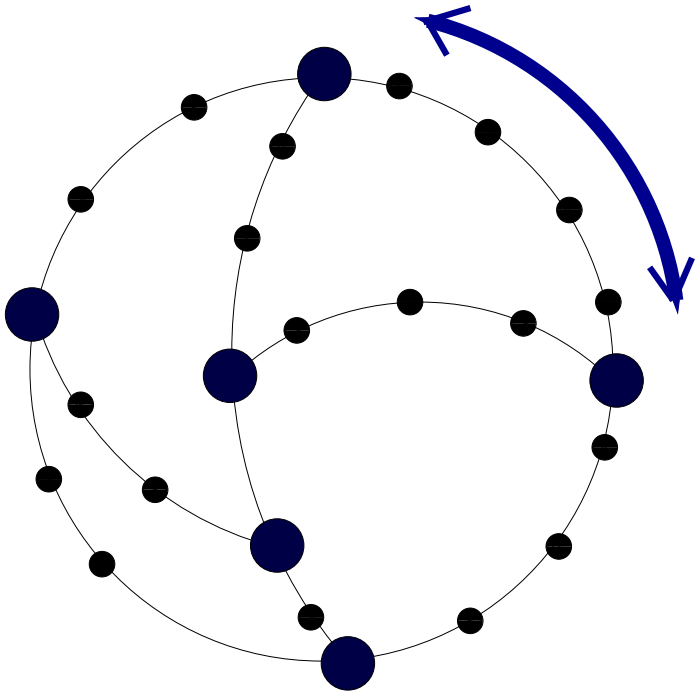
- A smooth graph is obtained from a connected graph by pruning recursively all vertices of degree 1.
- A 2-path (also known as Wright's path) is a path in a smooth graph where all internal vertices have degree 2.



2-paths

Def.

- A smooth graph is obtained from a connected graph by pruning recursively all vertices of degree 1.
- A 2-path (also known as Wright's path) is a path in a smooth graph where all internal vertices have degree 2.



2-paths statistics in critical random graphs using analytic combinatorics

$X_n =$ length of a random 2-path in $G(n, M)$ with $M = \frac{n}{2} + \mu n^{\frac{2}{3}}$.

$$p_k = \mathbb{P}(X_n = k) = \sum_{R>0} \underbrace{\mathbb{P}(G(n, M) \text{ has excess } R)}_{p_R(n, M)} \mathbb{P}(X_n = k | \text{ excess } = R)$$

2-paths statistics in critical random graphs using analytic combinatorics

$X_n =$ length of a random 2-path in $G(n, M)$ with $M = \frac{n}{2} + \mu n^{\frac{2}{3}}$.

$$p_k = \mathbb{P}(X_n = k) = \sum_{R>0} \underbrace{\mathbb{P}(G(n, M) \text{ has excess } R)}_{p_R(n, M)} \mathbb{P}(X_n = k | \text{ excess } = R)$$

$$\mathbb{E}\left(a^{X_n}\right) = \sum_{k=0}^{+\infty} p_k \mathbf{a}^k = \sum_{k=0}^{\infty} \sum_{R>0} p_R(n, M) \, \mathbb{P}(X_n = k | \text{ excess } = R) \, \mathbf{a}^k$$

2-paths statistics in critical random graphs using analytic combinatorics

$X_n =$ length of a random 2-path in $G(n, M)$ with $M = \frac{n}{2} + \mu n^{\frac{2}{3}}$.

$$p_k = \mathbb{P}(X_n = k) = \sum_{R>0} \underbrace{\mathbb{P}(G(n, M) \text{ has excess } R)}_{p_R(n, M)} \mathbb{P}(X_n = k | \text{ excess } = R)$$

$$\begin{aligned}
\mathbb{E}\left(a^{X_n}\right) &= \sum_{k=0}^{+\infty} p_k \mathbf{a}^k = \sum_{k=0}^{\infty} \sum_{R>0} p_R(n, M)\, \mathbb{P}(X_n = k | \text{ excess } = R)\, \mathbf{a}^k \\
&= \sum_{R>0} \left[\sum_{k=0}^{\infty} \mathbb{P}(X_n = k | \text{ excess } = R)\, \mathbf{a}^k \right] \quad p_R(n, M)
\end{aligned}$$

2-paths statistics in critical random graphs using analytic combinatorics

$X_n =$ length of a random 2-path in $G(n, M)$ with $M = \frac{n}{2} + \mu n^{\frac{2}{3}}$.

$$p_k = \mathbb{P}(X_n = k) = \sum_{R>0} \underbrace{\mathbb{P}(G(n, M) \text{ has excess } R)}_{p_R(n, M)} \mathbb{P}(X_n = k | \text{ excess } = R)$$

$$\begin{aligned}
\mathbb{E}\left(a^{X_n}\right) &= \sum_{k=0}^{+\infty} p_k \mathbf{a}^k = \sum_{k=0}^{\infty} \sum_{R>0} p_R(n, M)\, \mathbb{P}(X_n = k | \text{ excess } = R)\, \mathbf{a}^k \\
&= \sum_{R>0} \left[\sum_{k=0}^{\infty} \mathbb{P}(X_n = k | \text{ excess } = R)\, \mathbf{a}^k \right] \quad p_R(n, M) \\
&= \sum_{R>0} \left[\frac{n!\, [z^n]\, \frac{U^t}{t!} \,.e^V \,.\frac{(1-T)}{(1-\mathbf{a}\, T)} \,.E_R}{n!\, [z^n]\, \frac{U^t}{t!} \,.e^V \,.E_R} \right] \quad p_R(n, M)
\end{aligned}$$

2-paths statistics in critical random graphs using analytic combinatorics

$X_n =$ length of a random 2-path in $G(n, M)$ with $M = \frac{n}{2} + \mu n^{\frac{2}{3}}$.

$$p_k = \mathbb{P}(X_n = k) = \sum_{R>0} \underbrace{\mathbb{P}(G(n, M) \text{ has excess } R)}_{p_R(n, M)} \mathbb{P}(X_n = k | \text{ excess } = R)$$

$$\begin{aligned}
\mathbb{E}\left(a^{X_n}\right) &= \sum_{k=0}^{+\infty} p_k \mathbf{a}^k = \sum_{k=0}^{\infty} \sum_{R>0} p_R(n, M)\, \mathbb{P}(X_n = k | \text{ excess } = R)\, \mathbf{a}^k \\
&= \sum_{R>0} \left[\sum_{k=0}^{\infty} \mathbb{P}(X_n = k | \text{ excess } = R)\, \mathbf{a}^k \right] \quad p_R(n, M) \\
&= \sum_{R>0} \left[\frac{n!\, [z^n]\, \frac{U^t}{t!} . e^V . \frac{(1-T)}{(1-\mathbf{a}\, T)} . E_R}{n!\, [z^n]\, \frac{U^t}{t!} . e^V . E_R} \right] \quad p_R(n, M)
\end{aligned}$$

Theorem R., R., R. (2016)

The length X_n of a random 2-path in $G(n, M = \frac{n}{2} + \mu n^{\frac{2}{3}}))$ satisfies

$$\mathbb{E}(X_n(X_n - 1) \cdots (X_n - k + 1)) \sim \sqrt{2\pi} \sum_{R>0} e_R\, A(3R + k + \frac{1}{2}, \mu)\, n^{\frac{k}{3}}$$

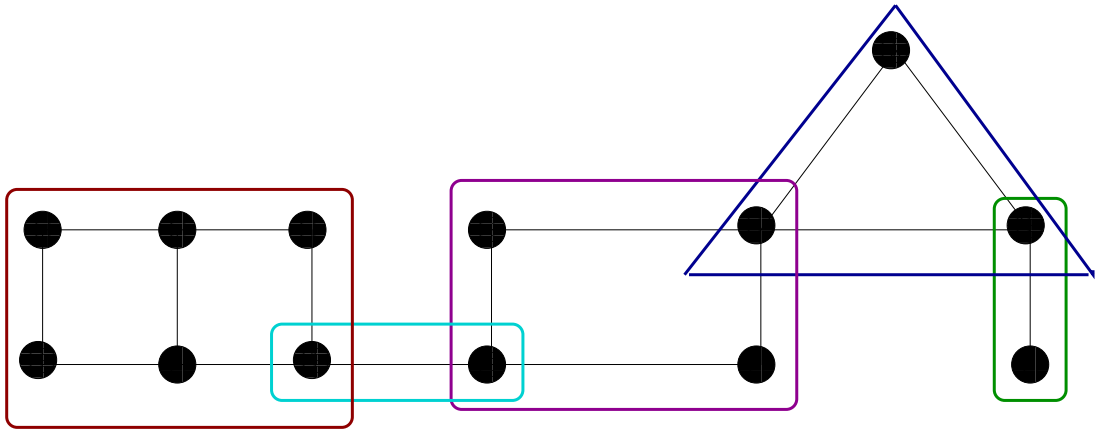
Outline of the talk

- Introduction, motivations & previous works.
- 2-path statistics.
- Largest block.
- Conclusion & Perspectives.

Blocks and maximum block

Def.

A biconnected component or block or 2-connected component is a maximal biconnected subgraph.



Works on circumference of (sub)critical random graphs

JANSON, ŁUCZAK, RUCIŃSKI (2000)

For any $\omega = \omega(n) \gg 1$, a.a.s it holds that

- **Subcritical:** when $M = \frac{1-\varepsilon}{2} n$ the circumference $\mathcal{C}$ of $G(n, M)$ is **at most** ω.
- when $M = \frac{n}{2} - s$ with $n^{2/3} \ll s \ll n$, $\frac{\mathcal{C}}{\omega} < \frac{n}{s} < \mathcal{C}\omega$.
- **Critical:** when $M = \frac{n}{2} + \mathcal{O}(n^{\frac{2}{3}})$, $\frac{\mathcal{C}}{\omega} < n^{\frac{1}{3}} < \mathcal{C}\omega$.
- In the **supercritical** phase, similar bounds are known (unless the graph is Hamiltonian).

Works on circumference of (sub)critical random graphs

JANSON, ŁUCZAK, RUCIŃSKI (2000)

For any $\omega = \omega(n) \gg 1$, a.a.s it holds that

- **Subcritical:** when $M = \frac{1-\varepsilon}{2} n$ the circumference $\mathcal{C}$ of $G(n, M)$ is **at most** ω.
- when $M = \frac{n}{2} - s$ with $n^{2/3} \ll s \ll n$, $\frac{\mathcal{C}}{\omega} < \frac{n}{s} < \mathcal{C}\omega$.
- **Critical:** when $M = \frac{n}{2} + \mathcal{O}(n^{\frac{2}{3}})$, $\frac{\mathcal{C}}{\omega} < n^{\frac{1}{3}} < \mathcal{C}\omega$.
- In the **supercritical** phase, similar bounds are known (unless the graph is Hamiltonian).

Theorem R., R., R. (2016)

If $n - 2M \gg n^{\frac{2}{3}}$, the maximum block size (and the circumference) of $G(n, M)$ satisfies

$$\mathbb{E}[\mathcal{C}] \sim c_1 \times \frac{n}{n - 2M}$$

where $c_1 = \int_0^\infty \left(1 - e^{-E_1(v)}\right) dv$ with $E_1(x) = \frac{1}{2} \int_x^\infty e^{-t} \frac{dt}{t}$.

The generating function trick for the MAX

For any discrete random variable X we have

$$\mathbb{E}[X] = \sum_{k \geq 0} k\, \mathbb{P}[X = k] = \sum_{k \geq 0} (1 - \mathbb{P}[X \leq k]) \ . \qquad (\star)$$

Let $\mathcal{F}$ be a family of graphs with exponential generating function (EGF) $F(z)$. Let $\mathcal{F}^{(k)}$ be a subfamily of graphs of $\mathcal{F}$ where some fixed size parameter $\mathcal{P}$ (ex: size of cycles, ...) is at most k. Let $F^{(k)}(z)$ be the EGF of $\mathcal{F}^{(k)}$.

Equation $(\star)$ can be translated into the following GF-version

$$\Xi(z) = \sum_{k \geq 0} \left(F(z) - F^{(k)}(z) \right)$$

such that the expectation of the **maximum** size of the parameter $\mathcal{P}$ is

$$\frac{n!\, [z^n]\, \Xi(z)}{n!\, [z^n]\, F(z)} \ .$$

Proof of the theorem

Using the last observation, the expectation of $\mathcal{C}$ is $\frac{n![z^n]}{\binom{\binom{n}{2}}{M}}\Xi(z)$

$$\sum_{k\geq 0}\left[1-\frac{n!}{\binom{\binom{n}{2}}{M}}[z^n]\frac{U(z)^{n-M+r}}{(n-M+r)!}\left(e^{V(z)-\sum_{i=k+1}^{\infty}\frac{T(z)^i}{2i}}\right)E_r^{[k]}(z)\right]$$

Proof of the theorem

Using the last observation, the expectation of $\mathcal{C}$ is $\frac{n![z^n]}{\binom{\binom{n}{2}}{M}}\Xi(z)$

$$\sum_{k\geq 0}\left[1-\frac{n!}{\binom{\binom{n}{2}}{M}}[z^n]\frac{U(z)^{n-M+r}}{(n-M+r)!}\left(e^{V(z)-\sum_{i=k+1}^{\infty}\frac{T(z)^i}{2i}}\right)E_r^{[k]}(z)\right]$$

Using the term by term inequalities

$$1\leq E_r^{[k]}(z)\leq E_r(z)\leq\frac{e_rT(z)}{(1-T(z))^{3r}}$$

saddle-point methods

$$\mathbb{E}[\mathcal{C}]\sim\sum_{k\geq 0}\left(1-\exp\left(-\frac{1}{2}\int_{(k+1)(1-\frac{2M}{n})}^{\infty}e^{-v}\frac{dv}{v}\right)\right)$$

and finally Euler Maclaurin on k, we get the result.

The maximum block size of critical random graphs

Inside the scaling window, we have the following.

Theorem R., R., R. (2016)

Let λ be any real constant and $M = \frac{n}{2}(1 + \lambda n^{-1/3})$. The maximum block-size of $G(n, M)$ verifies :

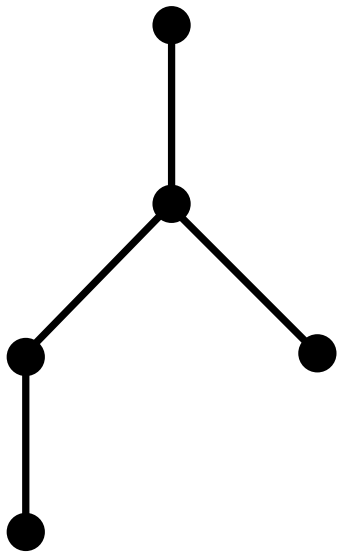
$$\mathbb{E}(\mathcal{C}) \sim c_2(\lambda)\, n^{1/3},$$

where

$$c_2(\lambda) = \frac{1}{\alpha} \int_0^\infty \left(1 - \sqrt{2\pi} \sum_{r \geq 0} \sum_{d \geq 0} A\left(3r + \frac{1}{2}, \lambda\right) e^{-E_1(u)}\, e_{r,d}\left(e^{-u}\right) \right) du$$

($E_1(x)$ is as in the previous theorem, α is the positive solution of $\lambda = \alpha^{-1} - \alpha$ and the $(e_{r,d})$ are polynomials with rational coefficients which can be computed recursively).

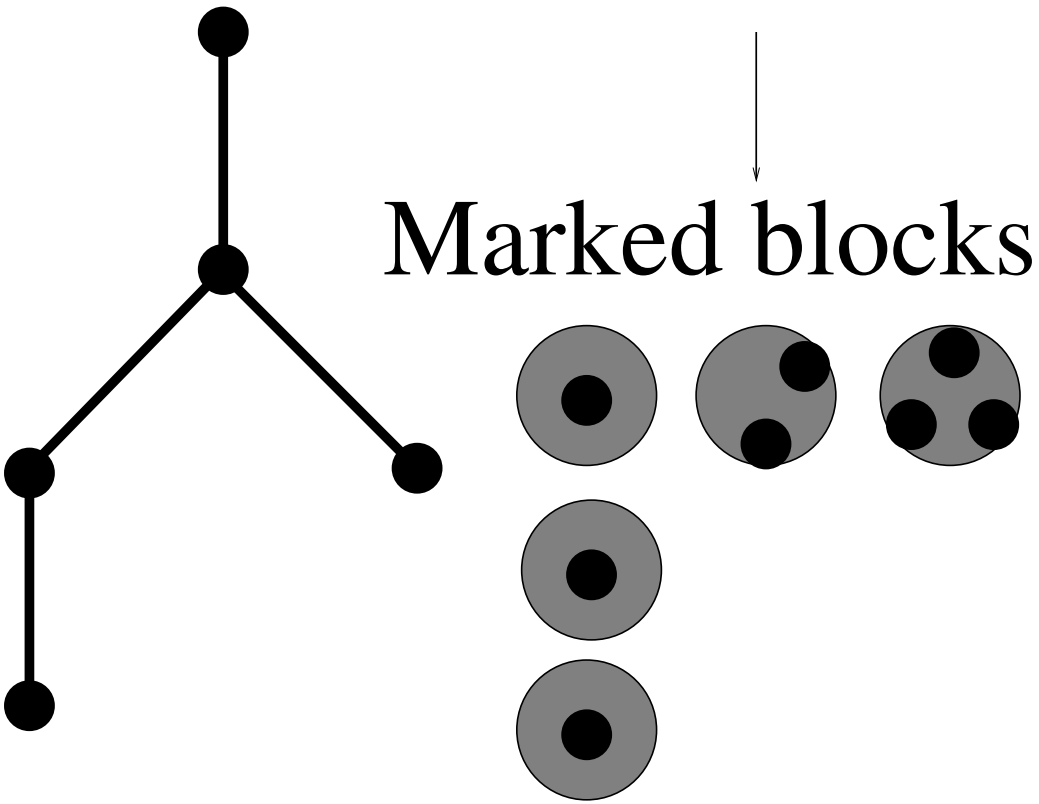
Ideas of the proof



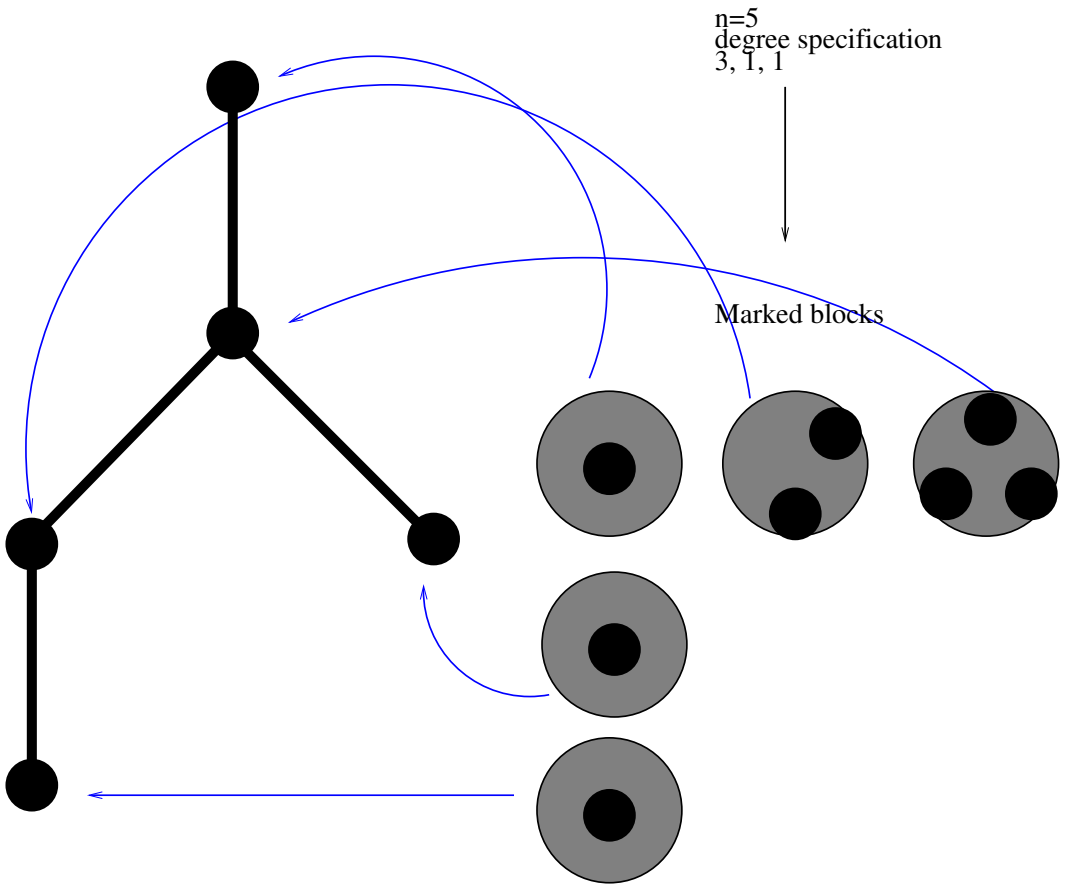
n=5
degree specification

3, 1, 1

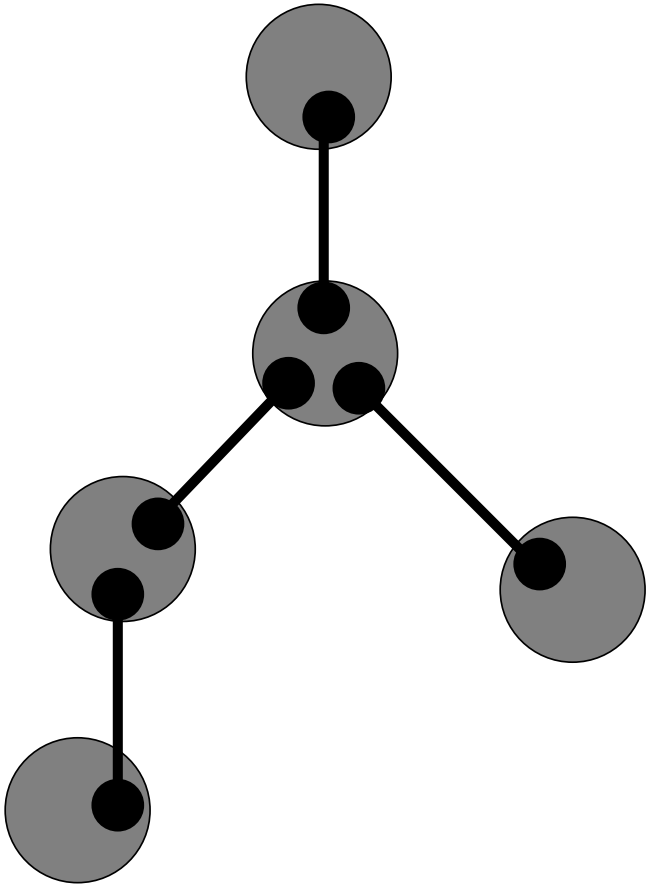
Ideas of the proof



Ideas of the proof



Ideas of the proof



Ideas of the proof

- Enumerate trees with given degree specification (use Prufer code).
- Use $cpr(s, d)$, i.e. CHAE, PALMER, ROBINSON 's formula for the number of **cubic block multigraphs** with s single edges and d double edges.
- Show that the EGF of smooth graphs of excess r with biconnected 3-regular kernels behaves as

$$B_r(z) \asymp \frac{b_r}{(1-z)^{3r}} \text{ with } b_r = \sum_{s+2d=3r} \frac{cpr(s, d)}{2^d\,(2r)!}$$

- Similarly for kernels with at most k vertices

$$B_r^{[k]}(z) \asymp b_r \frac{1-z^k}{(1-z)^{3r}}$$

- Use derivation to count similar s-**marked** objects

$$b_r^{\bullet s} = \frac{1}{s!} b_r \prod_{i=1}^{s} [3r + (s-i)] \text{ and } B_r^{\bullet s,\,[k]}(z) \asymp b_r^{\bullet s} \frac{1-z^k}{(1-z)^{3r+s}}\,.$$

- Plant the blocks in the trees with degree specification (**degree** $= \sharp$ **marked vertices**).

Outline of the talk

- Introduction, motivations & previous works.
- 2-path statistics.
- Largest block.
- Conclusion & Perspectives.

Remarks on the scale $\mathcal{O}(n^{\frac{1}{3}})$

Known results

- On circumference (cf. JANSON, ŁUCZAK, RUCIŃSKI) when $M = \frac{n}{2} + \mathcal{O}(n^{\frac{2}{3}})$, a.a.s $\frac{\mathcal{C}}{\omega} < n^{\frac{1}{3}} < \mathcal{C}\omega$.
- On diameter (cf. NACHMIAS, PERES) If $p = \frac{1}{n}$, for any ε there exists $A = A(\varepsilon) < \infty$ such that for all large n

$$\mathbb{P}\Big[\mathrm{diam}(C_1) \notin \Big[A^{-1} n^{\frac{1}{3}}, A\, n^{\frac{1}{3}}\Big]\Big] < \varepsilon\,.$$

Remarks on the scale $\mathcal{O}(n^{\frac{1}{3}})$

Known results

- On circumference (cf. JANSON, ŁUCZAK, RUCIŃSKI) when $M = \frac{n}{2} + \mathcal{O}(n^{\frac{2}{3}})$, a.a.s $\frac{\mathcal{C}}{\omega} < n^{\frac{1}{3}} < \mathcal{C}\omega$.
- On diameter (cf. NACHMIAS, PERES) If $p = \frac{1}{n}$, for any ε there exists $A = A(\varepsilon) < \infty$ such that for all large n

$$\mathbb{P}\Big[\mathrm{diam}(C_1) \notin \Big[A^{-1} n^{\frac{1}{3}},\, A\, n^{\frac{1}{3}}\Big]\Big] < \varepsilon\,.$$

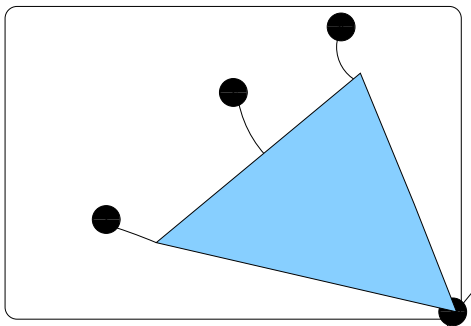
Using that the length X_n of a random 2-path in $G(n,\, M = \frac{n}{2} + \mu n^{\frac{2}{3}})$) verifies $\mathbb{E}(X_n(X_n-1)\cdots(X_n-k+1)) \sim \sqrt{2\pi}\sum_{R>0} e_R\, A(3R+k+\frac{1}{2},\mu)\, n^{\frac{k}{3}}$ and method of moments, it can be shown that **every distance** (including diameter, longuest path, circumference) scales as $\mathcal{O}(n^{\frac{1}{3}})$ w.h.p.

Example:

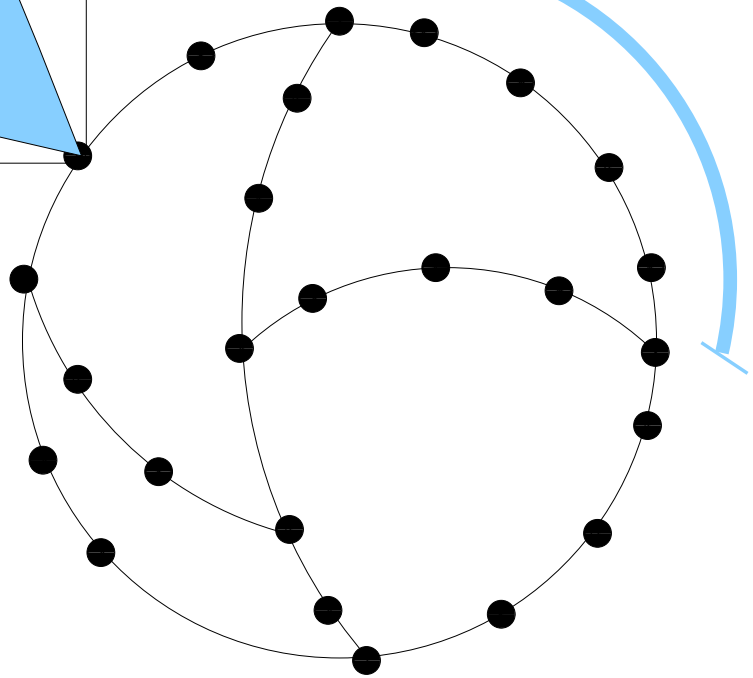
$\forall\varepsilon \qquad \exists B = B(\varepsilon) < \infty$ s.t. as $n \gg 1$

$$\mathbb{P}\Big[\text{longuest path} \notin \Big[B^{-1} n^{\frac{1}{3}},\, B\, n^{\frac{1}{3}}\Big]\Big] < \varepsilon\,.$$

Ideas of the proof



2–path ~ c $n^{1/3}$



pending trees of size $O(n^{1/3})$ and height $O(n^{1/6})$

Excess is bounded

Conclusion

Works done

- Random graphs distances are of order $\mathcal{O}(n^{\frac{1}{3}})$ inside their critical windows. (cf. also ADDAGIO-BERRY, BROUTIN, GOLDSMITH using other methods)
- Enumerative/analytic techniques to capture the expectation of largest block (cf. arxiv.org/abs/1605.04340).

Possible extensions

- Distribution of other extremal distance parameters
- Graphs/hypergraphs with constraints on degrees. (cf. works of DE PANAFIEU, RAMOS (2016))

THANK!