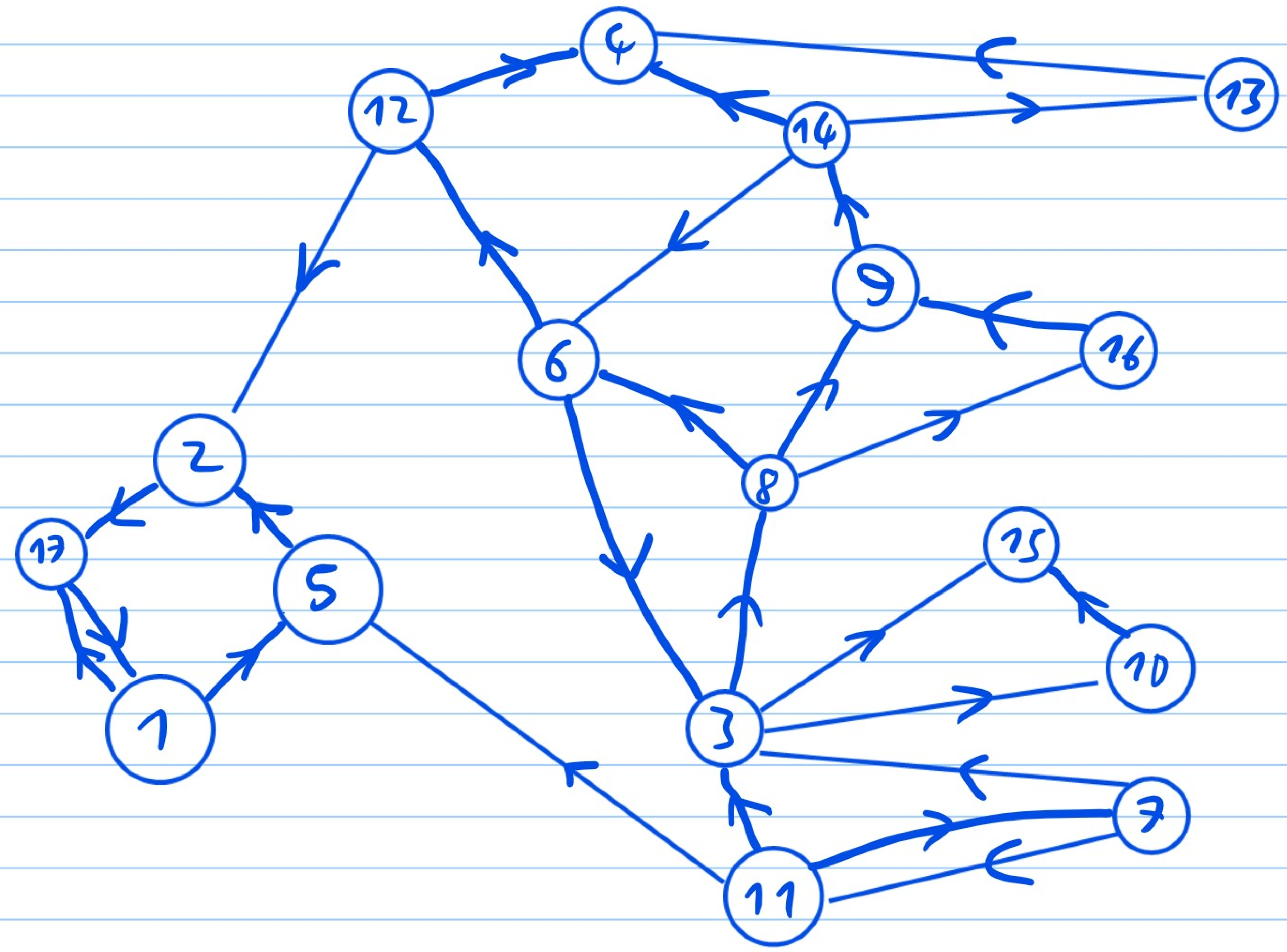
The scaling limit of a critical random directed graph

JW
Christina Goldschmidt

I Introduction

a) $\vec{G}(n,p)$



vertices: $\{1, 2, \dots, n\}$

edges: keep each of the $n(n-1)$ possible directed edges independently with probability p.

We're interested in strongly connected components (SCC)

x and y are in the same SCC if there are two paths: $x \to y$
$y \to x$

b) Phase transition for the SCC

Same as the Erdős–Rényi graph:

$p = \frac{c}{n}$ · $c > 1$ $\to$ unique giant component
· $c < 1$ $\to$ only small components

In fact, something more precise is known

Theorem: Łuczak–Seierstad ('09)

Write $p = \frac{1}{n} + \frac{d_n}{n^{4/3}}$, assume $d_n = o(n^{1/3})$

· If $d_n \to +\infty$ then: · largest SCC has size $\sim 4 n^{1/3} d_n^2$
· 2nd largest has size $O(n^{2/3}/d_n)$

· If $d_n \to -\infty$ then: · largest SCC has size $O\left(\frac{n^{2/3}}{|d_n|}\right)$

Q: What happens if d_n stays bounded or converges?

II Case of the Erdös-Rényi graph

$G(n,p)$ $\quad p = \frac{1}{n} + \frac{d}{n^{4/3}} + o\left(\frac{1}{n^{4/3}}\right)$ $\quad d \in \mathbb{R}$

Theorems:

- Aldous (97): let $Z_1^n \geq Z_2^n \geq \dots$ be the sizes of the connected components of $G(n,p)$. Then

$$\boxed{\frac{Z_i^n}{n^{2/3}} \xrightarrow{(d)} \sigma_i}$$

(in ℓ^2, as sequences)

- Addario-Berry, Broutin, Goldschmidt (12)

Let $(C_i^n)_i$ be the cc of $G(n,p)$ with

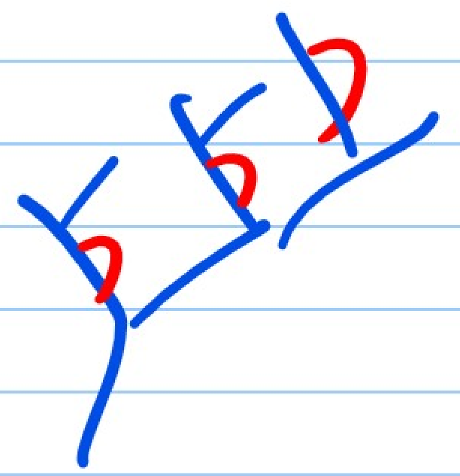
$\# C_i^n = Z_i^n$. Then

$$\boxed{\frac{C_i^n}{n^{1/3}} \xrightarrow{(d)} \mathcal{C}_i}$$

(in Gromov-Hausdorff, as a sequence)

The main takeaway is that the distance between two typical points is of order $n^{1/3}$.

(Side note: both the C_i^n and $\mathcal{C}_i$ have descriptions as "binary trees with a few additional edges")



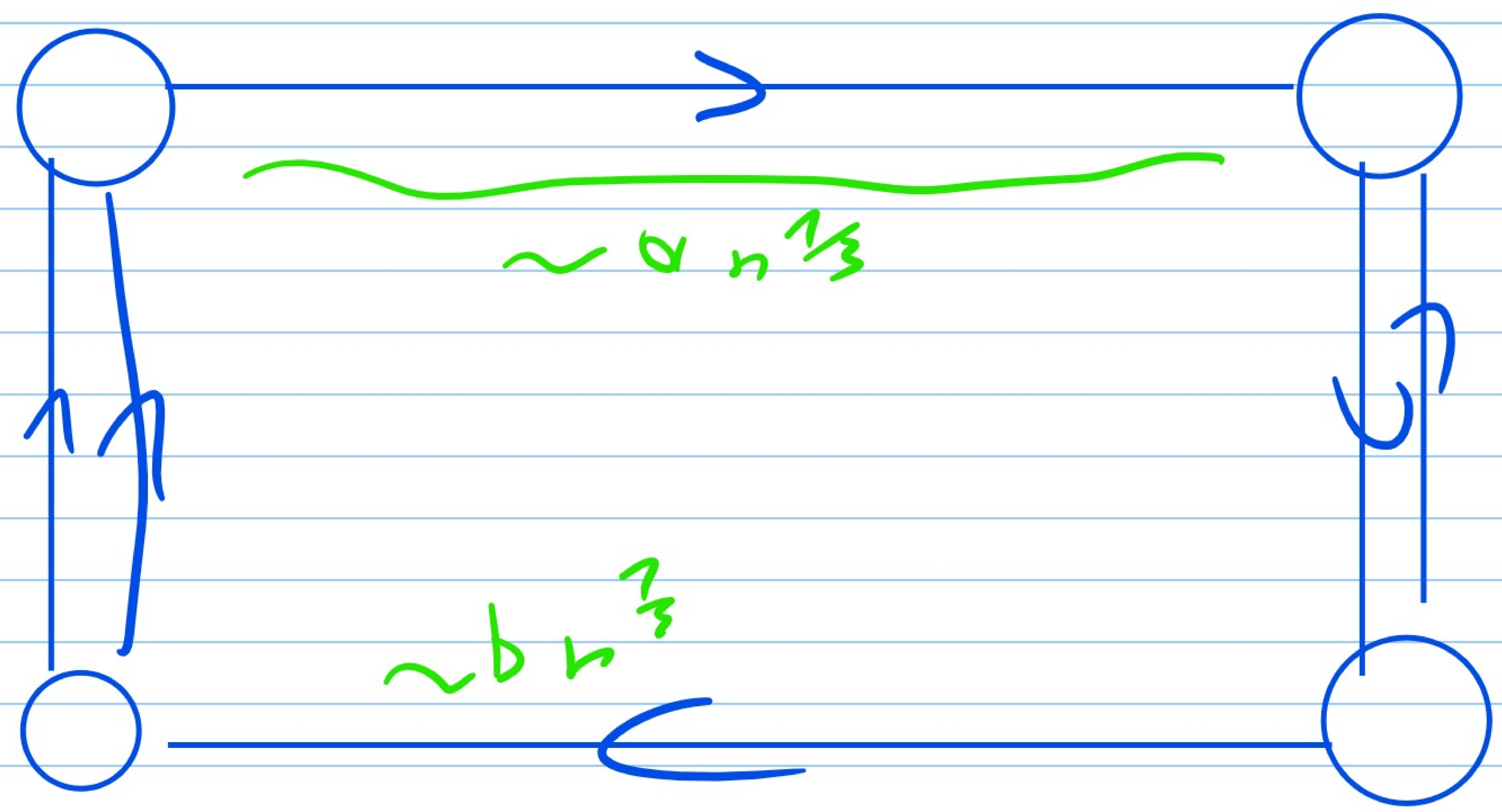
III Our results

$$p = \frac{1}{n} + \frac{\lambda}{n^{4/3}} + o\left(\frac{1}{n^{4/3}}\right)$$

a) Informally

Within the SCC:

- w.h.p. no vertices with degree ≥ 4
- number of vertices with degree $= 3$ is $\asymp 1$
- number of vertices with degree $= 2$ is $\asymp n^{1/3}$

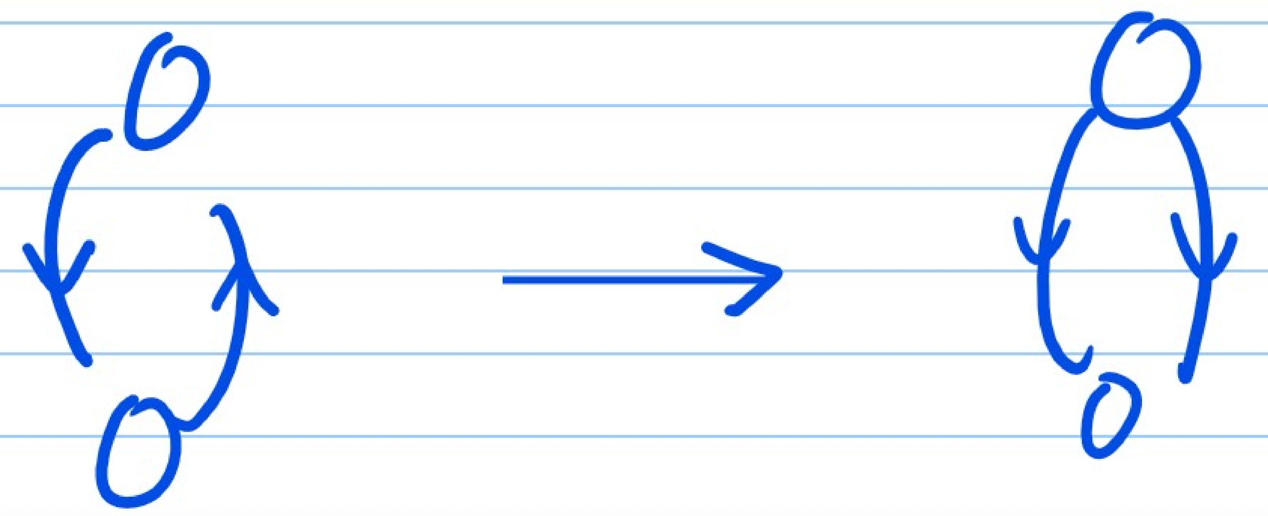


b) Metric directed multigraphs (MDM)

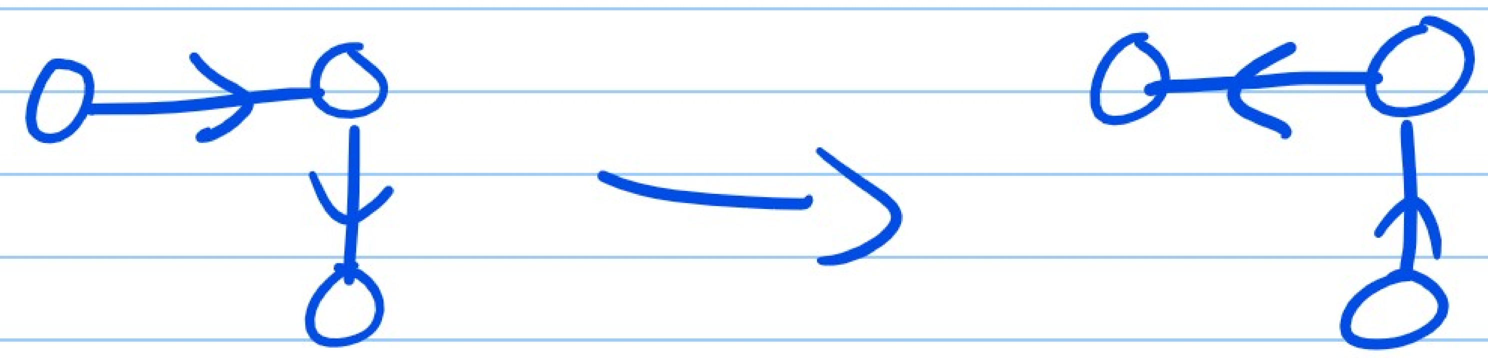
Definition: $\vec{\mathcal{G}}$ is the set of finite multigraphs where each edge has a direction and a length

- An isomorphism is a pair of bijections
  $f: V(X) \to V(Y)$
  $g: E(X) \to E(Y)$
  which preserve the structure

examples: How many isomorphisms?



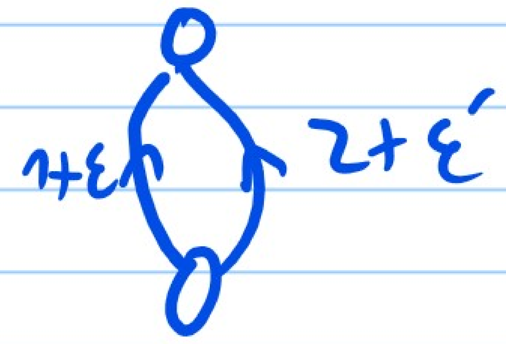
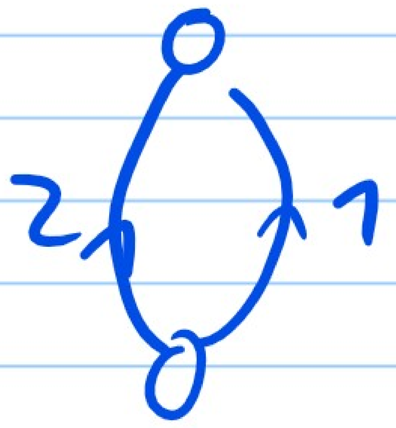
0



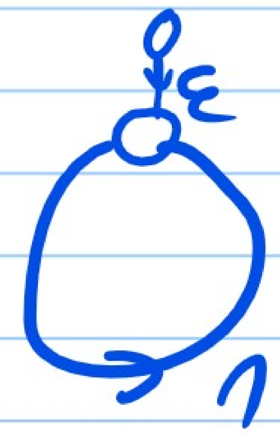
1

2

distance: $$d_{\overrightarrow{G}}(X,Y) = \inf_{\substack{(f,g) \\ \text{isom } X \to Y}} \sum_{e \in E(X)} \left| \ell_X(e) - \ell_Y(g(e)) \right|$$

$d_{\vec{\mathcal{G}}}\Big($  ,  $\Big) = \varepsilon + \varepsilon'$

<u>remark:</u> this metric is very rigid :

$d_{\vec{\mathcal{G}}}\Big($  ,  $\Big) = \infty$

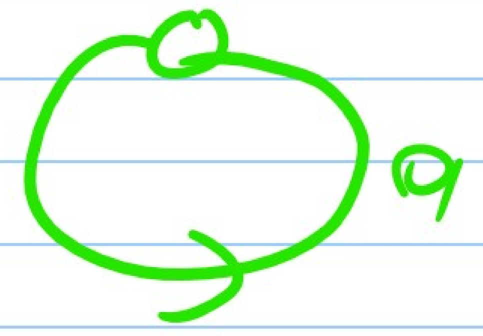
c) Main theorem:

Let $(C_i(n), n \in \mathbb{N})$ be the SCC of $\vec{G}(n,p)$, viewed as MDMs by removing the $d^0 = 2$ vertices. Then:

$$\boxed{\left(\frac{C_i(n)}{n^{1/3}}, i \in \mathbb{N}\right) \xrightarrow{d} (\mathcal{C}_i, i \in \mathbb{N})}$$

where $\mathcal{C} \in \vec{\mathcal{G}}^{\mathbb{N}}$ s.t.

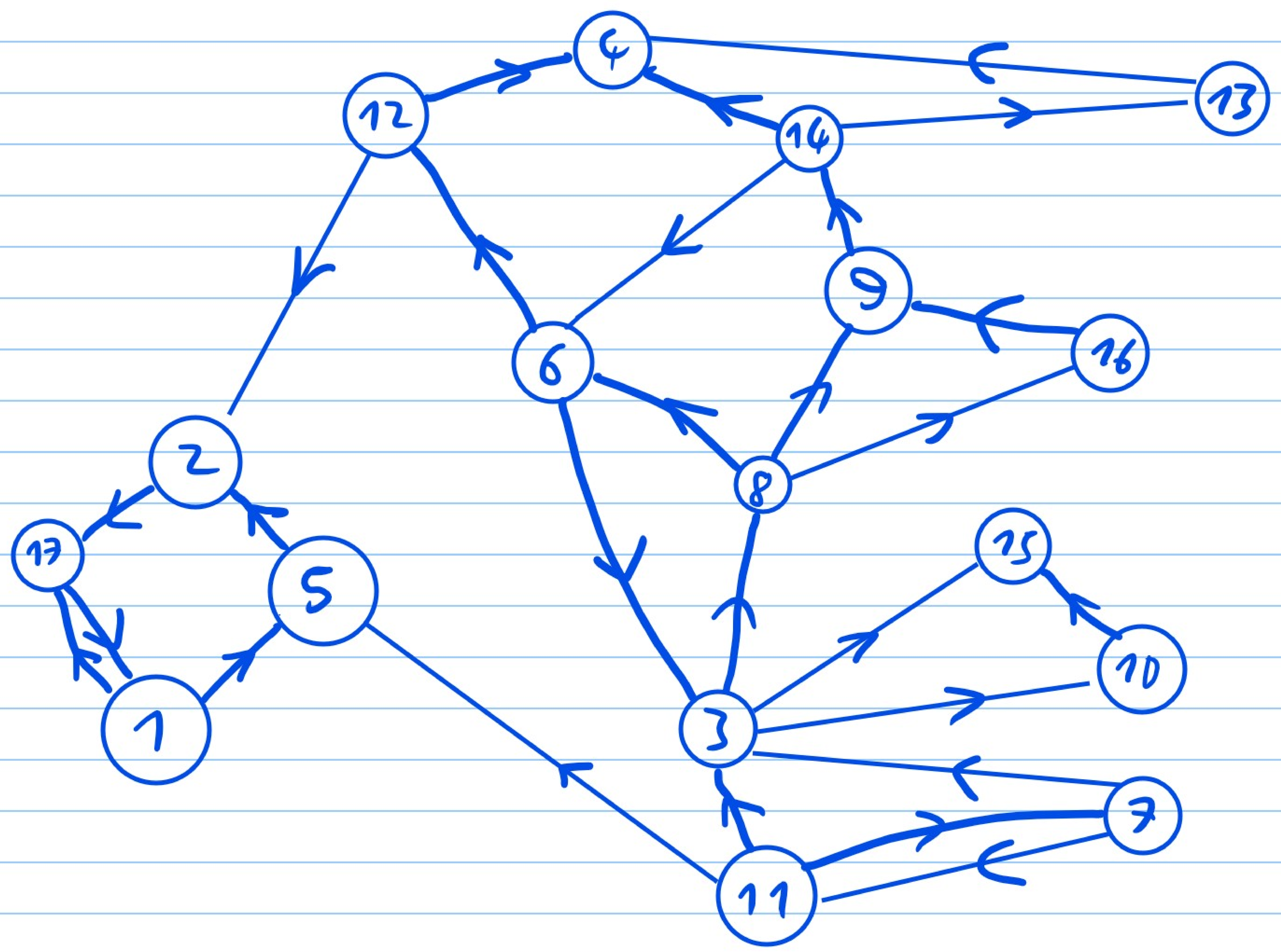
- finitely many terms are 3-regular
- the rest are cycles.

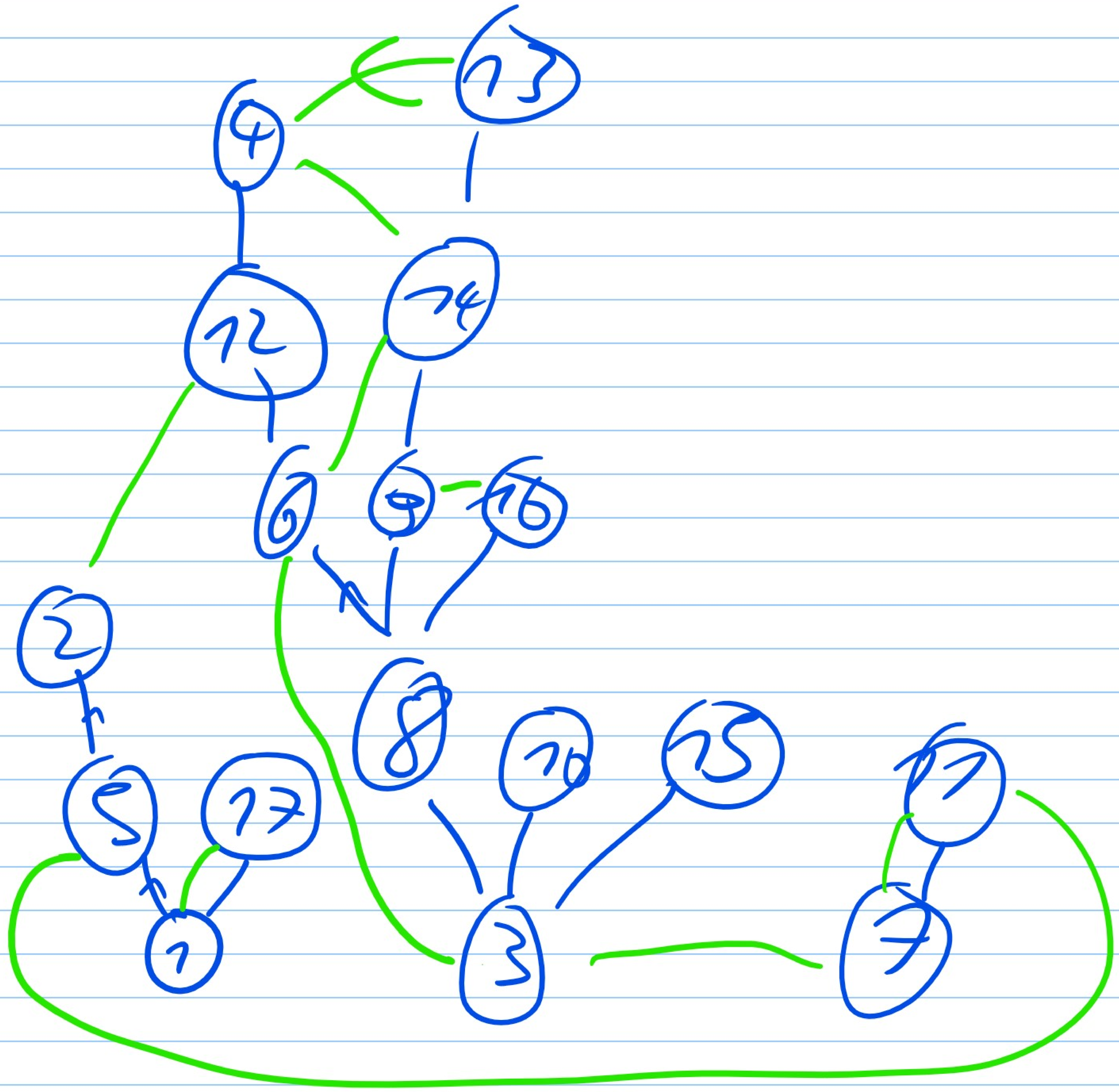


topology for sequences: $d(A,B) = \sum_{i=1}^{\infty} d_{\to_{\varphi}}(A_i, B_i)$

IV Exploration and structure

a) A version of depth-first search





b) Edge classification

— tree edge (forward)

— surplus edge (forward)

— back edge (backward)

Interaction between forward and backward

$\longrightarrow$ SCC.

V Limit behaviour of forest and surplus

prop: "Forward edges without arrows"

$$\overset{(d)}{=} \quad G(n,p)$$

Erdős-Rényi

So we know a lot about the trees and surplus edges. In particular:

- number of vertices in a tree $\sim n^{2/3}$
- typical distances are $\sim n^{1/3}$

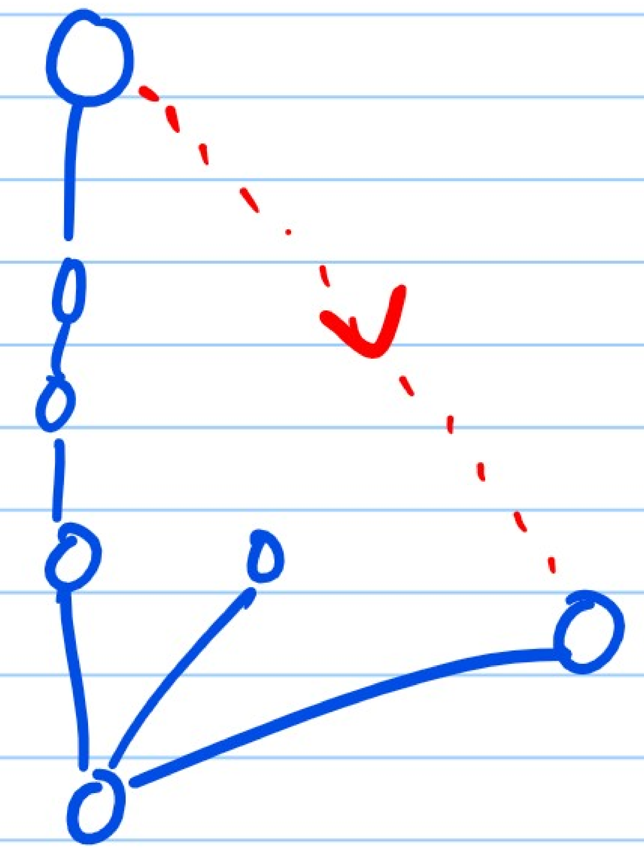
- surplus edges are $O(1)$ in number.

It turns out, surplus edges don't count!

prop: $\mathbb{P}[\text{a SCC of } \vec{G}(n,p) \text{ has a surplus edge}] \xrightarrow[n\to\infty]{} 0$

"pf":

$\mathbb{P}(\quad \text{counts})$



VI Limit of the back edges

a) problem?

In a single tree, there are $\frac{k(k-1)}{2}$ possible back edges, appearing independently with probability $p \sim \frac{1}{n}$.

But $k \approx n^{2/3}$ ...

$E[\# \text{ of back edges}] \approx \frac{1}{n} n^{4/3} \longrightarrow \infty$??

<u>b) solution!</u> Most of the BE don't matter!

How to find those which do matter:

- go around the tree
- keep the first ancestral BE
- keep any BE which is either ancestral or links into something we've already selected.

This turns out to converge to a PPP on the limiting continuum tree!

What we end up with:

Other works:

- Coulson ('19): estimates for size of largest component
- De Panafieu, Dovgal, Ralaivaosaona, Rasendrahasina, Wagner ('20): asymptotics for
  - probability of acyclic $\sim c/n$
  - ———— only singleton and cyclic SCCs. $\to c'$

c and c' explicit