Topology of Scale-Free Graphs

How Random Interaction Begets Holes

Chunyin Siu
Cornell University
cs2323@cornell.edu

Topology of Scale-Free Graphs — Homology and Homotopy

How Random Interaction Begets Holes

Chunyin Siu
Cornell University
cs2323@cornell.edu

So, preferential attachment…

(Stephen Coast
https://www.fractalus.com/steve/stuff/ipmap/)

So, preferential attachment…

- topological properties

(Stephen Coast
https://www.fractalus.com/steve/stuff/ipmap/)

So, preferential attachment…

- topological properties
- random fluctuation?

(Stephen Coast
https://www.fractalus.com/steve/stuff/ipmap/)

So, preferential attachment…

- topological properties
- random fluctuation?

- —> random topology

(Stephen Coast
https://www.fractalus.com/steve/stuff/ipmap/)

Agenda

random topology

Agenda

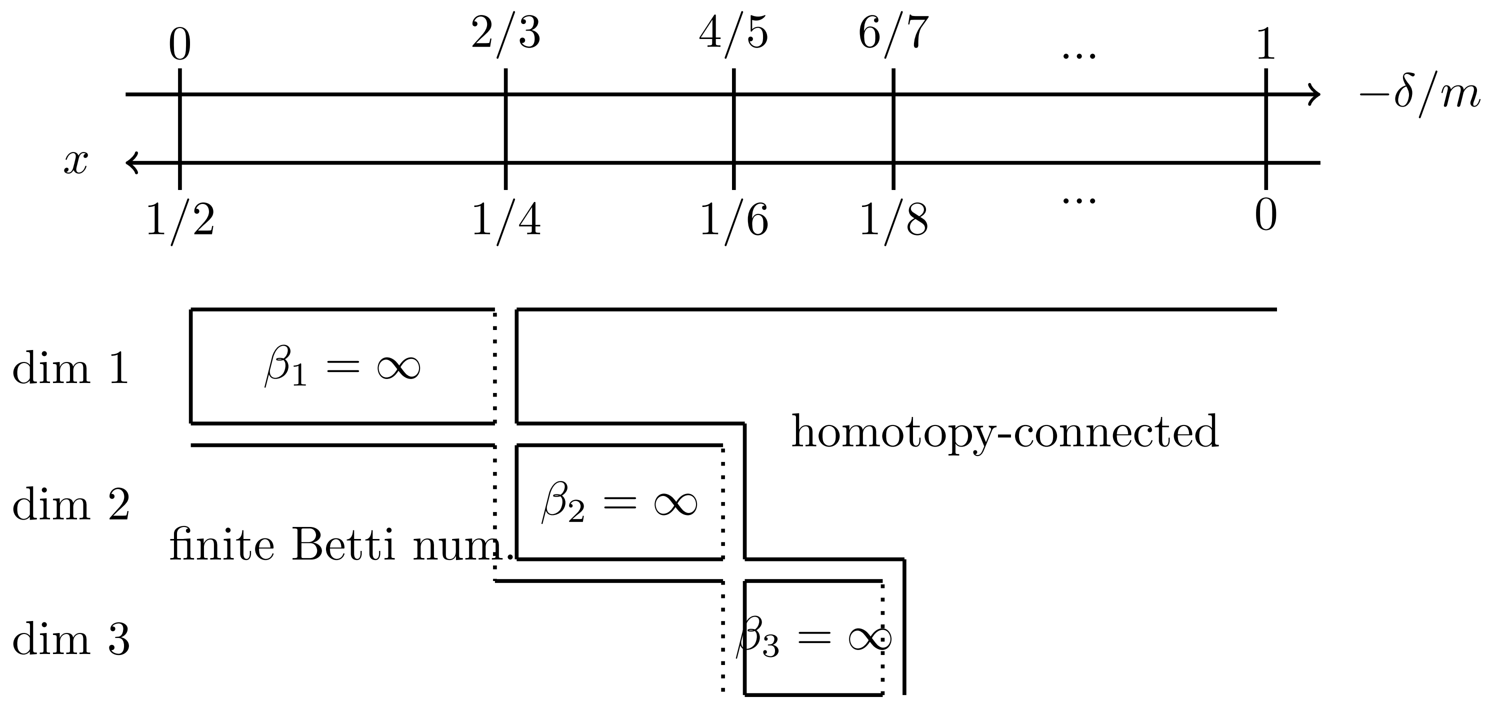
random topology

preferential attachment

Agenda

random topology

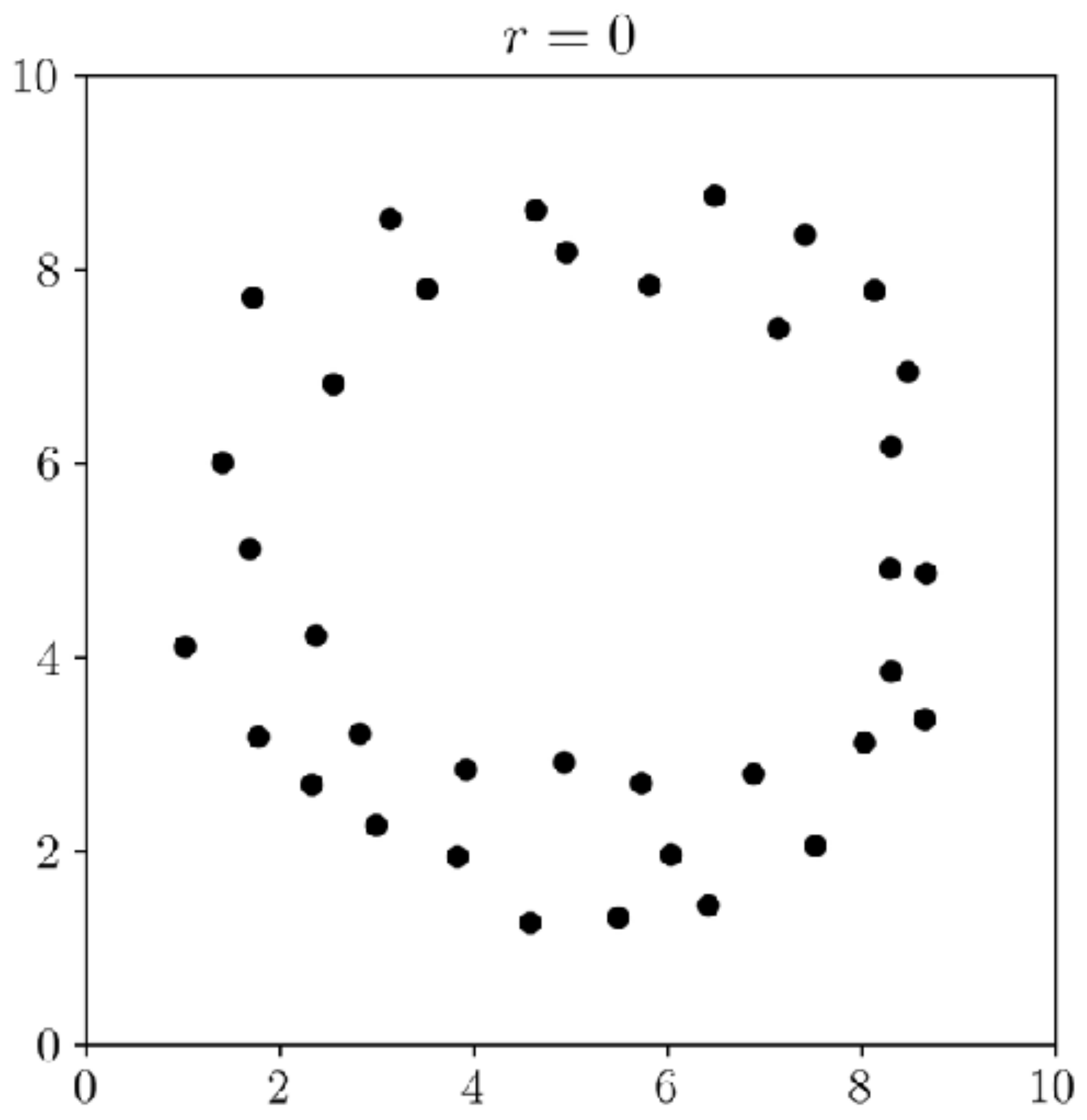
preferential attachment

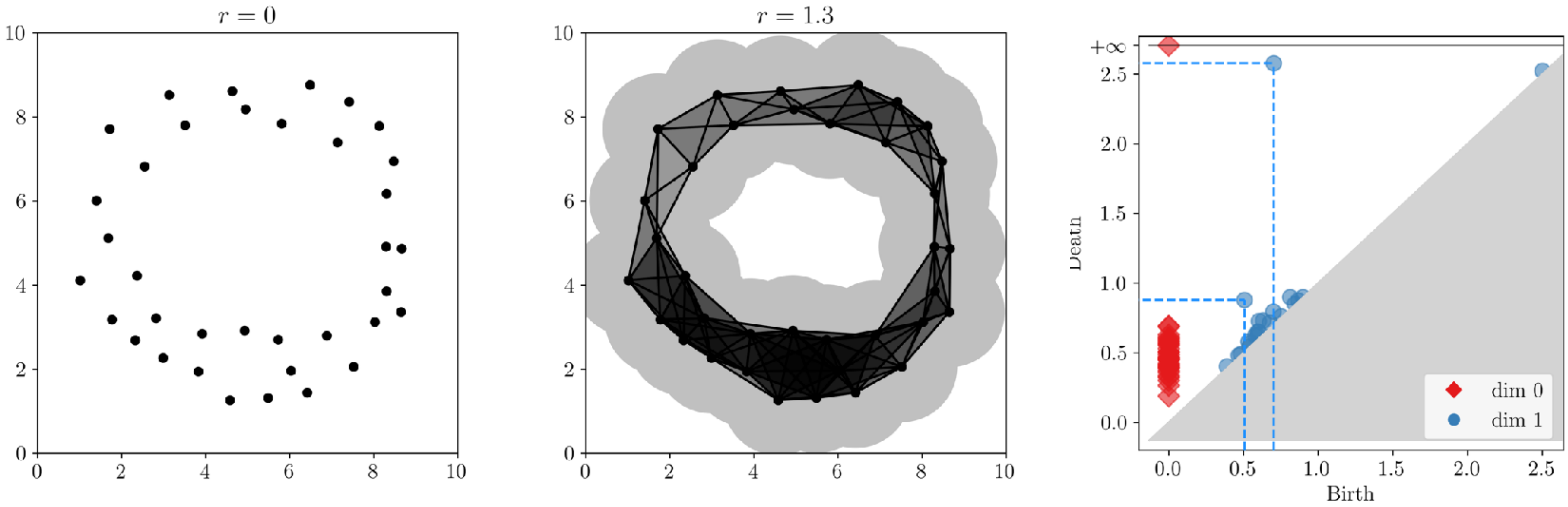


our result

I. A Probabilist’s Apology

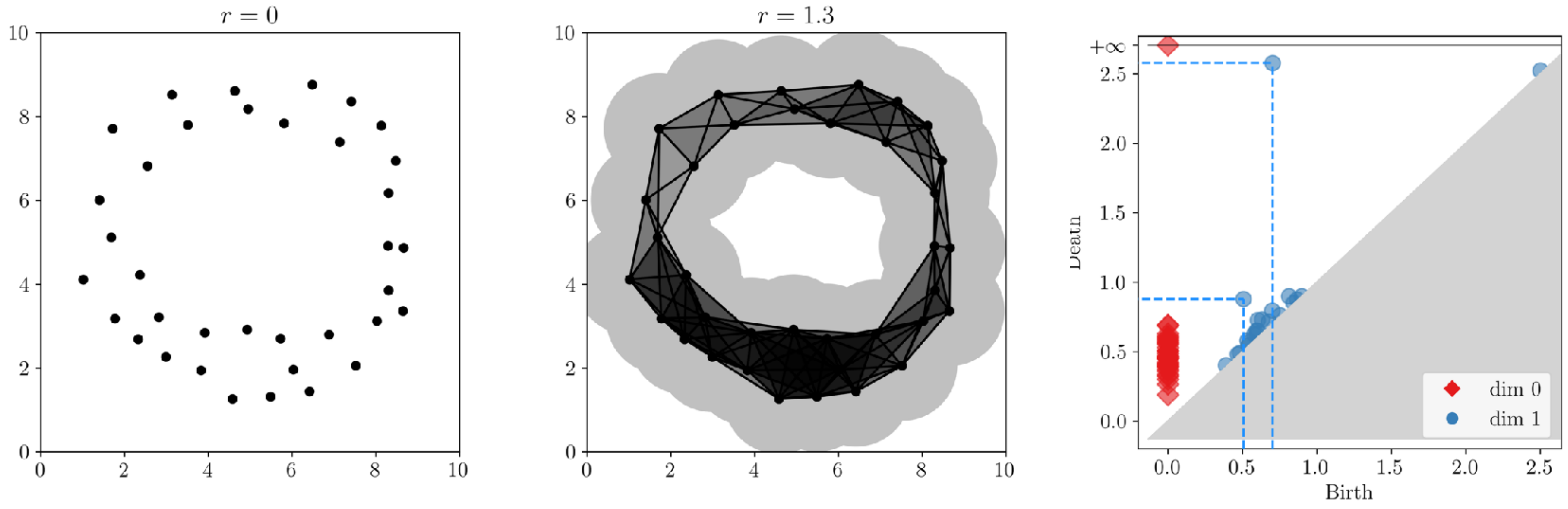
Why Random Topology and What we Know


$r = 0$
10
8
6
4
2
0
0
2
4
6
8
10

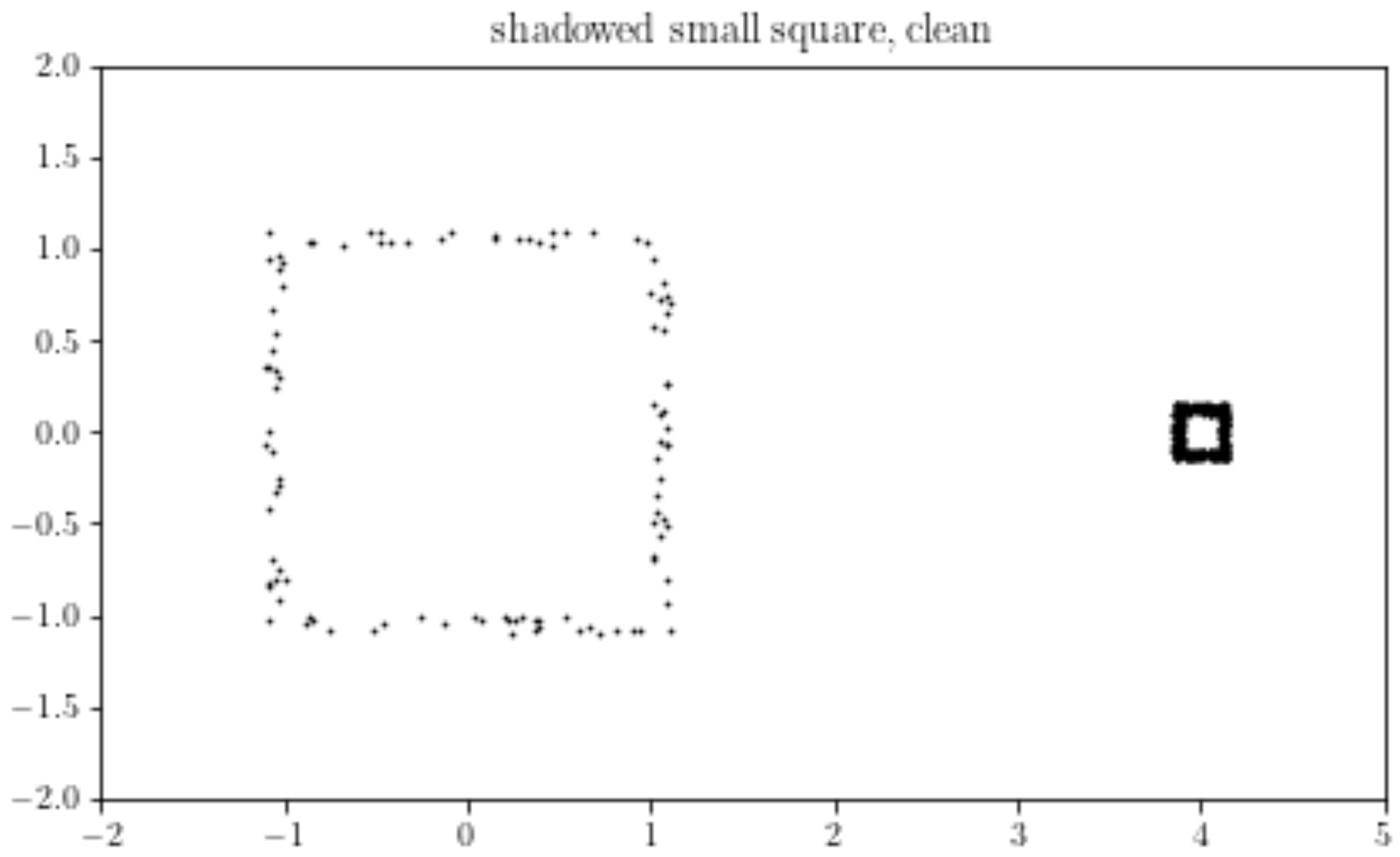


plots generated by Andrey Yao

Size is Signal

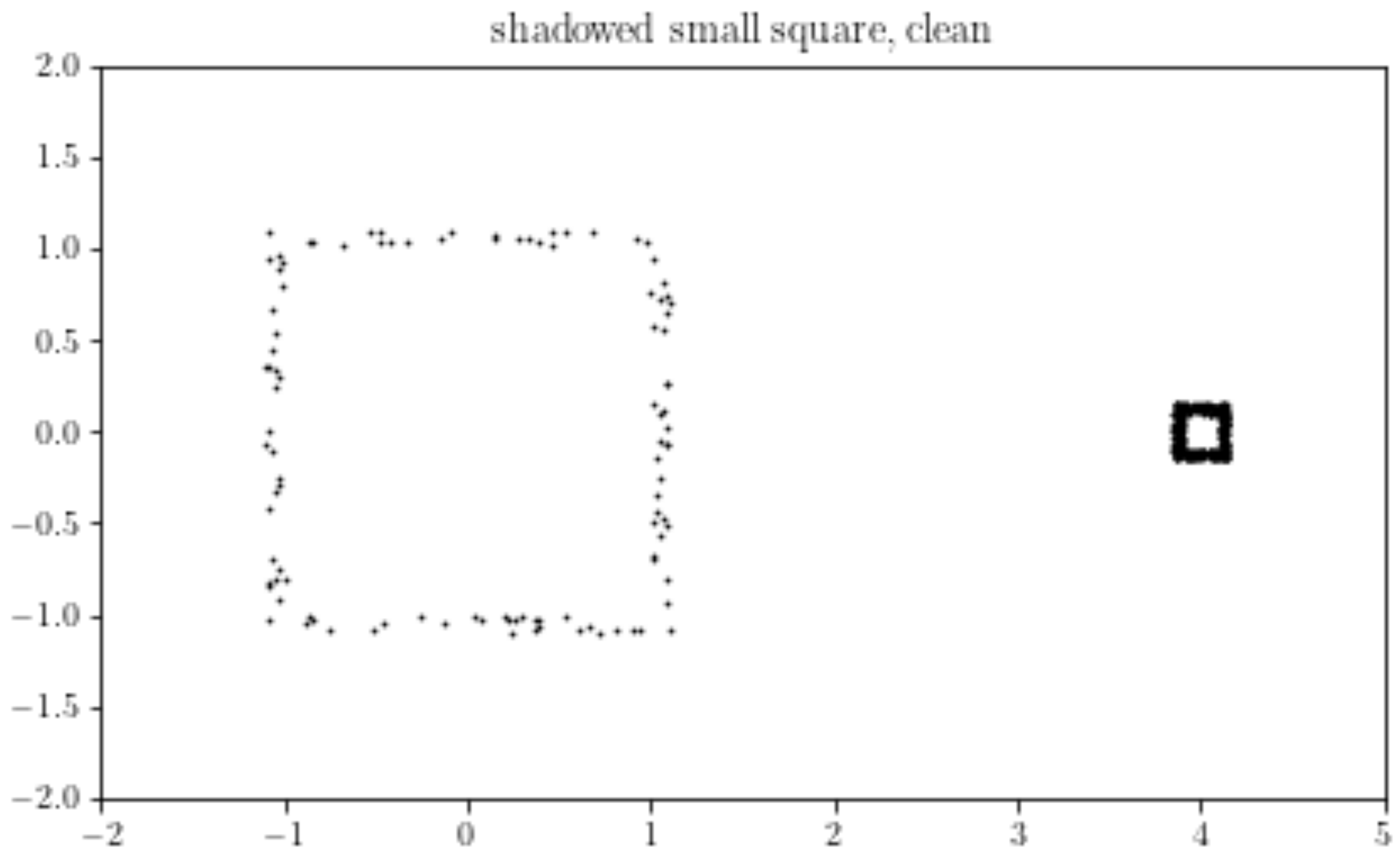


Or is it?

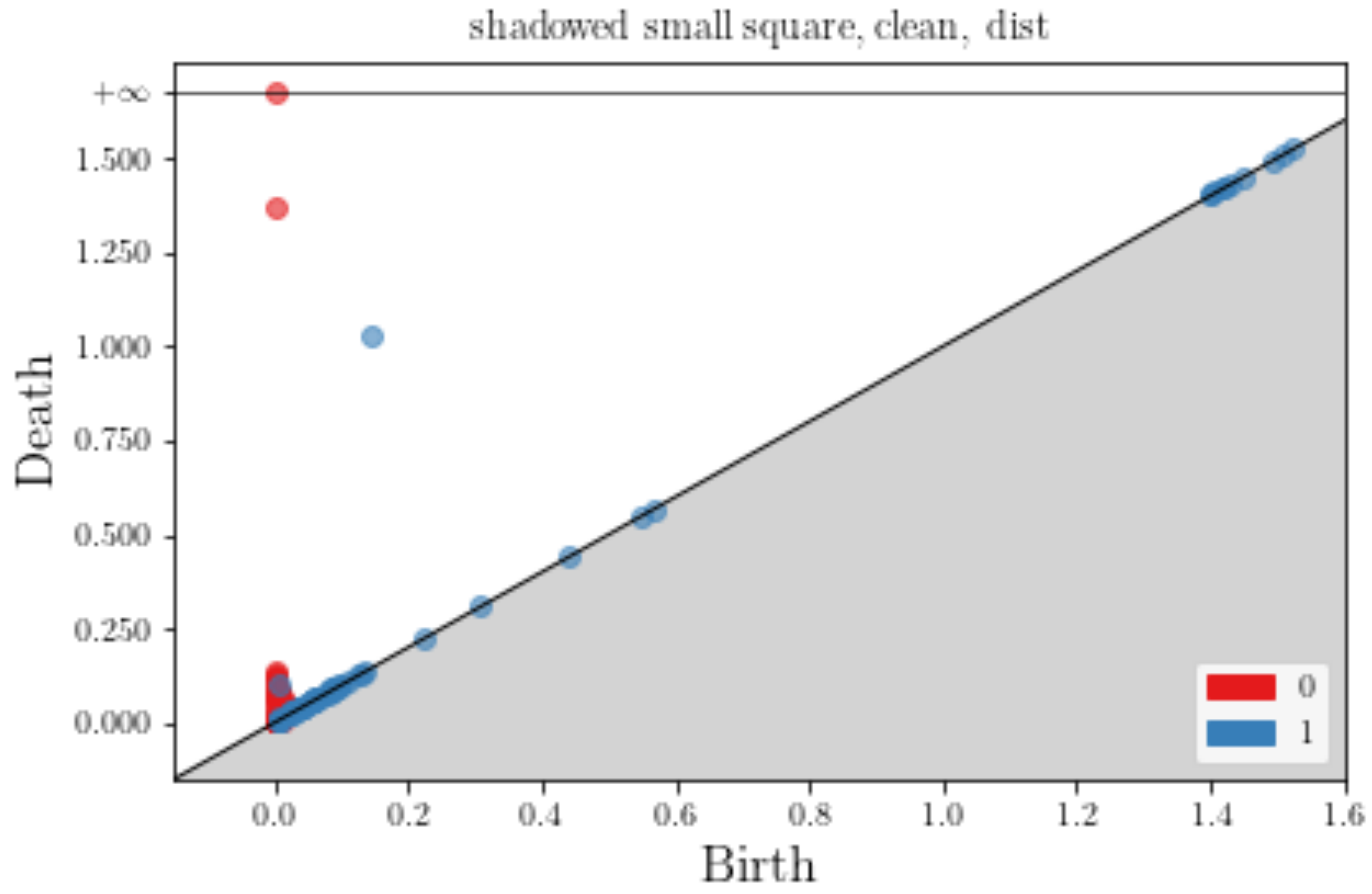


shadowed small square, clean
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2
−1
0
1
2
3
4
5

Or is it?



shadowed small square, clean
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2
−1
0
1
2
3
4
5



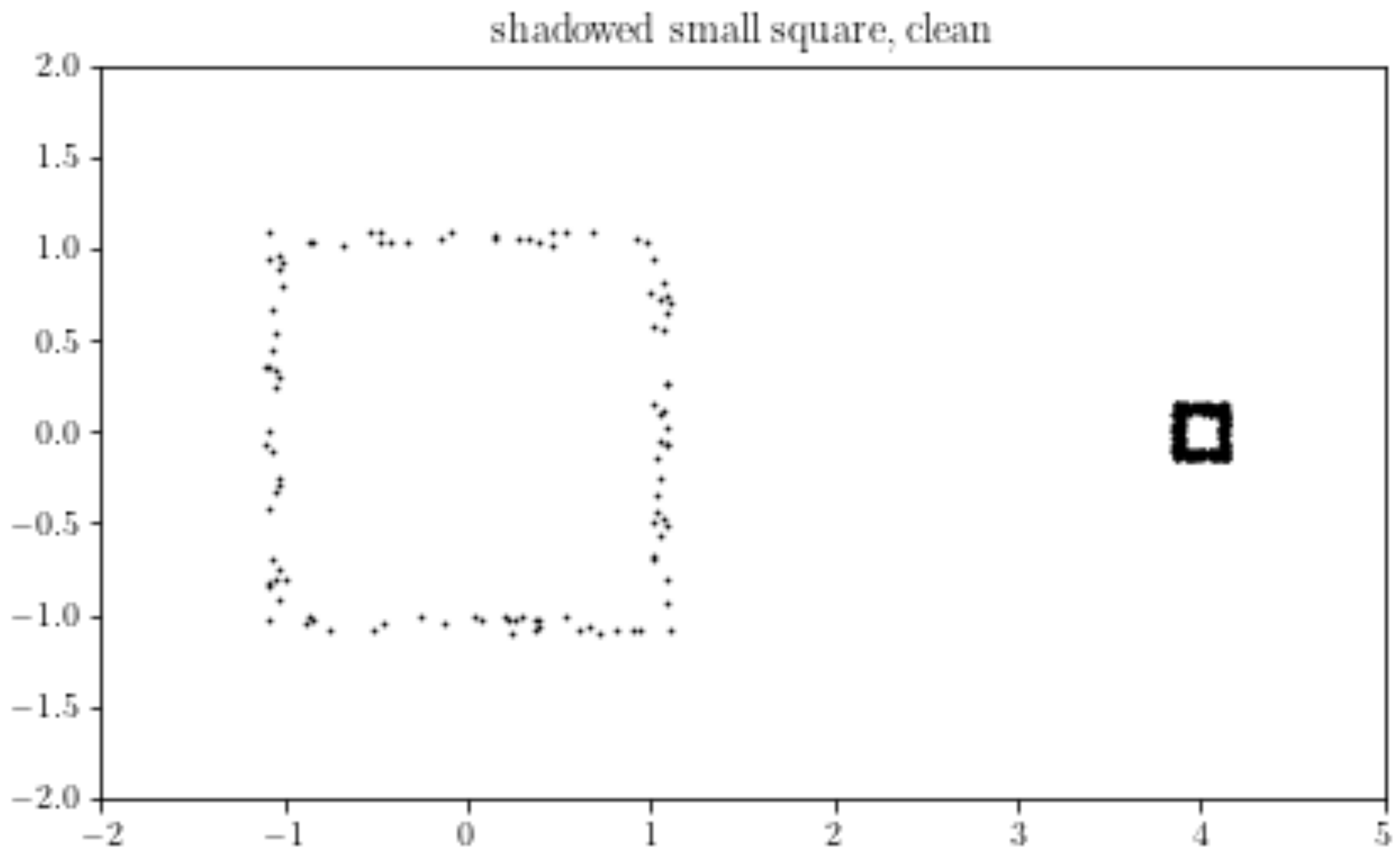
shadowed small square, clean, dist
+∞
1.500
1.250
1.000
0.750
0.500
0.250
0.000
Death
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
Birth
0
1

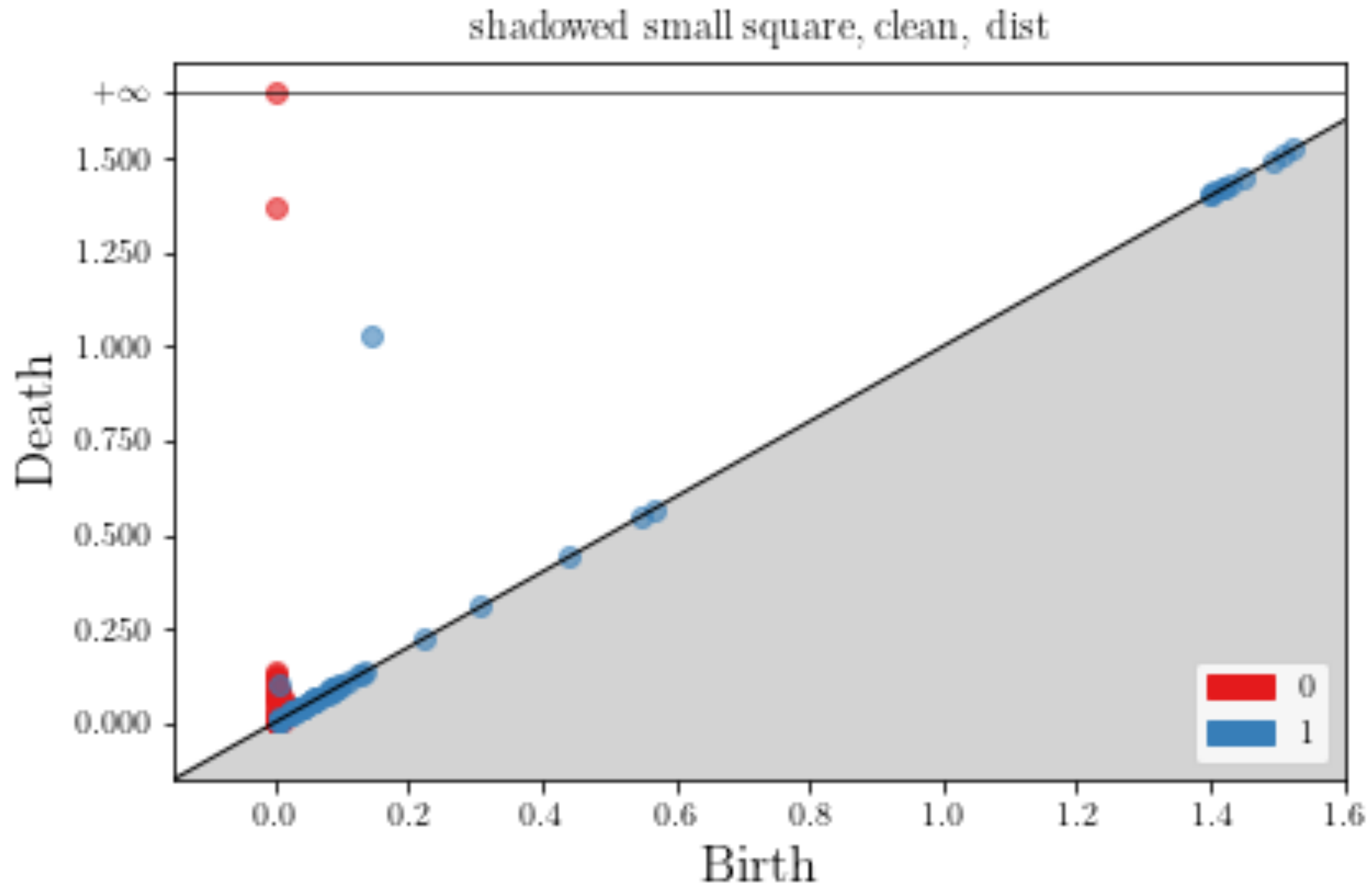
Size is Signal?

Surprise

~~Size~~ is Signal.

Random points don't do that.





Signal is what is not random.

Signal is what is not **random.**
So what is **random?**

What we know
[not meant to be complete]

What we know
[not meant to be complete]

- Erdos-Renyi clique complexes

What we know
[not meant to be complete]

- Erdos-Renyi clique complexes
    - Kahle 2009, 2014
    - Kahle and Meckes 2013
    - Costa et al 2015
    - Malen 2023
    - etc

What we know
[not meant to be complete]

- Erdos-Renyi clique complexes
    - Kahle 2009, 2014
    - Kahle and Meckes 2013
    - Costa et al 2015
    - Malen 2013
    - etc
- random geometric complexes

What we know
[not meant to be complete]

- Erdos-Renyi clique complexes
  - Kahle 2009, 2014
  - Kahle and Meckes 2013
  - Costa et al 2015
  - Malen 2013
  - etc
- random geometric complexes
  - Kahle 2011
  - Kahle and Meckes 2013
  - Yogeshwaran and Adler 2015
  - Bobrowski et al 2017
  - Hiraoka et al 2018
  - Thomas and Owada 2021a, b
  - Owada and Wei 2022
  - etc

II. Preferential Attachment

Beyond independence and homogeneity

Independent and identically distributed?

Independent and identically distributed?

(Stephen Coast
https://www.fractalus.com/steve/stuff/ipmap/)

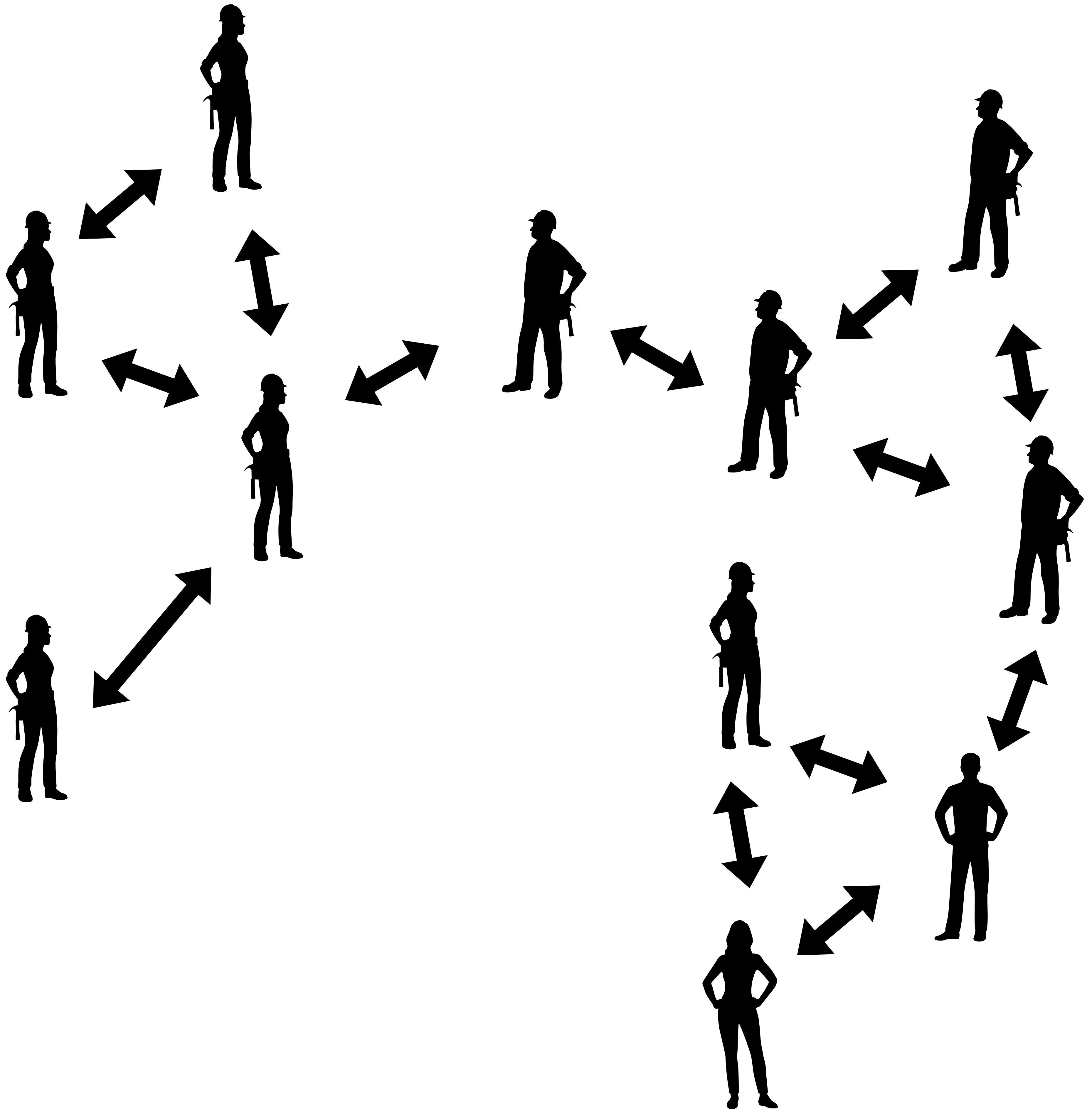
Preferential Attachment

[Albert and Barabasi 1999]

(Stephen Coast
https://www.fractalus.com/steve/stuff/ipmap/)

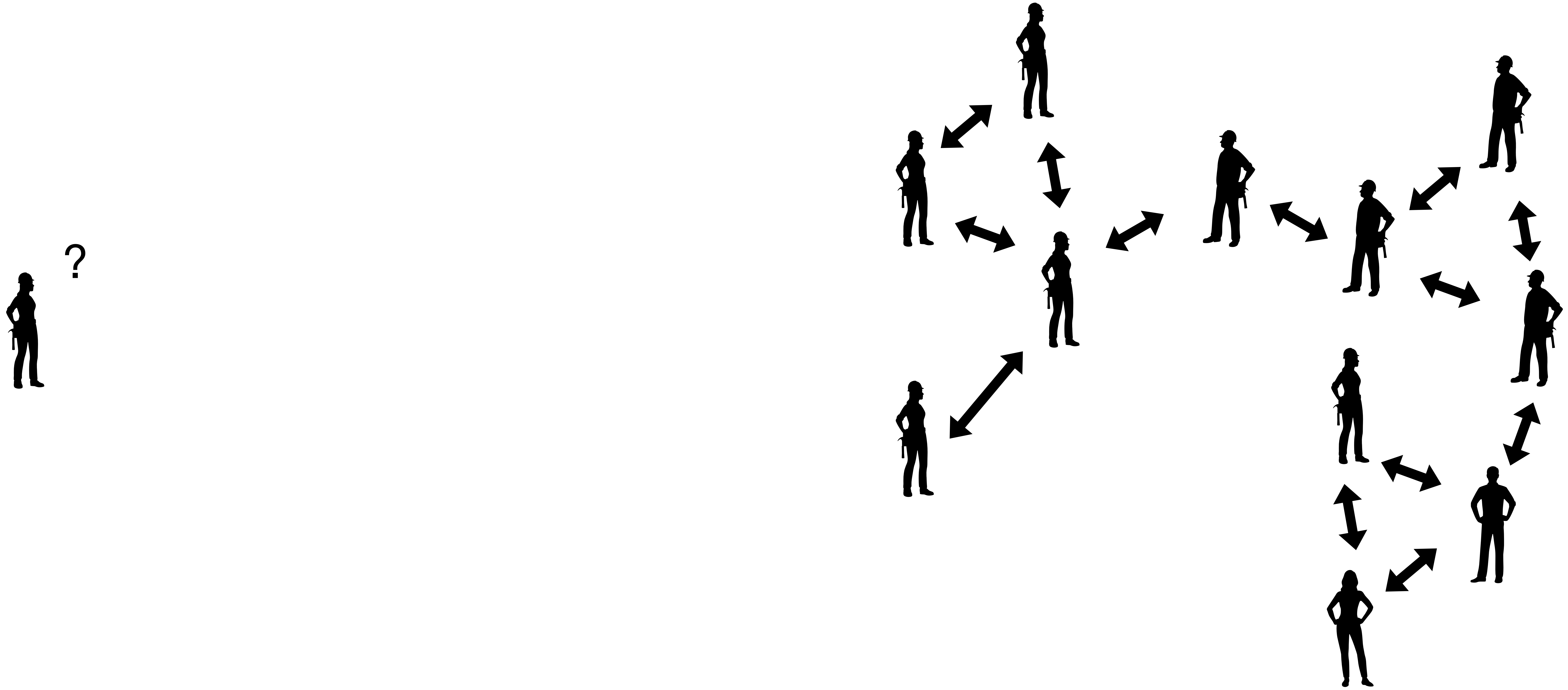
Preferential Attachment

[Albert and Barabasi 1999]



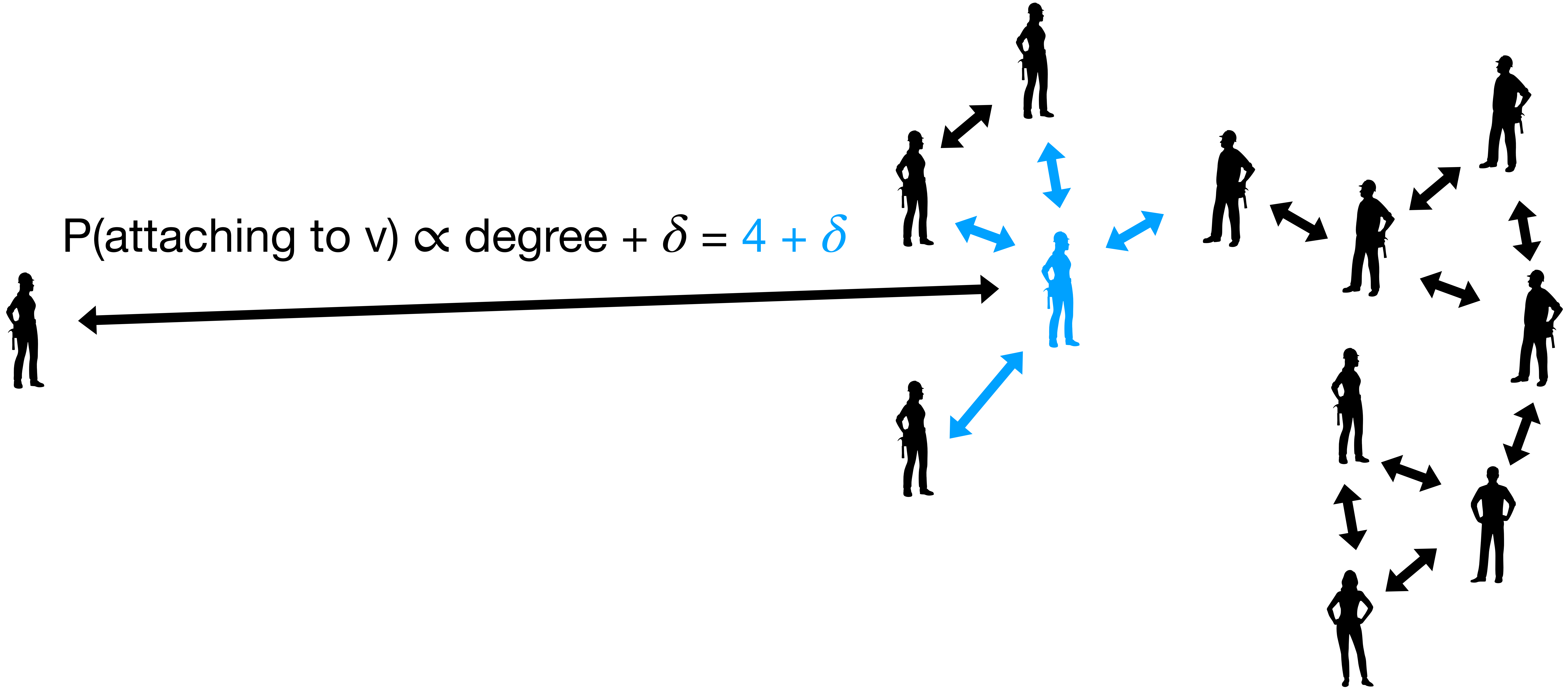
Preferential Attachment

[Albert and Barabasi 1999]



Preferential Attachment

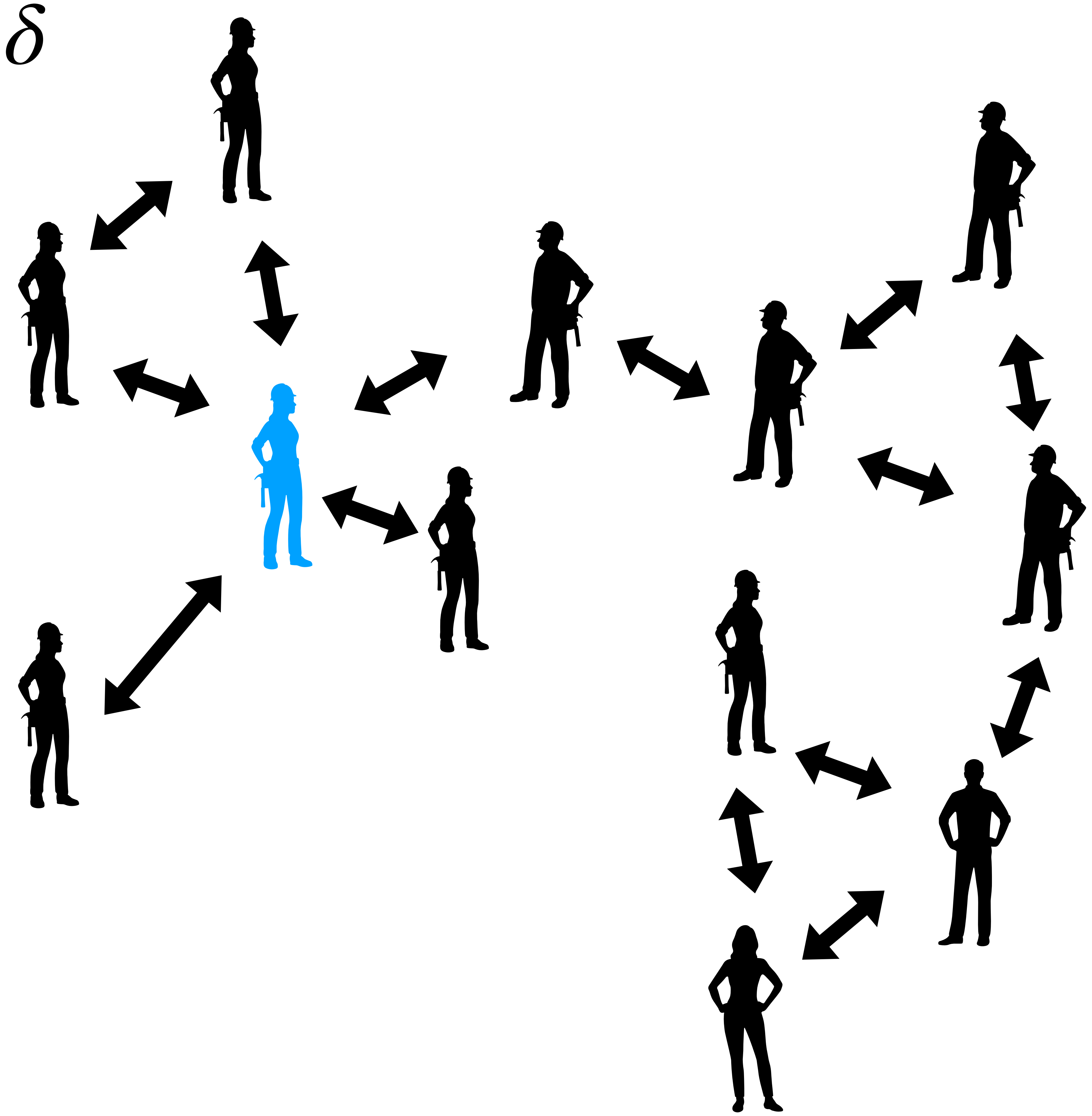
[Albert and Barabasi 1999]



Preferential Attachment

[Albert and Barabasi 1999]

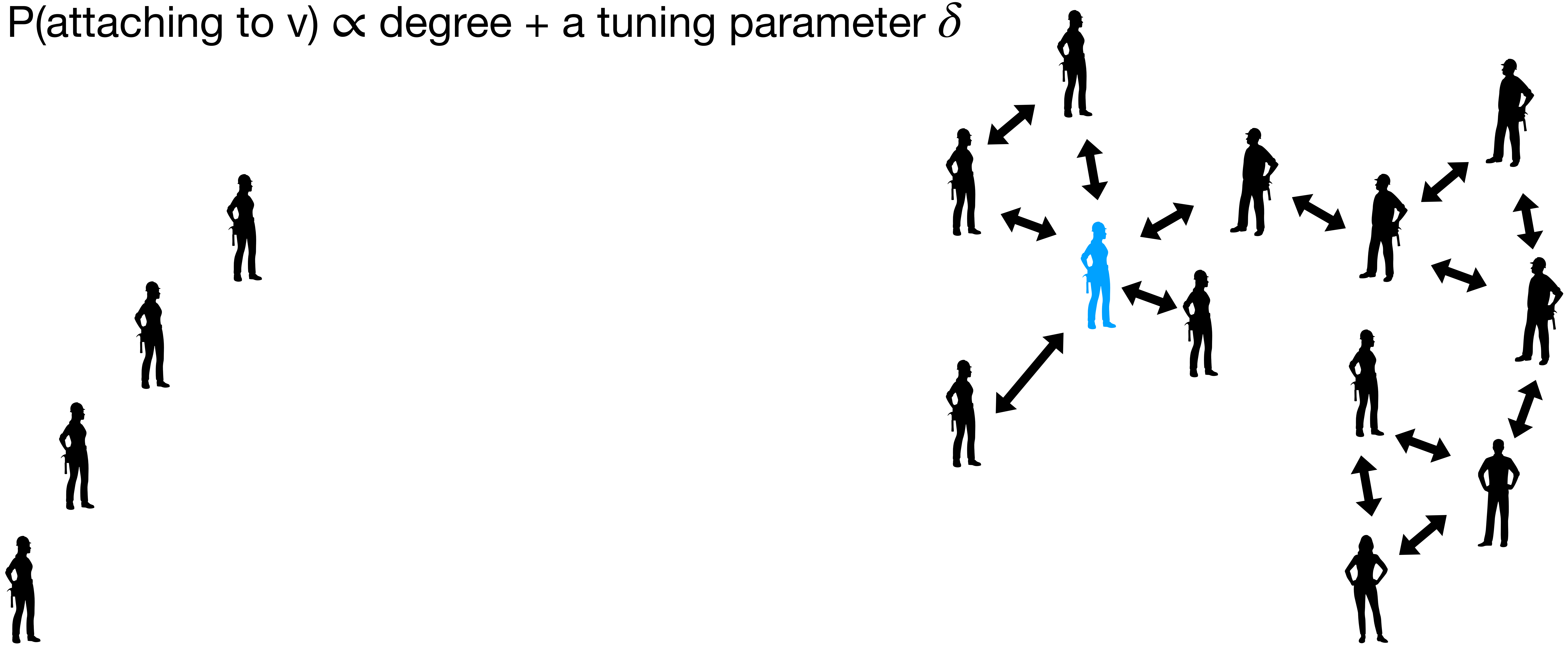
P(attaching to v) $\propto$ degree + a tuning parameter δ



Preferential Attachment

[Albert and Barabasi 1999]

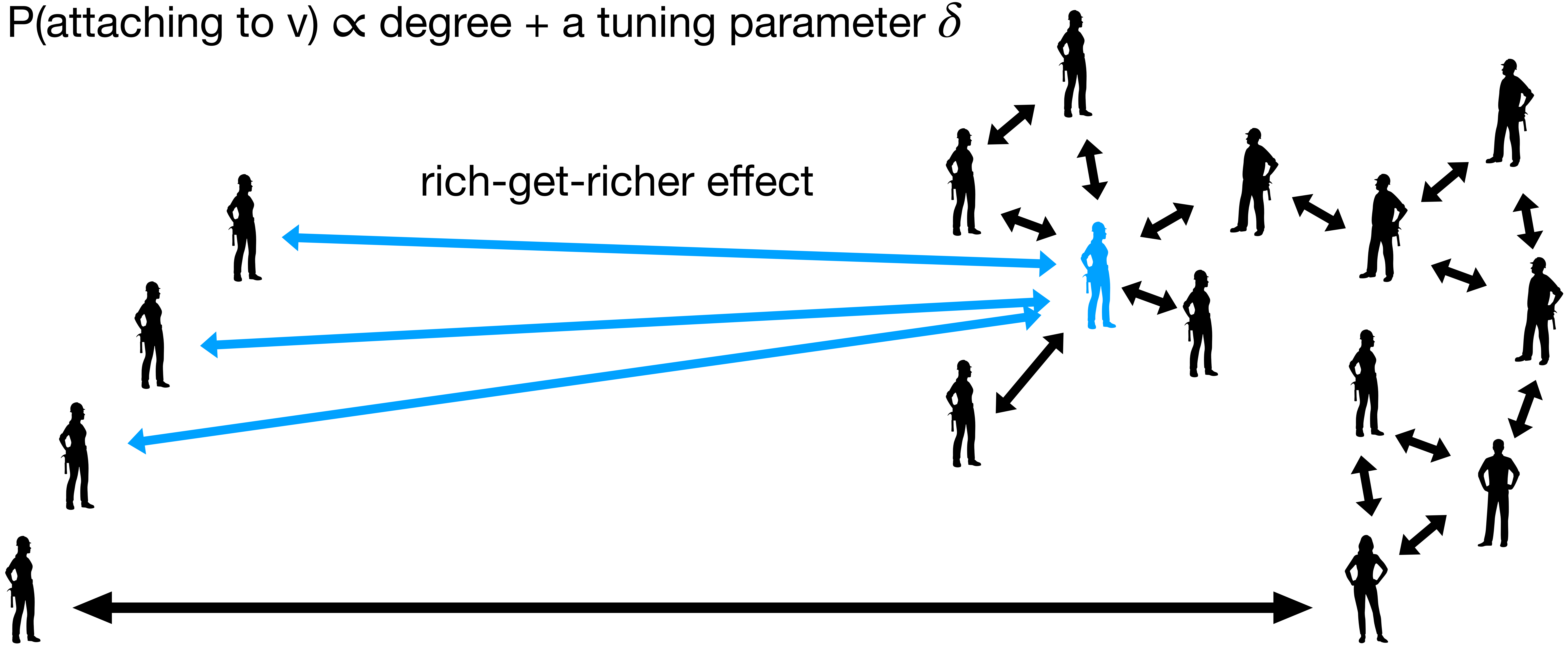
P(attaching to v) $\propto$ degree + a tuning parameter δ



Preferential Attachment

[Albert and Barabasi 1999]

P(attaching to v) $\propto$ degree + a tuning parameter δ



What do we know?

What do we know?

- degree distribution [Albert and Barabasi 1999]

What do we know?

- degree distribution [Albert and Barabasi 1999]
- triangle counts and clustering coefficient [Bollobas and Ridden 2002, Prokhorenkova et al 2013]

What do we know?

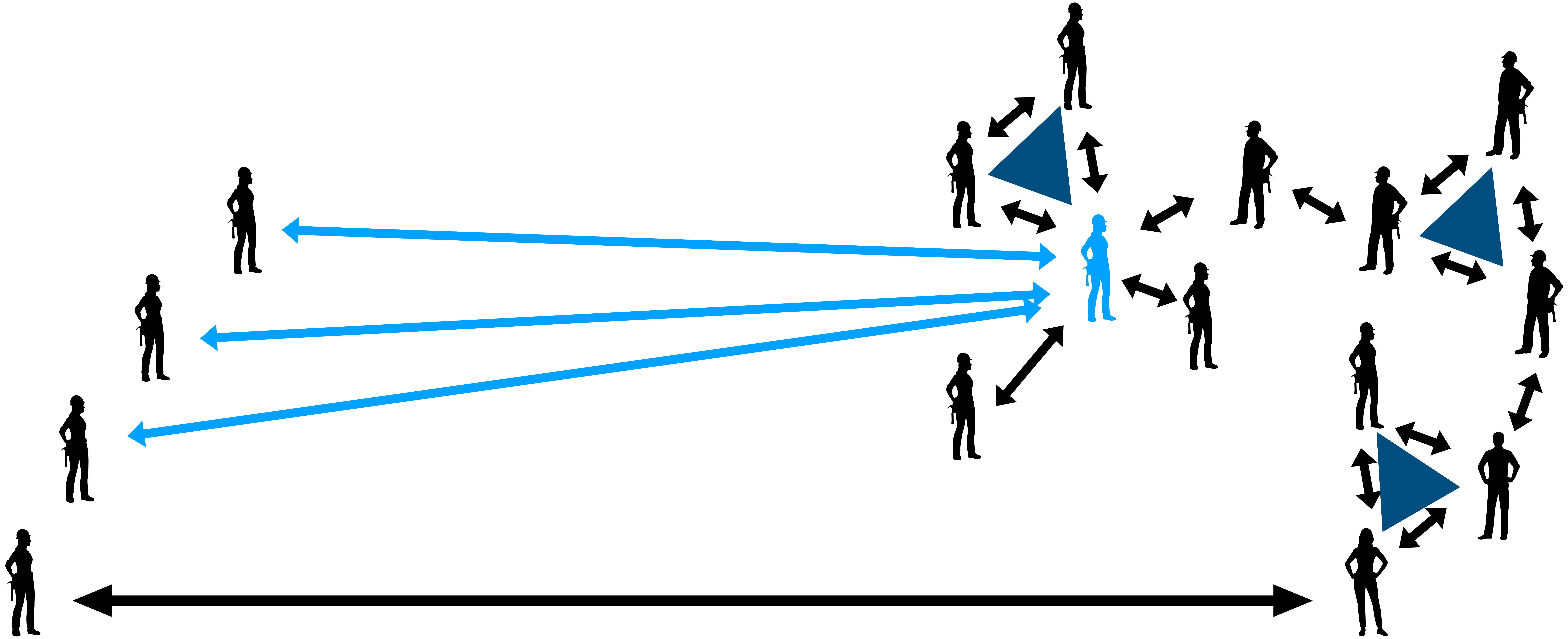
- degree distribution [Albert and Barabasi 1999]
- triangle counts and clustering coefficient [Bollobas and Ridden 2002, Prokhorenkova et al 2013]
- subgraph counts [Garavaglia and Steghuis 2019]

What do we know?

- degree distribution [Albert and Barabasi 1999]
- triangle counts and clustering coefficient [Bollobas and Ridden 2002, Prokhorenkova et al 2013]
- subgraph counts [Garavaglia and Steghuis 2019]
- and more…

Clique Complex

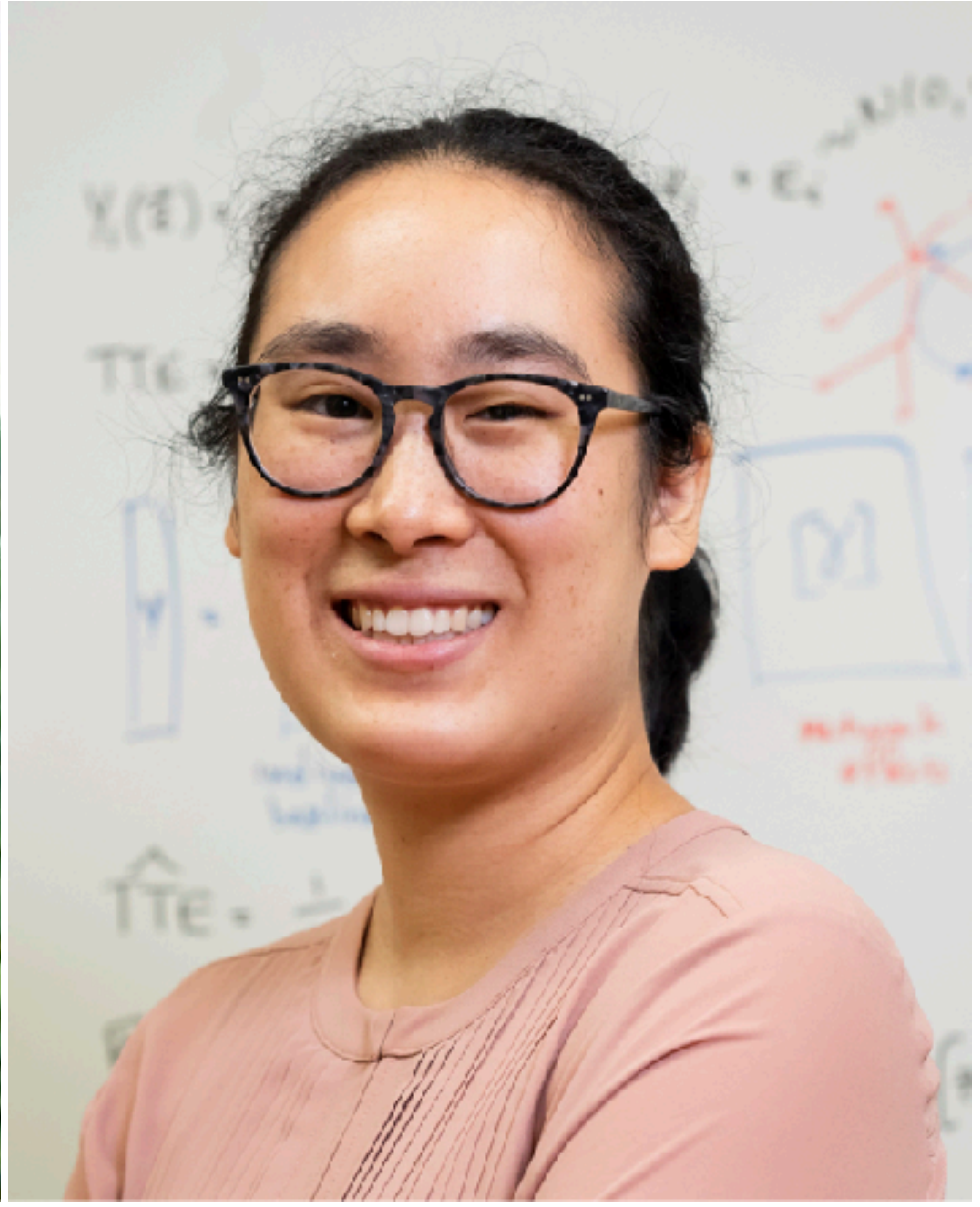
aka Flag Complex



III Topology of Preferential Attachment

My Lovely Collaborators

Avhan Misra



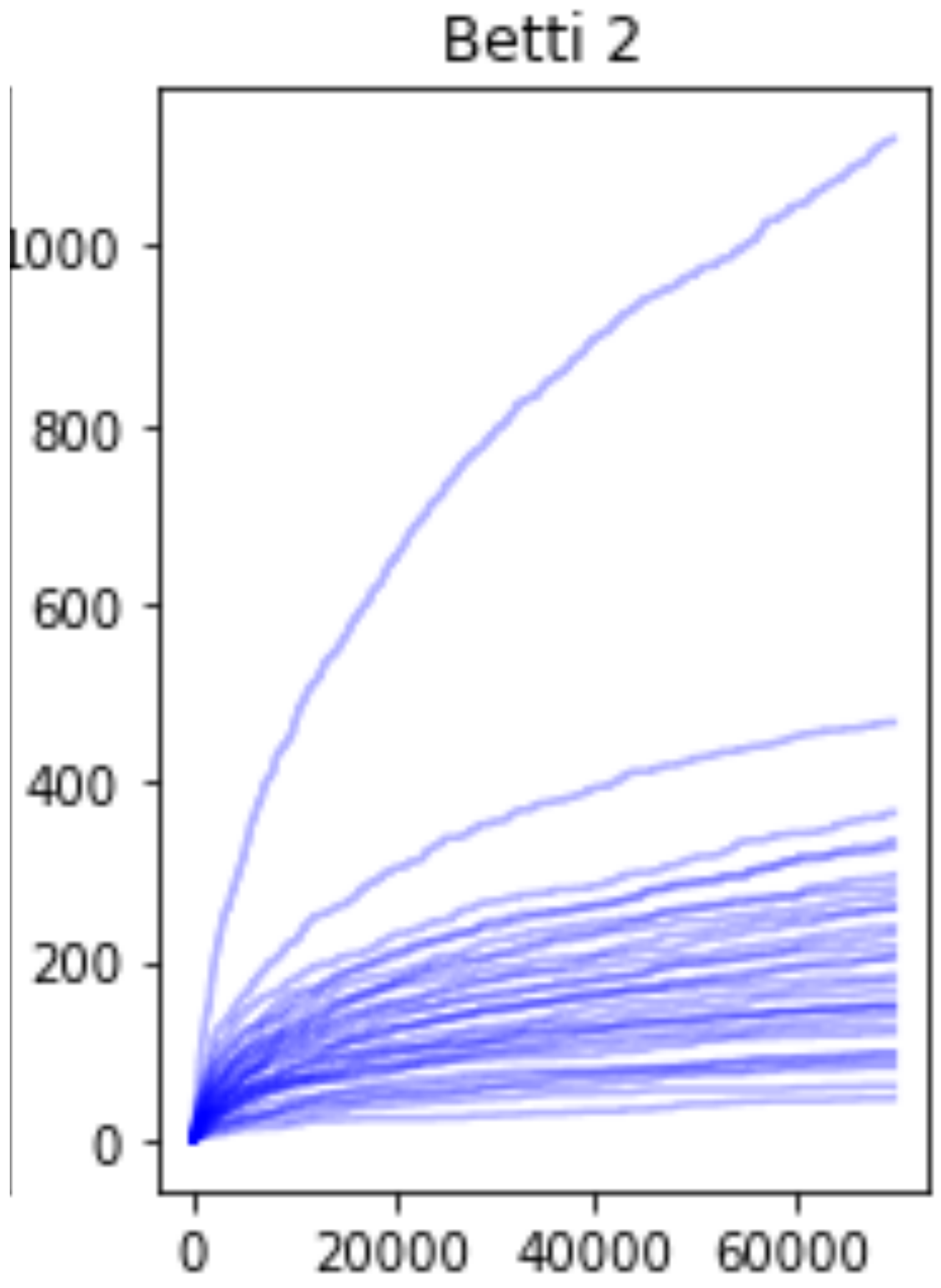
Christina Lee Yu

Gennady Samorodnitsky

Rongyi He (Caroline)

Betti Number β_q

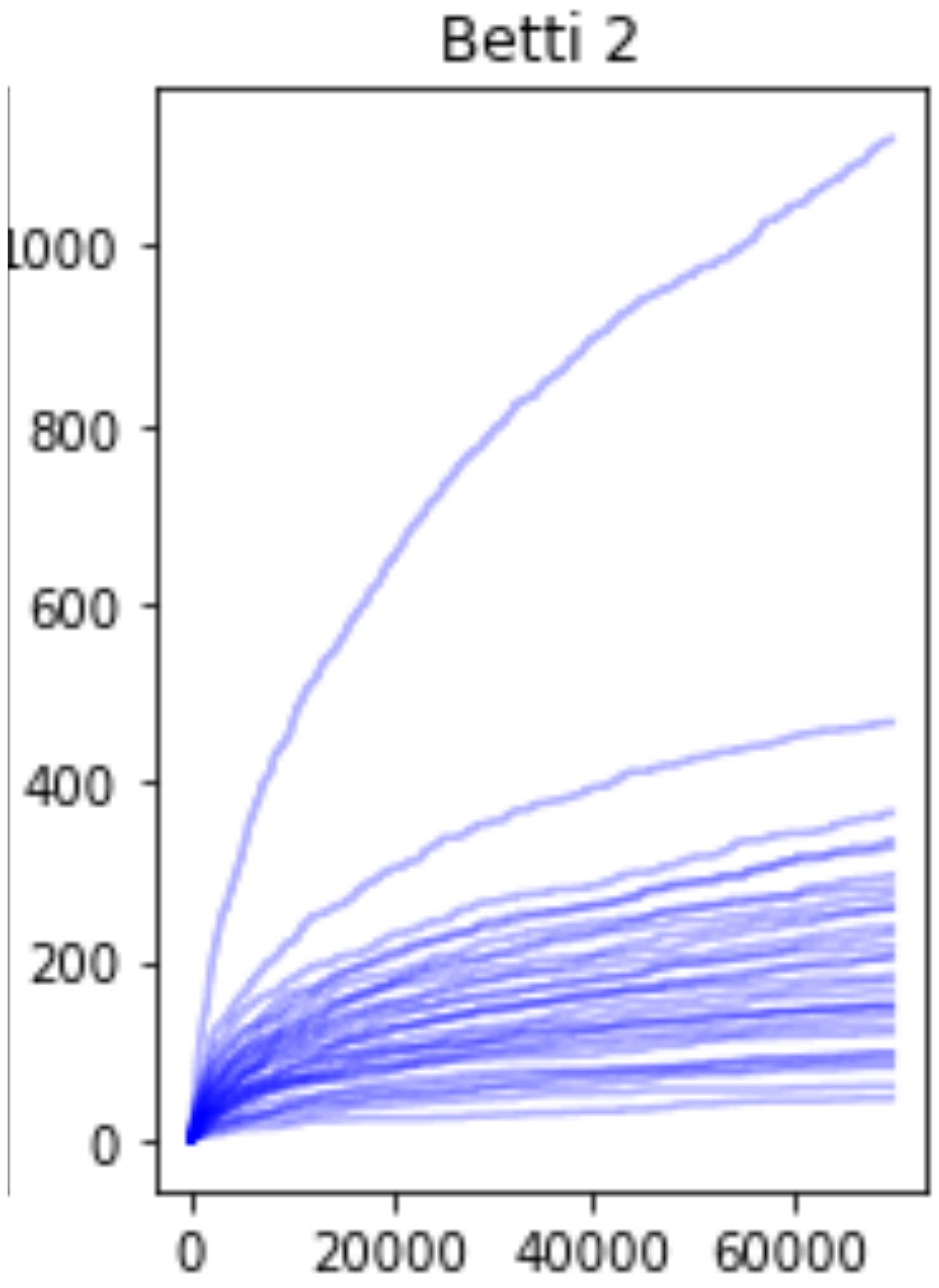
Betti Number β_q



Different curves, different random seeds.
All curves have the same model parameters.

Betti Number β_q

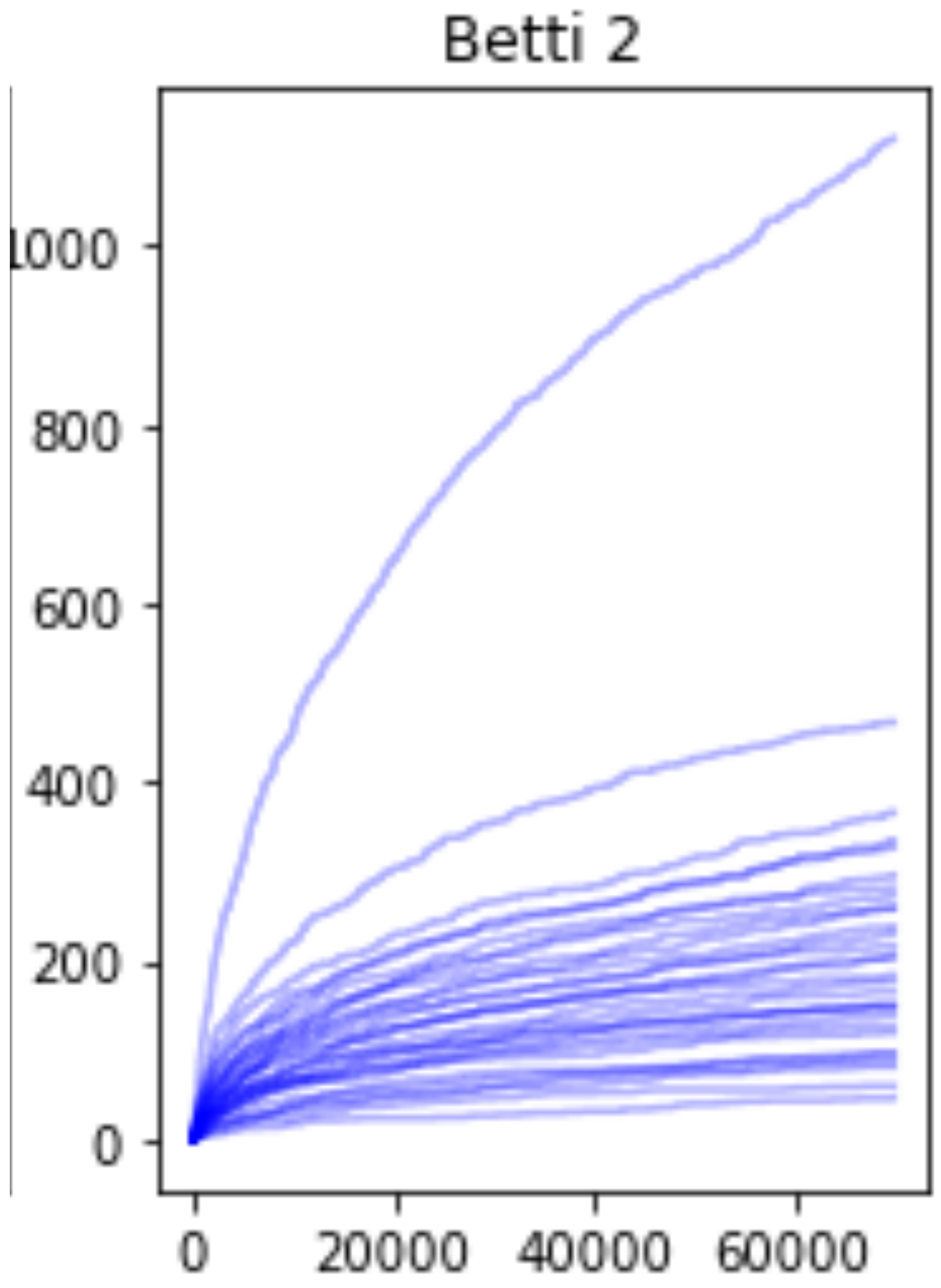
- increasing trend



Different curves, different random seeds.
All curves have the same model parameters.

Betti Number β_q

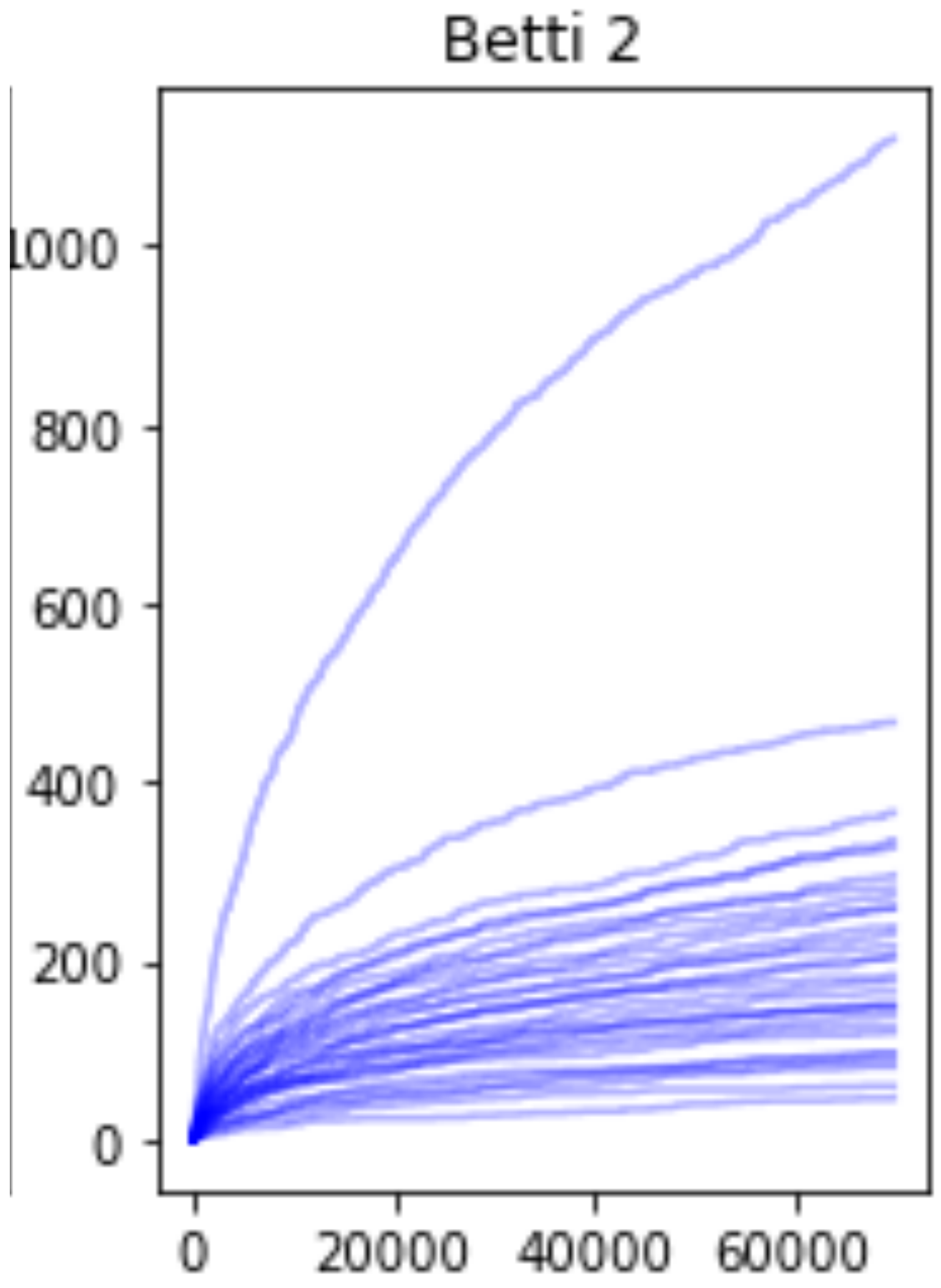
- increasing trend
- concave growth



Different curves, different random seeds.
All curves have the same model parameters.

Betti Number β_q

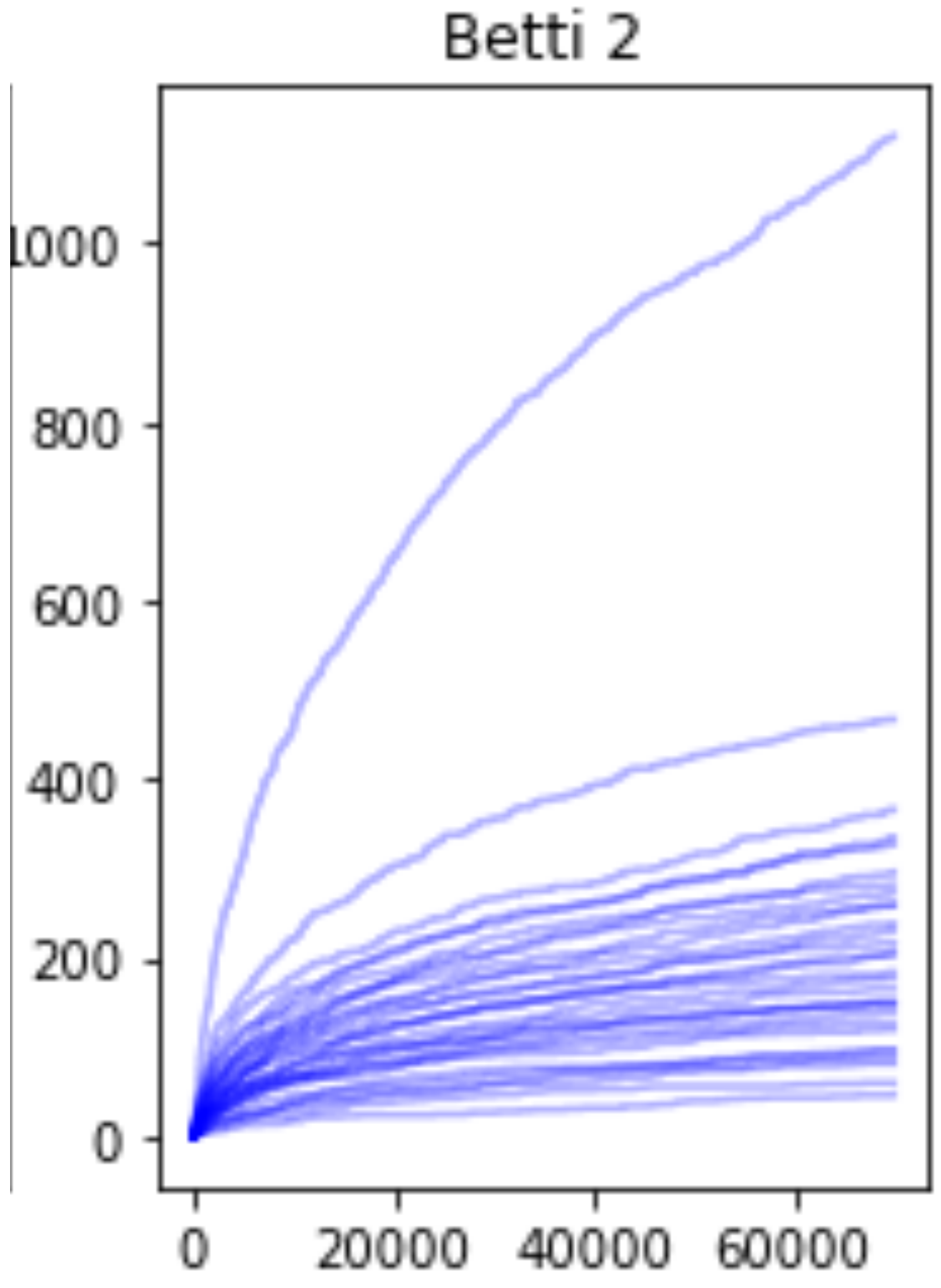
- increasing trend
- concave growth
- outlier



Different curves, different random seeds.
All curves have the same model parameters.

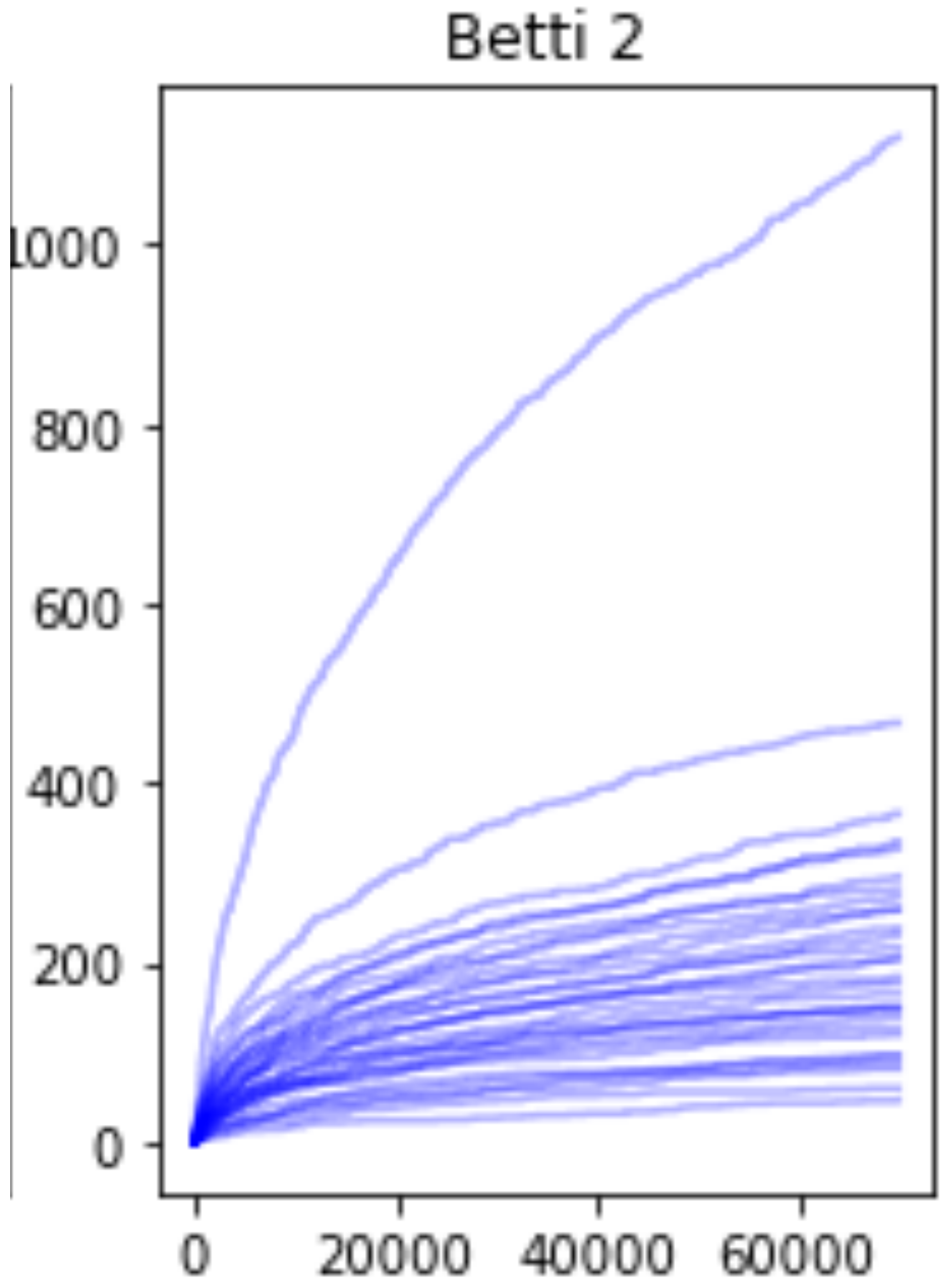
Betti Number β_q

- With probability at least $1-\varepsilon$,
- $c_\varepsilon(\text{num of nodes}^{1-4x}) \leq \beta_2 \leq C_\varepsilon(\text{num of nodes}^{1-4x})$
  - $x \in (0,1/2)$ decreases with the preferential attachment strength
    - $P[T \text{ attaches to } i] \propto T^{-x}$



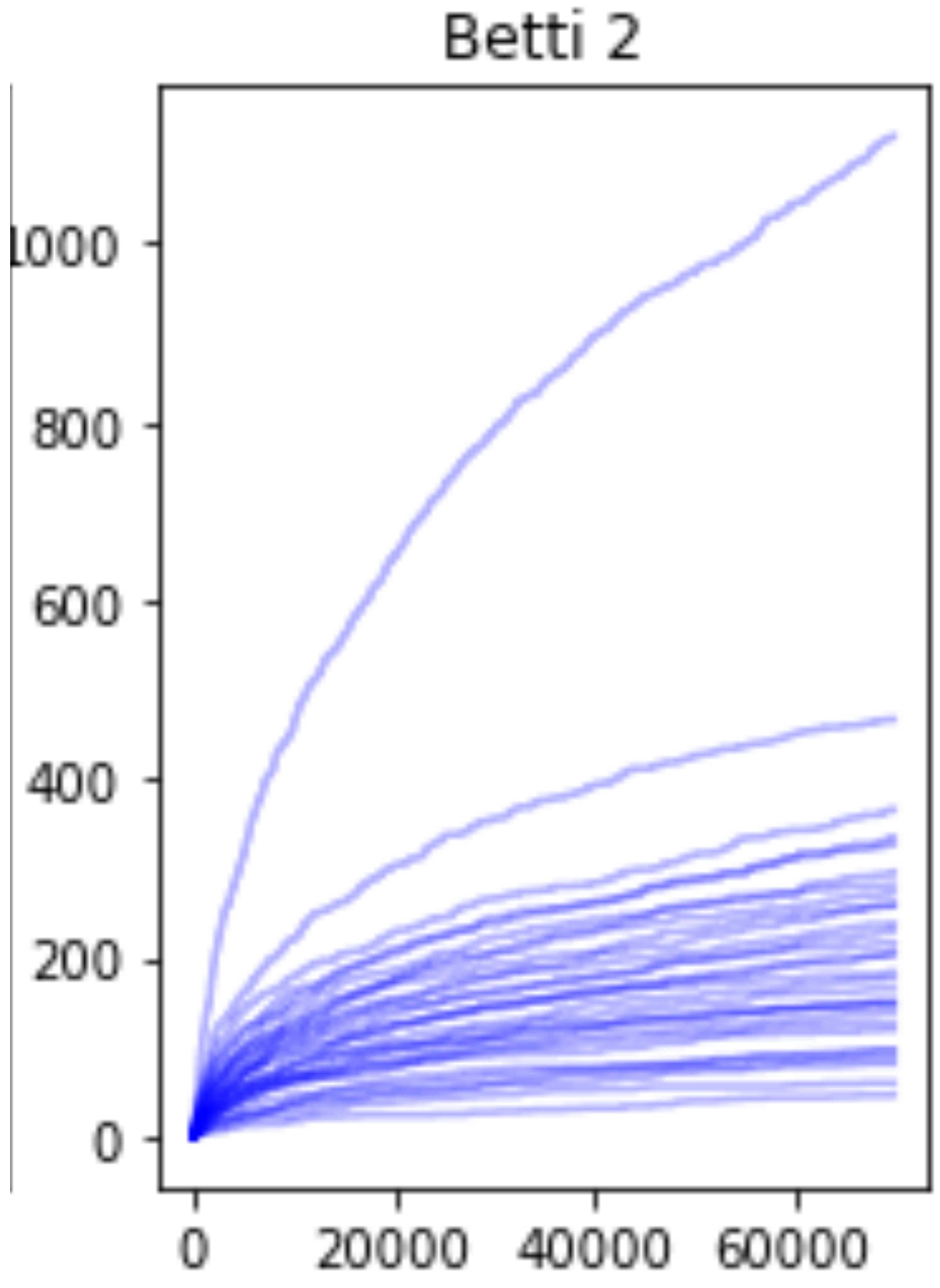
Betti Number β_q

- With probability at least $1 - \varepsilon$,
- $c_\varepsilon(\text{num of nodes}^{1-4x}) \leq \beta_2 \leq C_\varepsilon(\text{num of nodes}^{1-4x})$
  - $x \in (0,1/2)$ decreases with the preferential attachment strength
    - $P[T \text{ attaches to } i] \propto T^{-x}$
  - If $1 - 4x < 0$, then $\beta_2 \leq C_\varepsilon$.



Betti Number β_q

- With probability at least $1 - \varepsilon$,
- $c_\varepsilon(\text{num of nodes}^{1-4x}) \leq \beta_2 \leq C_\varepsilon(\text{num of nodes}^{1-4x})$
  - $x \in (0,1/2)$ decreases with the preferential attachment strength
    - $P[T \text{ attaches to } i] \propto T^{-x}$
  - If $1 - 4x < 0$, then $\beta_2 \leq C_\varepsilon$.
- $c_\varepsilon(\text{num of nodes}^{1-2qx}) \leq \beta_q \leq C_\varepsilon(\text{num of nodes}^{1-2qx})$ for $q \geq 2$.

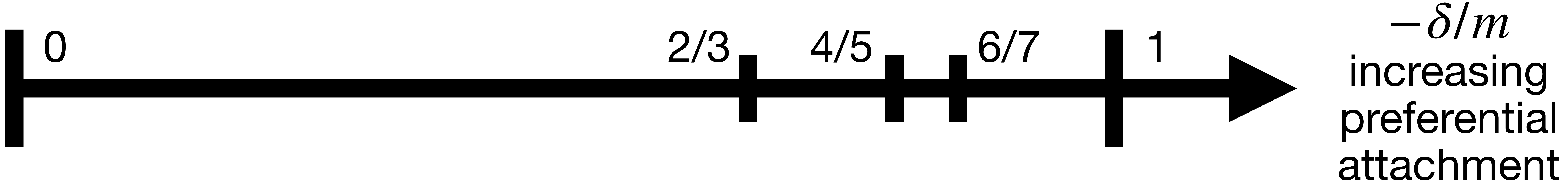


Phase transition

Recall

P(attaching to v) ∝ degree + δ

m = number of edges per new node

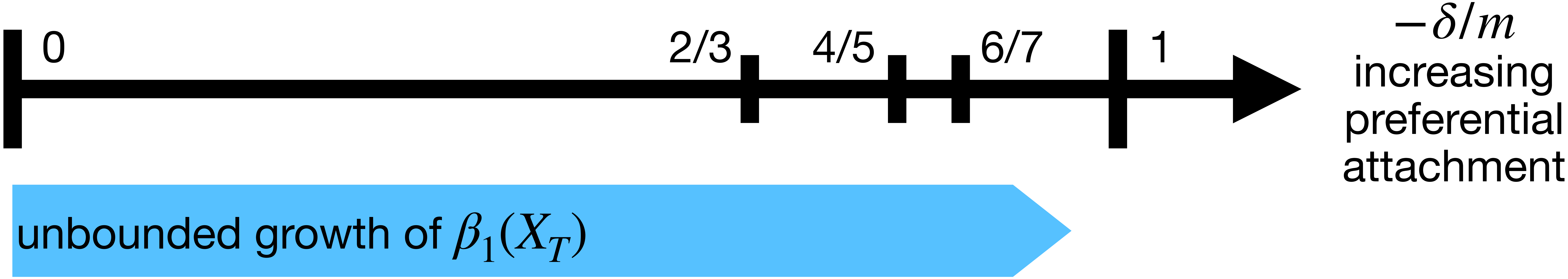


Phase transition

Recall

P(attaching to v) $\propto$ degree + δ

m = number of edges per new node

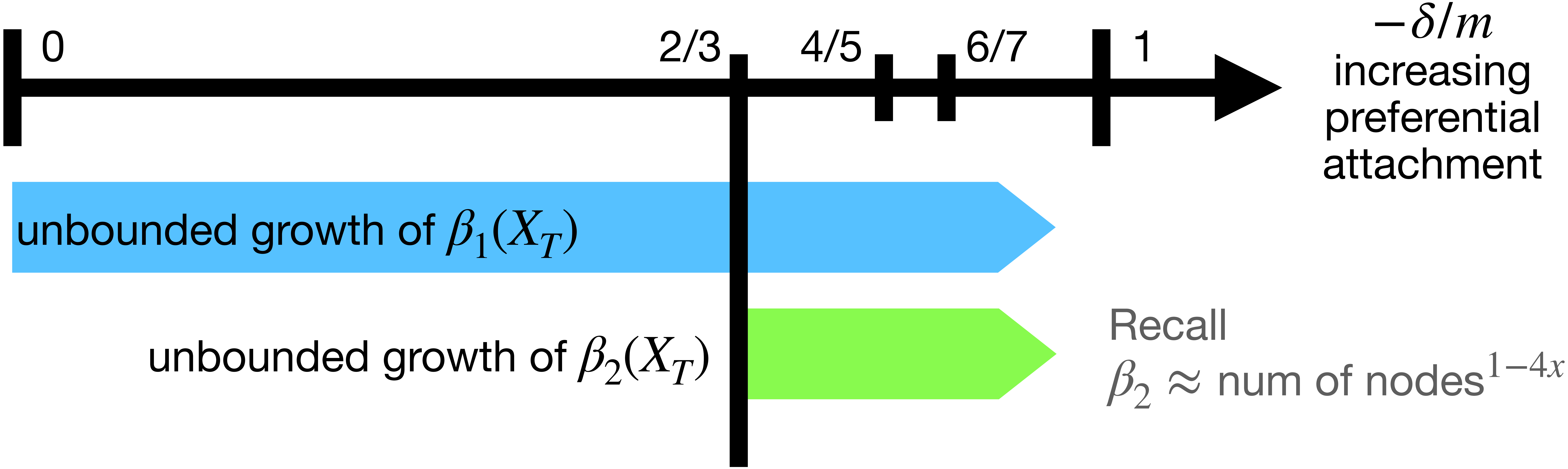


Phase transition

Recall

P(attaching to v) $\propto$ degree + δ

m = number of edges per new node



Recall

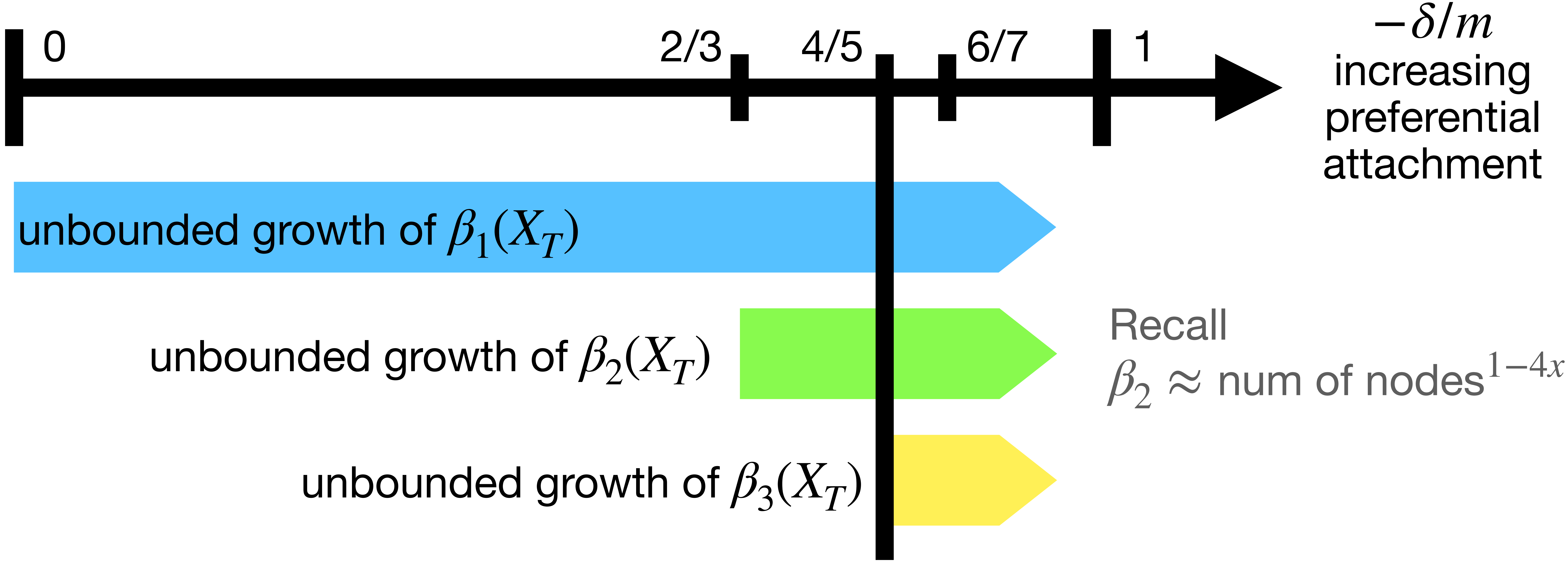
$\beta_2 \approx$ num of nodes^{1-4x}

Phase transition

Recall

P(attaching to v) ∝ degree + δ

m = number of edges per new node



Recall

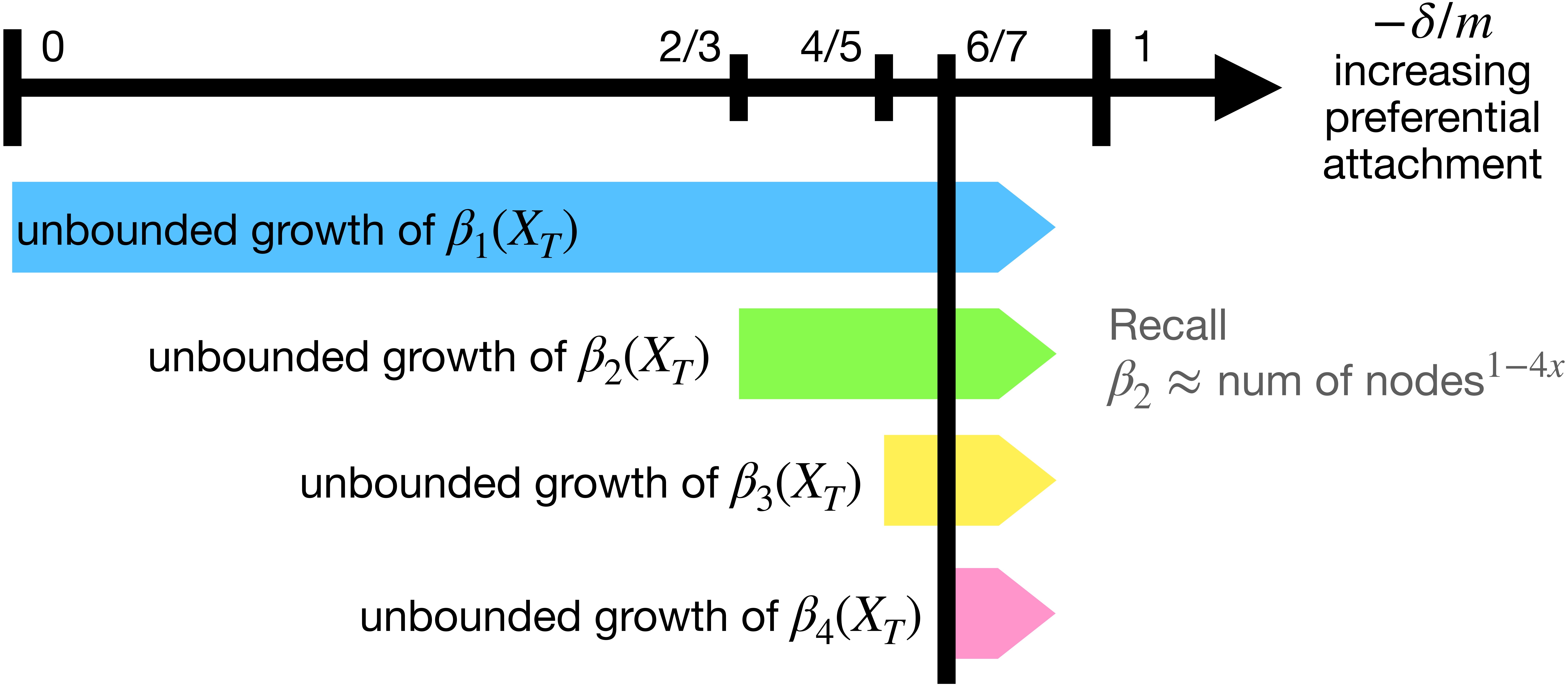
$\beta_2 \approx$ num of nodes^{1-4x}

Phase transition

Recall

P(attaching to v) ∝ degree + δ

m = number of edges per new node

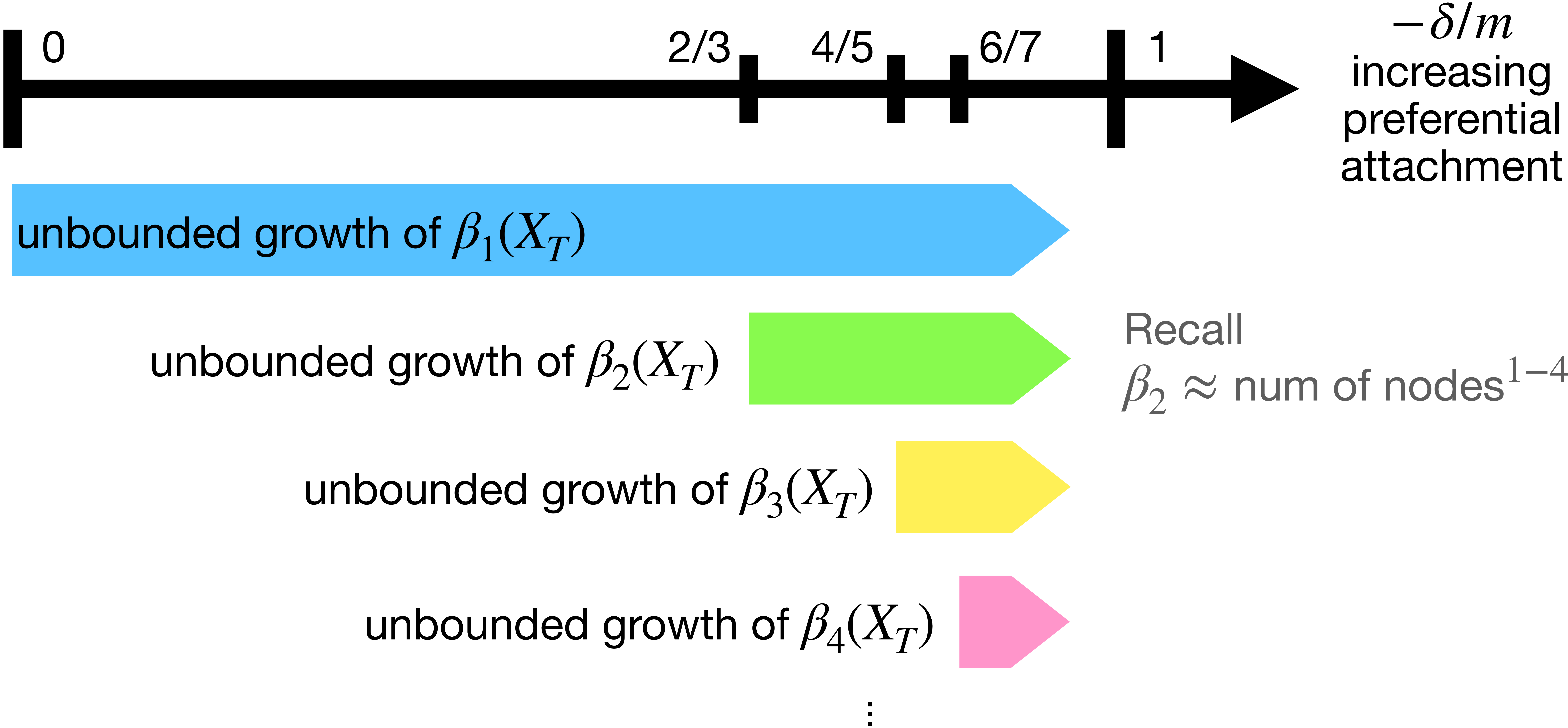


Recall

$\beta_2 \approx$ num of nodes^{1-4x}

Phase transition

Recall
P(attaching to v) ∝ degree + δ
m = number of edges per new node

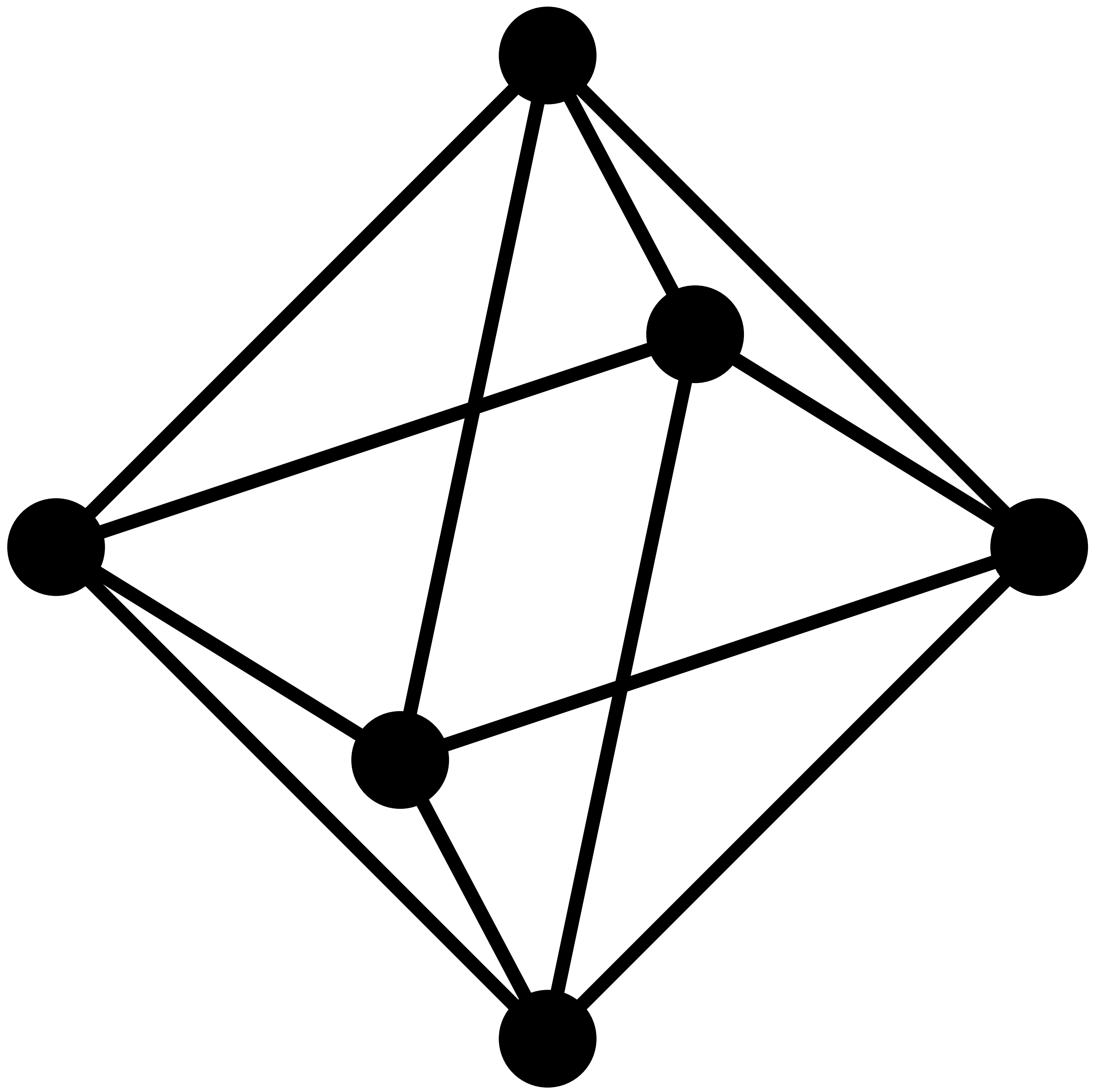


Recall
$\beta_2 \approx$ num of nodes^{1-4x}

Theorem: $\beta_2 \approx \text{num of nodes}^{1-4x}$

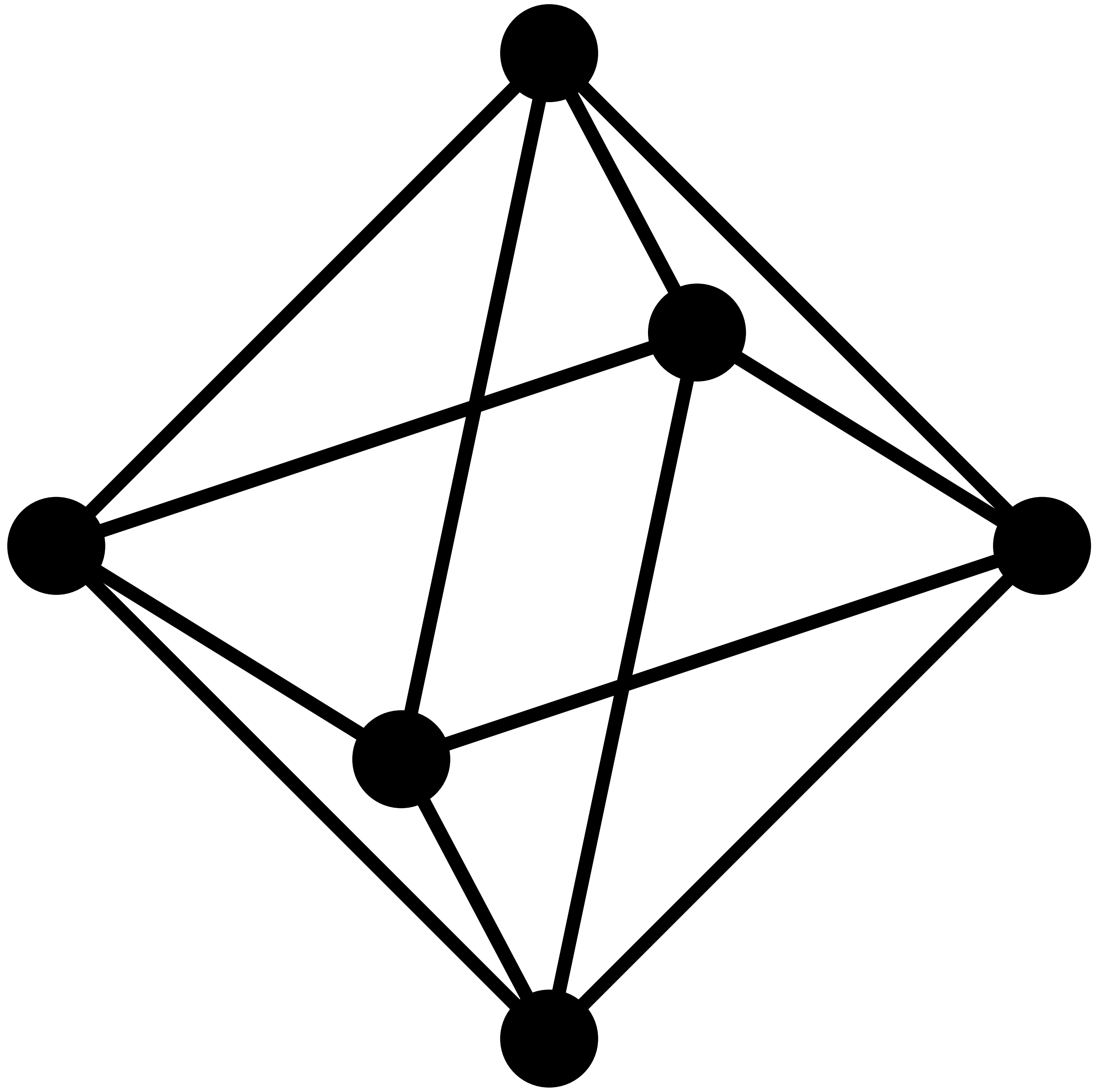
Proof?

Proof of $\beta_2 \approx$ num of nodes^{1-4x}

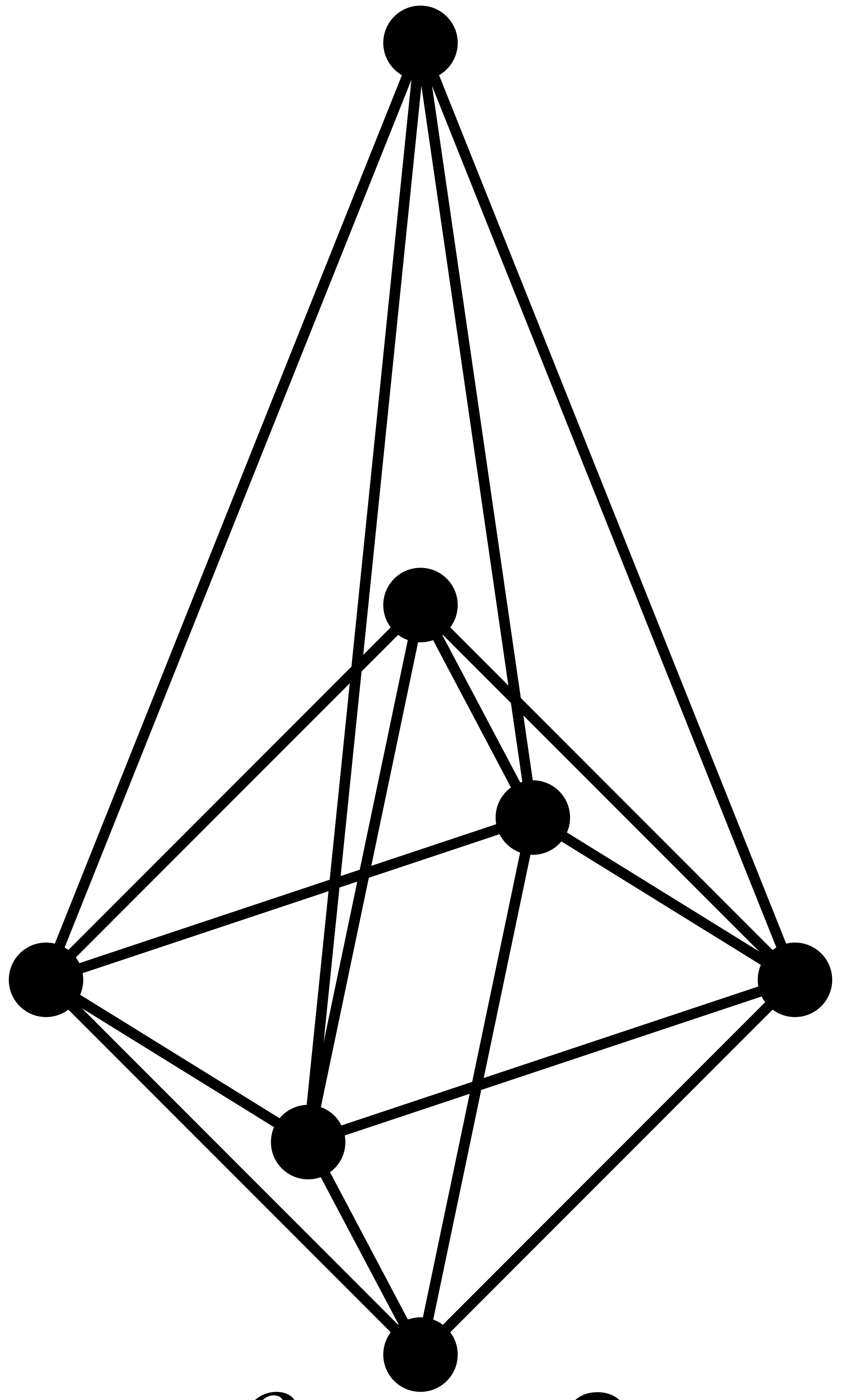


$\beta_2 = 1$

Proof of $\beta_2 \approx$ num of nodes^{1-4x}

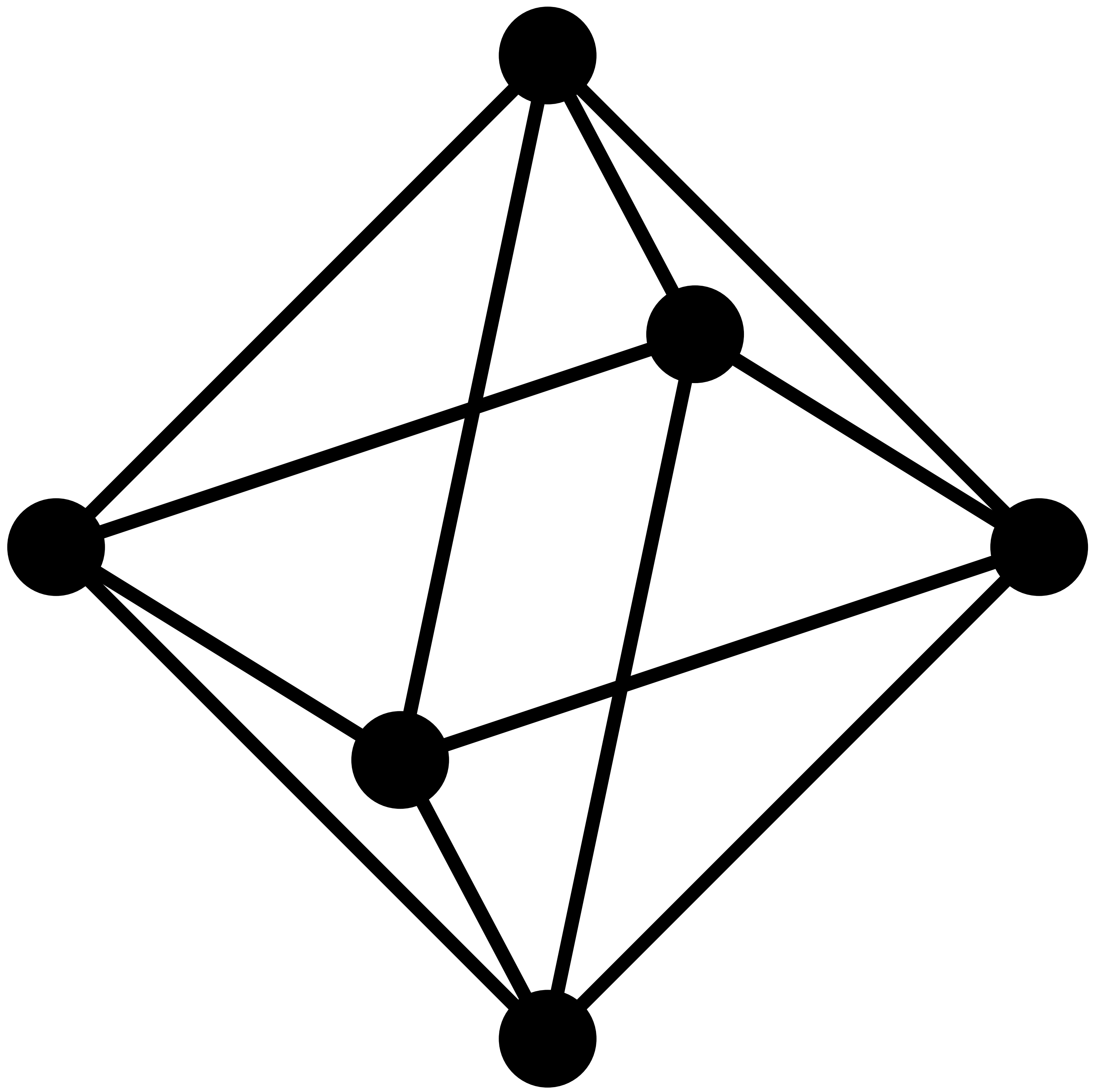


$\beta_2 = 1$

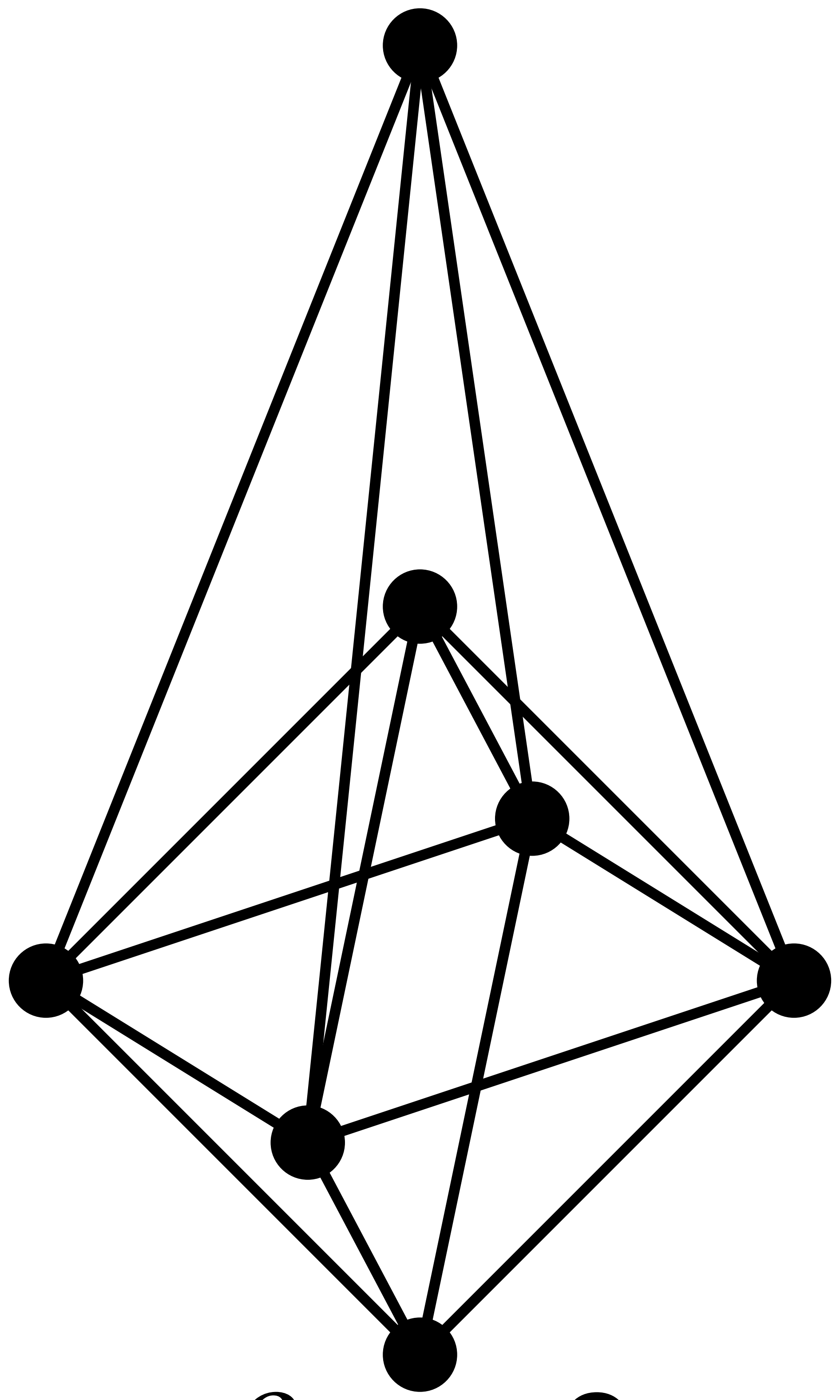


$\beta_2 = 2$

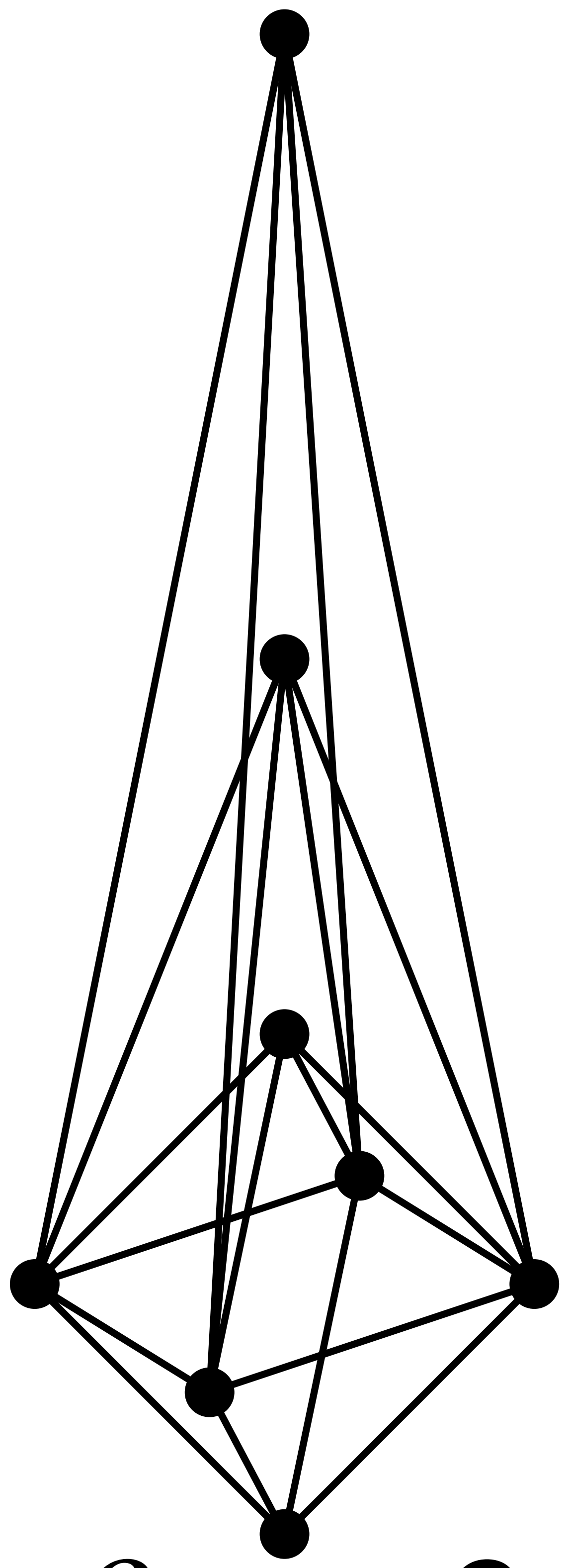
Proof of $\beta_2 \approx$ num of nodes^{1-4x}



$\beta_2 = 1$



$\beta_2 = 2$



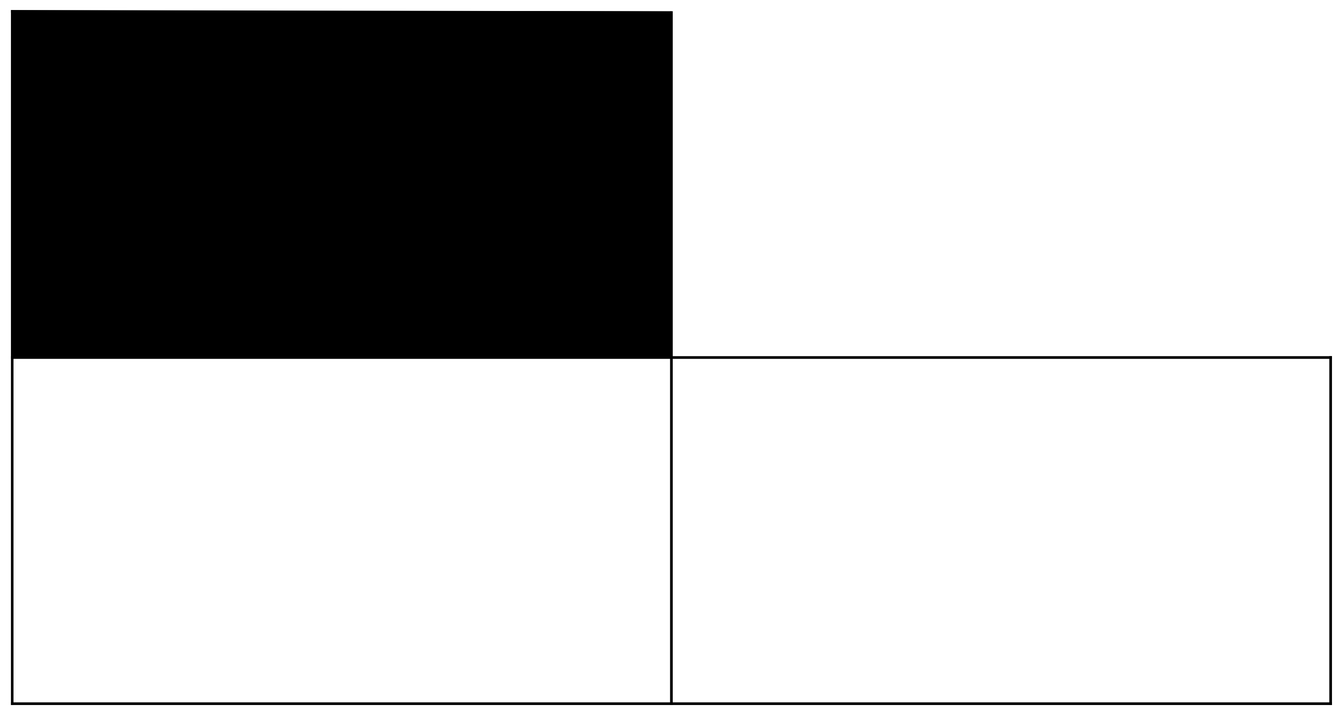
$\beta_2 = 3$

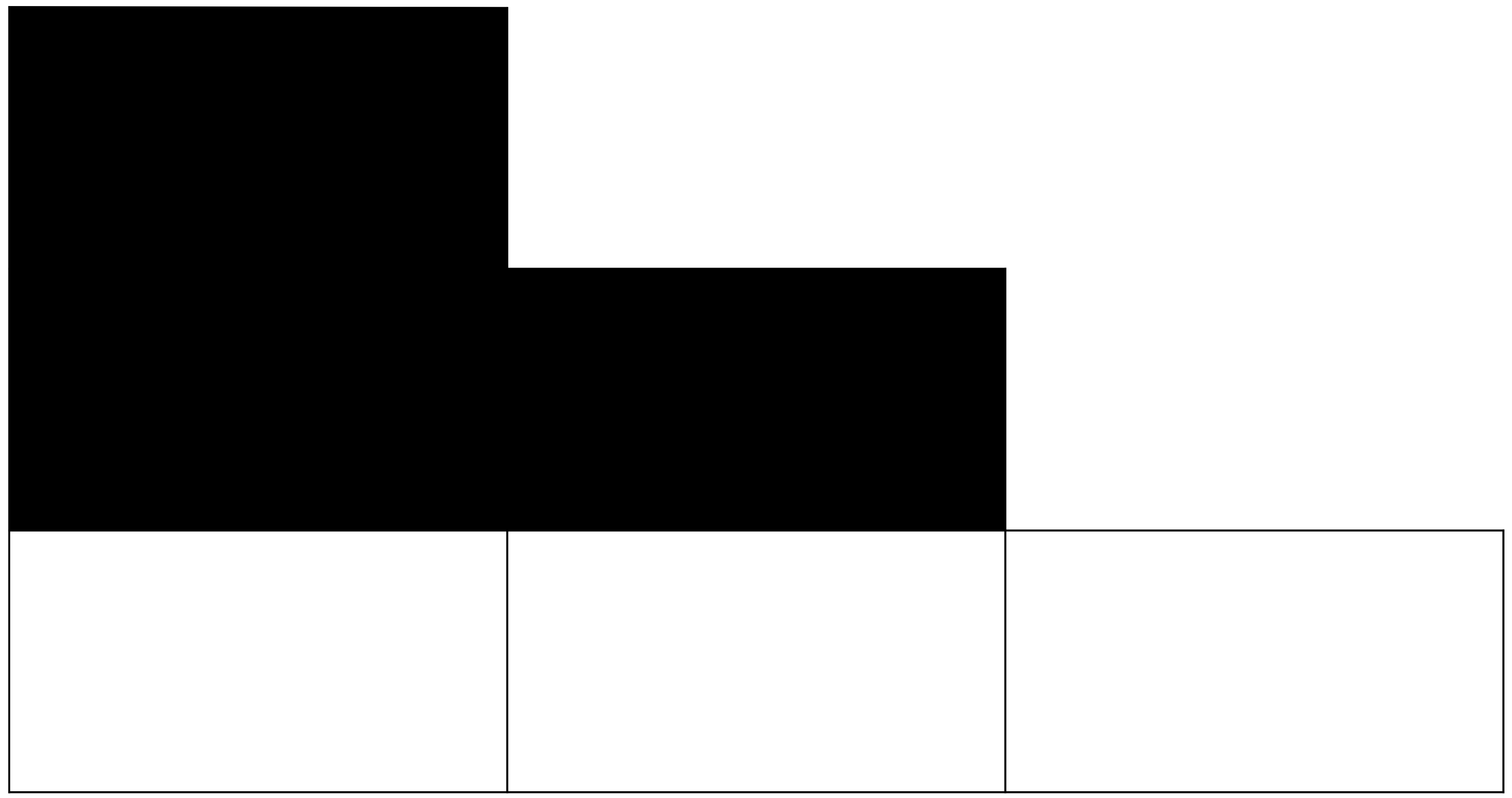
Homotopy-Connectivity?

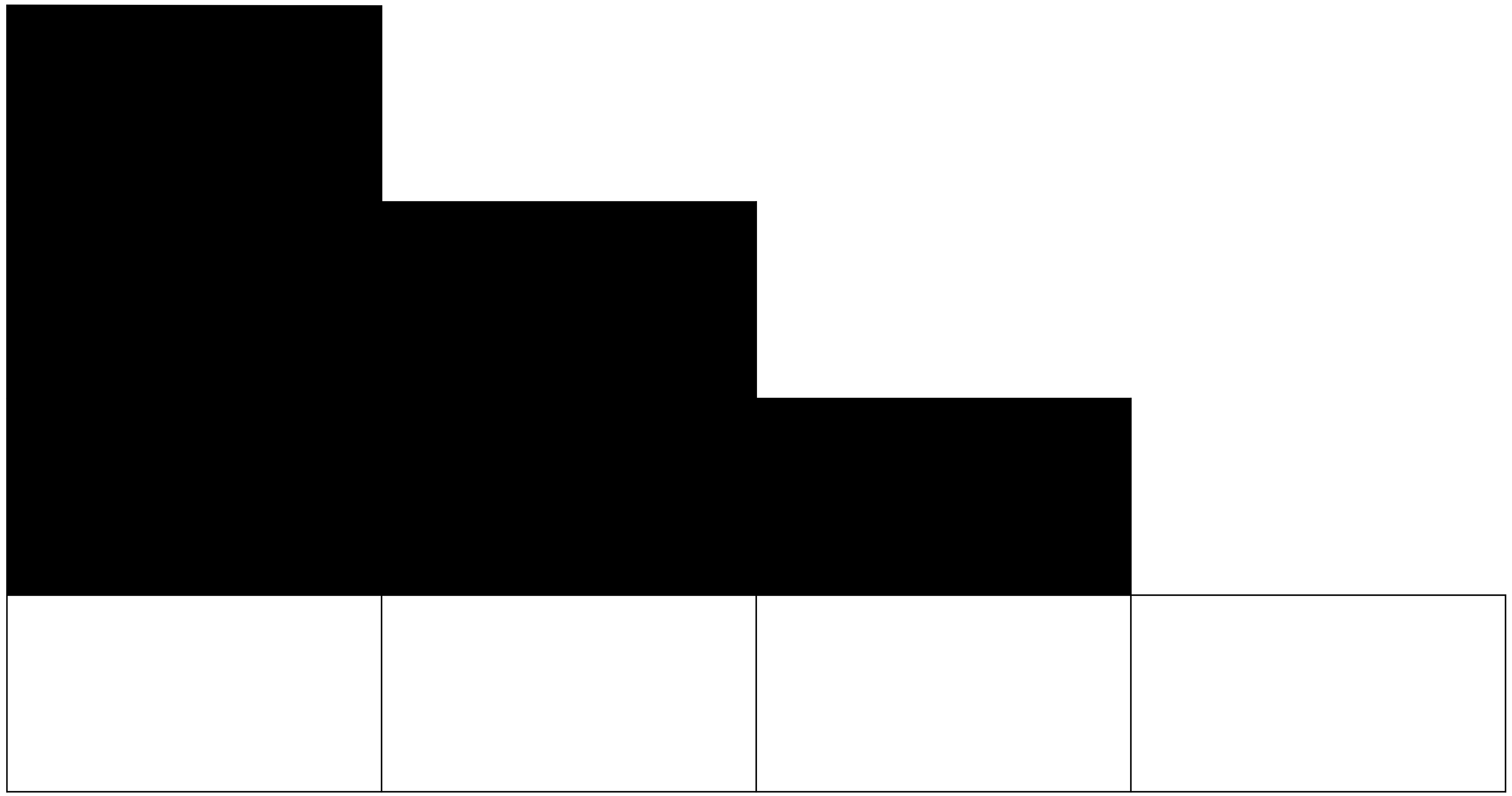
Homotopy-Connectivity?

$\beta_2 \approx$ num of nodes^{1-4x}

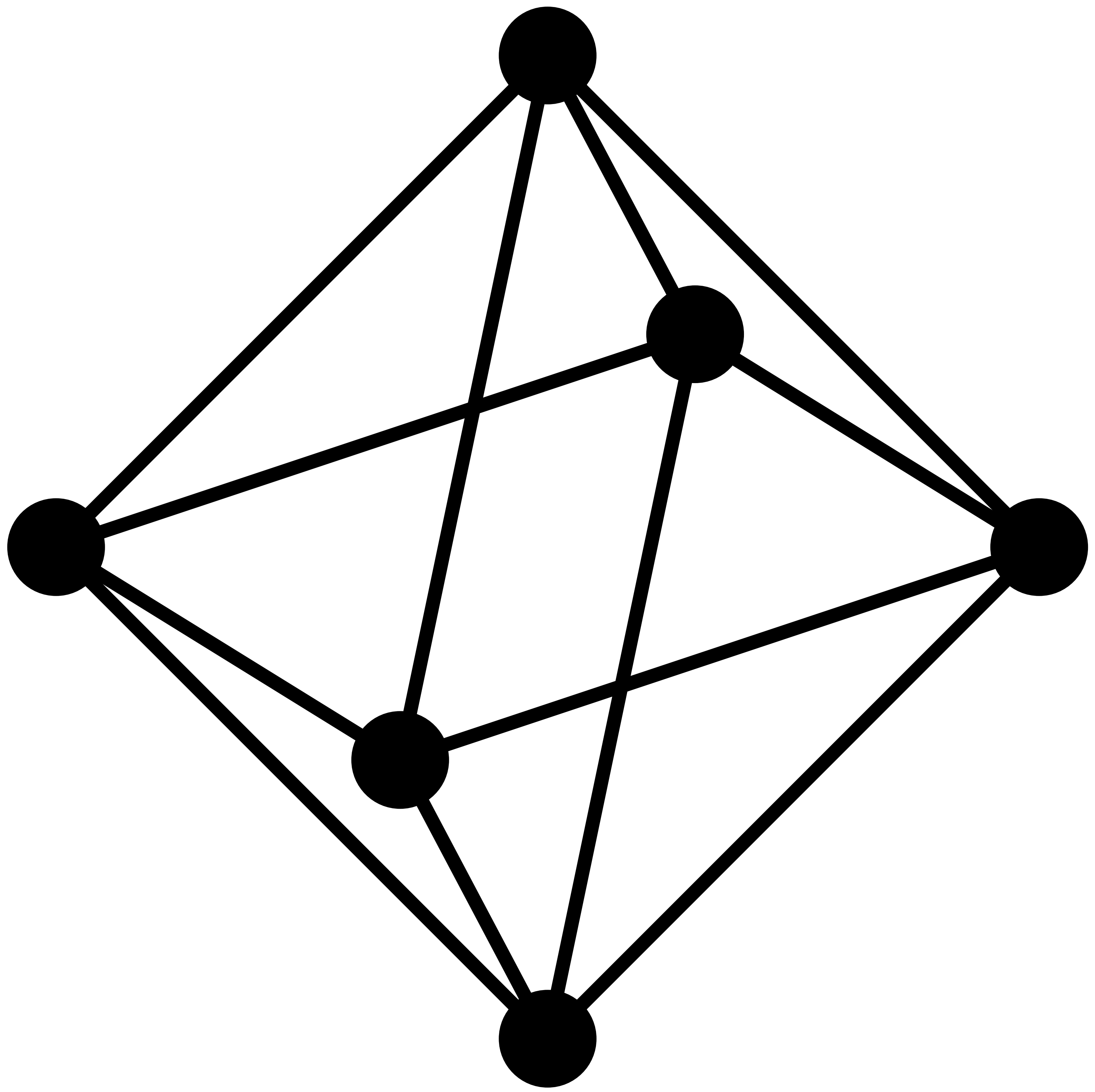
Pass to infinity



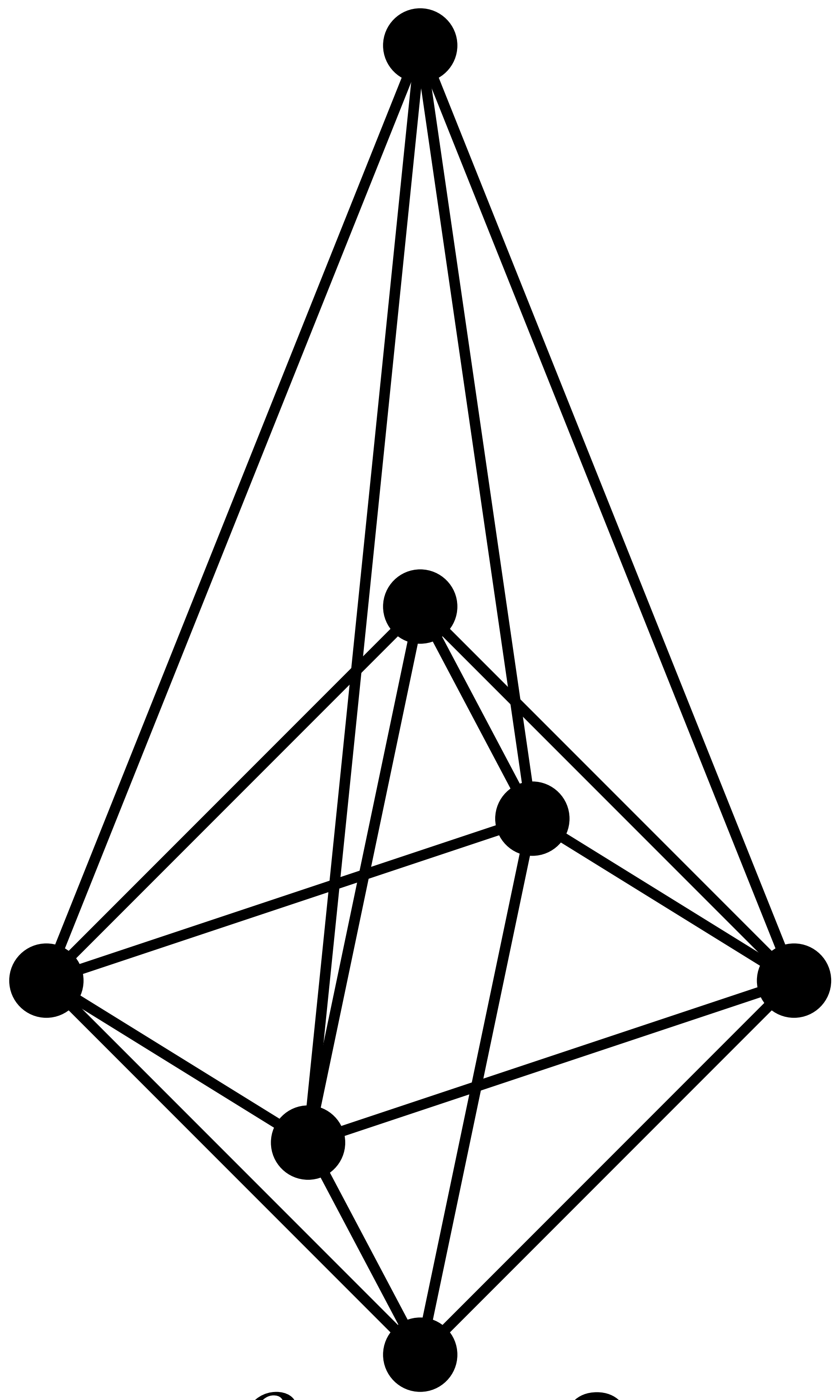




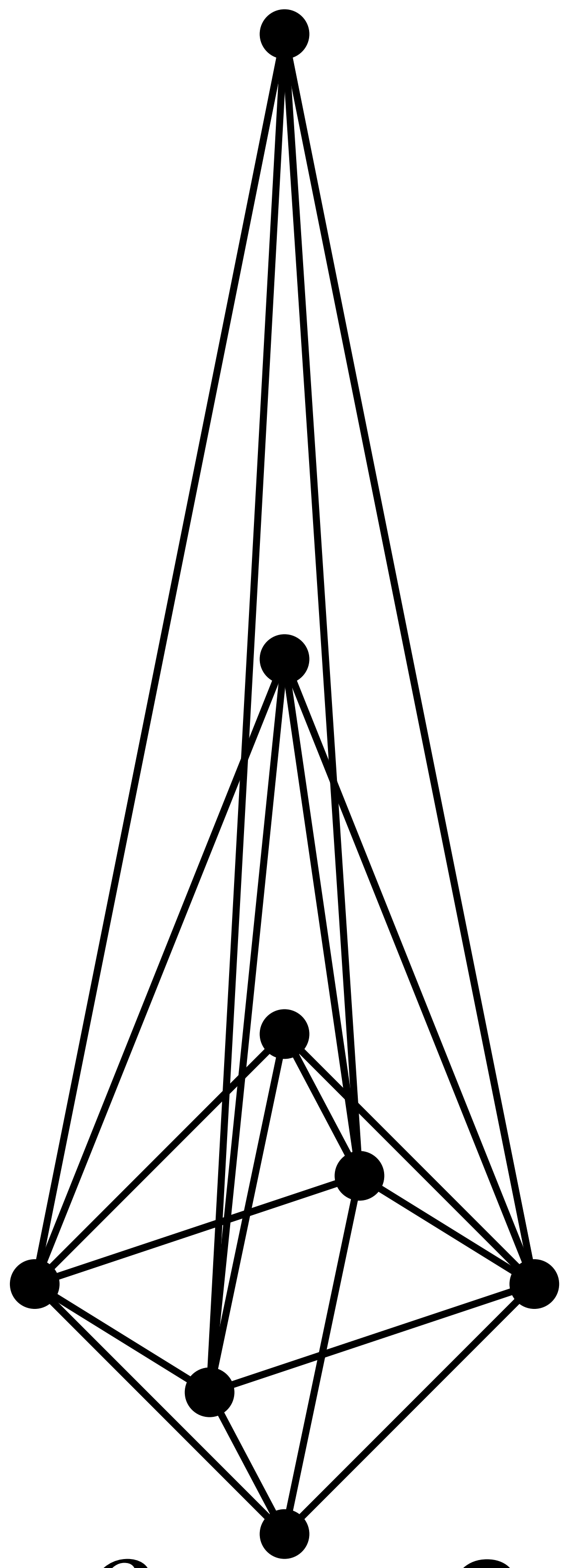
Will all of these be filled in at infinity?



$\beta_2 = 1$



$\beta_2 = 2$



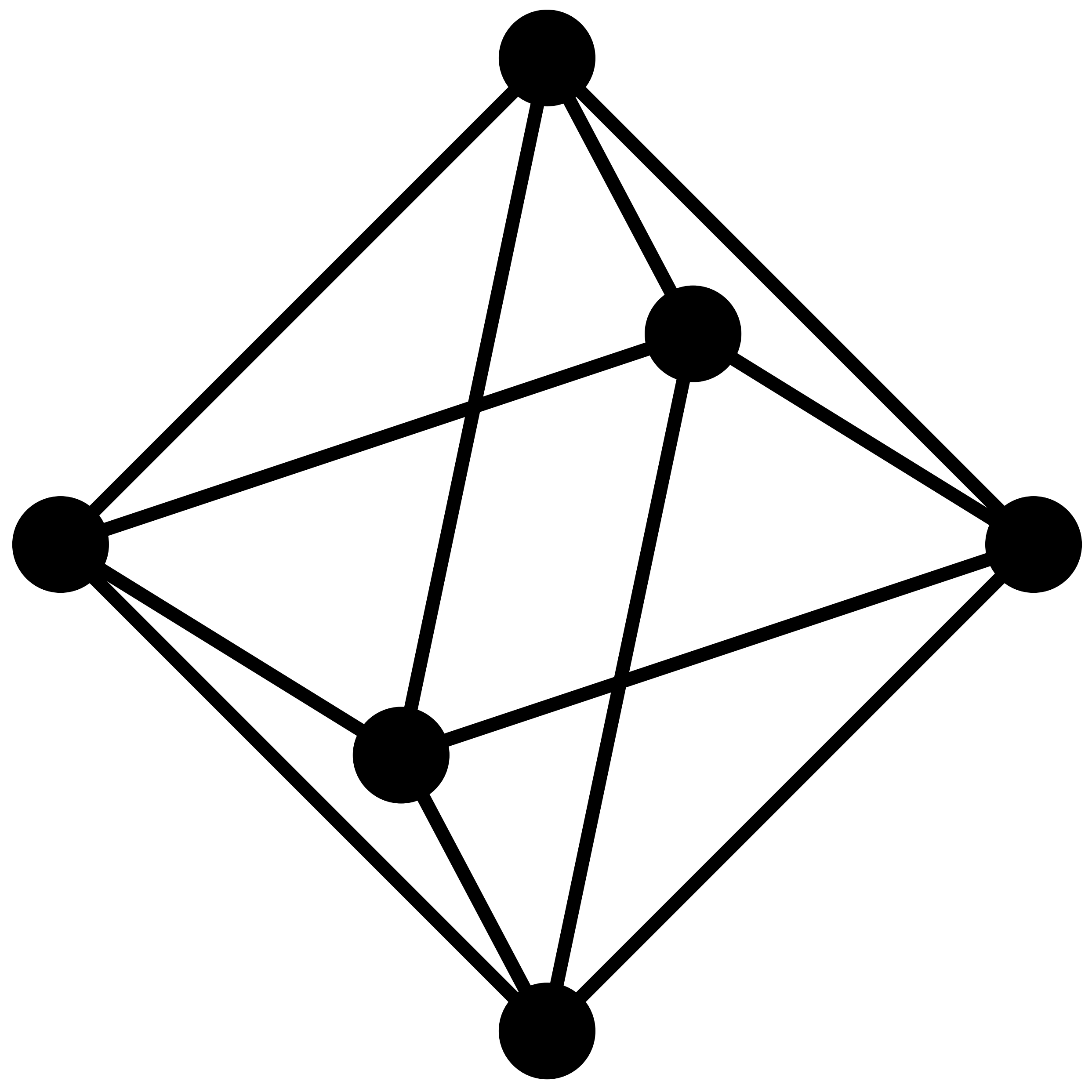
$\beta_2 = 3$

[Barmak 2023]

- A clique complex is q-homotopy-connected
- if every collection of $2(q + 1)$ nodes has a common neighbor.

[Barmak 2023]

- A clique complex is q-homotopy-connected
- if every collection of 2(q + 1) nodes has a common neighbor.



Homotopy-Connected

- Almost surely, the infinite preferential attachment complex
- is q-homotopy-connected if $x \leq \frac{1}{2(q+1)}$

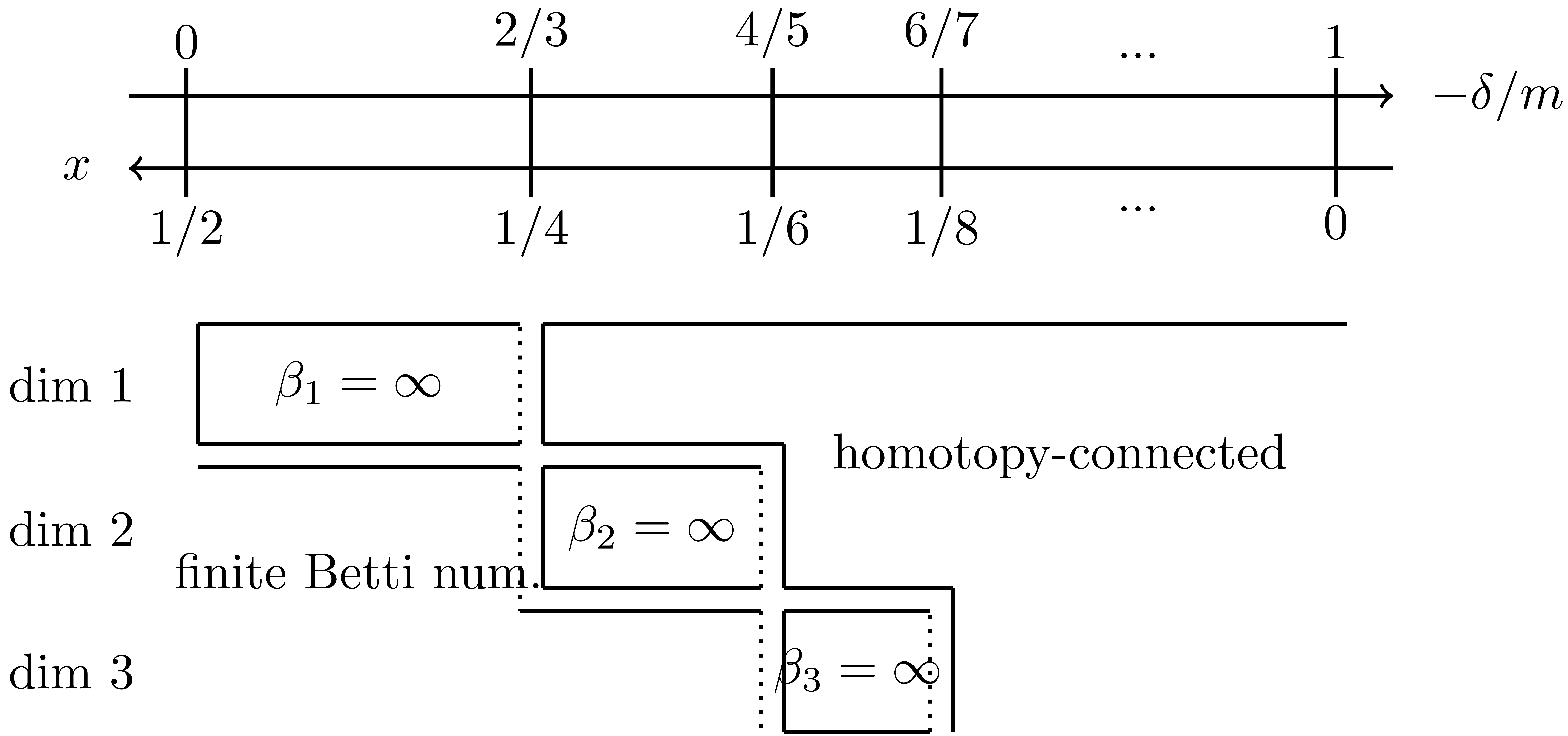
Recall:
$x \in (0,1/2)$ decreases with
the preferential attachment strength
$P[T$ attaches to $i] \propto T^{-x}$

Homotopy-Connected

- Almost surely, the infinite preferential attachment complex
- is q-homotopy-connected if $x \leq \dfrac{1}{2(q+1)}$
- has infinite Betti number at dimension q if $\dfrac{1}{2(q+1)} < x \leq \dfrac{1}{2q}$

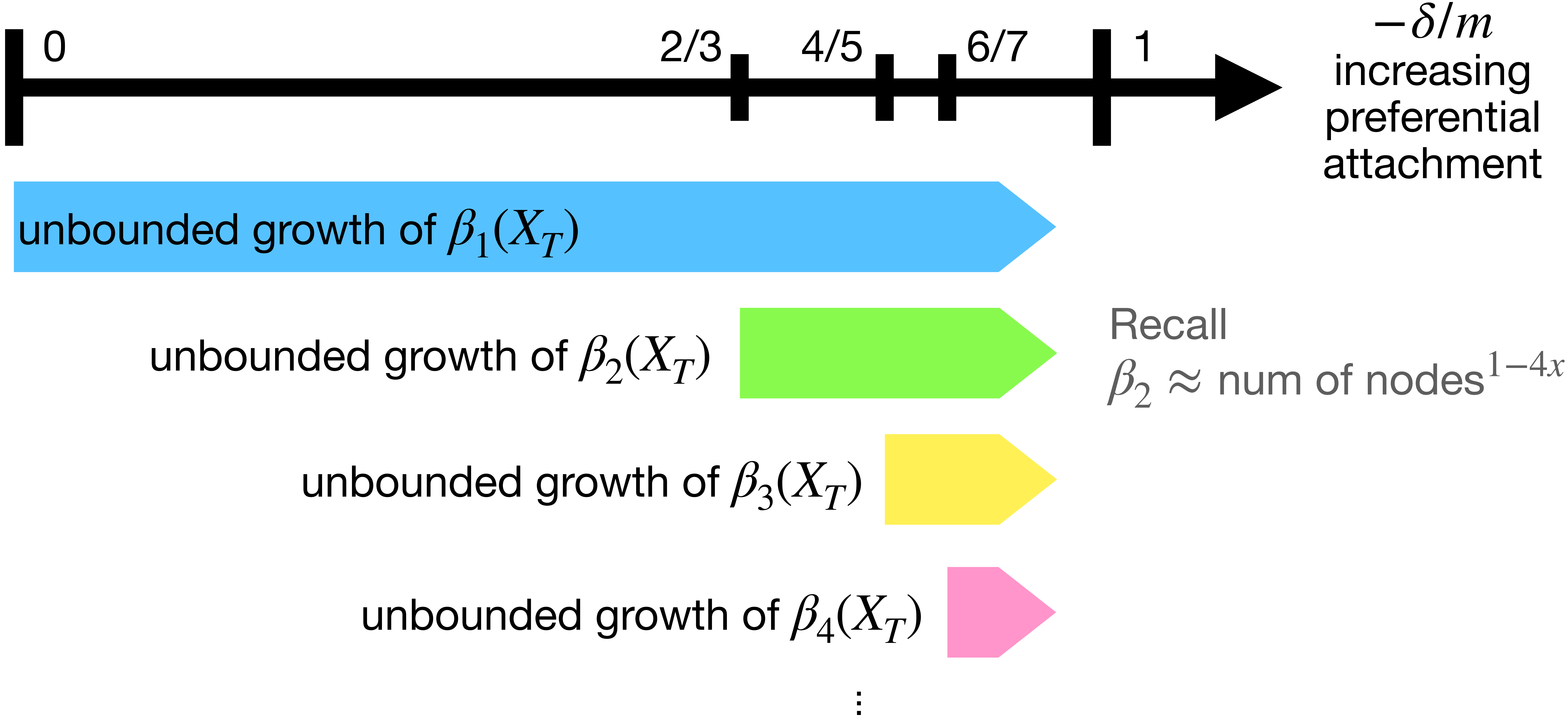
Recall:
$x \in (0,1/2)$ decreases with
the preferential attachment strength
$P[T \text{ attaches to } i] \propto T^{-x}$

Phase Transition



Phase transition

Recall
P(attaching to v) ∝ degree + δ
m = number of edges per new node



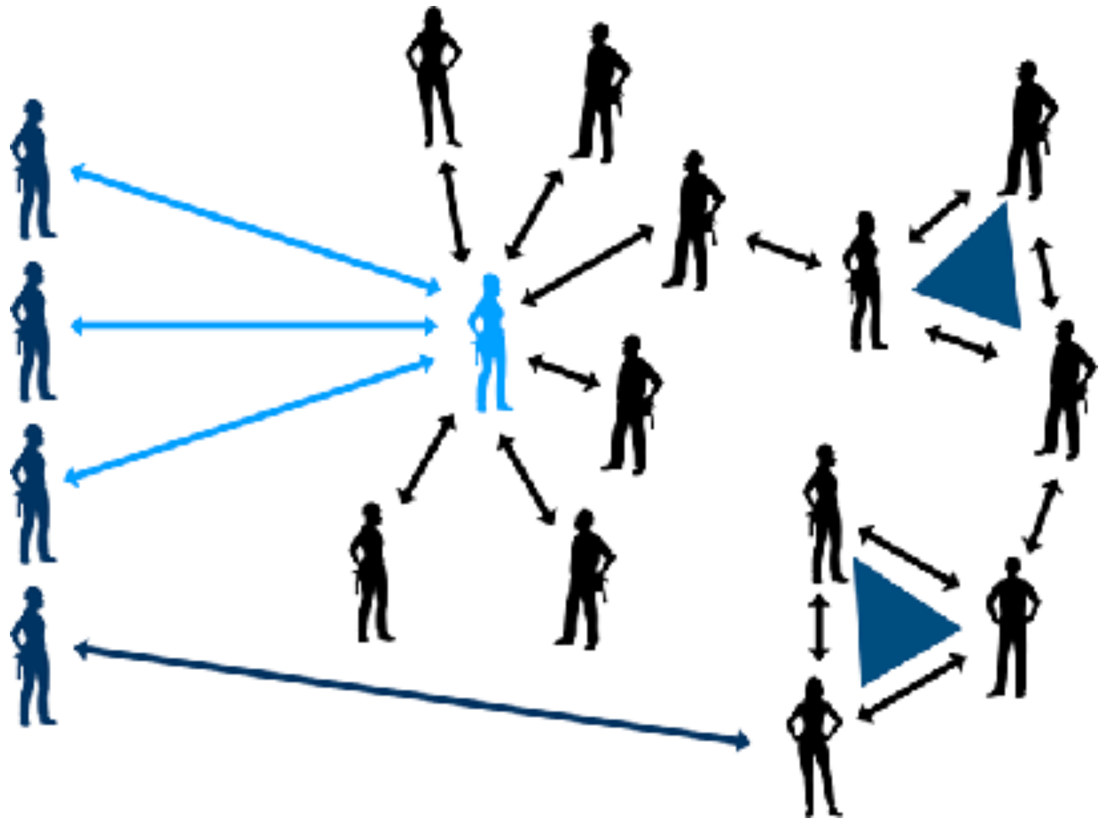
Recall
$\beta_2 \approx$ num of nodes^{1-4x}

- If the preferential attachment effect is strong enough,
- $\beta_q(X_T)$ grows sublinearly with high probability
- $\pi_q(X_\infty) \cong 0$ almost surely

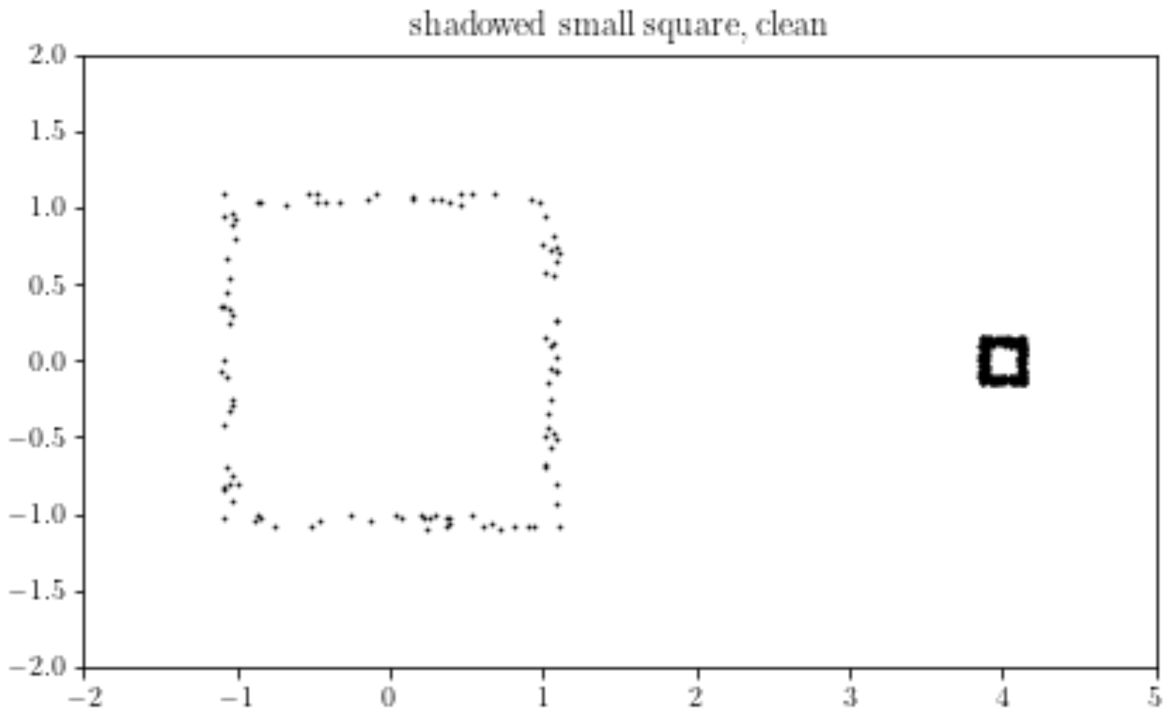
What did we learn today?

- Random topology is cool.
- Preferential attachment graph has interesting topology.
- More interesting things are waiting to be discovered.

Chunyin Siu cs2323@cornell.edu
Cornell University



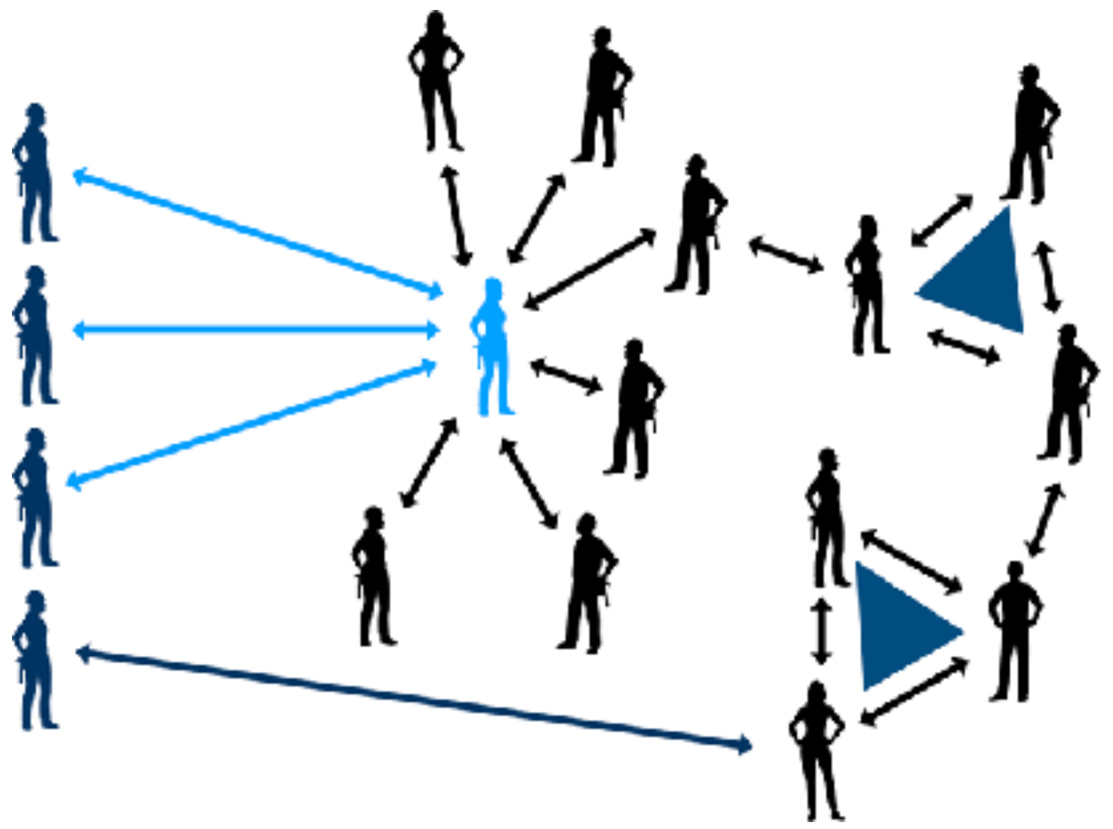
arxiv paper



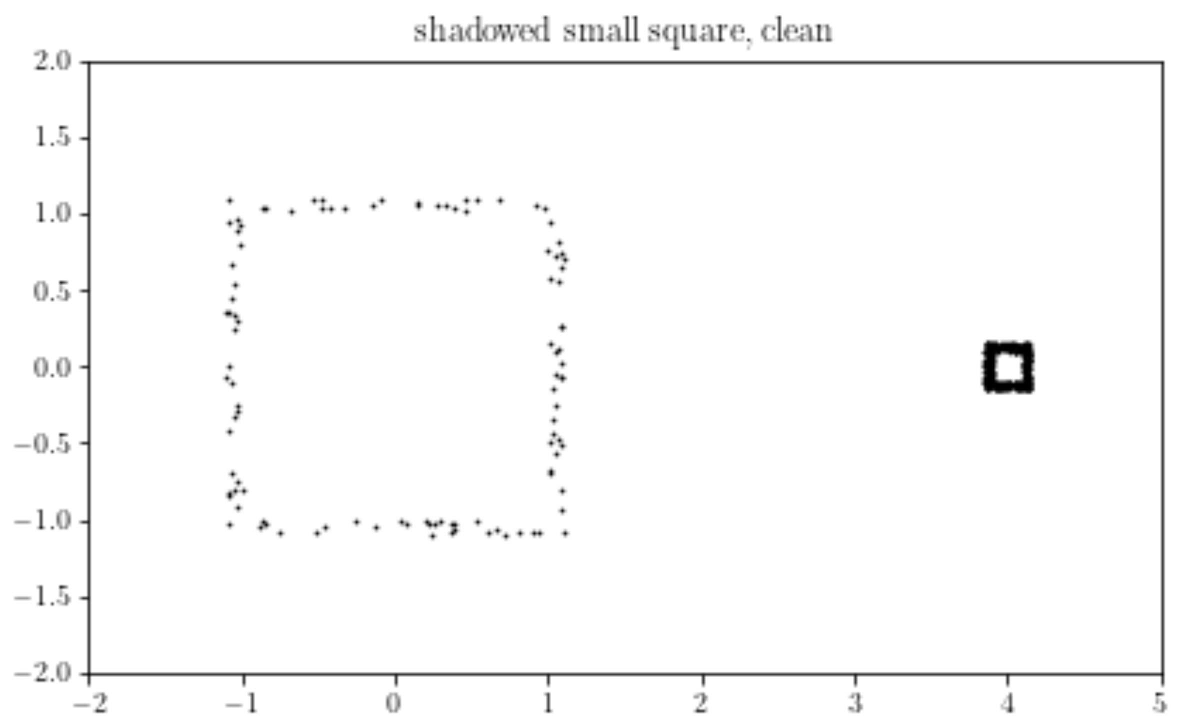
my video about small holes

Thank you!

Chunyin Siu **cs2323@cornell.edu**
Cornell University

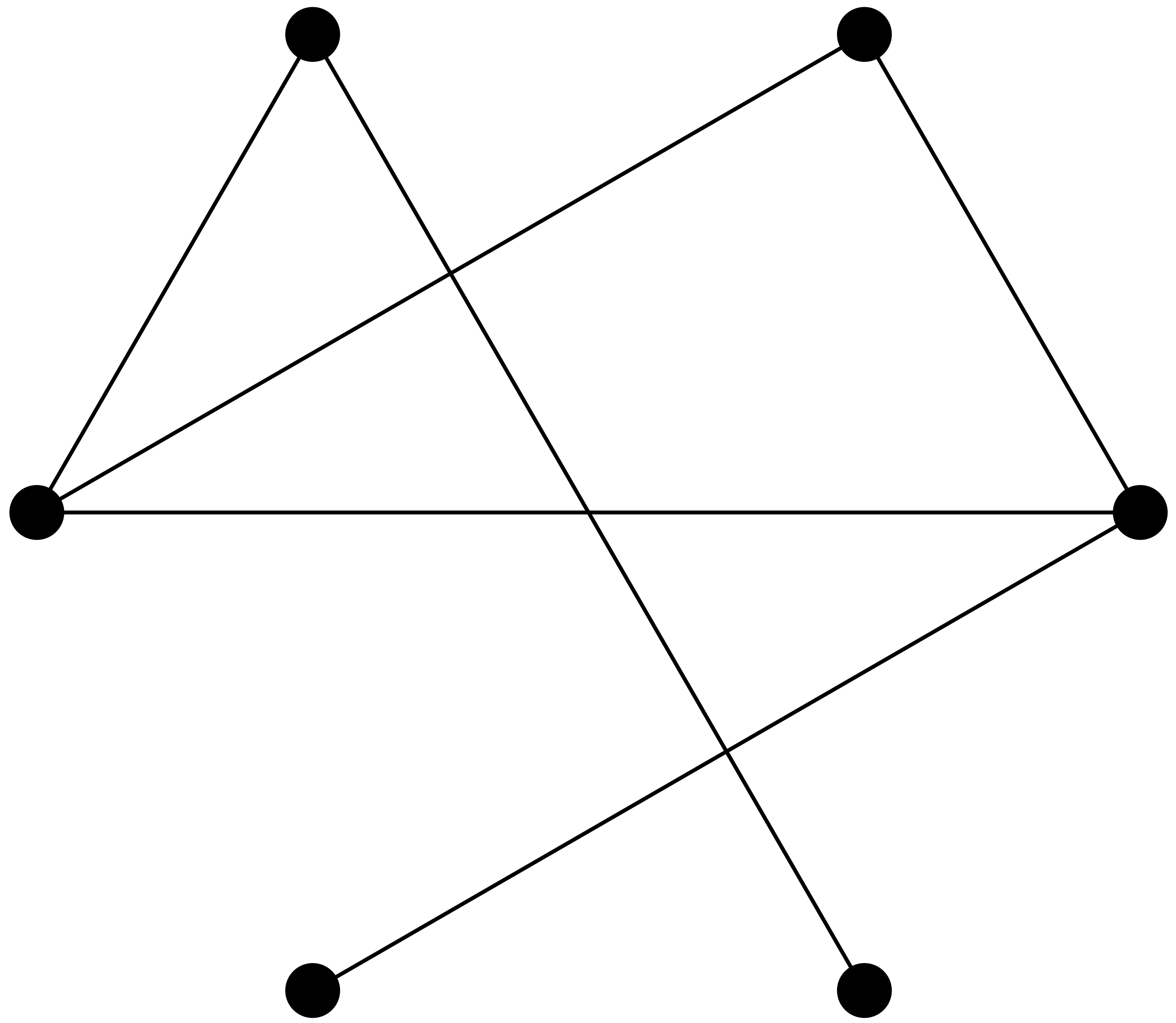


arxiv paper

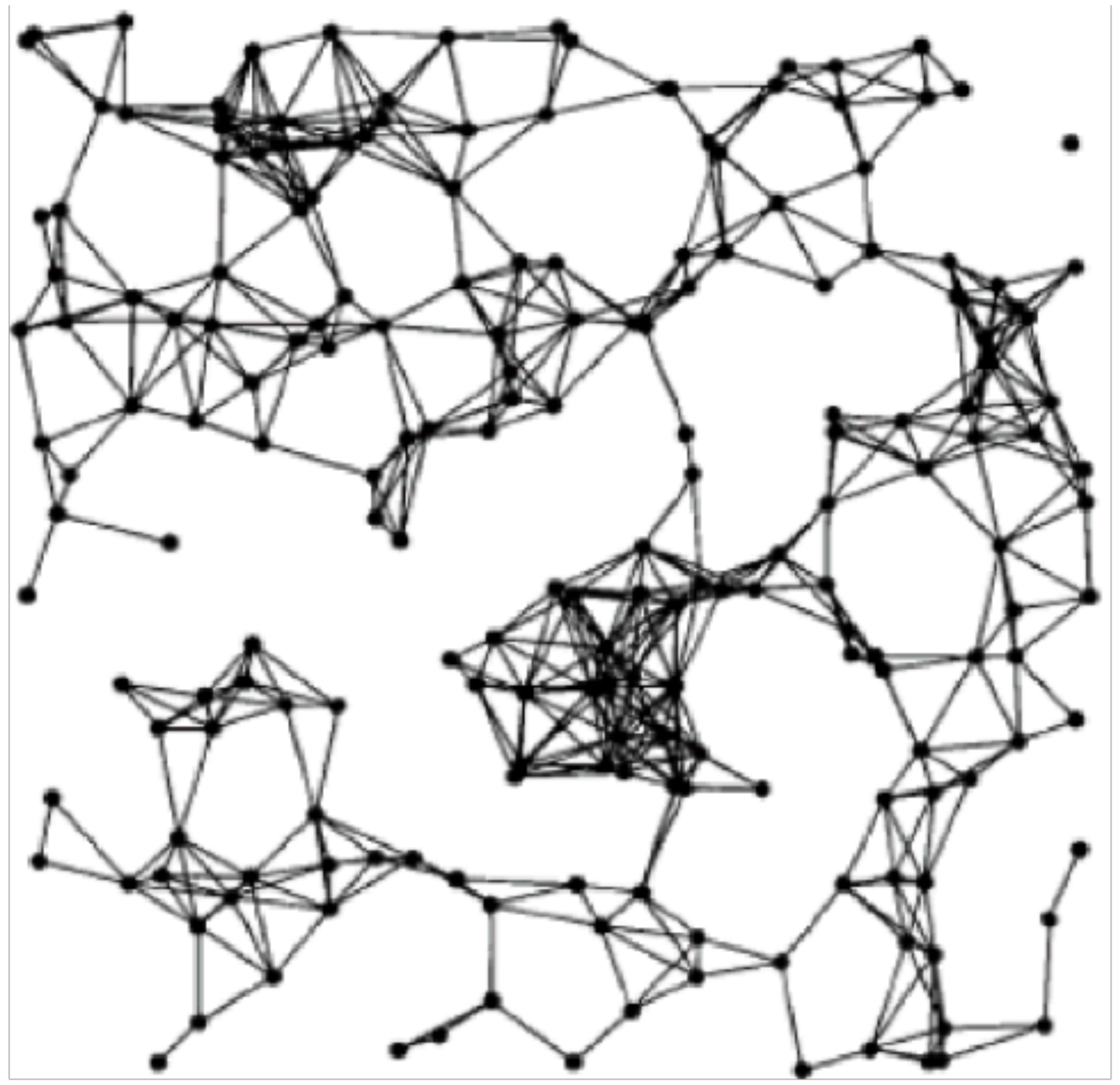


my video about small holes

Tapas of Random Topology

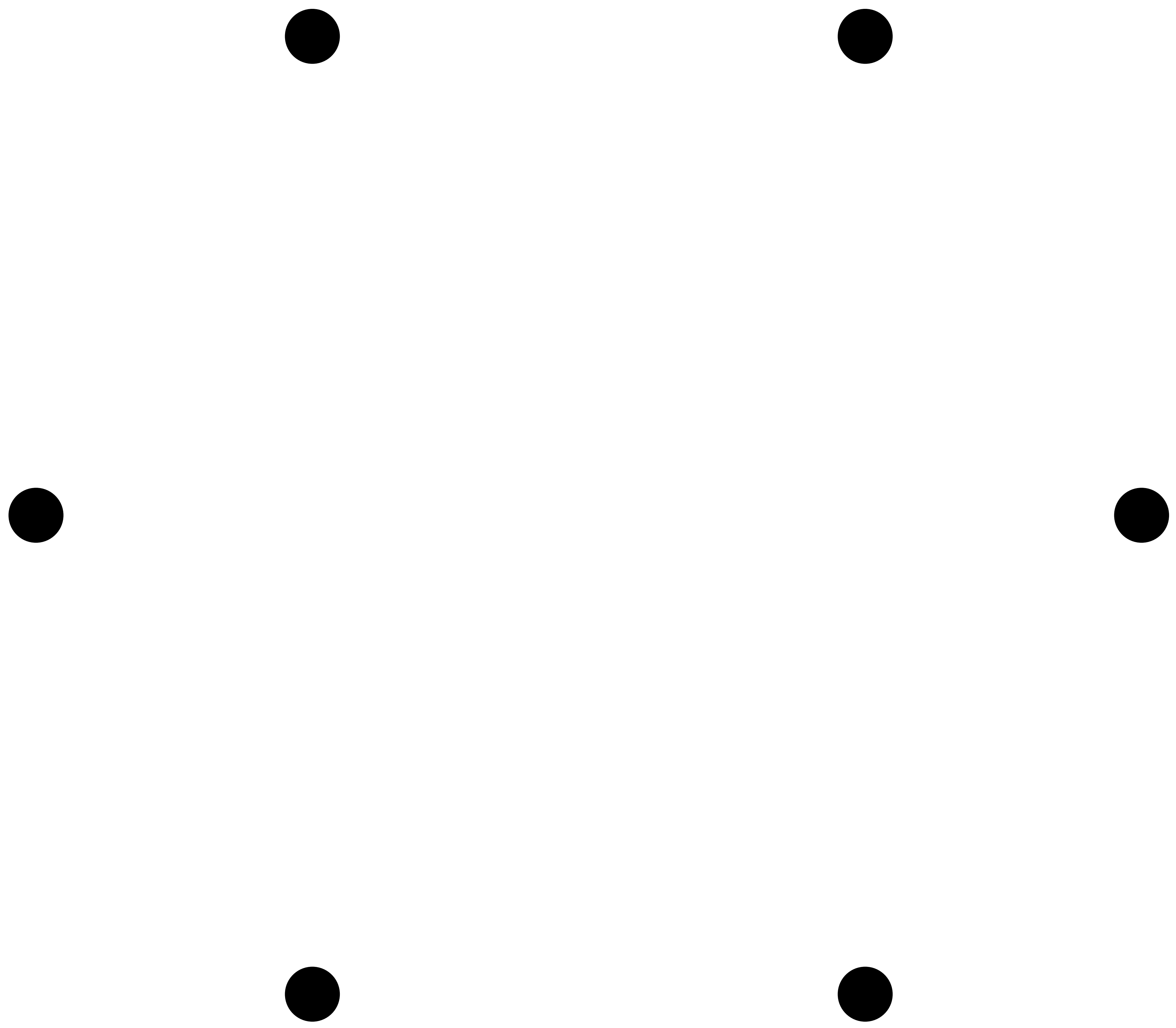


Erdo-Renyi Complexes

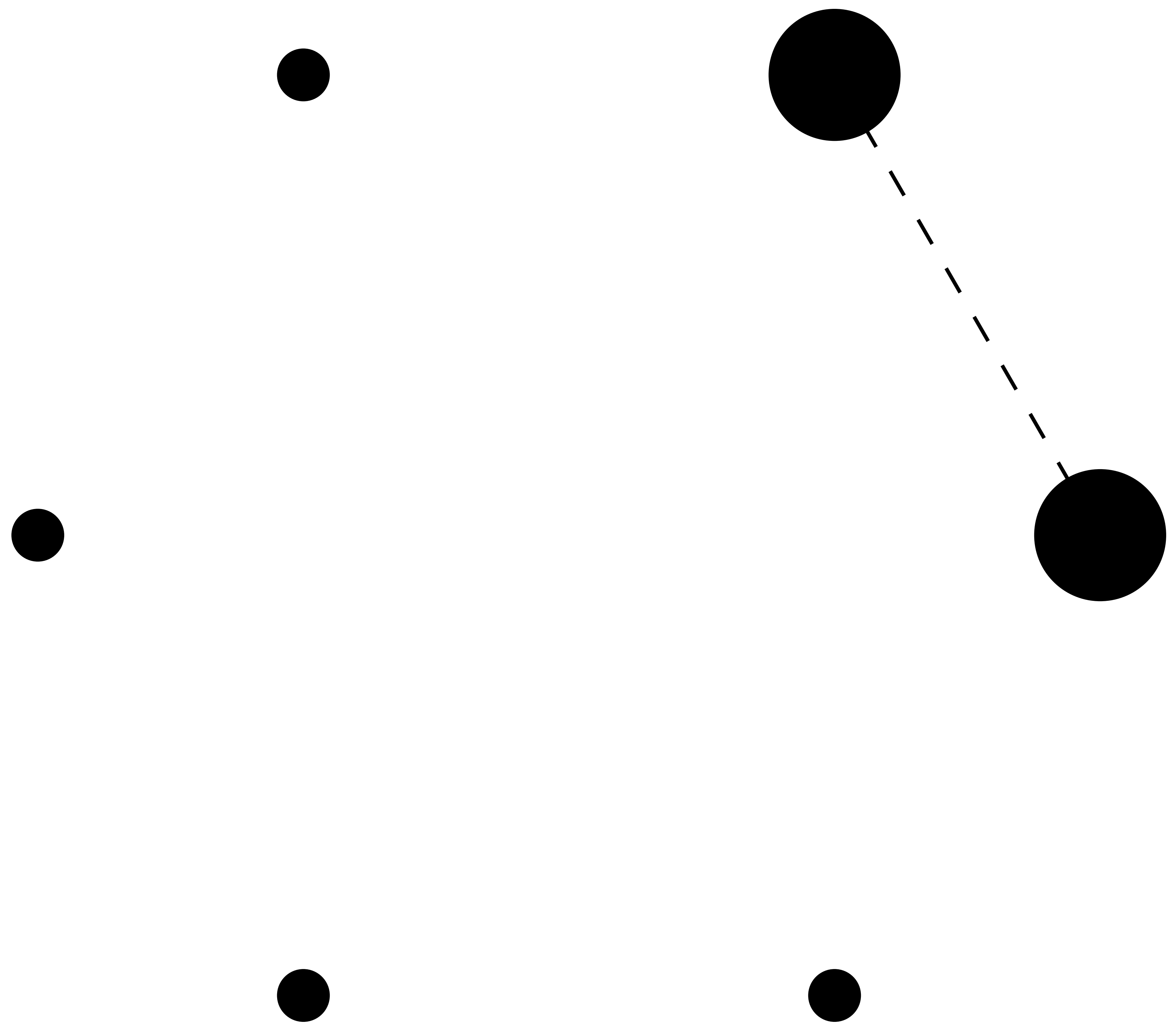


Geometric Complexes

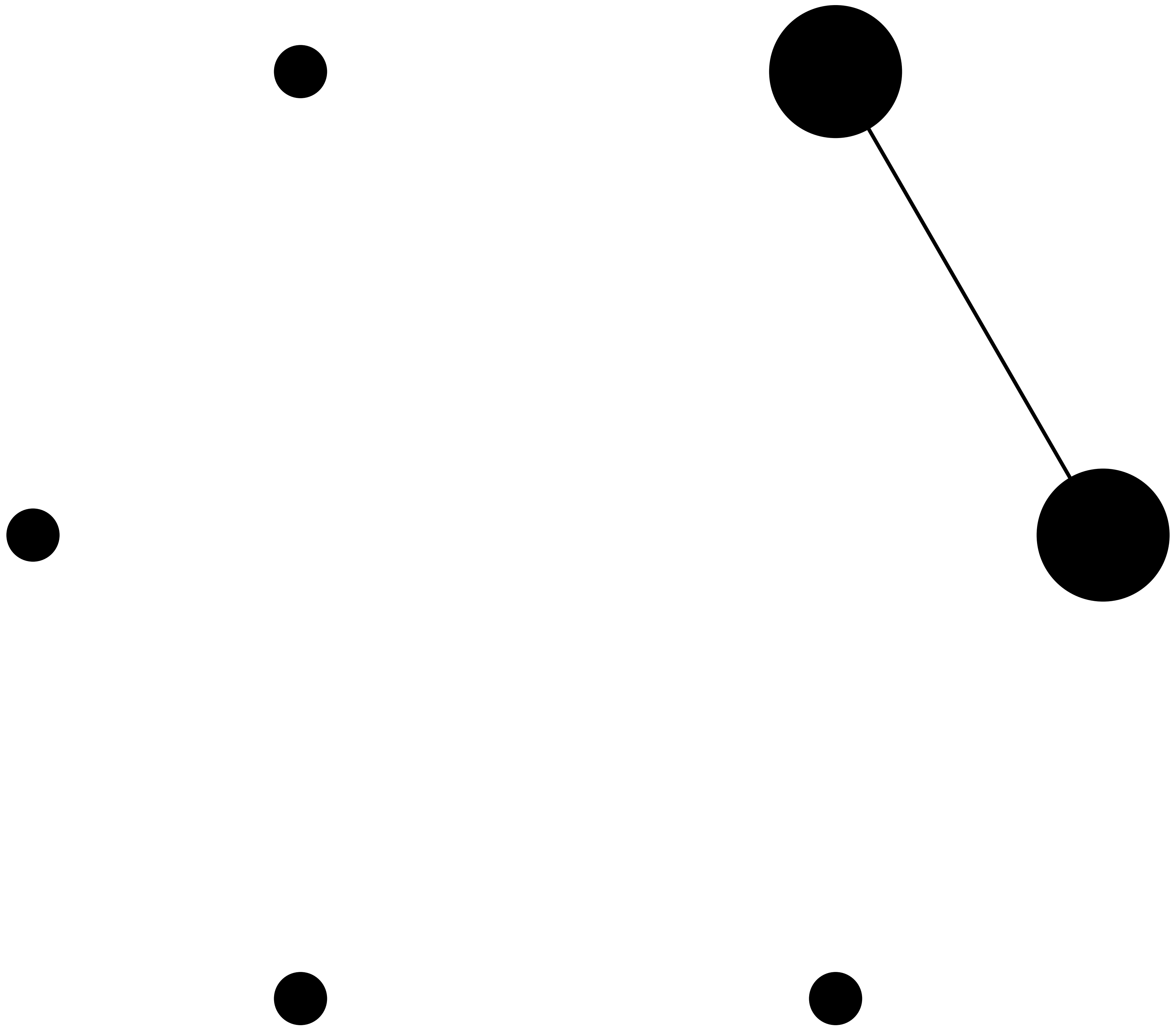
Erdos-Renyi graphs



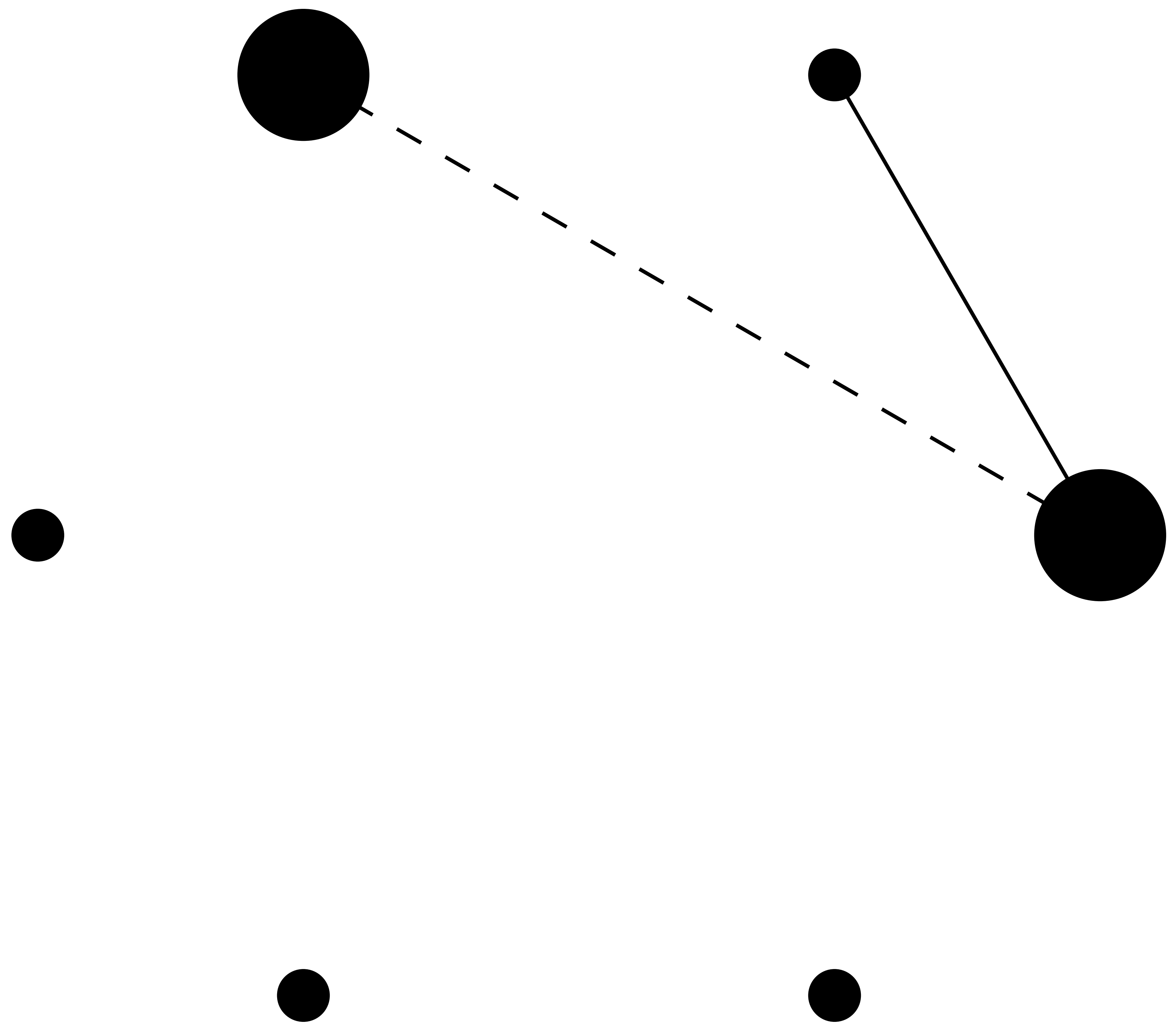
Erdos-Renyi graphs



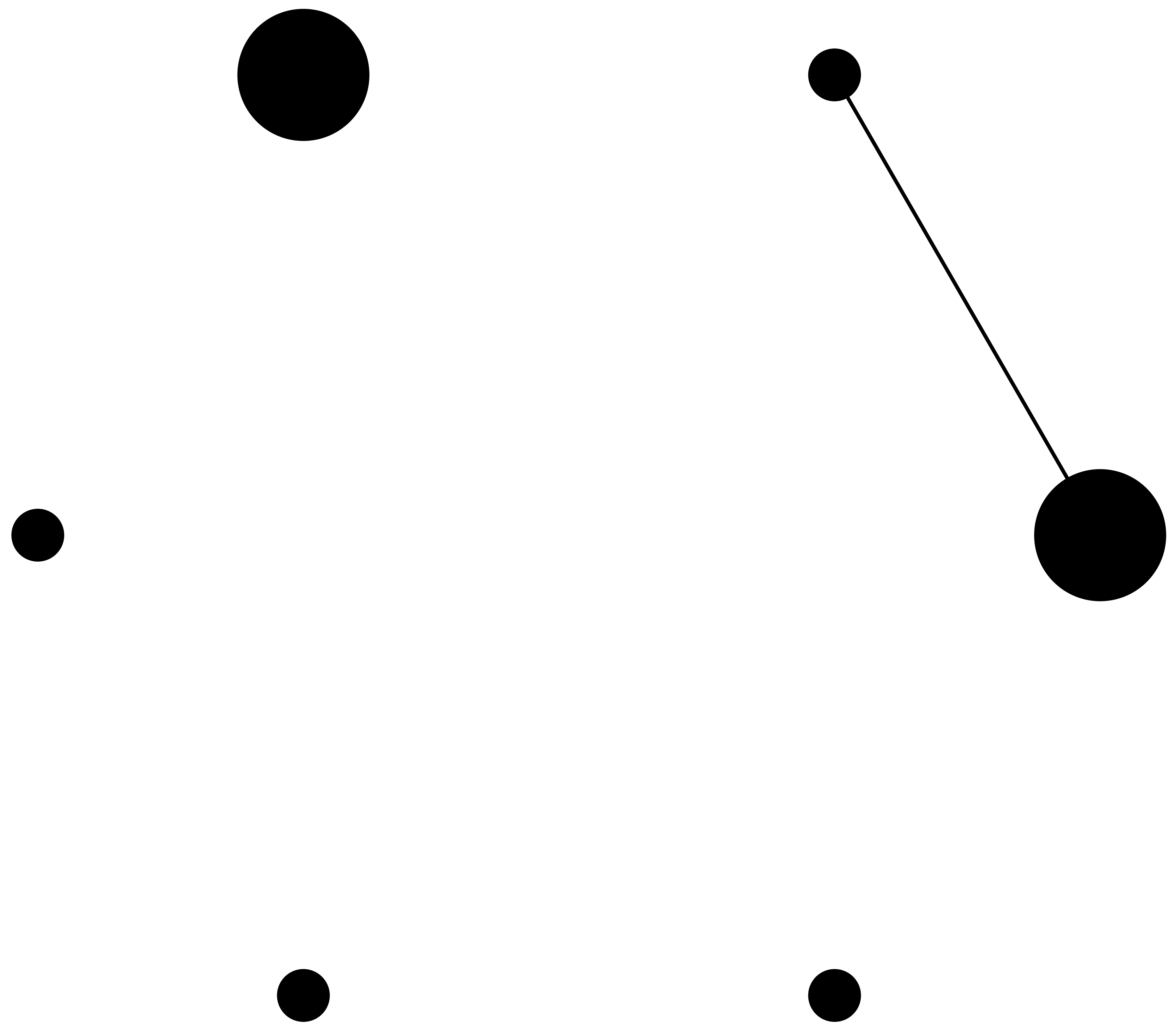
Erdos-Renyi graphs



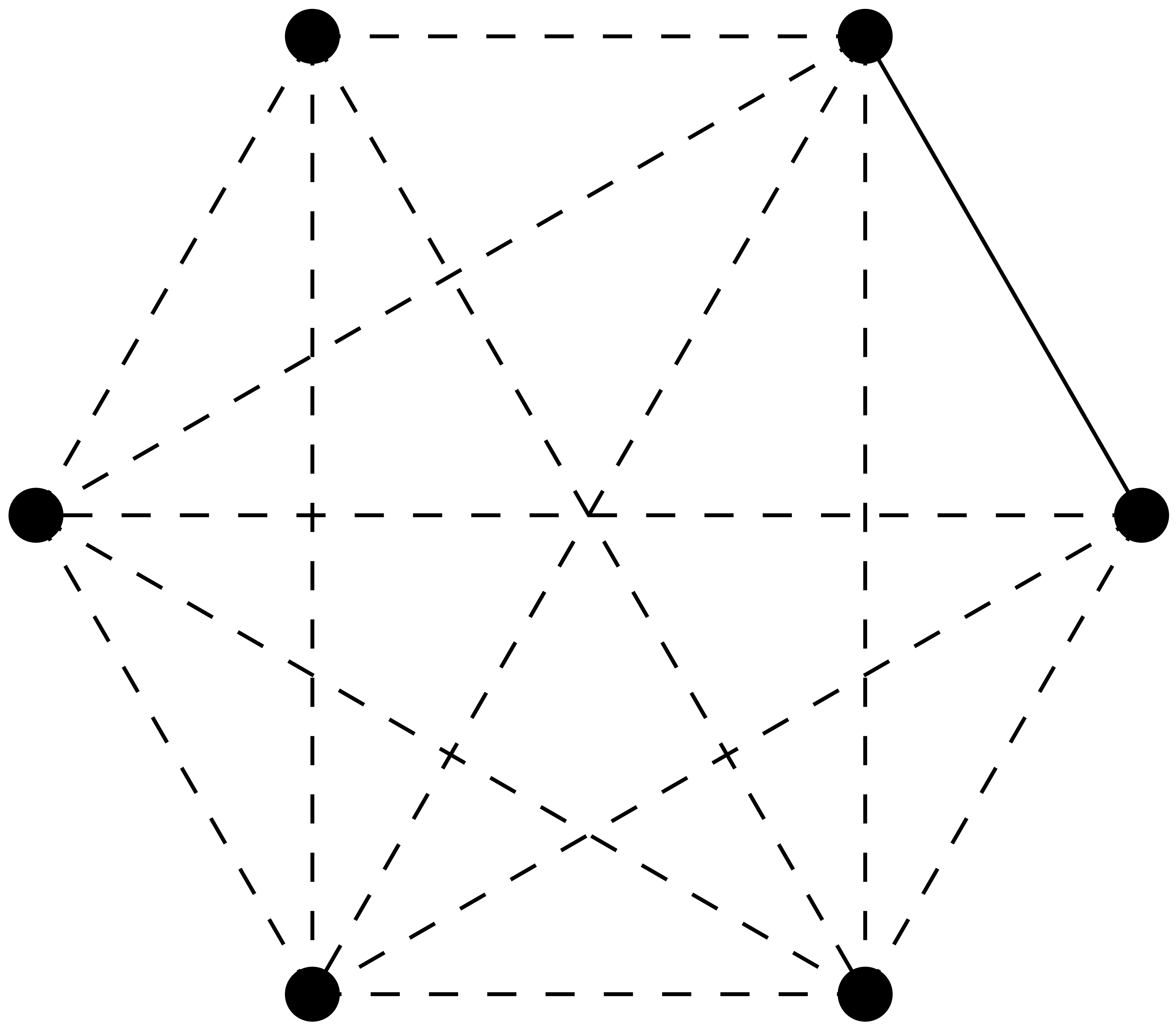
Erdos-Renyi graphs



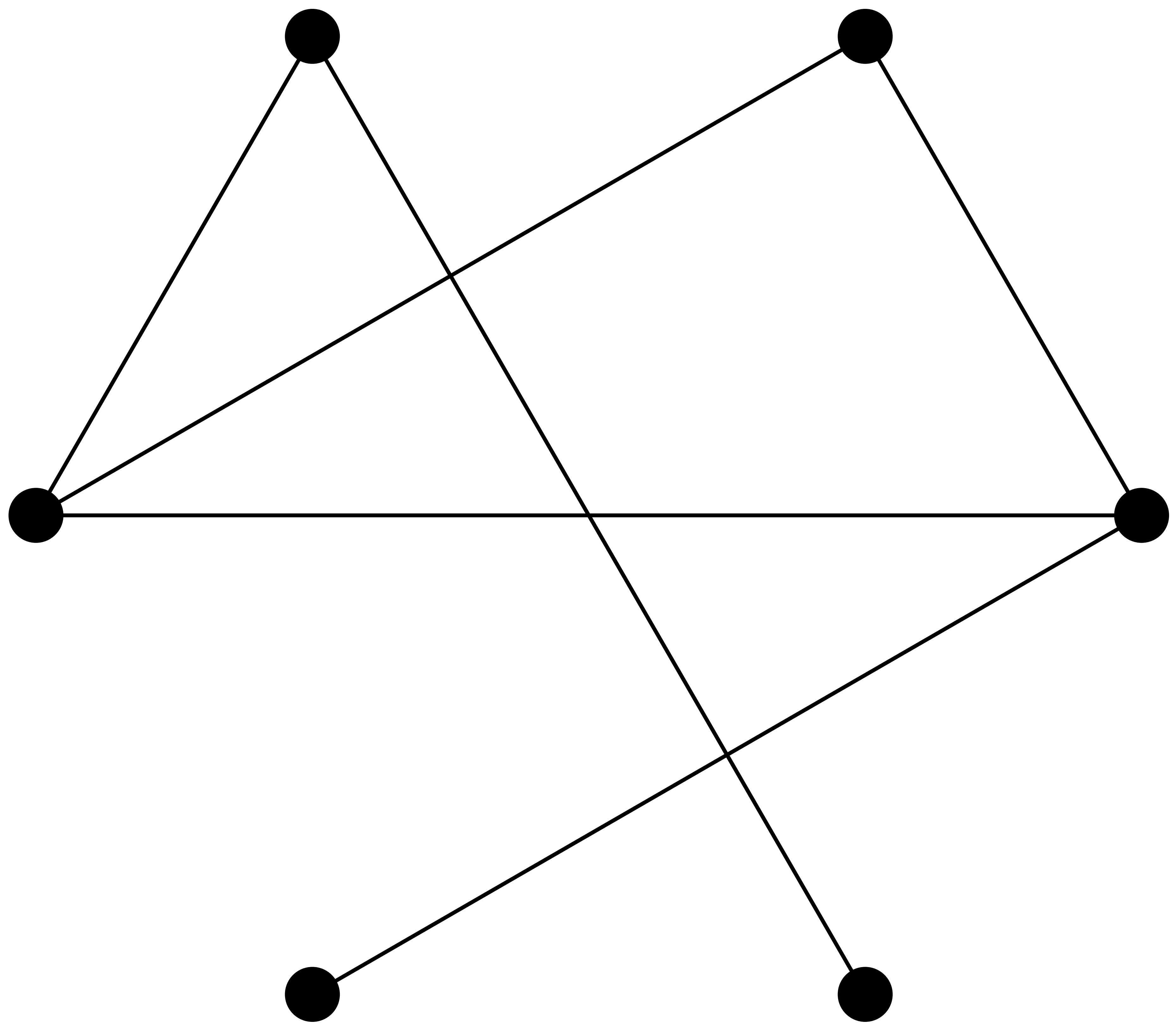
Erdos-Renyi graphs



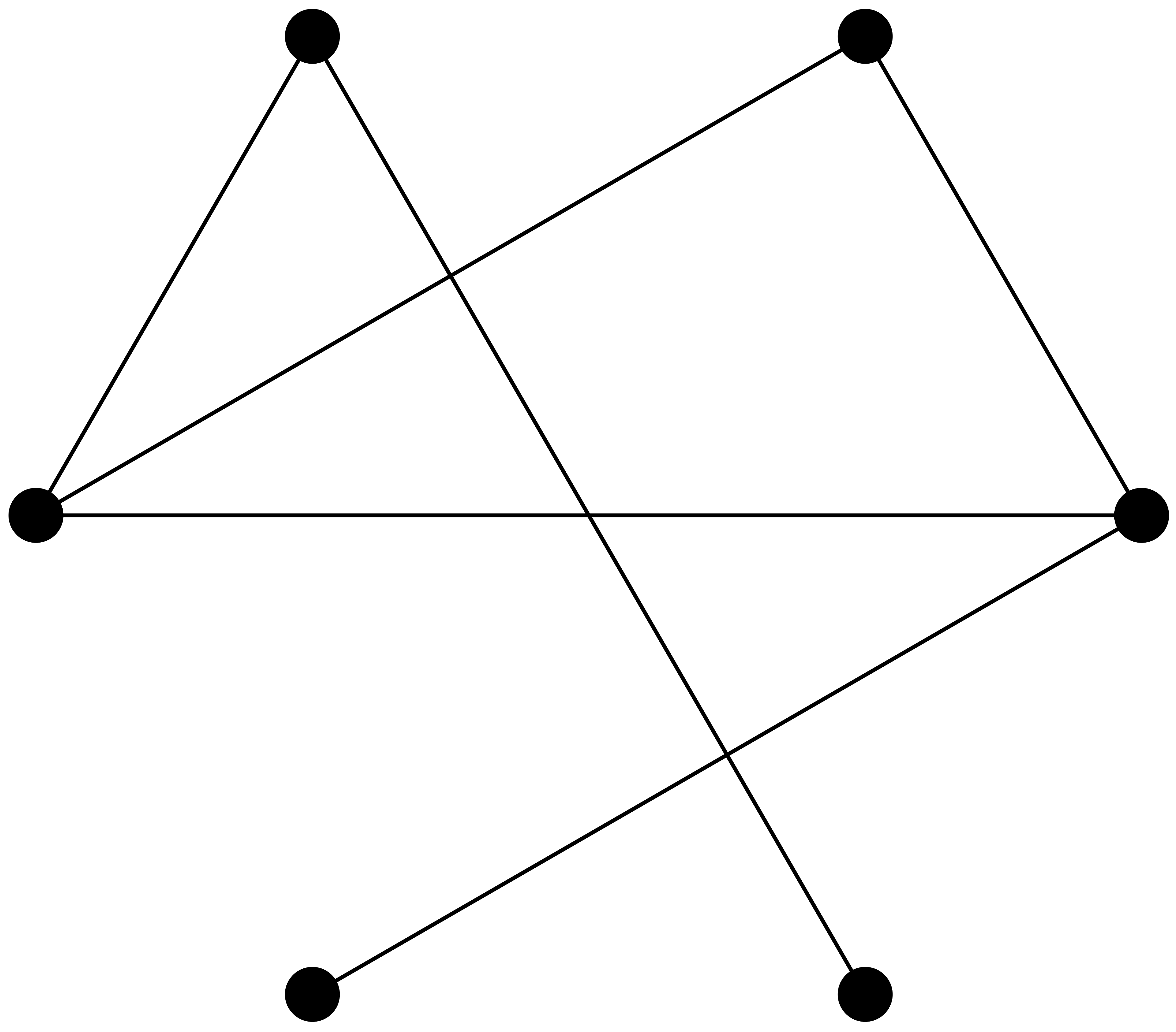
Erdos-Renyi graphs



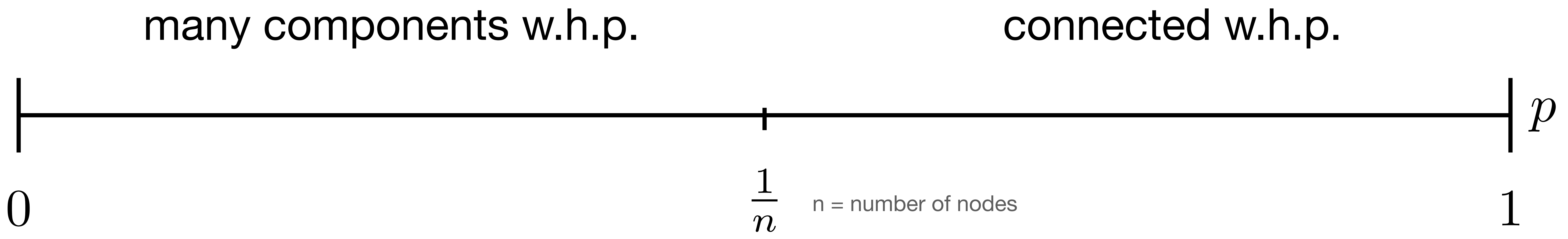
Erdos-Renyi graphs



Erdos-Renyi graphs

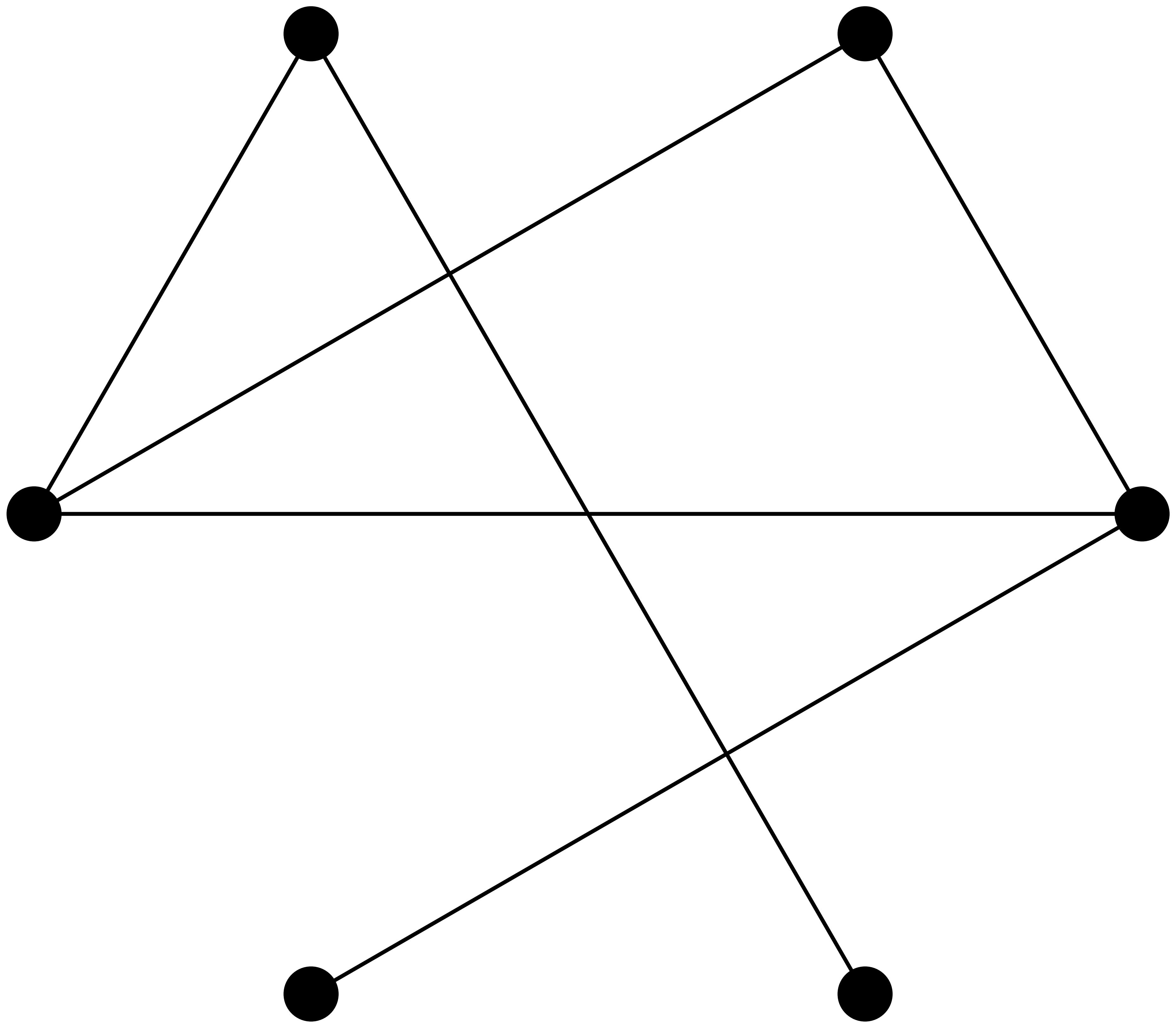


Phase Transition
[Erdos-Renyi 1960]

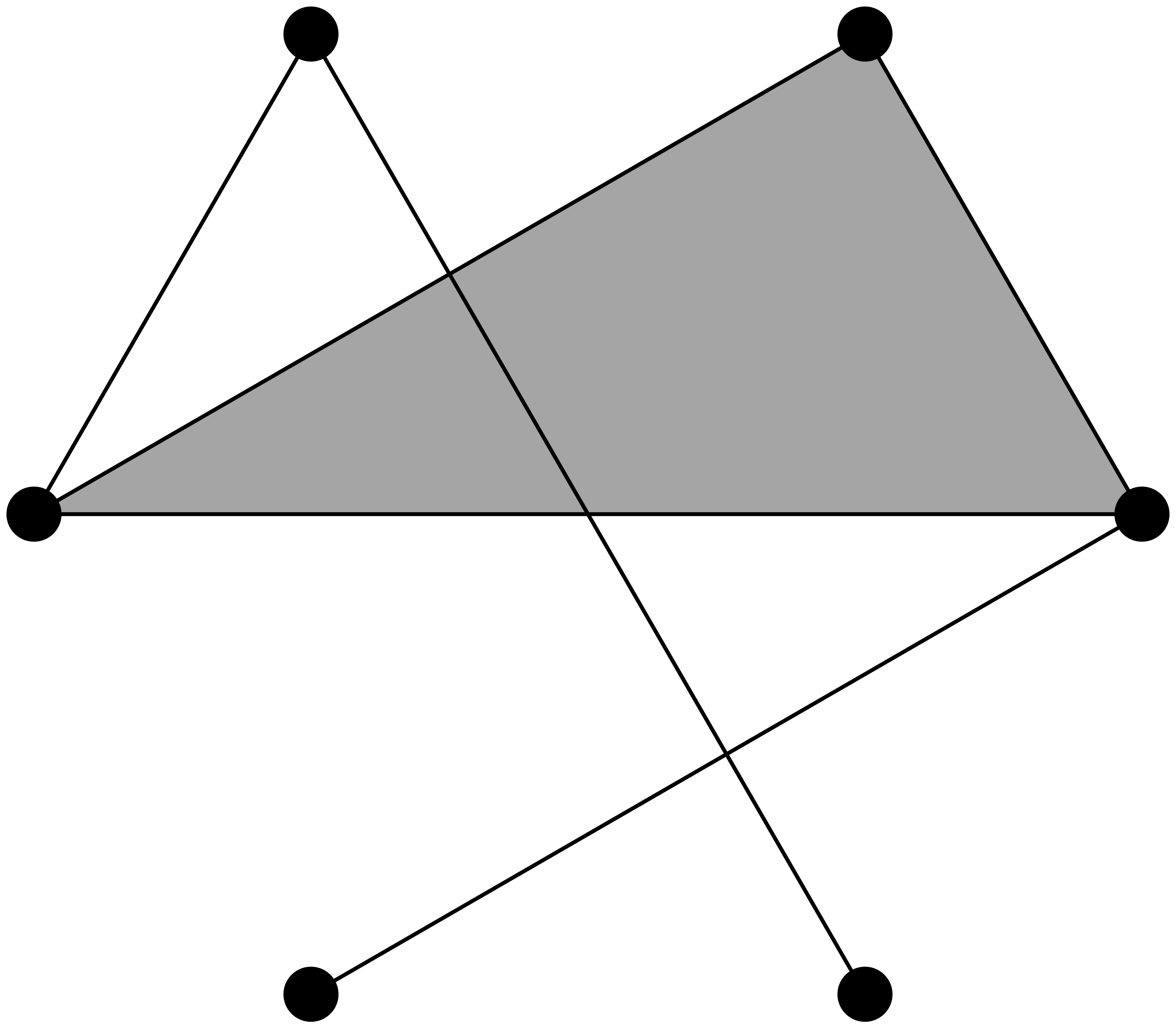


all log terms and constants forgone

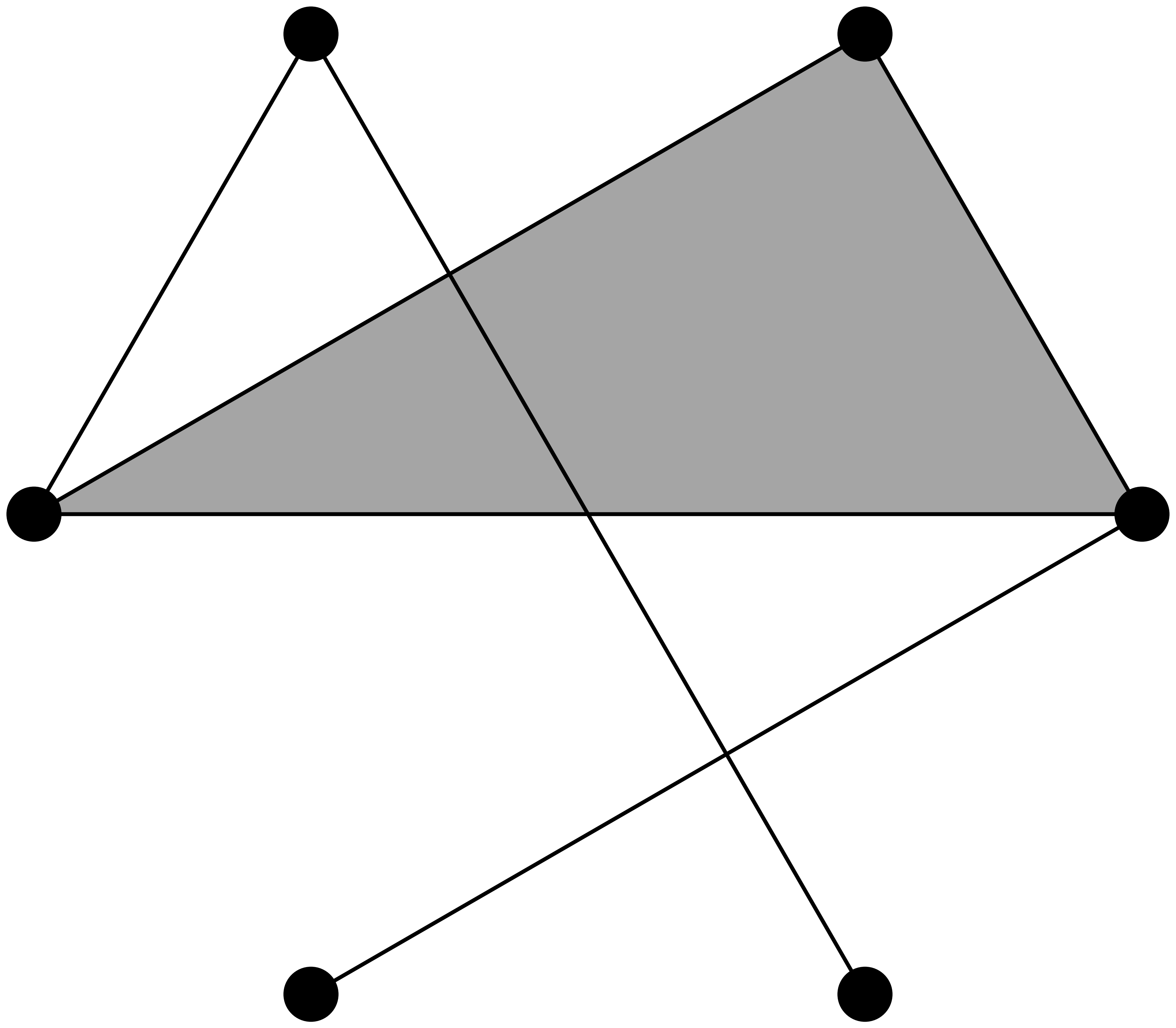
Erdos-Renyi Clique Complex

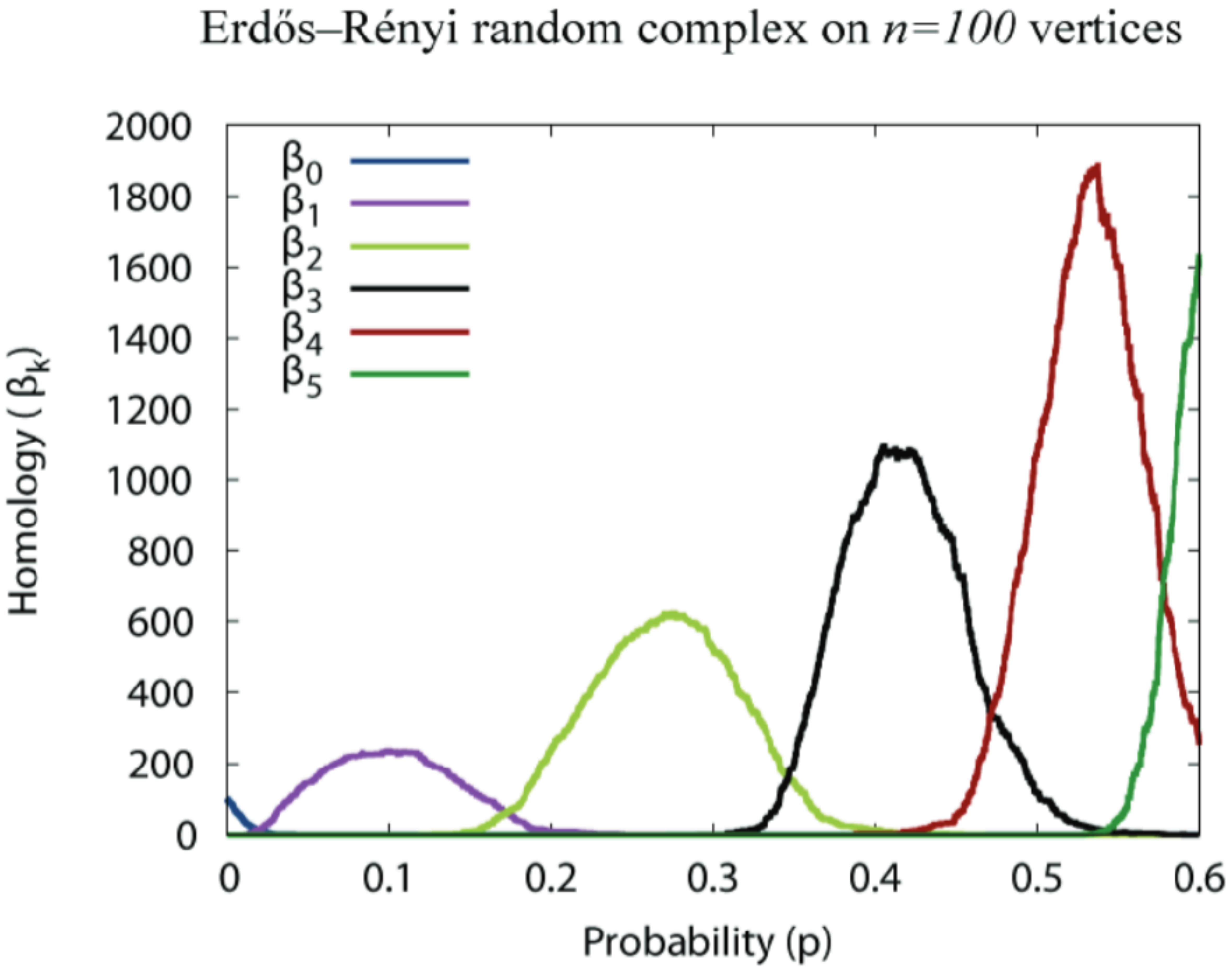


Erdos-Renyi Clique Complex



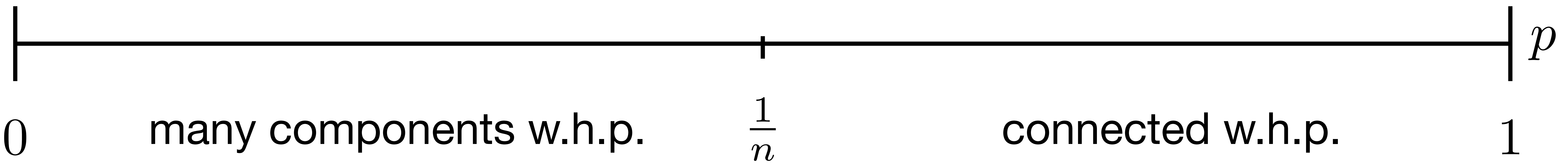
Betti Numbers





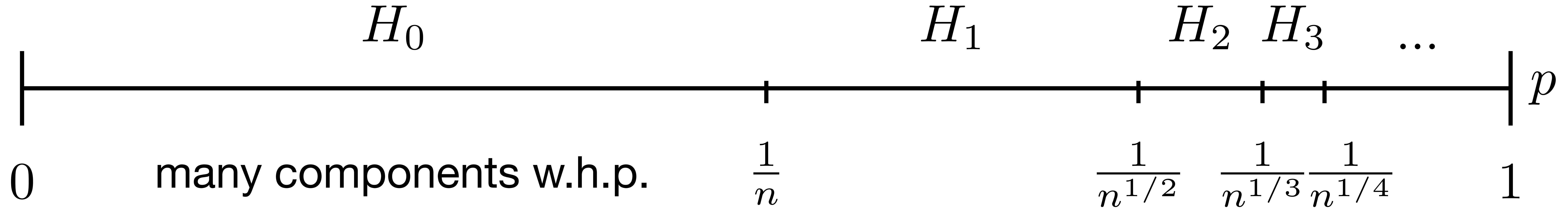
computation and plotting done by Zomorodian

Phase Transition
[Erdos-Renyi 1960]



n = number of nodes
all log terms and constants forgone

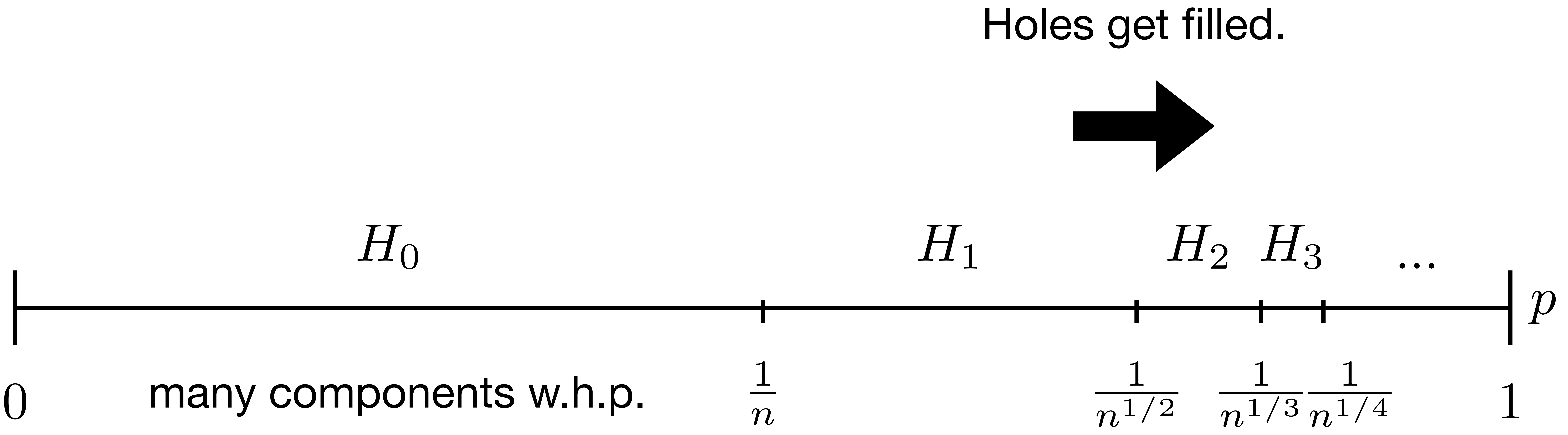
Phase Transition
[Kahle 2009, 2014]



n = number of nodes
all log terms and constants forgone

Phase Transition

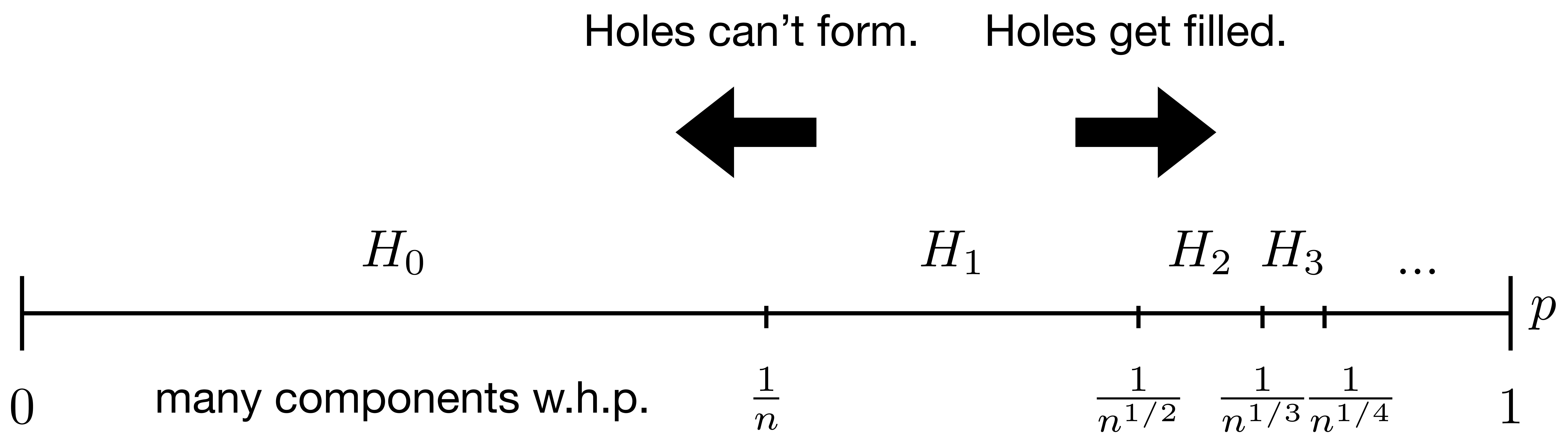
[Kahle 2009, 2014]



n = number of nodes
all log terms and constants forgone

Phase Transition

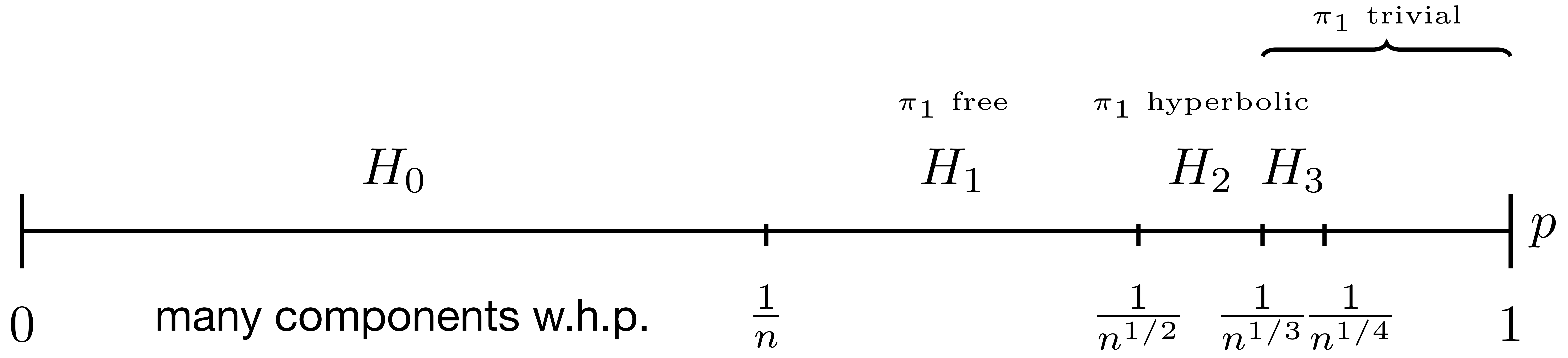
[Kahle 2009, 2014]



n = number of nodes
all log terms and constants forgone

Fundamental Group

[Kahle 2009, Babson 2012, Costa-Farber-Horak 2015]



all log terms and constants forgone

Geometric Complexes

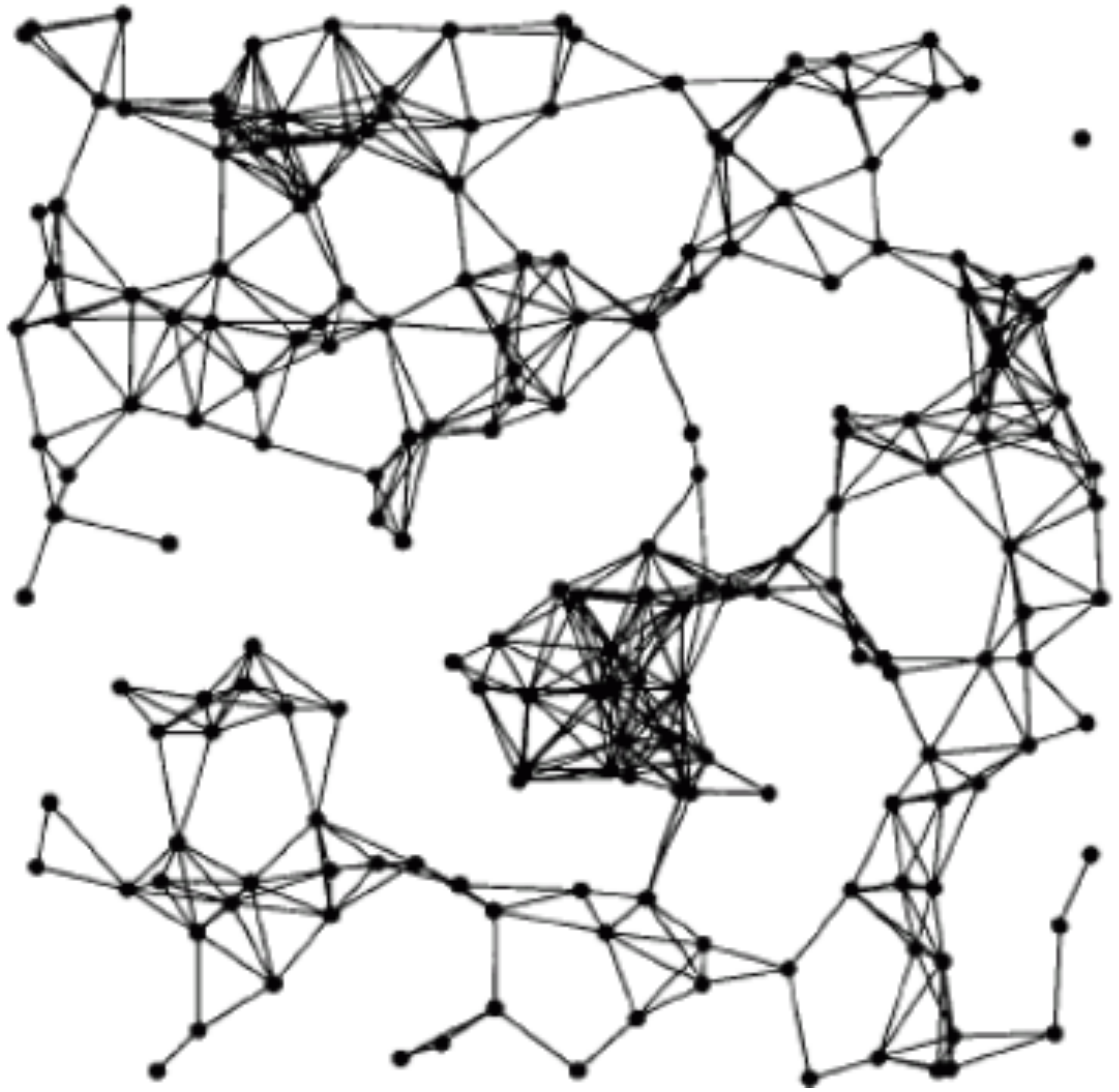


image credit: Penrose

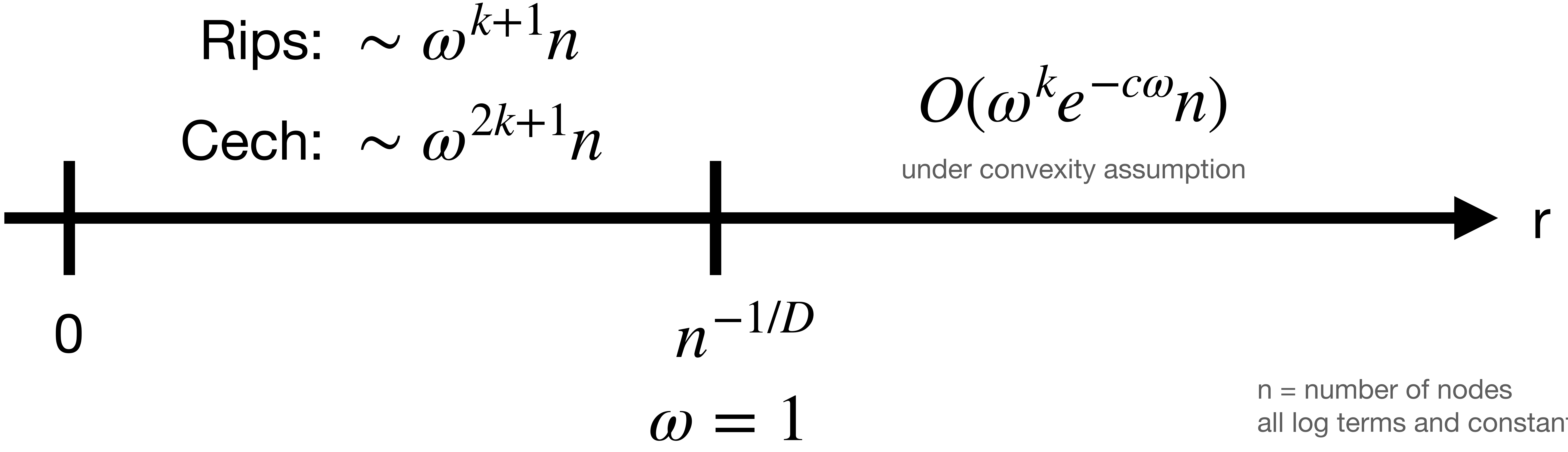
Expected Betti numbers at dimension k

- Let $\omega = nr^D$, where D is the ambient dimension

Expected Betti numbers at dimension k

[Kahle 2011]

- Let $\omega = nr^D$, where D is the ambient dimension



n = number of nodes
all log terms and constants forgone

Functional Convergence at dimension k?

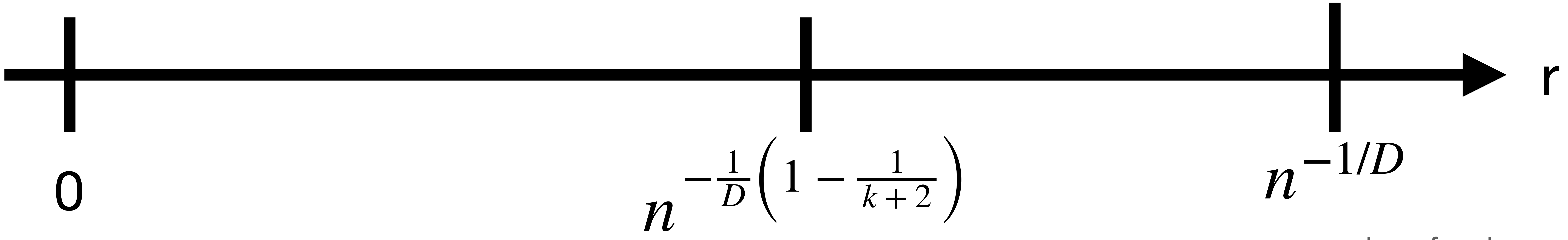
[Thomas and Owada 2020]

n = number of nodes
all log terms and constants forgone

Functional Convergence at dimension k?

[Thomas and Owada 2020]

- Cech: weak convergence in finite-dimensional sense

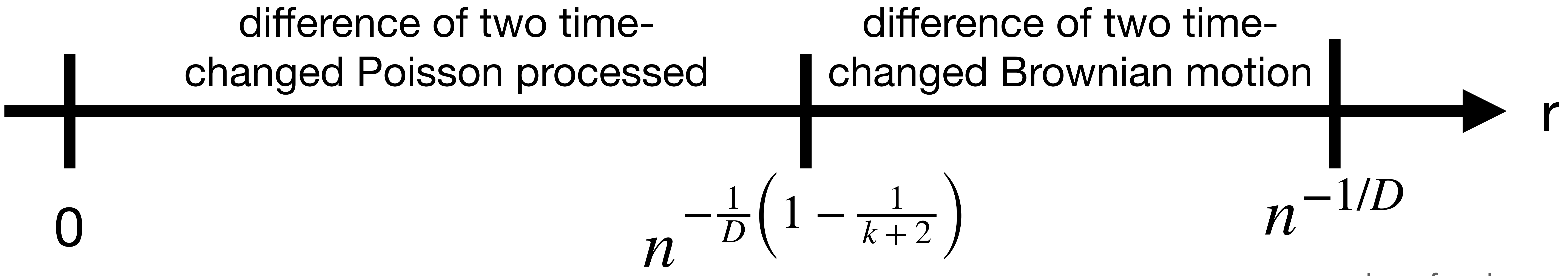


n = number of nodes
all log terms and constants forgone

Functional Convergence at dimension k?

[Thomas and Owada 2020]

- Cech: weak convergence in finite-dimensional sense

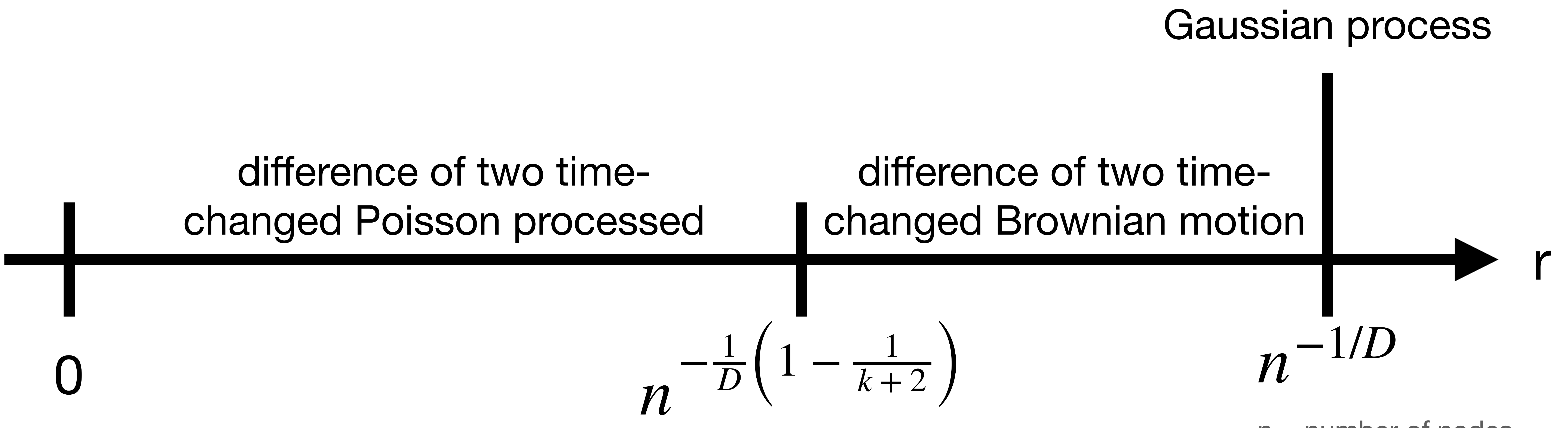


n = number of nodes
all log terms and constants forgone

Functional Convergence at dimension k?

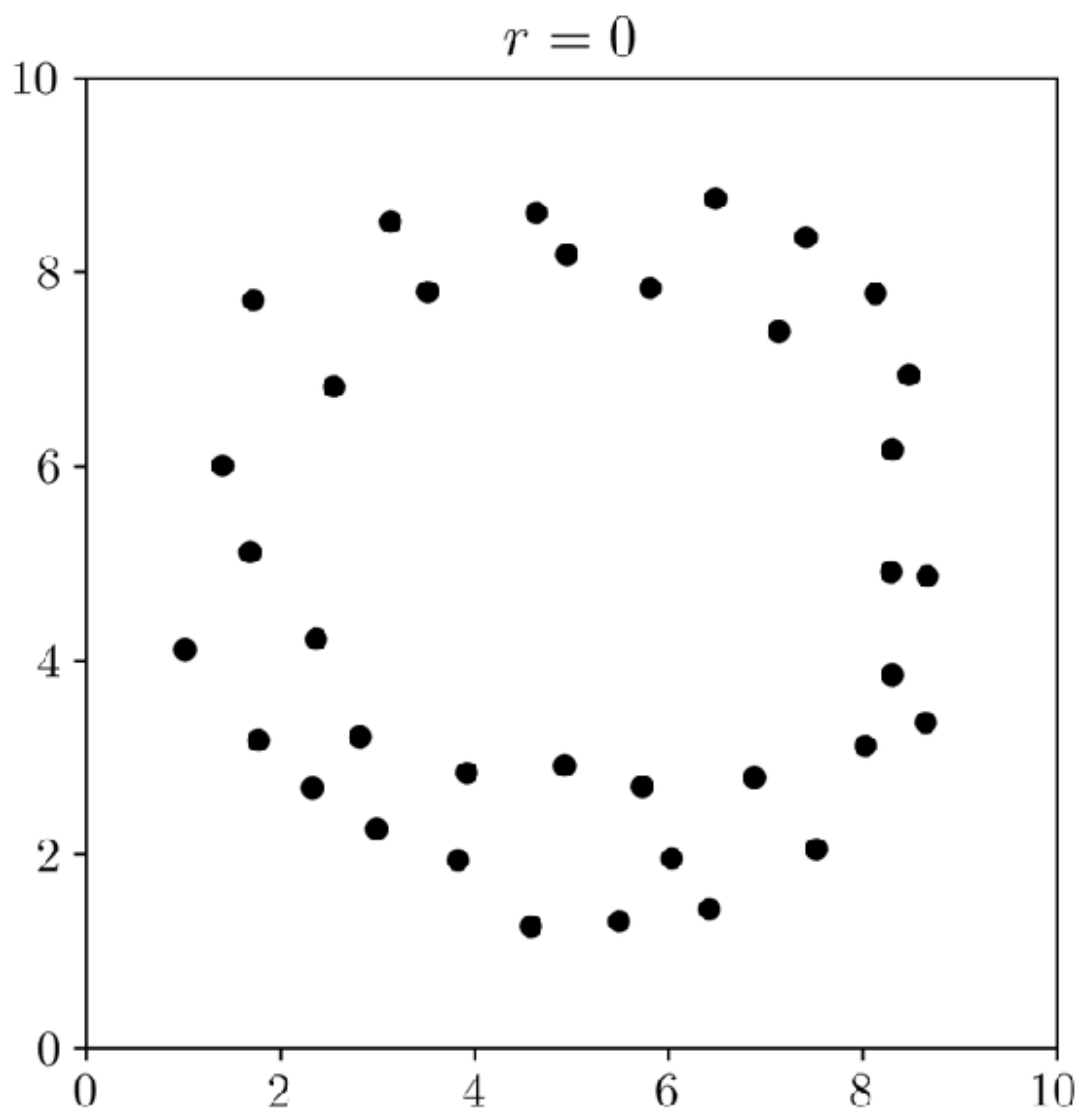
[Thomas and Owada 2020]

- Cech: weak convergence in finite-dimensional sense



n = number of nodes
all log terms and constants forgone

Maximally Persistent Cycles



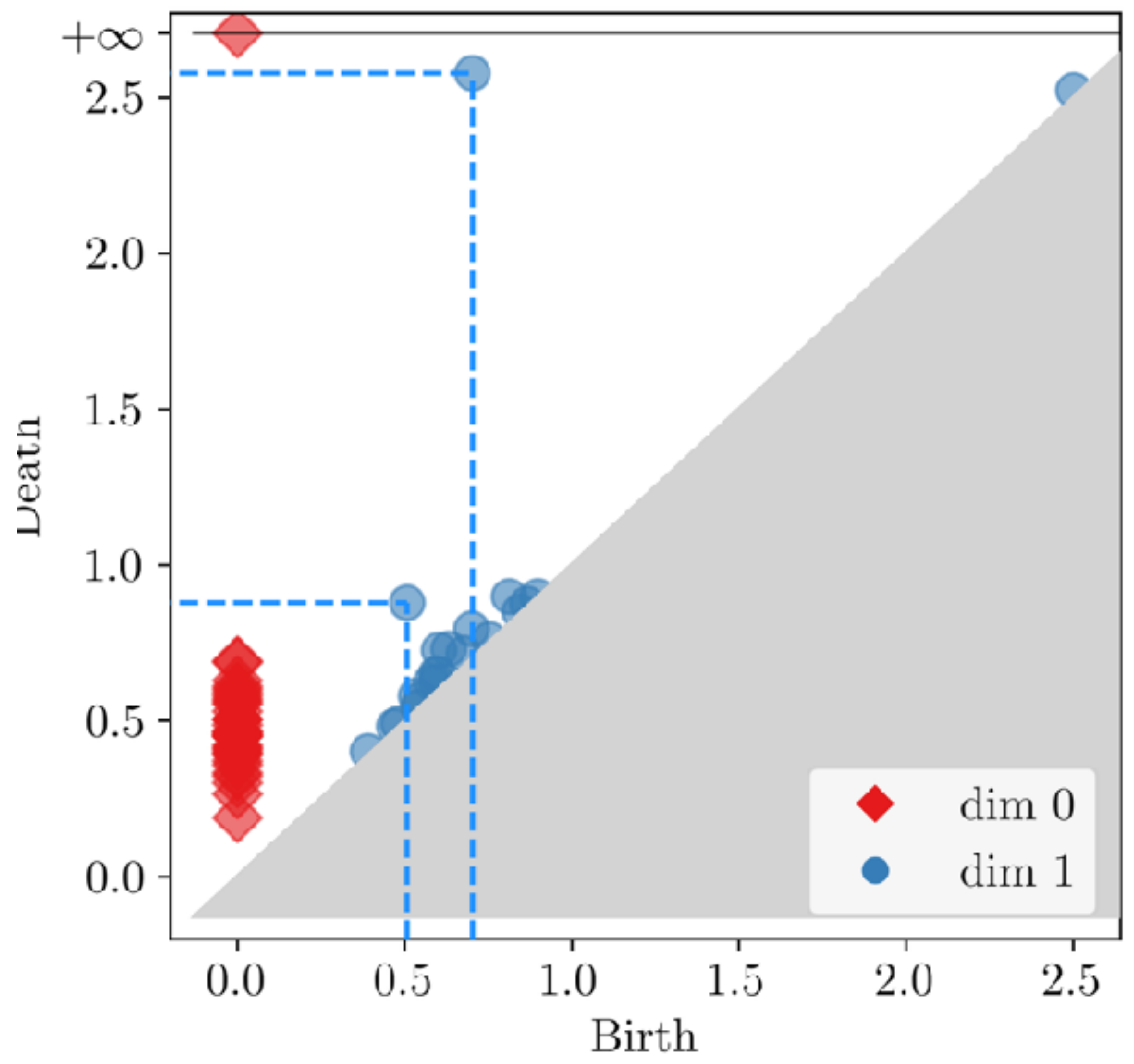


image credit: Andrey Yao

Maximally Persistent Cycles

n points in expectation

k-cycle

Maximally Persistent Cycles

[Bobrowski-Kahle-Skraba 2017]

n points in expectation

k-cycle

$$c\left(\frac{\log n}{\log\log n}\right)^{1/k} \leq \text{max persistence} \leq C\left(\frac{\log n}{\log\log n}\right)^{1/k}$$

a.a.s.

- 4 CPU cores
- 40 minutes for the Betti numbers
- 7.5 hours for bounds
- memory issues for larger graphs

Subtleties

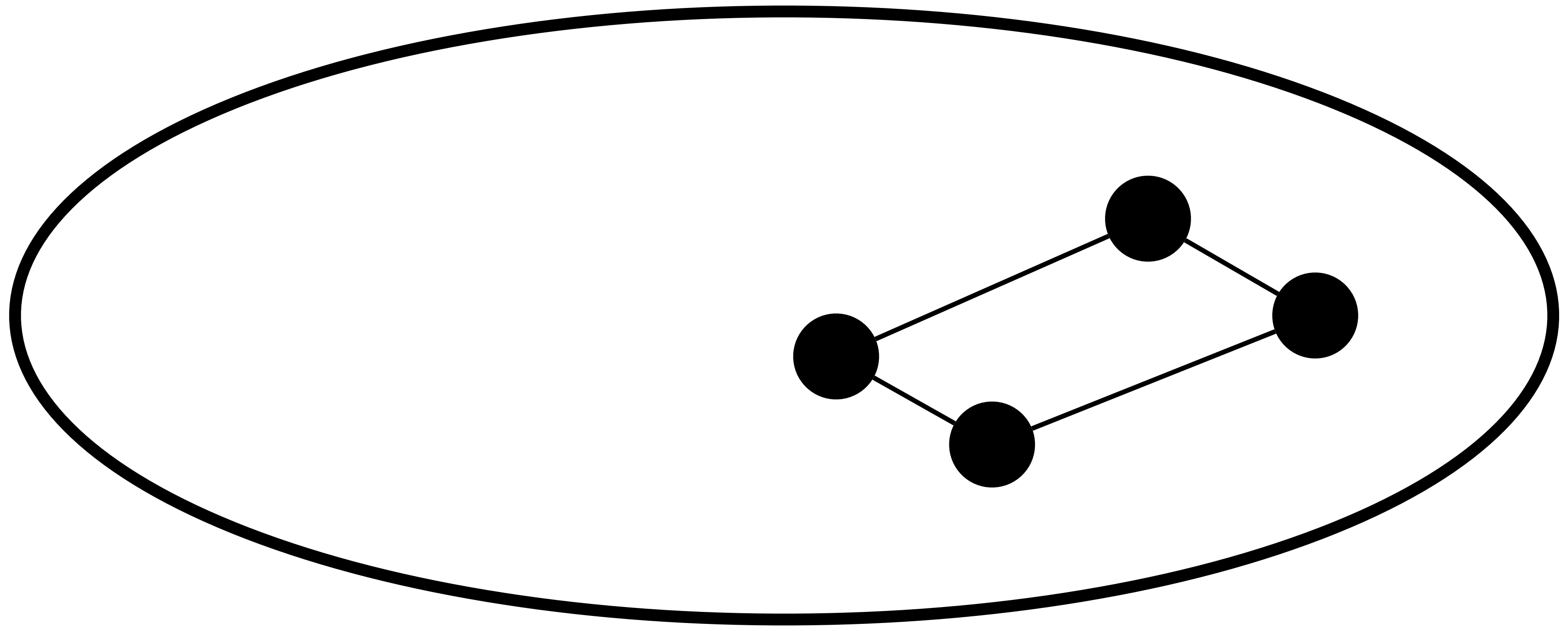
- Need homological algebra to relate Betti numbers with counts

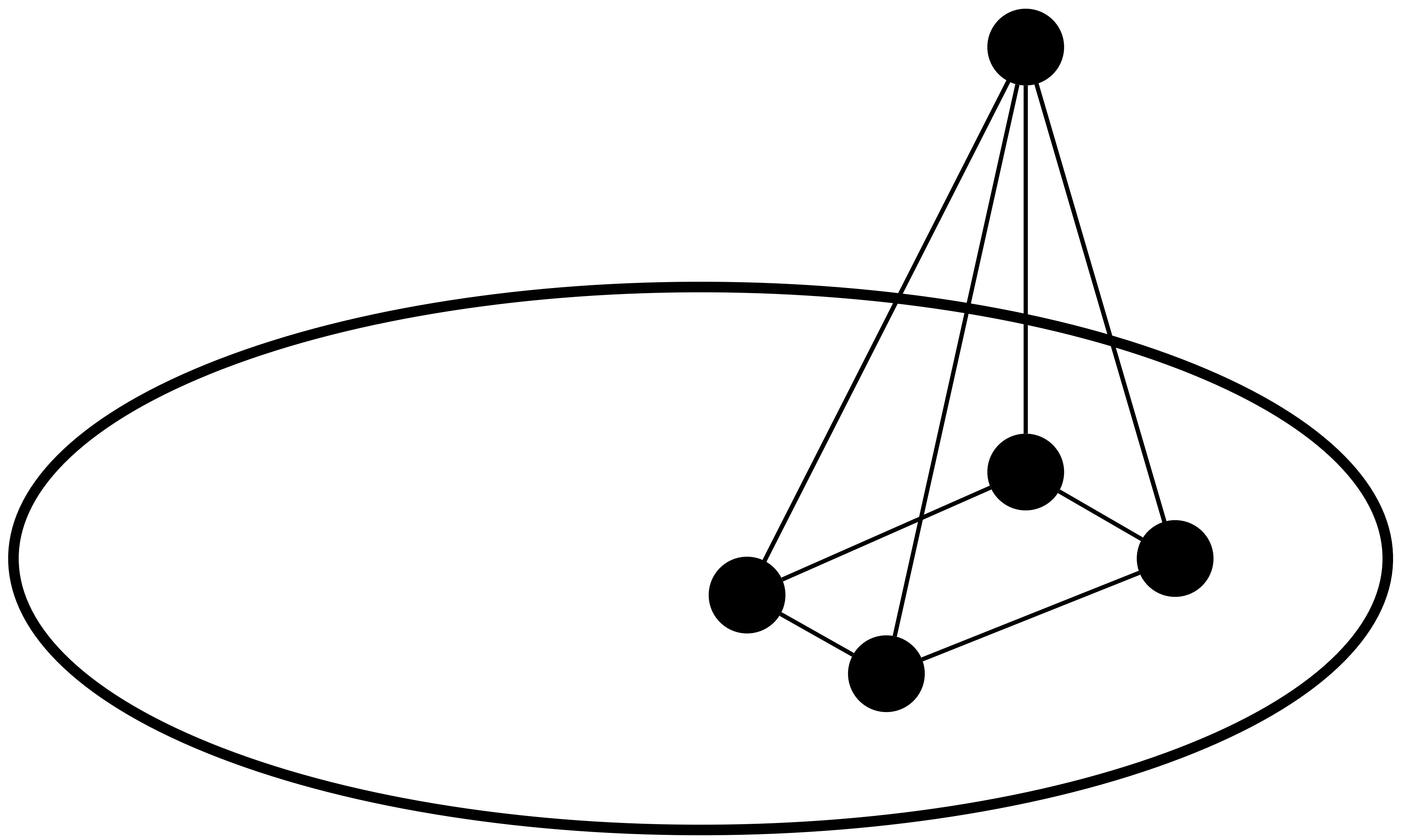
Subtleties

- Need homological algebra to relate Betti numbers with counts
  - adding a vertex = construct mapping cone

Subtleties

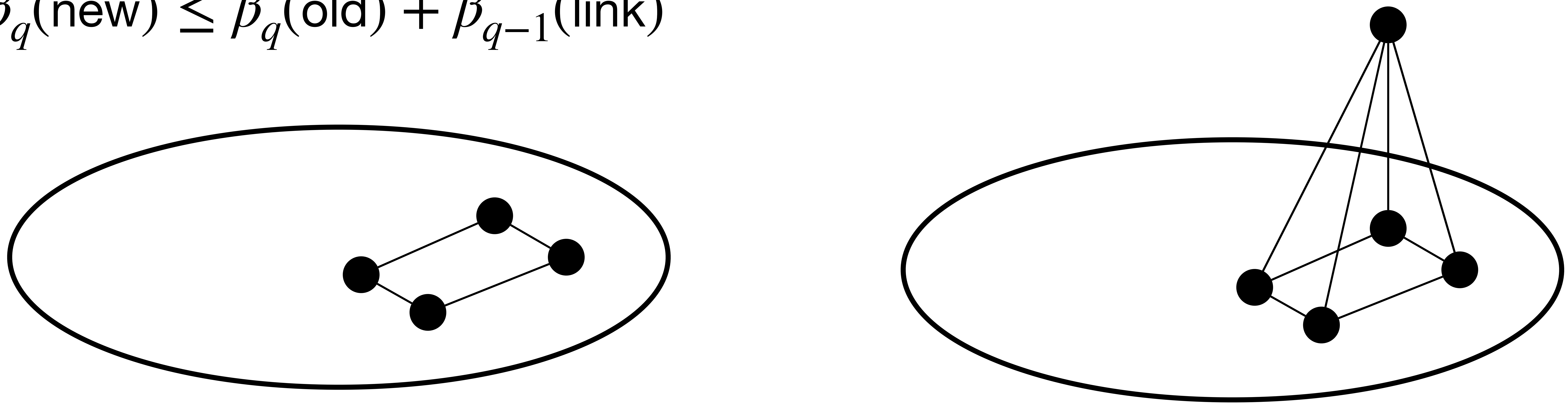
- Need homological algebra to relate Betti numbers with counts
  - adding a vertex = construct mapping cone





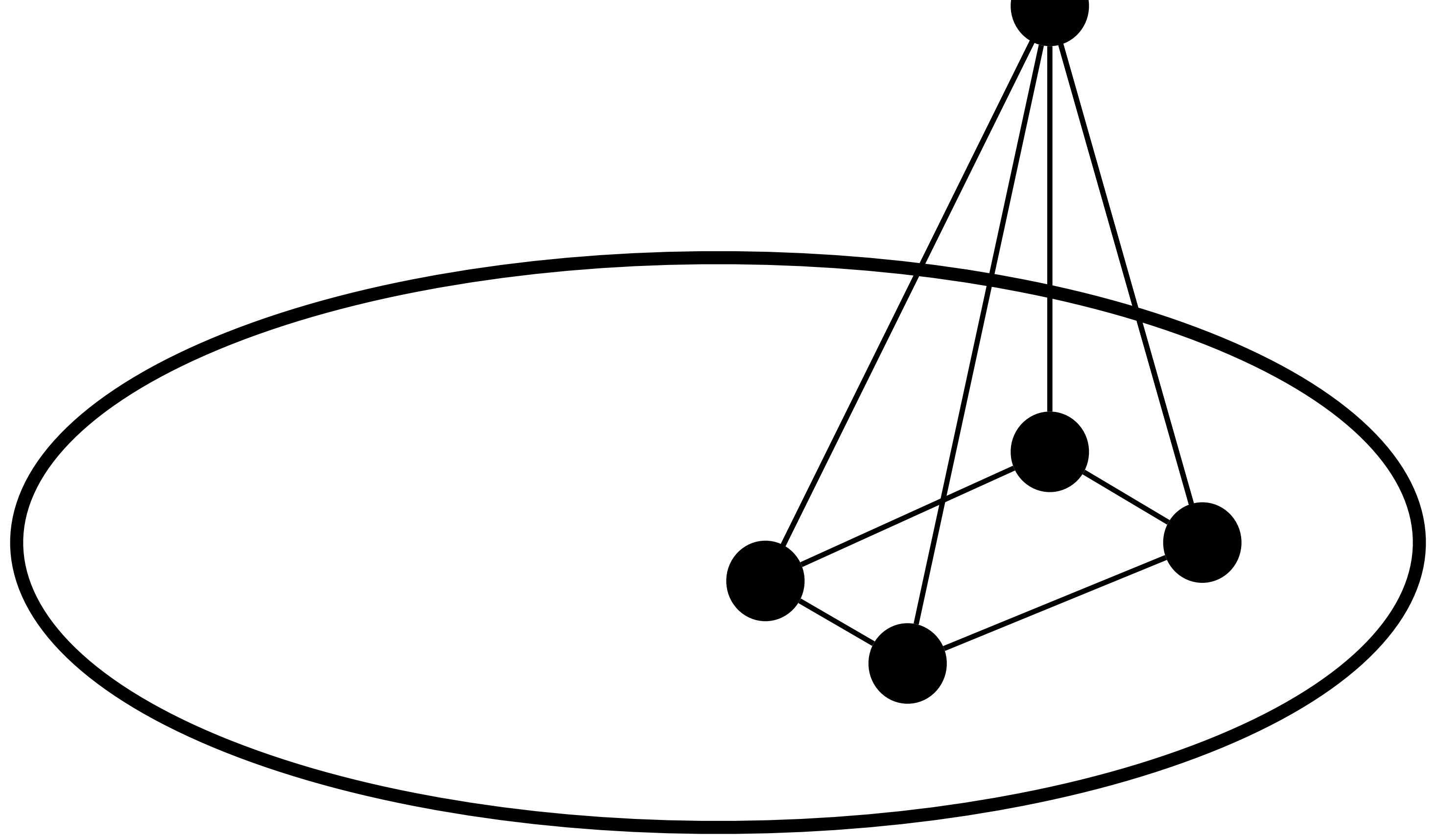
Subtleties

- Need homological algebra to relate Betti numbers with counts
  - adding a vertex = construct mapping cone
  - $\beta_q(\text{new}) \leq \beta_q(\text{old}) + \beta_{q-1}(\text{link})$



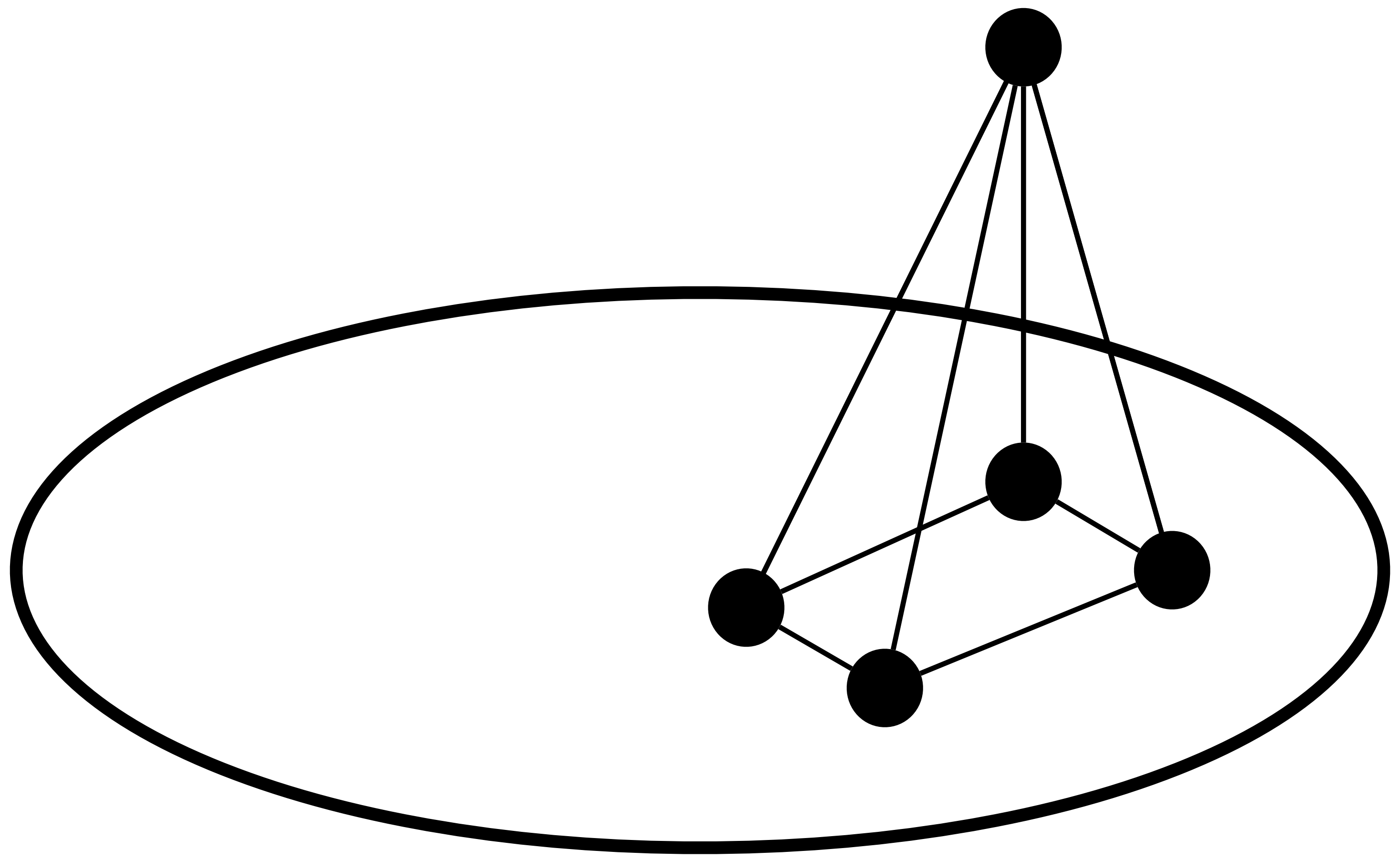
Subtleties

- Need homological algebra to relate Betti numbers with counts
  - $\beta_q(\text{new}) - \beta_q(\text{old}) \leq \beta_{q-1}(\text{link})$



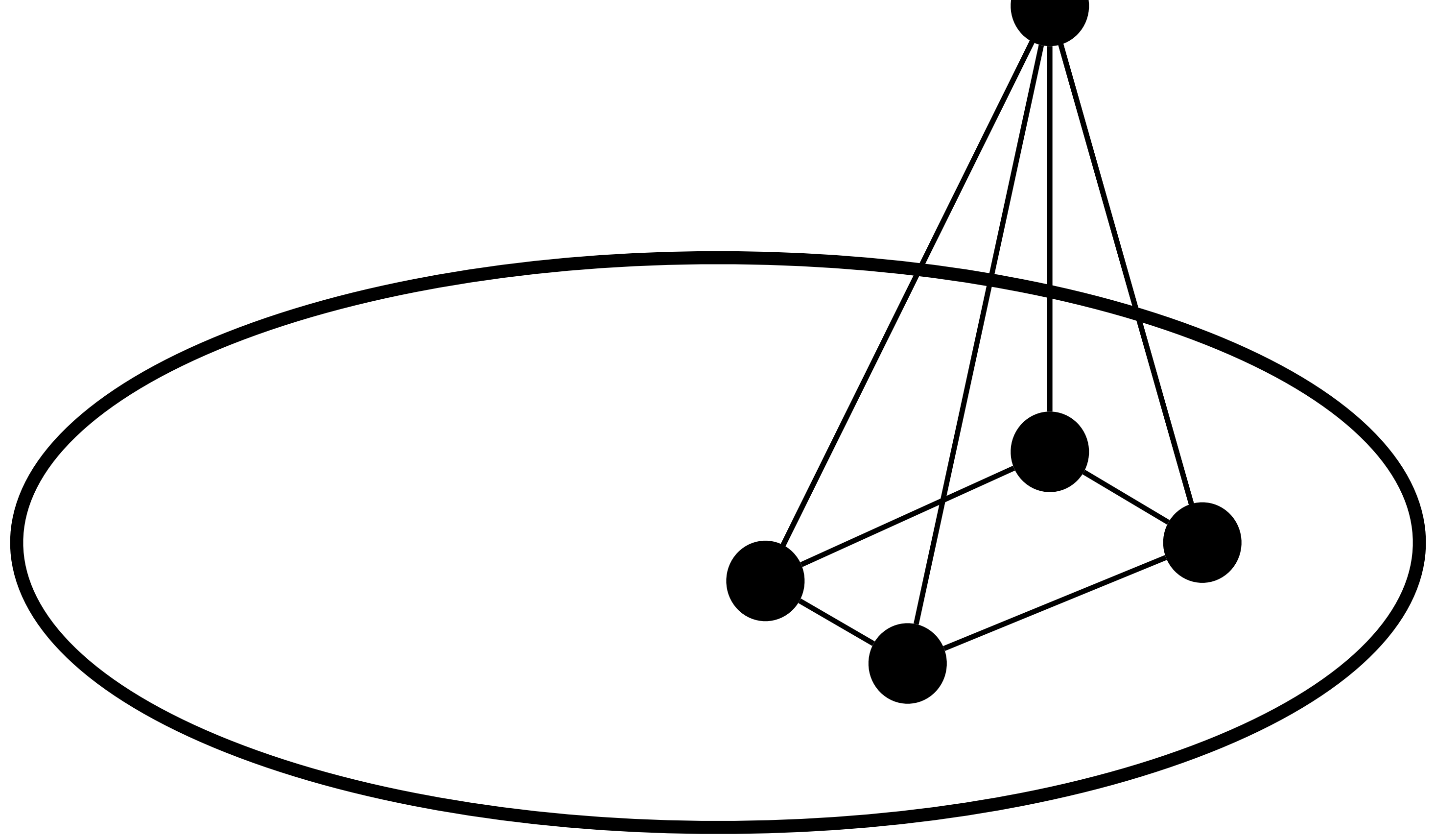
Subtleties

- Need homological algebra to relate Betti numbers with counts
  - $\beta_q(\text{new}) - \beta_q(\text{old}) \leq \beta_{q-1}(\text{link})$
- Identify the “square count” as the main term with minimal cycle results in [Gal 2005] and [Kahle 2009]



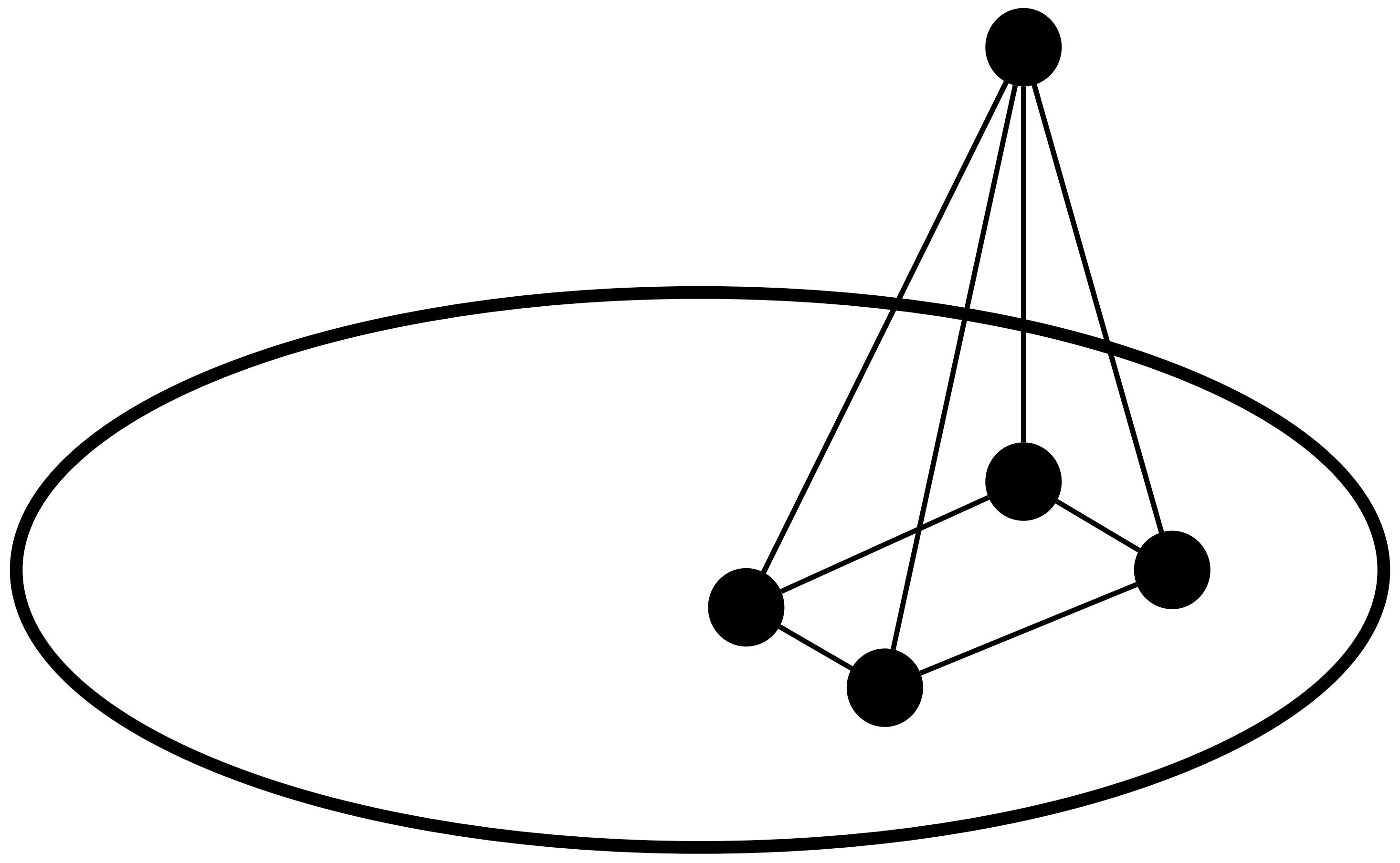
Subtleties

- Need homological algebra to relate Betti numbers with counts
  - $\beta_q(\text{new}) - \beta_q(\text{old}) \leq \beta_{q-1}(\text{link})$
- Identify the “square count” as the main term with minimal cycle results in [Gal 2005] and [Kahle 2009]
- Generalize minimal cycle results with homological algebra



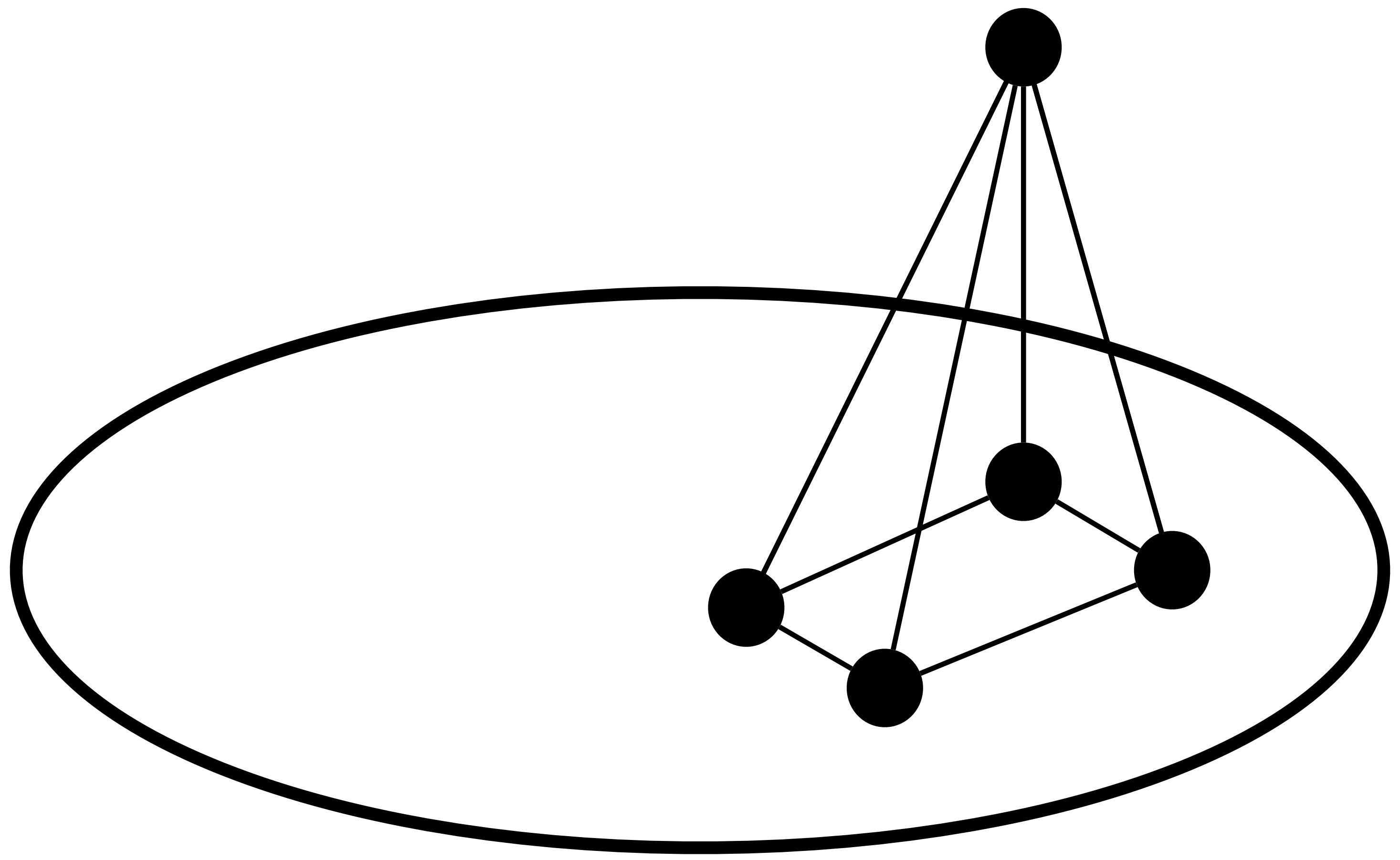
Subtleties

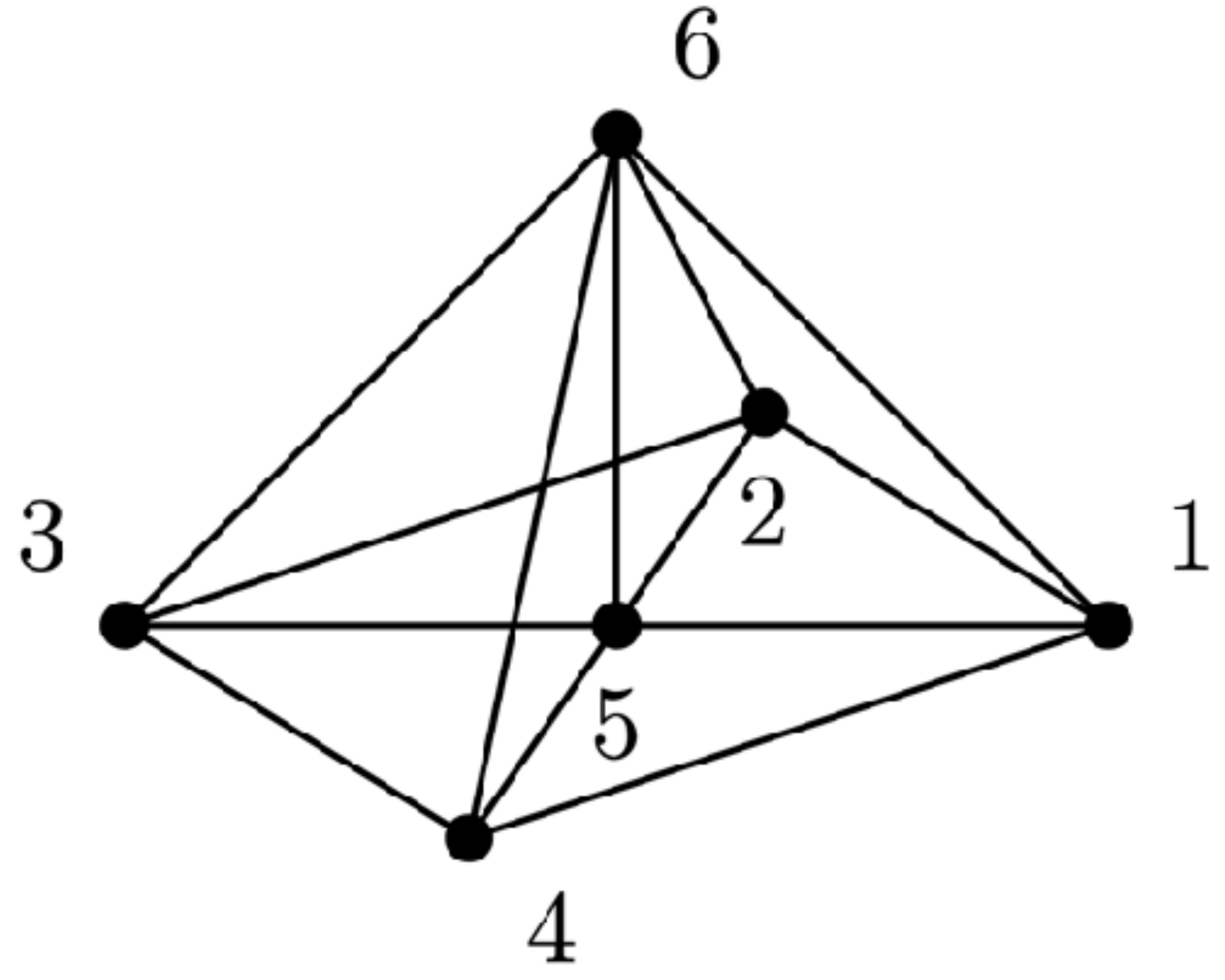
- Need homological algebra to relate Betti numbers with counts
  - $\beta_q(\text{new}) - \beta_q(\text{old}) \le \beta_{q-1}(\text{link})$
- Identify the "square count" as the main term with minimal cycle results in [Gal 2005] and [Kahle 2009]
- Generalize minimal cycle results with homological algebra
  - $1 - \beta_q(\text{link}, S^{q-1}) - \beta_q(\text{link}) \le \beta_q(\text{new}) - \beta_q(\text{old}) \le \beta_{q-1}(\text{link})$



Subtleties

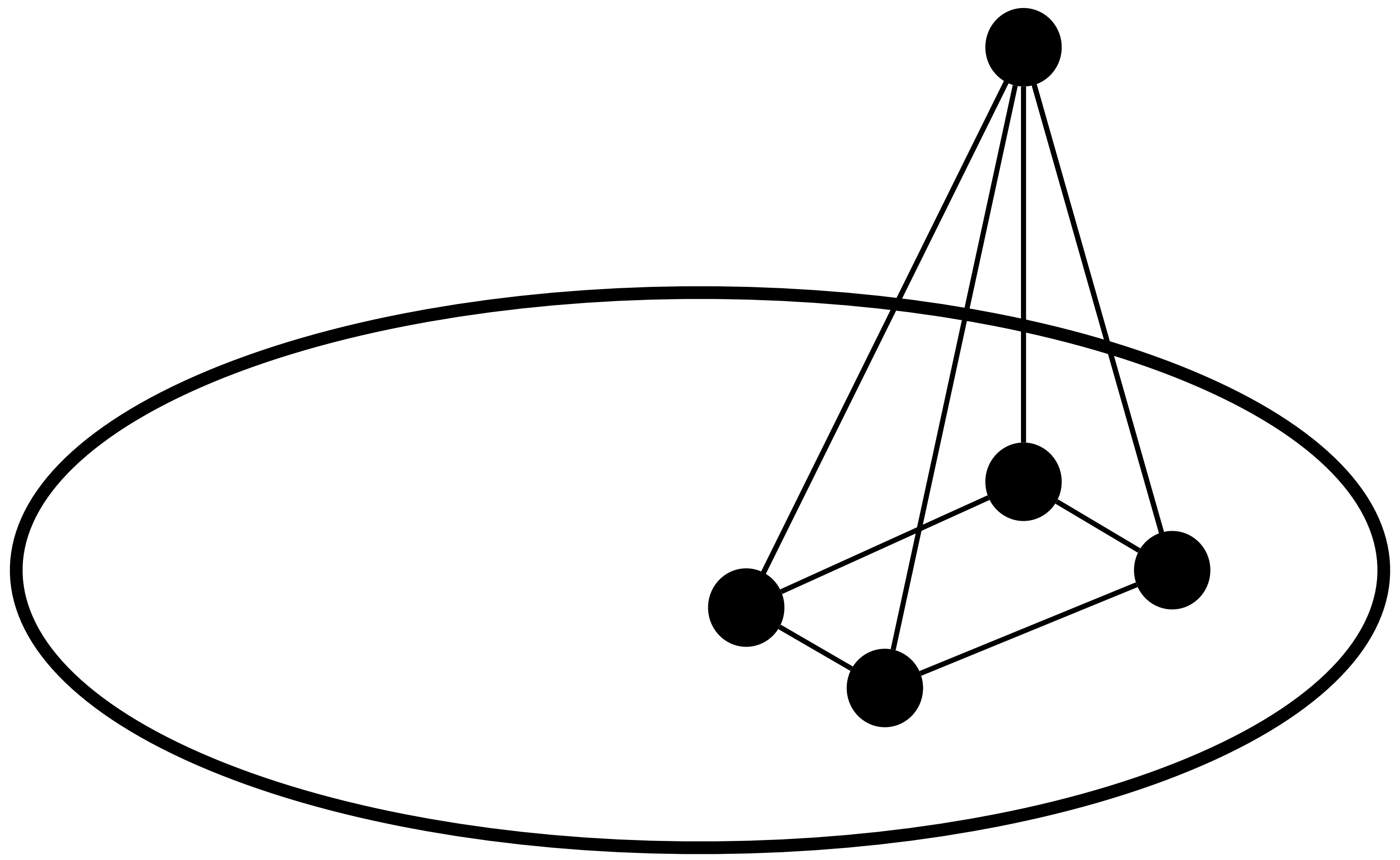
- Need homological algebra to relate Betti numbers with counts
  - $\beta_q(\text{new}) - \beta_q(\text{old}) \leq \beta_{q-1}(\text{link})$
- Identify the "square count" as the main term with minimal cycle results in [Gal 2005] and [Kahle 2009]
- Generalize minimal cycle results with homological algebra
  - $1 - \beta_q(\text{link}, S^{q-1}) - \beta_q(\text{link}) \leq \beta_q(\text{new}) - \beta_q(\text{old}) \leq \beta_{q-1}(\text{link})$

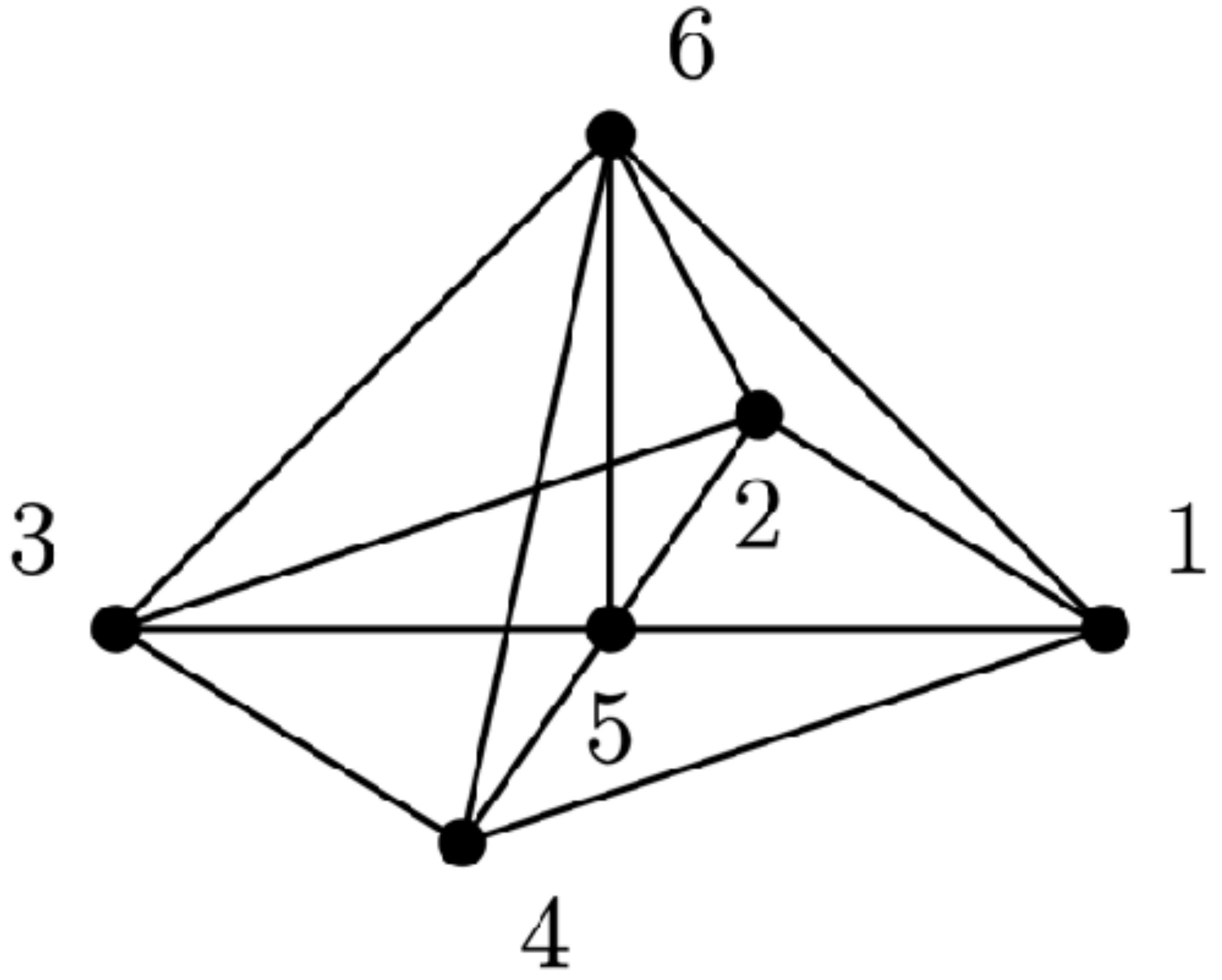




Subtleties

- Need homological algebra to relate Betti numbers with counts
  - $\beta_q(\text{new}) - \beta_q(\text{old}) \leq \beta_{q-1}(\text{link})$
- Identify the “square count” as the main term with minimal cycle results in [Gal 2005] and [Kahle 2009]
- Generalize minimal cycle results with homological algebra
  - $1 - \beta_q(\text{link}, S^{q-1}) - \beta_q(\text{link}) \leq \beta_q(\text{new}) - \beta_q(\text{old}) \leq \beta_{q-1}(\text{link})$
- Apply graph counting result in [Garavaglia and Stegehuis 2019] on a large class of subgraphs



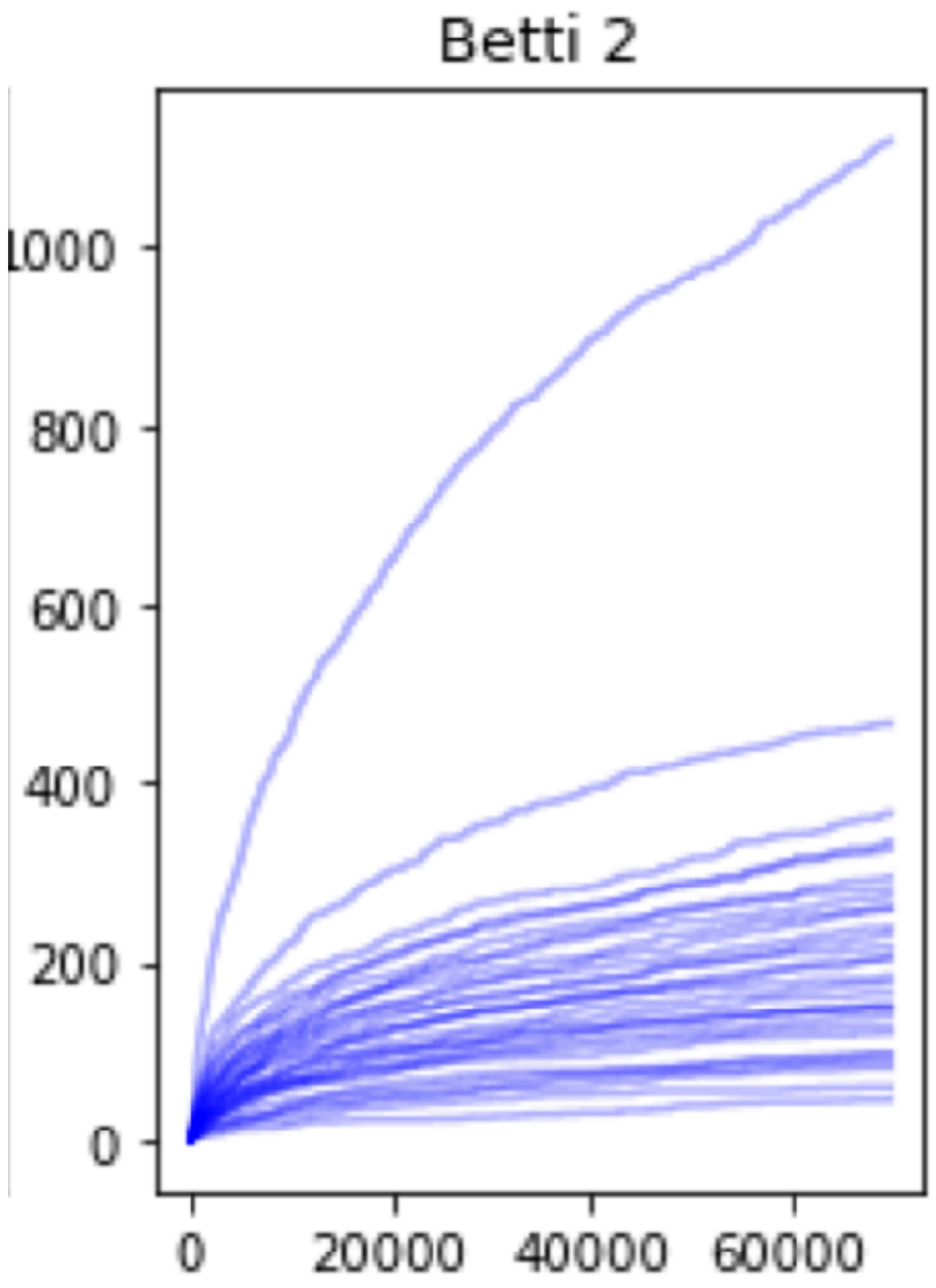


Theorem: $E[\beta_2] \approx$ num of nodes^{1-4x}

In practice???

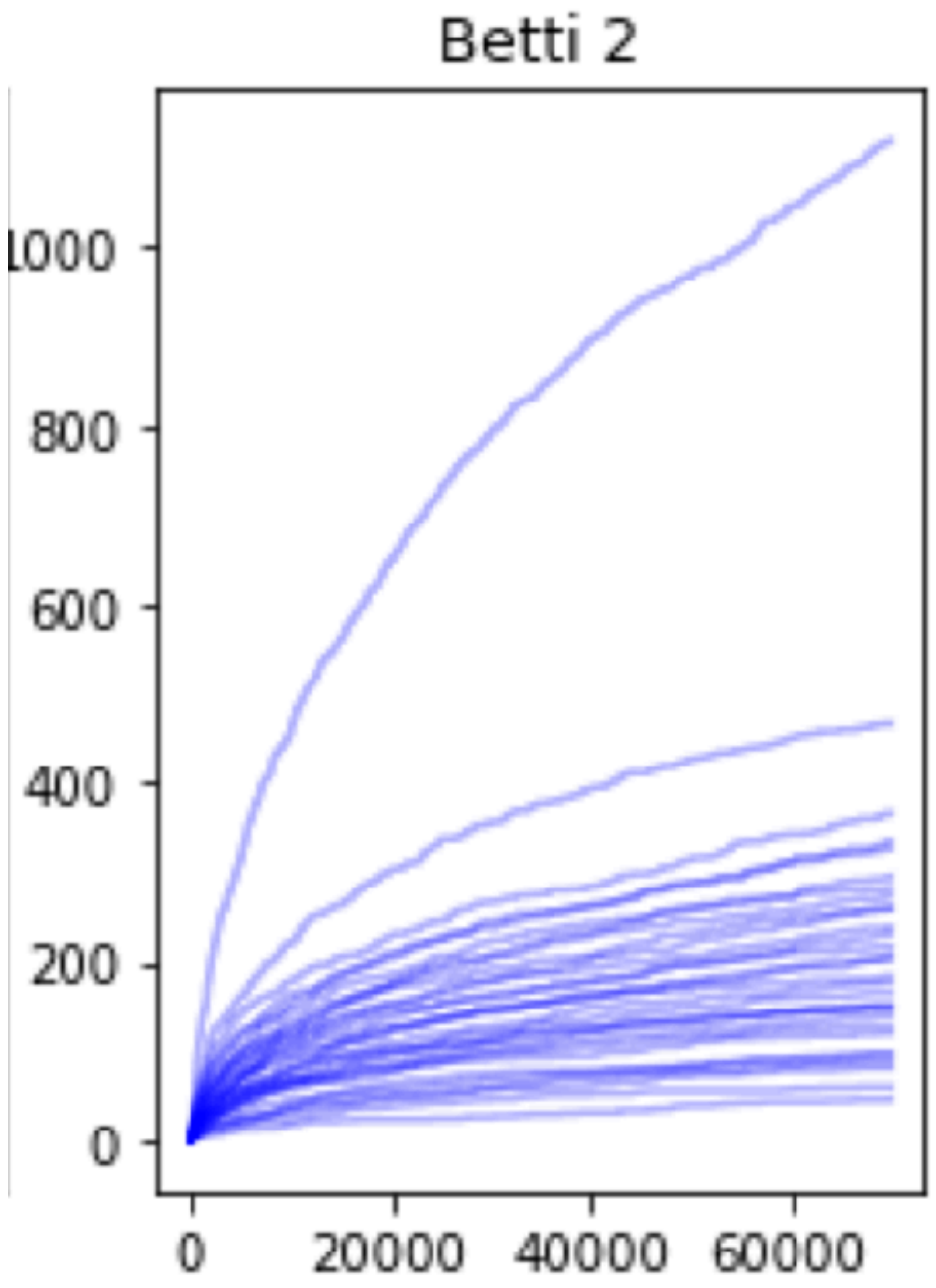
$$E[\beta_2] \approx \textbf{num of nodes}^{1-4x}$$

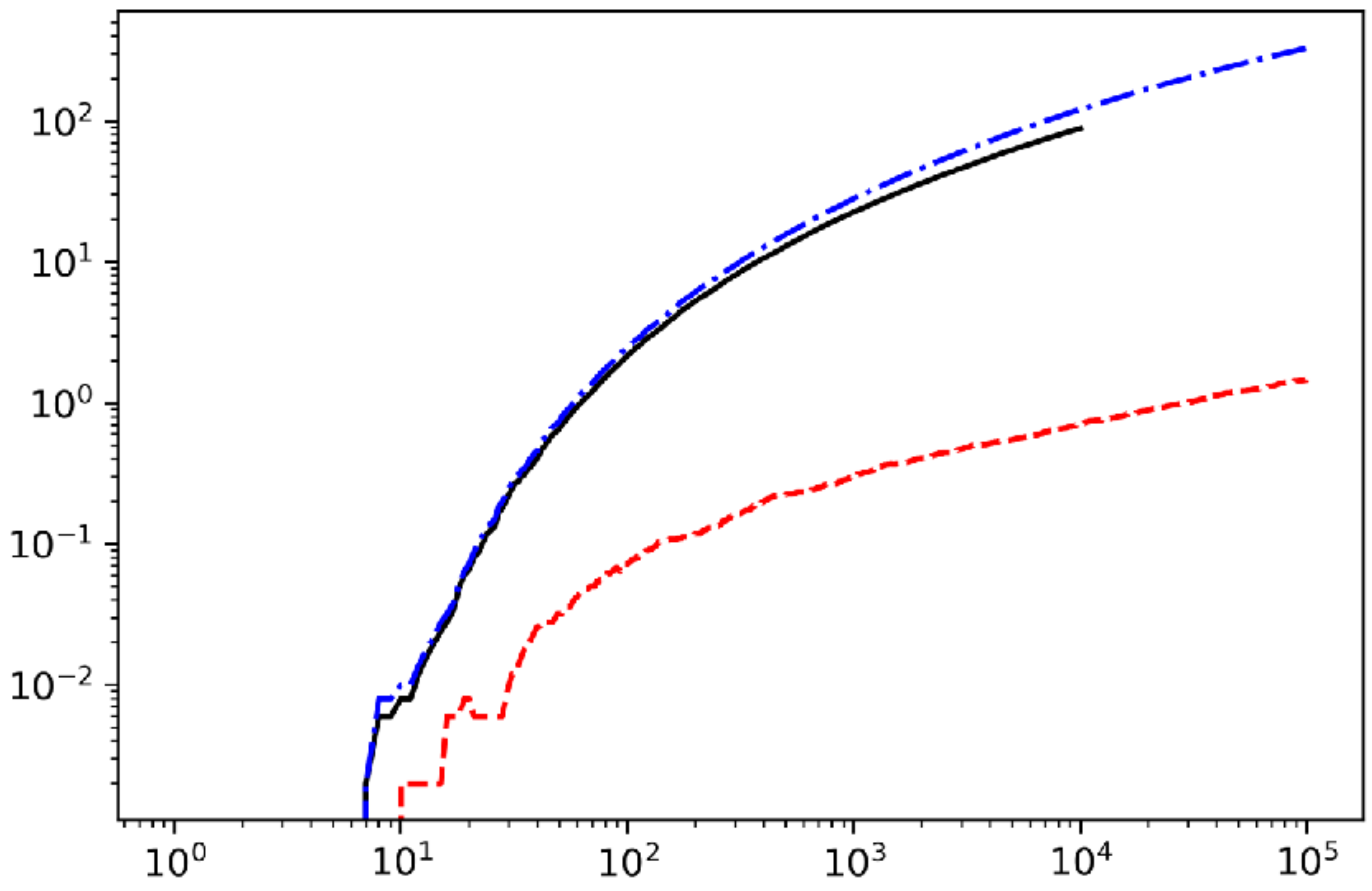
$$\log E[\beta_2] \approx (1-4x)\log(\textbf{num of nodes})$$



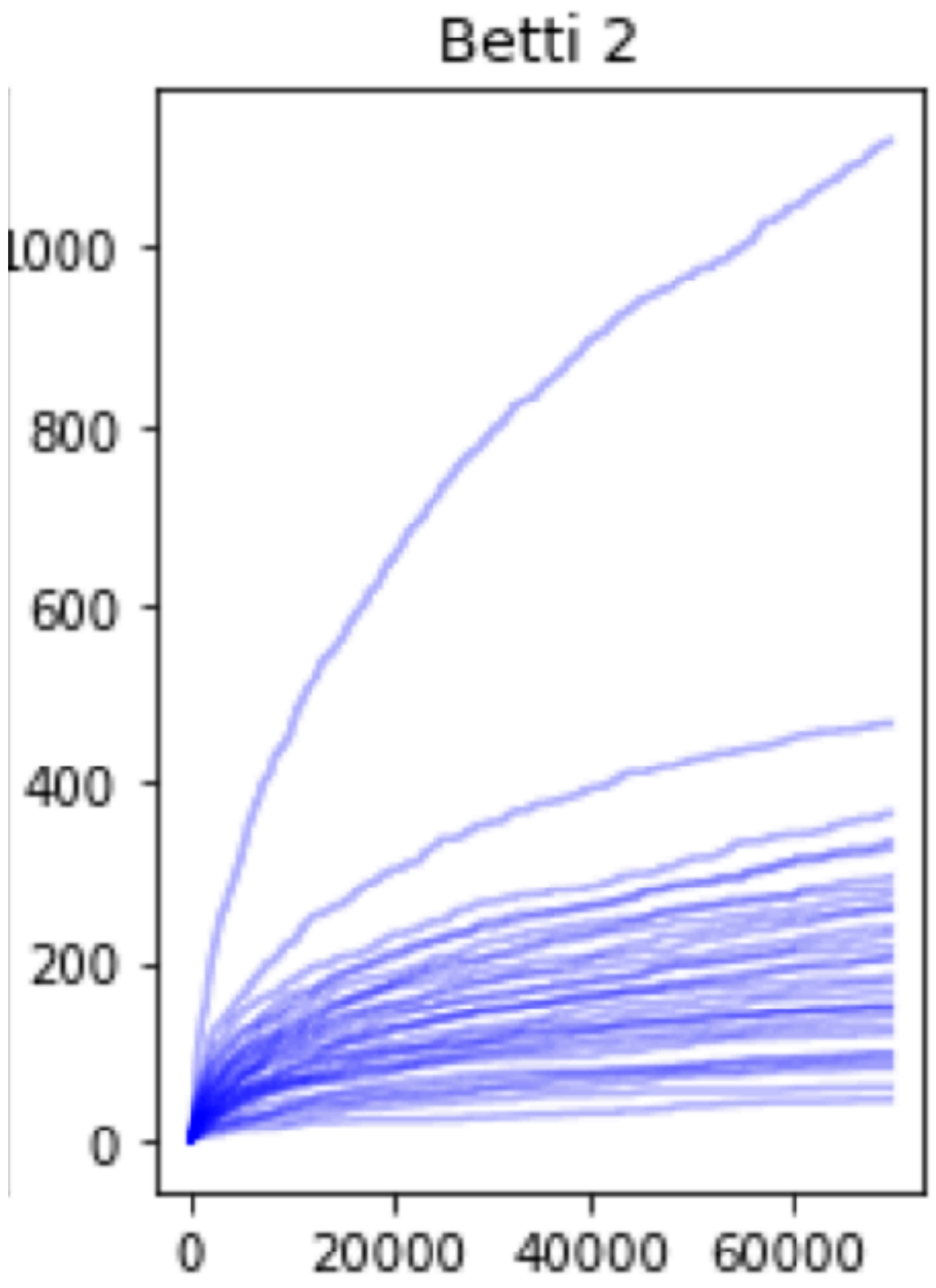
$$E[\beta_2] \approx \textbf{num of nodes}^{1-4x}$$

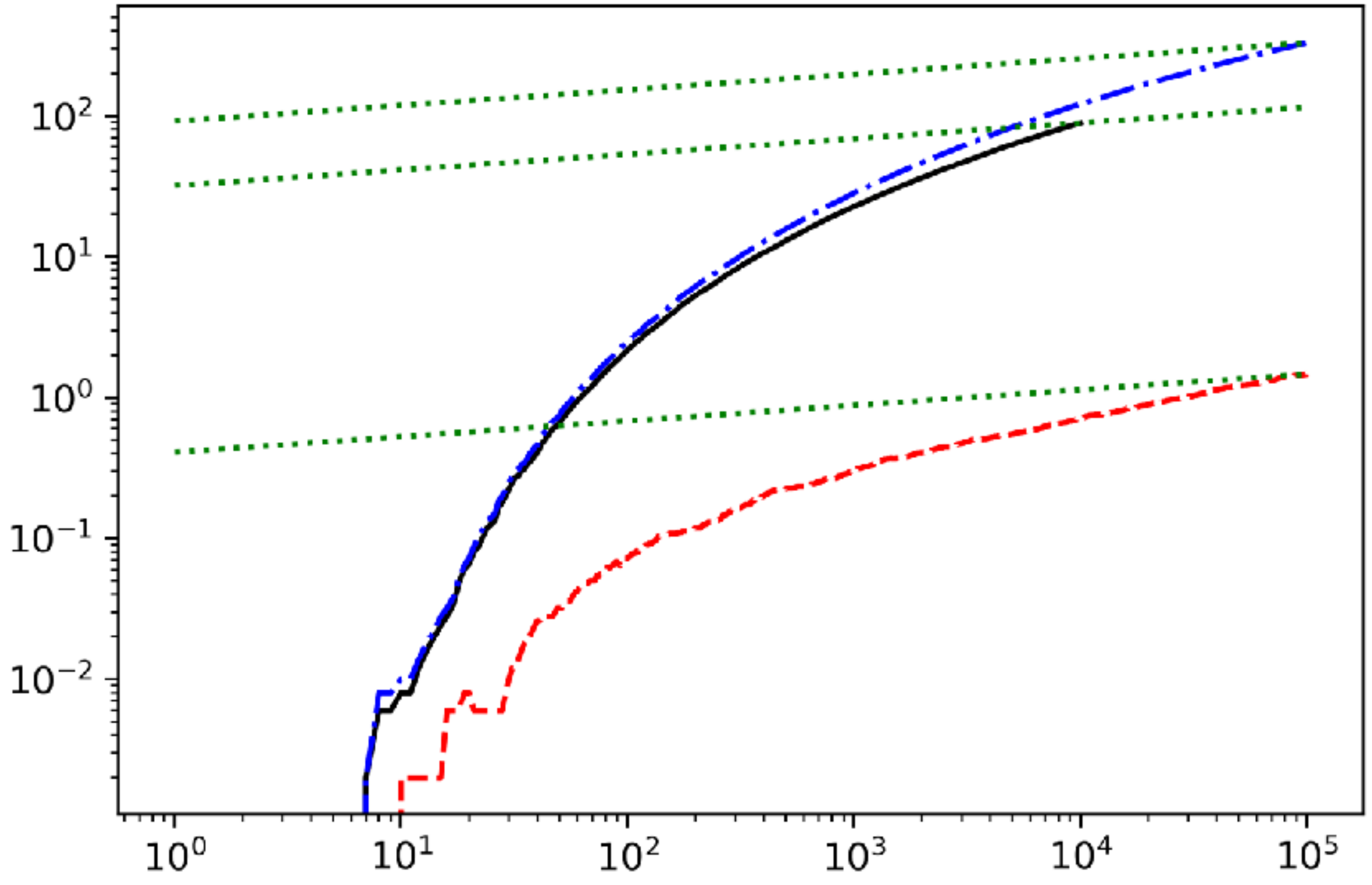
$$\log E[\beta_2] \approx (1-4x)\log(\textbf{num of nodes})$$





$E[\beta_2] \approx$ **num of nodes**$^{1-4x}$





V. What lies ahead

orders of magnitude of Betti numbers

homotopy connectedness

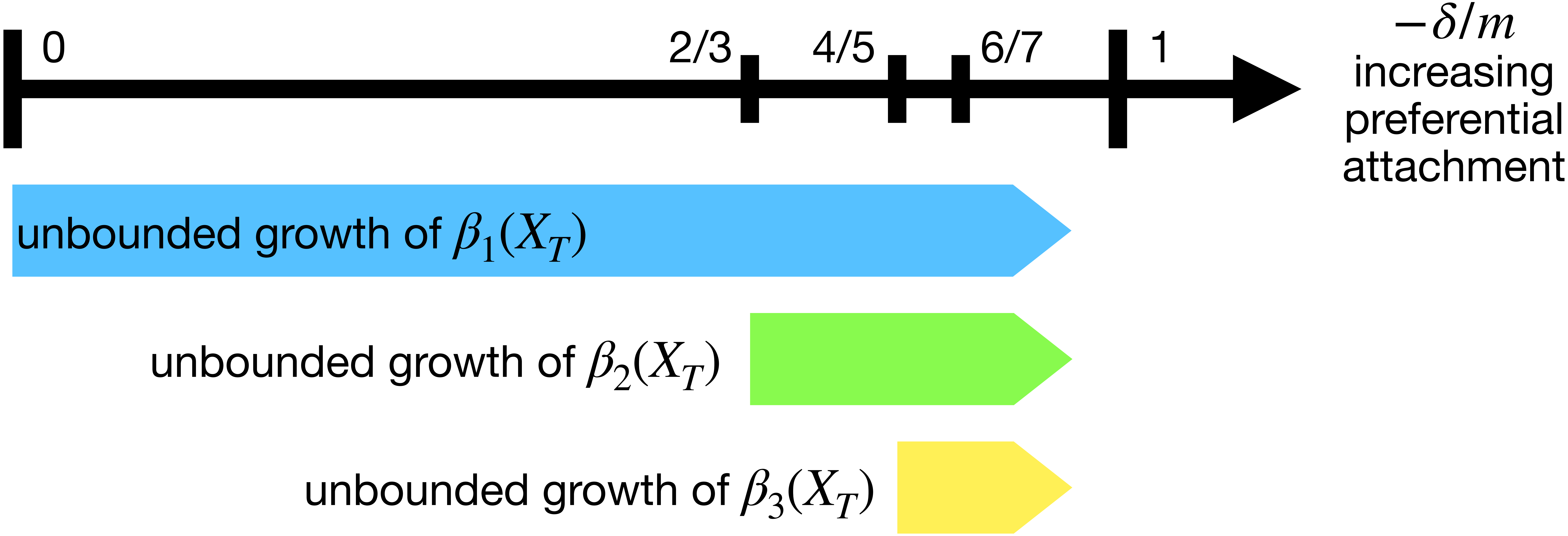
orders of magnitude of Betti numbers

homotopy connectedness

parameter estimation?

simplicial preferential attachment?

other non-homogeneous complexes?



0
2/3
4/5
6/7
1
$-\delta/m$
increasing preferential attachment
unbounded growth of $\beta_1(X_T)$
unbounded growth of $\beta_2(X_T)$
unbounded growth of $\beta_3(X_T)$