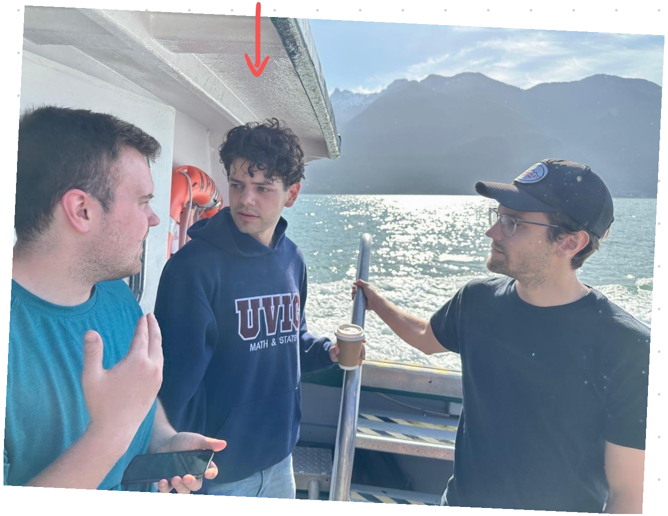
Heights and diameters of random trees and graphs

Based on joint work with

Gabriel Crudele



McGill

Serte Donderwinkel

university of groningen

The political and social roles of learned and professional societies

Louigi Addario-Berry

December $2^1 3^1$, $3^4 5^2$

McGill University

Keffiyehs

Keffiyehs

Keffiyehs

Should gay sex be punishable by prison?

"Gay people should no more be put in prison for having consensual sex than straight people should."

– David L Hull, "The social responsibility of professional societies".

Neutrality in academia

In 2022, “88 percent of professors at Canadian universities identified themselves as either ‘somewhat’ or ‘very’ left-leaning compared to only 12 percent who self-identify as right-leaning”.

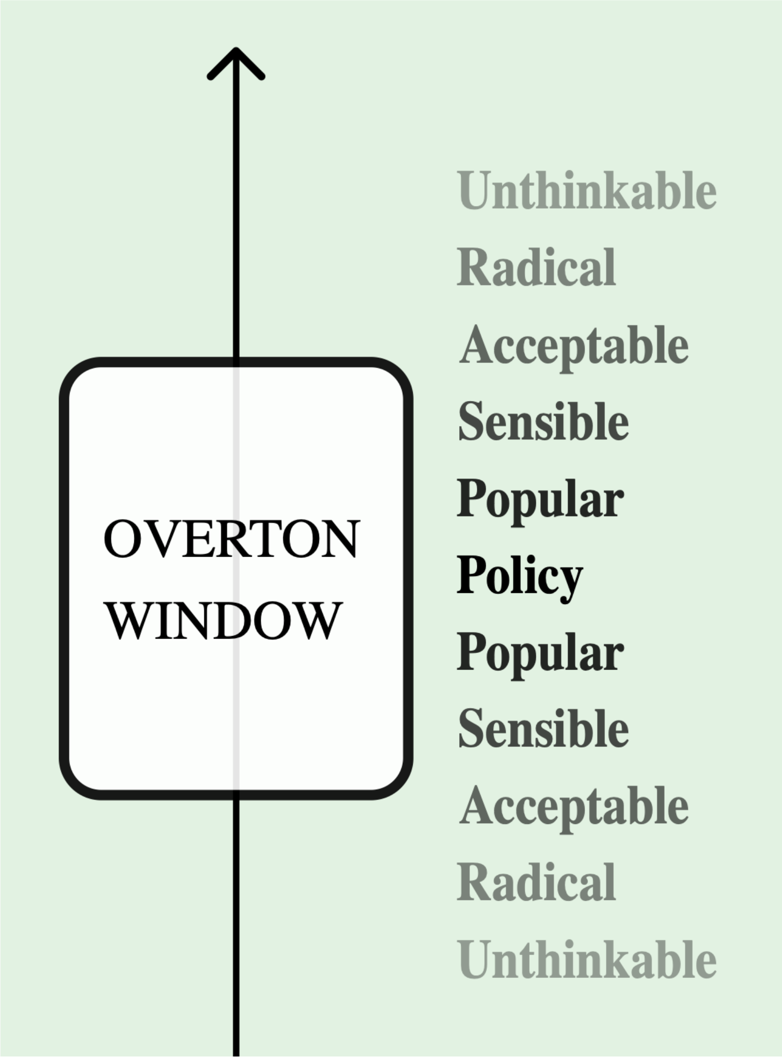
Political and ideological neutrality

"I recommend that [NSERC, SSHRC and CIHR's] enabling legislation be amended to enshrine a commitment to political and ideological neutrality.

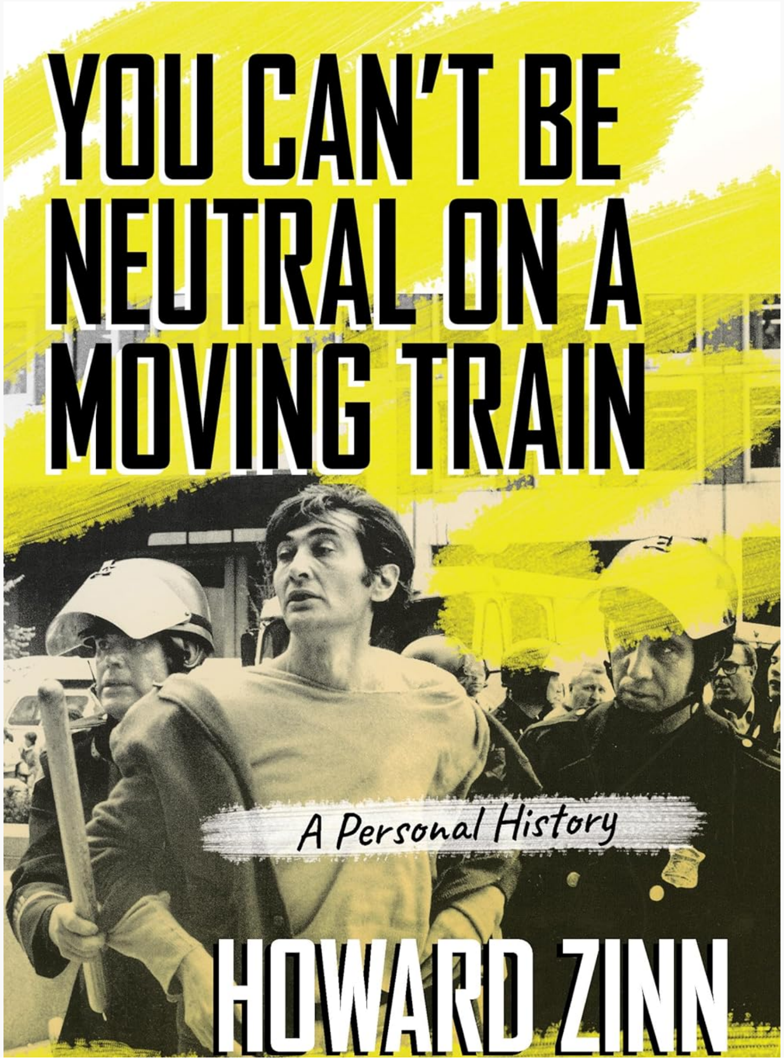
To ensure such neutrality, I also recommend that all references to EDI be removed from agency guidelines and criteria. This includes the elimination of many of SSHRC's EDI-focused grants."

- Dave Snow

The Overton window



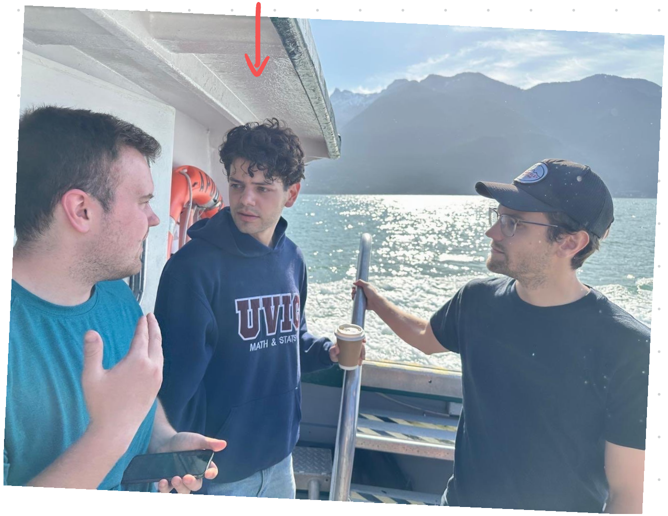
For whom is neutrality an option?



Heights and diameters of random trees and graphs

Based on joint work with

Gabriel Crudele



McGill

Serte Donderwinkel

university of groningen

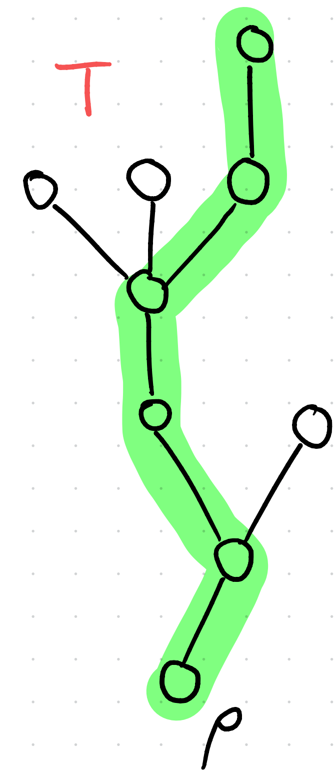
Trees

A **tree** is a connected acyclic graph

A **rooted tree** has a distinguished **root vertex**, ρ.

The **height** of rooted tree T is

$$\max(\mathrm{dist}(\rho, v) : v \text{ a vertex of } T)$$



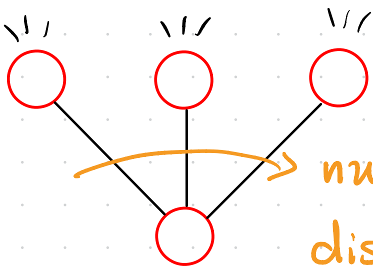
height(T) = 5

Bienaymé Trees*

* Family trees of branching processes

Offspring distributed according to a random variable X.

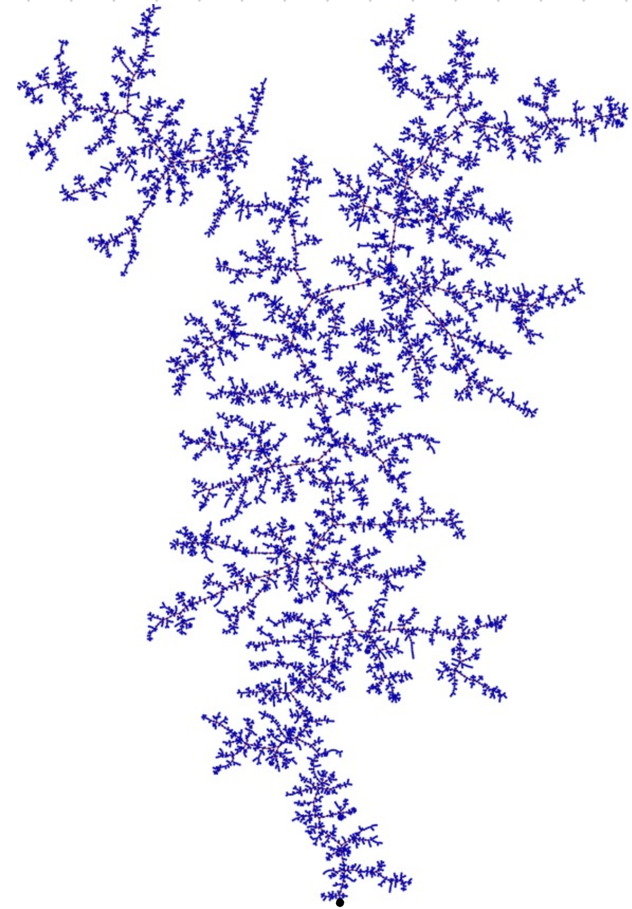
children reproduce independently.



number of children distributed as X.

This gives a random tree T

How does height(T) behave?



Bienaymé Trees*

Offspring distributed according to a random variable X.

This gives a random tree T

How does height (T) behave?

Trichotomy

- $\mathbb{E}X > 1$

  ↪ $\boxed{\mathbb{P}(ht(T) = \infty) > 0}$ (Classical)

Bienaymé Trees*

Offspring distributed according to a random variable X.

This gives a random tree T

How does height(T) behave?

Trichotomy

- $\mathbb{E}X > 1$

  $\hookrightarrow$ $\boxed{\mathbb{P}(ht(T) = \infty) > 0}$ (Classical)

- $\mathbb{E}X = \lambda < 1$, $\mathbb{E}(X^2) < \infty$

  $\hookrightarrow$ $\boxed{\mathbb{P}(ht(T) \geq h) \sim c \cdot \lambda^h}$ Kolmogorov (1938)

Bienaymé Trees*

Offspring distributed according to a random variable X.

This gives a random tree T

How does height(T) behave?

Trichotomy

- $\mathbb{E}X > 1$

  $\hookrightarrow$ $\mathbb{P}(ht(T) = \infty) > 0$ (Classical)

- $\mathbb{E}X = \lambda < 1$, $\mathbb{E}(X^2) < \infty$

  $\hookrightarrow$ $\mathbb{P}(ht(T) \geq h) \sim c \cdot \lambda^h$ Kolmogorov (1938)

- $\mathbb{E}X = 1$, $\mathbb{E}(X^2) < \infty$

  $\mathbb{P}(ht(T) \geq h) \sim \frac{2}{Var(X)} \cdot \frac{1}{h}$ Kolmogorov (1938)

  Kesten-Ney-Spitzer (1966)

Distances in random graphs

For a graph G, the diameter of G is

$$\mathrm{diam}(G) = \max\left(\mathrm{dist}(u,v) : u,v \text{ vertices of } G \text{ in the same connected component of } G\right)$$

Distances in random graphs

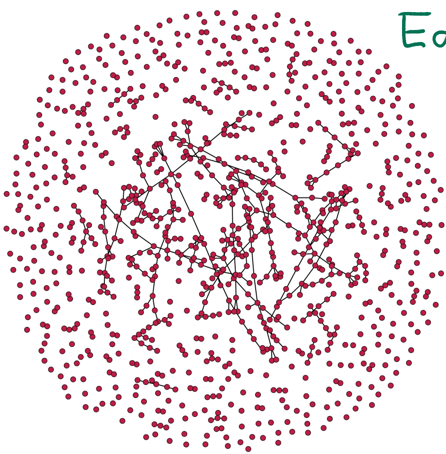
For a graph G, the diameter of G is

$$\operatorname{diam}(G) = \max(\operatorname{dist}(u,v) : u,v \text{ vertices of } G \text{ in the same connected component of } G)$$

Erdős-Rényi Random Graph

$G(n,p) \rightsquigarrow n$ vertices labeled $\{1,\dots,n\}$

Each edge independently present with probability p.



Distances in random graphs

For a graph G, the diameter of G is

$$\mathrm{diam}(G) = \max(\mathrm{dist}(u,v) : u,v \text{ vertices of } G \text{ in the same connected component of } G)$$

Erdős-Rényi Random Graph

$G(n,p) \rightsquigarrow n$ vertices labeled $\{1,\ldots,n\}$

Each edge independently present with probability p.

The diameter of $G(n,p)$ is essentially fully understood.

For a vertex v of G, $\mathcal{C}(v) =$ connected component of G containing v.

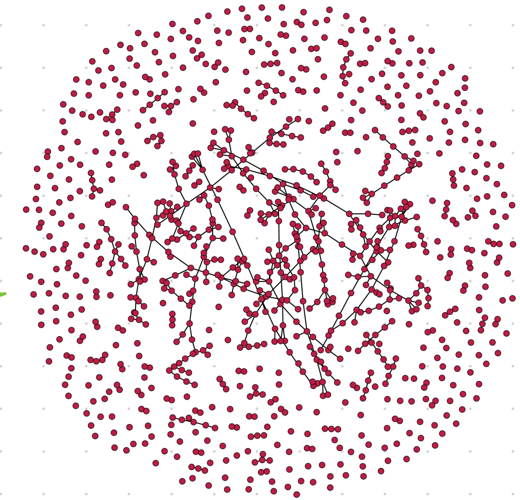
Diameter of $G(n,p)$

One notable range: $p = p(n) = \frac{\lambda}{n}$

$\lambda < 1$ diam$(C(1)) \asymp 1$ in expectation

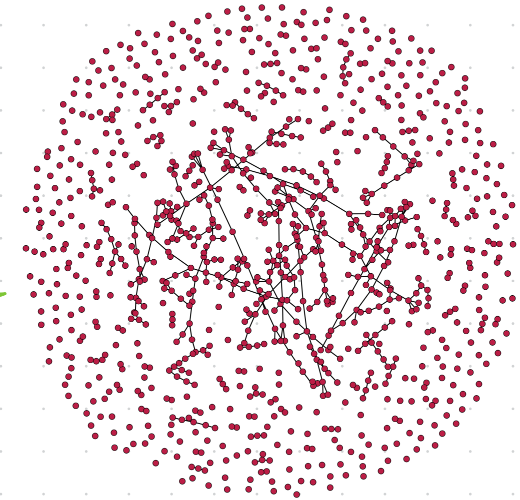
diam$(G(n,p)) \asymp \log n$

(Łuczak 1998)



Diameter of $G(n,p)$

One notable range: $p = p(n) = \frac{\lambda}{n}$



$\lambda < 1$ $\quad \mathrm{diam}(\mathcal{C}(1)) \asymp 1$ in expectation

$\mathrm{diam}(G(n,p)) \asymp \log n$

(Łuczak 1998)

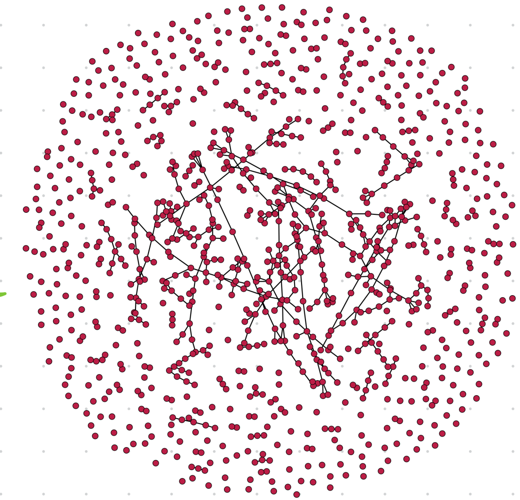
$\lambda > 1$ $\quad \mathrm{diam}(G(n,p)) \asymp \log n \asymp \mathrm{diam}(\mathcal{C}(1))$

(Chung-Lu 2001; Fernholz-Ramachandran 2007;

Riordan-Wormald 2010)

Diameter of $G(n,p)$

One notable range: $p = p(n) = \frac{\lambda}{n}$



$\lambda < 1$ $\quad \mathrm{diam}(\mathcal{C}(1)) \asymp 1$ in expectation

$\mathrm{diam}(G(n,p)) \asymp \log n$

(Łuczak 1998)

$\lambda > 1$ $\quad \mathrm{diam}(G(n,p)) \approx \log n \asymp \mathrm{diam}(\mathcal{C}(1))$

(Chung-Lu 2001; Fernholz-Ramachandran 2007;

Riordan-Wormald 2010)

$\lambda = 1$ $\quad \mathrm{diam}(\mathcal{C}(1)) \asymp 1$, $\mathrm{diam}(G(n,p)) \asymp n^{1/3}$

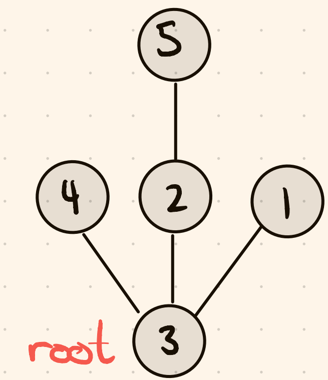
(AB-Broutin-Reed 2006; Nachmias-Peres 2007)

FIXED-DEGREE MODELS

Trees with given child sequences.

Say rooted tree T has child sequence $c = (c_1, \dots, c_n)$ if

- $V(T) = [n] := \{1, 2, \dots, n\}$
- Vertex i has c_i children for $i \in [n]$

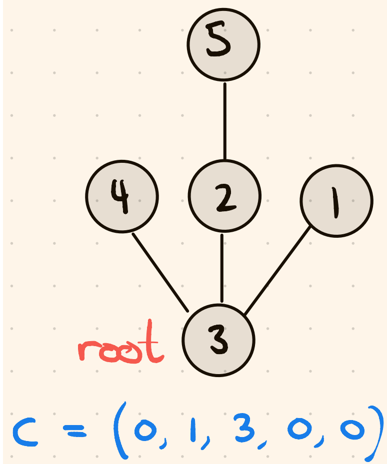


$c = (0, 1, 3, 0, 0)$

Trees with given child sequences.

Say rooted tree T has child sequence $c = (c_1, \ldots, c_n)$ if

- $V(T) = [n] := \{1, 2, \ldots, n\}$
- Vertex i has c_i children for $i \in [n]$



Define $\mathcal{T}_c = \{T : T$ is a rooted tree with child sequence $c\}$

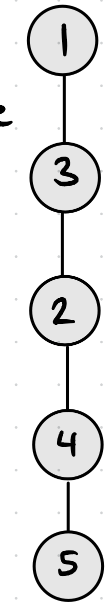
Let $T \in_u \mathcal{T}_c$.

be a uniformly random sample from

Question: How does the random variable $\text{height}(T)$ behave?

Heights of trees with given child sequences.

Let $T \in_u \mathcal{T}_c$. **Question:** How does the random variable $\mathrm{height}(T)$ behave?

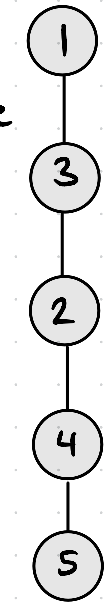


If $c = (0, \underbrace{1, \ldots, 1}_{n-1 \text{ times}})$ then T is a path, $\mathrm{height}(T) = n-1$.

Nodes with exactly one child "vertically stretch" trees in essentially deterministic ways.

Heights of trees with given child sequences.

Let $T \in_u \mathcal{T}_c$. **Question:** How does the random variable height(T) behave?



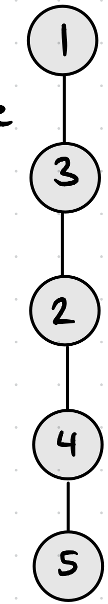
If $c = (0, \underbrace{1, \ldots, 1}_{n-1 \text{ times}})$ then T is a path, height$(T) = n-1$.

Nodes with exactly one child "vertically stretch" trees in essentially deterministic ways.

Theorem (A-B, Donderwinkel 2022). Fix $\varepsilon > 0$, fix a child sequence $c = (c_1, \ldots, c_n)$ s.t. a) $\mathcal{T}_c \neq \emptyset$; b) $n_1 := \#\{i \in [n] : c_i = 1\} \leq (1-\varepsilon)n$.
Let $T \in_u \mathcal{T}_c$. Then $\mathbb{E}\,\mathrm{diam}(T) = O_\varepsilon(\sqrt{n})$.

Heights of trees with given child sequences.

Let $T \in_u \mathcal{T}_c$. **Question:** How does the random variable height(T) behave?



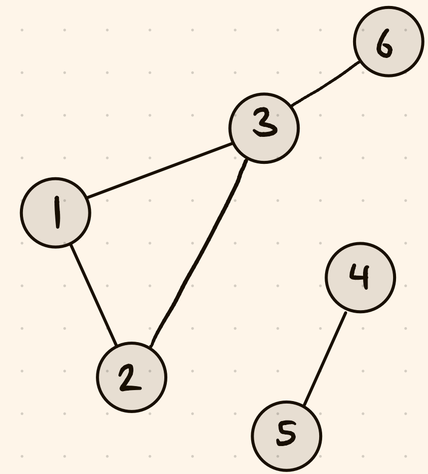
If $c = (0, \underbrace{1, \ldots, 1}_{n-1 \text{ times}})$ then T is a path, height$(T) = n-1$.

Nodes with exactly one child "vertically stretch" trees in essentially deterministic ways.

Theorem (A-B, Donderwinkel 2022). Fix $\varepsilon > 0$, fix a child sequence $c = (c_1, \ldots, c_n)$ s.t. a) $\mathcal{T}_c \neq \emptyset$; b) $n_1 := \#\{i \in [n] : c_i = 1\} \leq (1-\varepsilon)n$. Let $T \in_u \mathcal{T}_c$. Then $\mathbb{E}\,\mathrm{diam}(T) = O_\varepsilon(\sqrt{n})$. We'll see this bound is tight.

Graphs with given degree sequences.

Say graph G has degree sequence $d=(d_1,\dots,d_n)$ if $V(G)=[n]$ and $\deg_G(i)=d_i$ for $i\in[n]$.



$d=(2,2,3,1,1,1)$

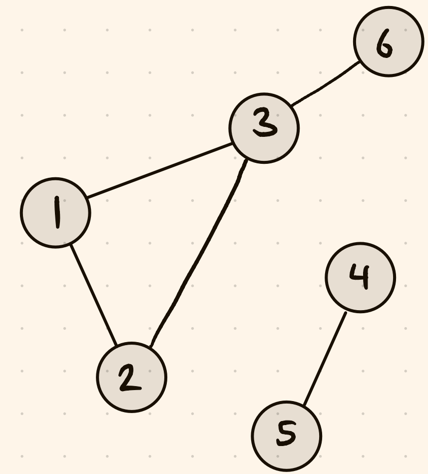
Graphs with given degree sequences.

Say graph G has degree sequence $d = (d_1, \dots, d_n)$ if $V(G) = [n]$ and $\deg_G(i) = d_i$ for $i \in [n]$.

Let $\mathcal{G}_d = \{\text{Graphs } G : G \text{ has degree sequence } d\}$

Let $\mathcal{C}_d = \{G \in \mathcal{G}_d : G \text{ is connected}\}$.

Let $G \in_u \mathcal{G}_d$ or let $G \in_u \mathcal{C}_d$.



$d = (2, 2, 3, 1, 1, 1)$

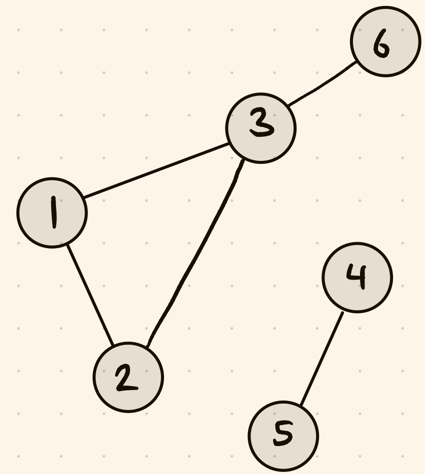
Graphs with given degree sequences.

Say graph G has degree sequence $d=(d_1,\dots,d_n)$ if $V(G)=[n]$ and $\deg_G(i)=d_i$ for $i\in[n]$.

Let $\mathcal{G}_d=\{\text{Graphs } G : G \text{ has degree sequence } d\}$

Let $\mathcal{C}_d=\{G\in\mathcal{G}_d : G \text{ is connected}\}$.

Let $G\in_u\mathcal{G}_d$ or let $G\in_u\mathcal{C}_d$.



$d=(2,2,3,1,1,1)$

Question: How does the random variable $\operatorname{diam}(G)$ behave?

Graphs with given degree sequences.

Say graph G has degree sequence $d=(d_1,\dots,d_n)$ if $V(G)=[n]$ and $\deg_G(i)=d_i$ for $i\in[n]$.

Let $\mathcal{G}_d=\{\text{Graphs } G : G \text{ has degree sequence } d\}$

Let $\mathcal{C}_d=\{G\in\mathcal{G}_d : G \text{ is connected}\}$.

Let $G\in_u \mathcal{G}_d$ or let $G\in_u \mathcal{C}_d$.

Question: How does the random variable $\mathrm{diam}(G)$ behave?

If $d_1+\cdots+d_{n-1}=2(n-1)$ then the elements of $\mathcal{C}_d$ are the trees with deg. seq. $(d_1,\dots,d_n)$

Diameters of graphs with given degree sequences.

Let $G \in_u \mathcal{G}_d$ or $G \in_u \mathcal{C}_d$. How does the random variable $\mathrm{diam}(G)$ behave?

If $d = (\underbrace{2, 2, \ldots, 2}_{n \text{ times}})$ then

$G \in_u \mathcal{G}_d$ has $\mathbb{E}[\mathrm{diam}(G)] \asymp n$. (G looks like the graph of a random permutation.)

$G \in_u \mathcal{C}_d$ has $\mathrm{diam}(G) = \lfloor n/2 \rfloor$ (G is a cycle of length n)

Diameters of graphs with given degree sequences.

Let $G \in_u \mathcal{G}_d$ or $G \in_u \mathcal{C}_d$. How does the random variable diam(G) behave?

If $d = (\underbrace{2, 2, \ldots, 2}_{n \text{ times}})$ then

$G \in_u \mathcal{G}_d$ has $\mathbb{E}[\mathrm{diam}(G)] \asymp n$. ($G$ looks like the graph of a random permutation.)

$G \in_u \mathcal{C}_d$ has $\mathrm{diam}(G) = \lfloor n/2 \rfloor$ (G is a cycle of length n)

Theorem (A-B, Crudele 2025). Fix $\varepsilon > 0$, fix a degree sequence $d = (d_1, \ldots, d_n)$ s.t. a) $\mathcal{G}_d \neq \emptyset$; b) $n_2 := \#\{i \in [n] : d_i = 2\} \leq (1-\varepsilon)n$.

Let $G \in_u \mathcal{C}_d$ or $G \in_u \mathcal{G}_d$. Then $\mathbb{E}\,\mathrm{diam}(G) = O_\varepsilon(\sqrt{n})$.

Diameters of graphs with given degree sequences.

Let $G \in_u \mathcal{G}_d$ or $G \in_u \mathcal{C}_d$. How does the random variable $\mathrm{diam}(G)$ behave?

If $d = (\underbrace{2, 2, \ldots, 2}_{n \text{ times}})$ then

$G \in_u \mathcal{G}_d$ has $\mathbb{E}[\mathrm{diam}(G)] \asymp n$. (G looks like the graph of a random permutation.)

$G \in_u \mathcal{C}_d$ has $\mathrm{diam}(G) = \lfloor n/2 \rfloor$ (G is a cycle of length n)

Theorem (A-B, Crudele 2025). Fix $\varepsilon > 0$, fix a degree sequence $d = (d_1, \ldots, d_n)$ s.t. a) $\mathcal{G}_d \neq \emptyset$; b) $n_2 := \#\{i \in [n] : d_i = 2\} \leq (1-\varepsilon)n$.

Let $G \in_u \mathcal{C}_d$. Then $\mathbb{E}\,\mathrm{diam}(G) = O_\varepsilon(\sqrt{n})$. This bound is also tight

Minimum degree 3

Theorem (A-B, Donderwinkel 2022). Fix $\varepsilon > 0$, fix a child sequence $c = (c_1, \dots, c_n)$ s.t. a) $\mathcal{T}_c \neq 0$; b) $n_1 := \#\{i \in [n] : c_i = 1\} \leq (1-\varepsilon)n$. Let $T \in_u \mathcal{T}_c$. Then $\mathbb{E}\,\mathrm{diam}(T) = O_\varepsilon(\sqrt{n})$.

Theorem (A-B, Crudele 2025). Fix $\varepsilon > 0$, fix a degree sequence $d = (d_1, \dots, d_n)$ s.t. a) $\mathcal{G}_d \neq \phi$; b) $n_2 := \#\{i \in [n] : d_i = 2\} \leq (1-\varepsilon)n$. Let $G \in_u \mathcal{G}_d$. Then $\mathbb{E}\,\mathrm{diam}(G) = O_\varepsilon(\sqrt{n})$.

Minimum degree 3

Theorem (A-B, Donderwinkel 2022). Fix $\varepsilon > 0$, fix a child sequence $c = (c_1, \dots, c_n)$ s.t. a) $\mathcal{T}_c \neq \emptyset$; b) $n_1 := \#\{i \in [n] : c_i = 1\} \leq (1-\varepsilon)n$. Let $T \in_u \mathcal{T}_c$. Then $\mathbb{E}\,\mathrm{diam}(T) = O_\varepsilon(\sqrt{n})$.

Theorem (A-B, Crudele 2025). Fix $\varepsilon > 0$, fix a degree sequence $d = (d_1, \dots, d_n)$ s.t. a) $\mathcal{G}_d \neq \emptyset$; b) $n_2 := \#\{i \in [n] : d_i = 2\} \leq (1-\varepsilon)n$. Let $G \in_u \mathcal{G}_d$. Then $\mathbb{E}\,\mathrm{diam}(G) = O_\varepsilon(\sqrt{n})$.

Theorem (A-B, Crudele 2025) Fix a degree sequence $d = (d_1, \dots, d_n)$ s.t. a) $\mathcal{G}_d \neq \emptyset$; b) $d_i \geq 3$ for all i. Let $G \in_u \mathcal{G}_d$. Then $\mathbb{P}(G \text{ connected}) = 1 - O(\frac{1}{n})$; $\mathbb{P}(\mathrm{diam}(G) > 62 \log n) = O(\frac{1}{n})$

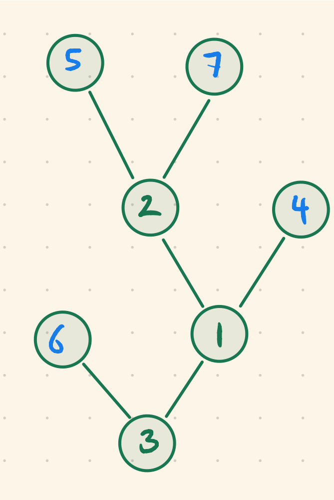
SOME PROOF IDEAS

Sampling uniform binary trees

Let $c = (\underbrace{2, 2, \ldots, 2}_{k \text{ times}}, \underbrace{0, \ldots, 0}_{k+1 \text{ times}})$.

Then $T = \mathcal{T}_c = \{$rooted binary trees with internal vertices $1, \ldots, k$,
leaves $k+1, \ldots, 2k+1\}$

To sample $T \in_u \mathcal{T}_c$:



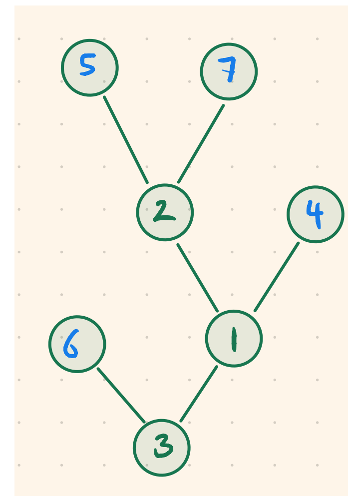
Sampling uniform binary trees

Let $c = (\underbrace{2, 2, \ldots, 2}_{k \text{ times}}, \underbrace{0, \ldots, 0}_{k+1 \text{ times}})$.

Then $T = T_c$ = rooted binary trees with internal vertices $1, \ldots, k$, leaves $k+1, \ldots, 2k+1$ }

To sample $T \in_u T_c$:

- Choose a sequence $S = (s_1, \ldots, s_{2k})$ with 2 copies of each integer from $1, \ldots, k$, uniformly at random.



Sampling uniform binary trees

Let $c = (\underbrace{2, 2, \ldots, 2}_{k \text{ times}}, \underbrace{0, \ldots, 0}_{k+1 \text{ times}})$.

Then $T = T_c =$ rooted binary trees with internal vertices $1, \ldots, k$, leaves $k+1, \ldots, 2k+1 \}$

To sample $T \in_u T_c$:

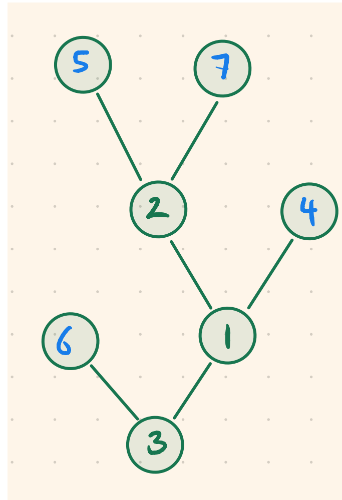
- Choose a sequence $S = (s_1, \ldots, s_{2k})$ with 2 copies of each integer from $1, \ldots, k$, uniformly at random.
- Let $i(1) < i(2) < \cdots < i(k)$ be the locations of the repeated entries (so $2 \le i(1) \le k+1$ and $i(k) = 2k$)
- Build a tree T with branches $s_1, s_2, \ldots, s_{i(1)-1}, k+1$

  $s_{i(1)}, \ldots, s_{i(2)-1}, k+2$

  $\vdots$

  $s_{i(k-1)}, \ldots, s_{i(k)-1}, 2k$,

  and $s_{2k}, 2k+1$



Sampling uniform binary trees

Let $c = (\underbrace{2,2,\dots,2}_{k \text{ times}}, \underbrace{0,\dots,0}_{k+1 \text{ times}})$.

Then $T = T_c$ = rooted binary trees with internal vertices $1,\dots,k$,
leaves $k+1,\dots,2k+1\}$

To sample $T \in_u T_c$:

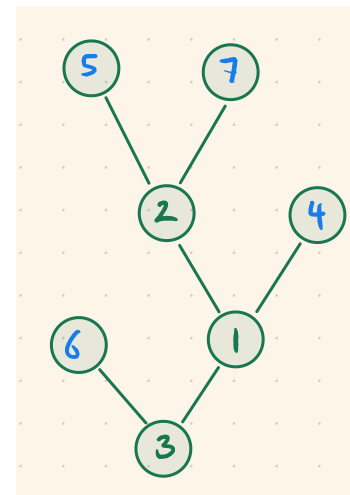
- Choose a sequence $S = (s_1,\dots,s_{2k})$ with 2 copies of each integer from $1,\dots,k$, uniformly at random.
- Let $i(1) < i(2) < \dots < i(k)$ be the locations of the repeated entries (so $2 \leq i(1) \leq k+1$ and $i(k) = 2k$)
- Build a tree T with branches $s_1, s_2, \dots, s_{i(1)-1}, k+1$

  $s_{i(1)}, \dots, s_{i(2)-1}, k+2$

  $\vdots$

  $s_{i(k-1)}, \dots, s_{i(k)-1}, 2k$,

  and $s_{2k}, 2k+1$



Sampling uniform binary trees

To sample $T \in_u \mathcal{T}_c$:

- Choose a sequence $S = (s_1, \dots, s_{2k})$ with 2 copies of each integer from $1, \dots, k$, uniformly at random.
- Let $i(1) < i(2) < \dots < i(k)$ be the locations of the repeated entries (so $2 \le i(1) \le k+1$ and $i(k) = 2k$)
- Build a tree T with branches

$$s_1, s_2, \dots, s_{i(1)-1}, k+1$$

$$s_{i(1)}, \dots, s_{i(2)-1}, k+2$$

$$\vdots$$

$$s_{i(k-1)}, \dots, s_{i(k)-1}, 2k,$$

and $s_{2k}, 2k+1$

Example

$$S = (1, 2, 3, \underset{6}{\underline{1}}, 5, \underset{7}{\underline{5}}, 4, \underset{8}{\underline{2}}, \underset{9}{\underline{4}}, \underset{10}{\underline{3}}) \qquad k = 5$$

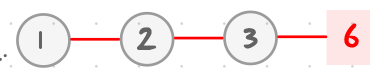
Sampling uniform binary trees

To sample $T \in_u \mathcal{T}_c$:

- Choose a sequence $S = (s_1, \ldots, s_{2k})$ with 2 copies of each integer from $1, \ldots, k$, uniformly at random.

- Let $i(1) < i(2) < \cdots < i(k)$ be the locations of the repeated entries (so $2 \le i(1) \le k+1$ and $i(k) = 2k$)

- Build a tree T with branches



$$s_1, s_2, \ldots, s_{i(1)-1}, k+1$$

$$s_{i(1)}, \ldots, s_{i(2)-1}, k+2$$

$$\vdots$$

$$s_{i(k-1)}, \ldots, s_{i(k)-1}, 2k,$$

and $s_{2k}, 2k+1$

Example

$$S = (1, 2, 3, \underset{6}{\underline{1}}, 5, \underset{7}{\underline{5}}, 4, \underset{8}{\underline{2}}, \underset{9}{\underline{4}}, \underset{10}{\underline{3}})$$

Sampling uniform binary trees

To sample $T \in_u \mathcal{T}_c$:

- Choose a sequence $S = (s_1, \dots, s_{2k})$ with 2 copies of each integer from $1, \dots, k$, uniformly at random.
- Let $i(1) < i(2) < \dots < i(k)$ be the locations of the repeated entries (so $2 \leq i(1) \leq k+1$ and $i(k) = 2k$)
- Build a tree T with branches

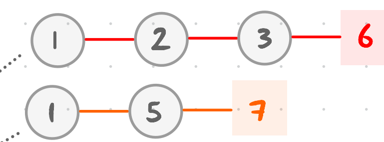
$$s_1, s_2, \dots, s_{i(1)-1}, k+1$$

$$s_{i(1)}, \dots, s_{i(2)-1}, k+2$$

$$\vdots$$

$$s_{i(k-1)}, \dots, s_{i(k)-1}, 2k,$$

and $s_{2k}, 2k+1$

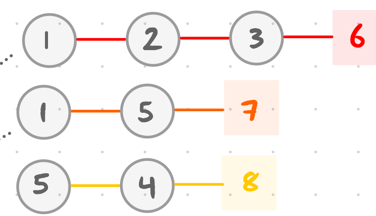


Example $S = (1, 2, 3, \underset{6}{\underline{1}}, 5, \underset{7}{\underline{5}}, 4, \underset{8}{\underline{2}}, \underset{9}{\underline{4}}, \underset{10}{\underline{3}})$

Sampling uniform binary trees

To sample $T \in_u \mathcal{T}_c$:

- Choose a sequence $S = (s_1, \dots, s_{2k})$ with 2 copies of each integer from $1, \dots, k$, uniformly at random.
- Let $i(1) < i(2) < \dots < i(k)$ be the locations of the repeated entries (so $2 \le i(1) \le k+1$ and $i(k) = 2k$)
- Build a tree T with branches



$$s_1, s_2, \dots, s_{i(1)-1}, k+1$$
$$s_{i(1)}, \dots, s_{i(2)-1}, k+2$$
$$\vdots$$
$$s_{i(k-1)}, \dots, s_{i(k)-1}, 2k,$$
and $s_{2k}, 2k+1$

Example

$$S = (1, 2, 3, \underset{6}{\underline{1}}, 5, \underset{7}{\underline{5}}, 4, \underset{8}{\underline{2}}, \underset{9}{\underline{4}}, \underset{10}{\underline{3}})$$

Sampling uniform binary trees

To sample $T \in_u \mathcal{T}_c$:

- Choose a sequence $S = (s_1, \dots, s_{2k})$ with 2 copies of each integer from $1, \dots, k$, uniformly at random.
- Let $i(1) < i(2) < \dots < i(k)$ be the locations of the repeated entries (so $2 \le i(1) \le k+1$ and $i(k) = 2k$)
- Build a tree T with branches

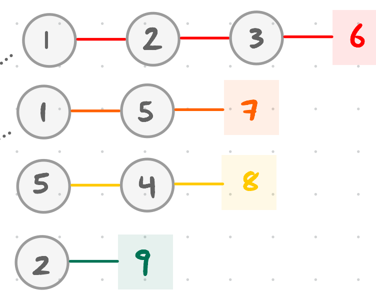
$$s_1, s_2, \dots, s_{i(1)-1}, k+1$$

$$s_{i(1)}, \dots, s_{i(2)-1}, k+2$$

$$\vdots$$

$$s_{i(k-1)}, \dots, s_{i(k)-1}, 2k,$$

and $s_{2k}, 2k+1$



Example

$$S = (1, 2, 3, \underset{6}{\underline{1}}, 5, \underset{7}{\underline{5}}, 4, \underset{8}{\underline{2}}, \underset{9}{\underline{4}}, \underset{10}{\underline{3}})$$

Sampling uniform binary trees

To sample $T \in_u \mathcal{T}_c$:

- Choose a sequence $S = (s_1, \dots, s_{2k})$ with 2 copies of each integer from $1, \dots, k$, uniformly at random.

- Let $i(1) < i(2) < \dots < i(k)$ be the locations of the repeated entries (so $2 \leq i(1) \leq k+1$ and $i(k) = 2k$)

- Build a tree T with branches

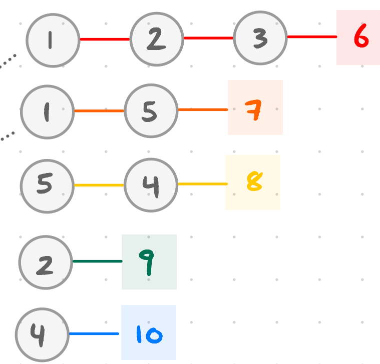
$s_1, s_2, \dots, s_{i(1)-1}, k+1$

$s_{i(1)}, \dots, s_{i(2)-1}, k+2$

$\vdots$

$s_{i(k-1)}, \dots, s_{i(k)-1}, 2k,$

and $s_{2k}, 2k+1$



Example

$$S = (1, 2, 3, \underset{6}{\underline{1}}, 5, \underset{7}{\underline{5}}, 4, \underset{8}{\underline{2}}, \underset{9}{\underline{4}}, \underset{10}{\underline{3}})$$

Sampling uniform binary trees

To sample $T \in_u \mathcal{T}_c$:

- Choose a sequence $S = (s_1, \dots, s_{2k})$ with 2 copies of each integer from $1, \dots, k$, uniformly at random.
- Let $i(1) < i(2) < \dots < i(k)$ be the locations of the repeated entries (so $2 \leq i(1) \leq k+1$ and $i(k) = 2k$)
- Build a tree T with branches

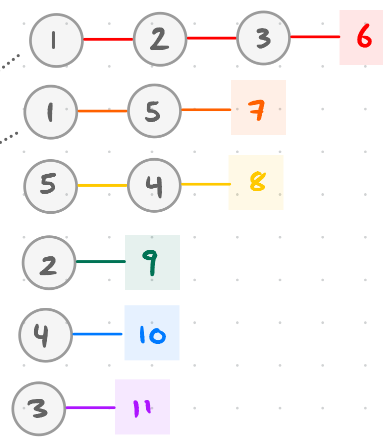
$s_1, s_2, \dots, s_{i(1)-1}, k+1$

$s_{i(1)}, \dots, s_{i(2)-1}, k+2$

$\vdots$

$s_{i(k-1)}, \dots, s_{i(k)-1}, 2k$,

and $s_{2k}, 2k+1$



NB The sequence S can be recovered from the tree T.

Example

$$S = (1, 2, 3, \underset{6}{\underline{1}}, 5, \underset{7}{\underline{5}}, 4, \underset{8}{\underline{2}}, \underset{9}{\underline{4}}, \underset{10}{\underline{3}}) \quad 11$$

Sampling uniform binary trees

To sample $T \in_u T_c$:

- Choose a sequence $S = (s_1, \dots, s_{2k})$ with 2 copies of each integer from $1, \dots, k$, uniformly at random.
- Let $i(1) < i(2) < \dots < i(k)$ be the locations of the repeated entries (so $2 \le i(1) \le k+1$ and $i(k) = 2k$)
- Build a tree T with branches

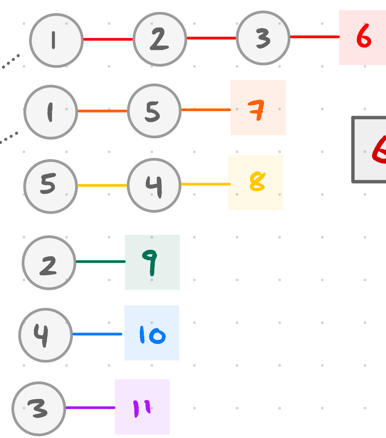
$s_1, s_2, \dots, s_{i(1)-1}, k+1$

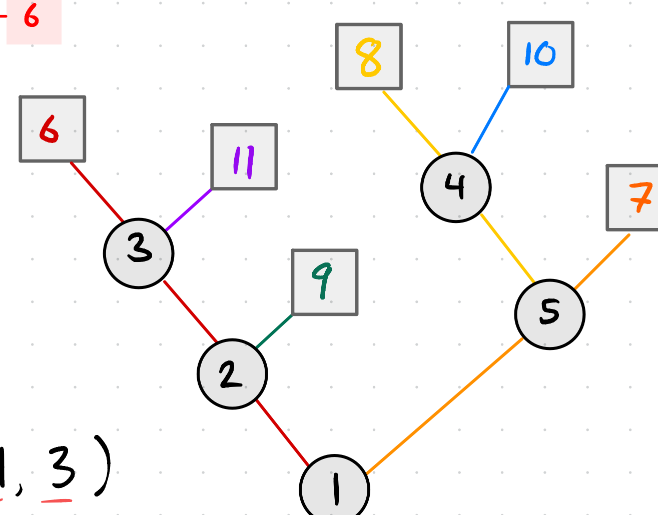
$s_{i(1)}, \dots, s_{i(2)-1}, k+2$

$\vdots$

$s_{i(k-1)}, \dots, s_{i(k)-1}, 2k$,

and $s_{2k}, 2k+1$





Example $S = (1, 2, 3, \underset{6}{\underline{1}}, 5, \underset{7}{\underline{5}}, 4, \underset{8}{\underline{2}}, \underset{9}{\underline{4}}, \underset{10}{\underline{3}})$

Sampling uniform binary trees

To sample $T \in_u \mathcal{T}_c$:

- Choose a sequence $S=(s_1,\dots,s_{2k})$ with 2 copies of each integer from $1,\dots,k$, uniformly at random.
- Let $i(1)<i(2)<\dots<i(k)$ be the locations of the repeated entries (so $2 \le i(1) \le k+1$ and $i(k)=2k$)
- Build a tree T with branches

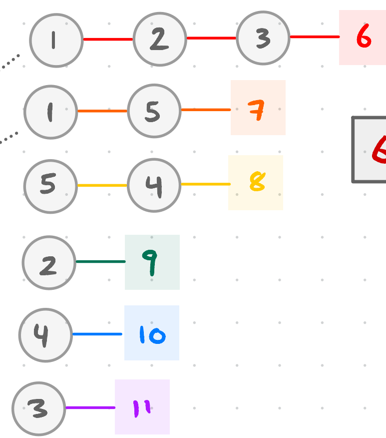
$s_1, s_2, \dots, s_{i(1)-1}, k+1$

$s_{i(1)}, \dots, s_{i(2)-1}, k+2$

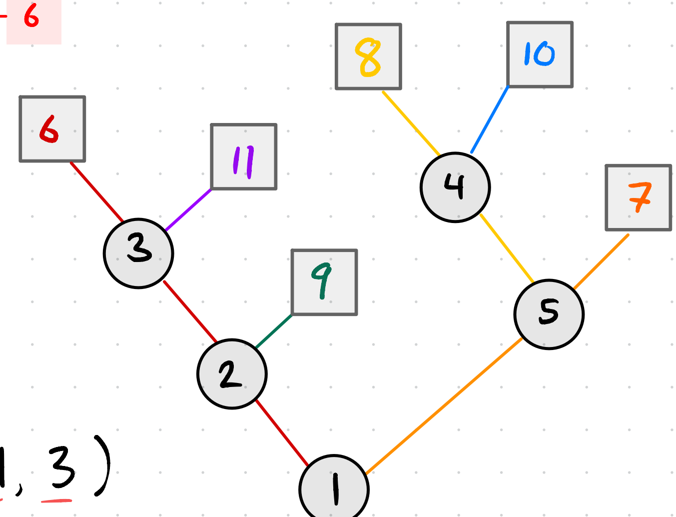
$\vdots$

$s_{i(k-1)}, \dots, s_{i(k)-1}, 2k,$

and $s_{2k}, 2k+1$



NB The sequence S can be recovered from the tree T.



Example

$$S=(1,2,3,\underset{6}{\underline{1}},5,\underset{7}{\underline{5}},4,\underset{8}{\underline{2}},\underset{9}{\underline{4}},\underset{10}{\underline{3}})$$

The length of the first branch

$S = (s_1, \ldots, s_{2k})$ has 2 copies of each integer from $1, \ldots, k$, uniformly at random.

$i(1)$ = index of first repeated entry.

The length of the first branch

$S = (s_1, \dots, s_{2k})$ has 2 copies of each integer from $1, \dots, k$, uniformly at random.

$i(1)$ = index of first repeated entry.

First branch is $s_1, s_2, \dots, s_{i(1)-1}, k+1$, so $\text{dist}_T(\text{root}, k+1) = i(1) - 1$

The length of the first branch

$S=(s_1,\dots,s_{2k})$ has 2 copies of each integer from $1,\dots,k$, uniformly at random.

$i(1)$ = index of first repeated entry.

First branch is $s_1, s_2, \dots, s_{i(1)-1}, k+1$, so $\mathrm{dist}_T(\mathrm{root}, k+1) = i(1)-1$

Then
$$\mathbb{P}(\mathrm{dist}(\mathrm{root}, k+1) \geq \ell) = \mathbb{P}(i(1) \geq \ell+1)$$
$$= \mathbb{P}(s_1, \dots, s_\ell \text{ are all distinct})$$
$$= \prod_{i=2}^{\ell}\left(1-\frac{i-1}{2k+1-(i-1)}\right) = e^{-\frac{\ell^2}{4k}+O\left(\frac{\ell^3}{k^2}\right)}$$

Taylor expansion of log.

The length of the first branch

$S = (s_1, \ldots, s_{2k})$ has 2 copies of each integer from $1, \ldots, k$, uniformly at random.

$i(1)$ = index of first repeated entry.

First branch is $s_1, s_2, \ldots, s_{i(1)-1}, k+1$, so $\mathrm{dist}_T(\mathrm{root}, k+1) = i(1) - 1$

Then
$$\mathbb{P}(\mathrm{dist}(\mathrm{root}, k+1) \geq \ell) = \mathbb{P}(i(1) \geq \ell + 1)$$
$$= \mathbb{P}(s_1, \ldots, s_\ell \text{ are all distinct})$$
$$= \prod_{i=2}^{\ell} \left(1 - \frac{i-1}{2k+1-(i-1)}\right) = e^{-\frac{\ell^2}{4k} + O\left(\frac{\ell^3}{k^2}\right)}$$

Taylor expansion of log.

It follows that $\boxed{\dfrac{\mathrm{dist}(\mathrm{root}, k+1)}{2\sqrt{k}} \xrightarrow{\mathrm{dist}} \mathrm{Rayleigh}}$ as $k \to \infty$.

The height of the (binary) tree

Example $S = (1, 2, 3, \underset{6}{\underline{1}}, 5, \underset{7}{\underline{5}}, 4, \underset{8}{\underline{2}}, \underset{9}{\underline{4}}, \underset{10}{\underline{3}})$

T has branches

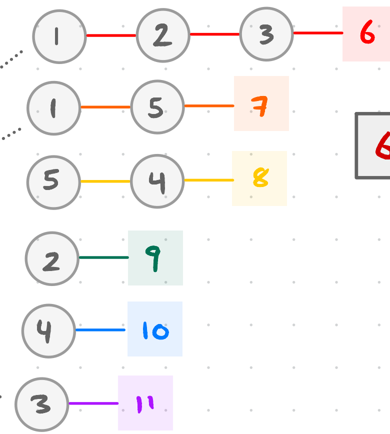
$s_1, s_2, \ldots, s_{i(1)-1}, k+1$

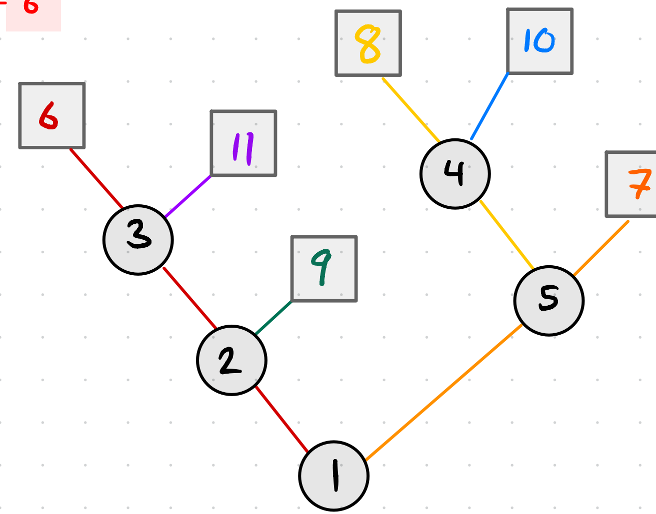
$s_{i(1)}, \ldots, s_{i(2)-1}, k+2$

$\vdots$

$s_{i(k-1)}, \ldots, s_{i(k)-1}, 2k$,

and $s_{2k}, 2k+1$





$$\frac{\mathrm{dist}(\mathrm{root}, k+1)}{2\sqrt{k}} \xrightarrow{\mathrm{dist}} \text{Rayleigh}$$

Repetitions become more frequent later in the sequence $\rightsquigarrow$ $\mathrm{height}(T) \approx \mathrm{dist}(\mathrm{root}, k+1) \approx \sqrt{k}$

Stochastic Comparison

For random variables X, Y, say $\overbrace{X \preccurlyeq_{st} Y}^{X \text{ is stochastically smaller than } Y}$ if

$$\mathbb{P}(X \geq h) \leq \mathbb{P}(Y \geq h) \text{ for all } h \in \mathbb{R}.$$

Stochastic Comparison

For random variables X, Y, say $\underbrace{X \preccurlyeq_{st} Y}_{X \text{ is stochastically smaller than } Y}$ if

$$\mathbb{P}(X \geq h) \leq \mathbb{P}(Y \geq h) \quad \text{for all } h \in \mathbb{R}.$$

Example $\text{Uniform}\{1, \ldots 6\} \preccurlyeq_{st} \text{Uniform}\{1, \ldots, 20\}$

$$\text{Bernoulli}(\tfrac{1}{2}) \not\preccurlyeq_{st} \text{Normal}(0,1)$$

$$\not\succcurlyeq_{st}$$

$$\mathbb{P}(\text{Bernoulli}(\tfrac{1}{2}) < 0) = 0 < \mathbb{P}(\text{Normal}(0,1) < 0)$$

$$\mathbb{P}(\text{Bernoulli}(\tfrac{1}{2}) < 2) = 1 > \mathbb{P}(\text{Normal}(0,1) < 2)$$

Stochastic Comparison

X is stochastically smaller than Y

For random variables X, Y, say $X \leqslant_{st} Y$ if

$$\mathbb{P}(X \geqslant h) \leqslant \mathbb{P}(Y \geqslant h) \text{ for all } h \in \mathbb{R}.$$

Example $\text{Uniform}\{1, \dots 6\} \leqslant_{st} \text{Uniform}\{1, \dots, 20\}$

$$\text{Bernoulli}(\tfrac{1}{2}) \not\leqslant_{st} \text{Normal}(0,1)$$

$$\not\geqslant_{st}$$

$$\mathbb{P}(\text{Bernoulli}(\tfrac{1}{2}) < 0) = 0 < \mathbb{P}(\text{Normal}(0,1) < 0)$$

$$\mathbb{P}(\text{Bernoulli}(\tfrac{1}{2}) < 2) = 1 > \mathbb{P}(\text{Normal}(0,1) < 2)$$

Fact $X \leqslant_{st} Y$ if and only if X, Y can be generated on a common probability space such that $\mathbb{P}(X \leqslant Y) = 1$

Lower degrees make taller trees.

Define a partial order $\succcurlyeq$ on child sequences via its covering relation, as follows:

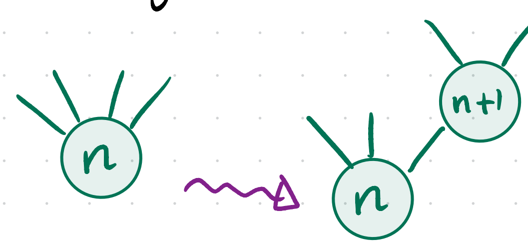
Lower degrees make taller trees.

Define a partial order $\succcurlyeq$ on child sequences via its covering relation, as follows:

Given child sequences $c = (c_1, \ldots, c_n)$ and $c' = (c'_1, \ldots, c'_{n+1})$ say c covers c' if (possibly after permuting entries)

a) $c'_1 = c_1,\ c'_2 = c'_2,\ \ldots,\ c'_{n-1} = c_{n-1}$

b) $c'_n + c'_{n+1} = c_n + 1$



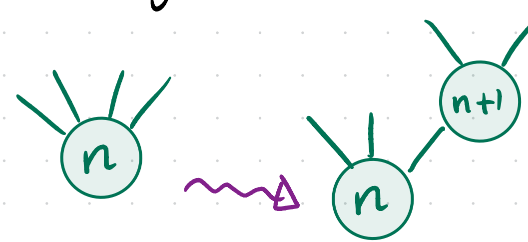
Lower degrees make taller trees.

Define a partial order $\succcurlyeq$ on child sequences via its covering relation, as follows:

Given child sequences $c=(c_1,\dots,c_n)$ and $c'=(c'_1,\dots,c'_{n+1})$ say c covers c' if (possibly after permuting entries)

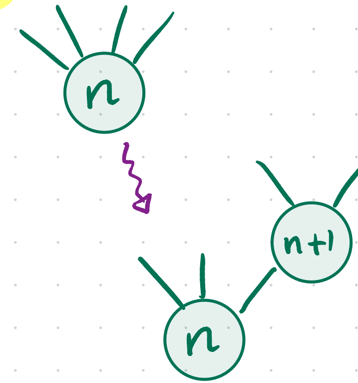
a) $c'_1=c_1,\ c'_2=c'_2,\ \dots,\ c'_{n-1}=c_{n-1}$

b) $c'_n+c'_{n+1}=c_n+1$



Theorem (A-B, Donderwinkel 2022)
If $c \succcurlyeq c'$ and $T \in_u \mathcal{T}_c$, $T' \in_u \mathcal{T}_{c'}$ then
$$\text{height}(T) \preccurlyeq_{st} \text{height}(T')$$

Binary trees are the tallest*



Theorem (A-B, Donderwinkel 2022)

If $c \succcurlyeq c'$ and $T \in_u \mathcal{T}_c$, $T' \in_u \mathcal{T}_{c'}$ then

$$\text{height}(T) \preccurlyeq_{st} \text{height}(T')$$

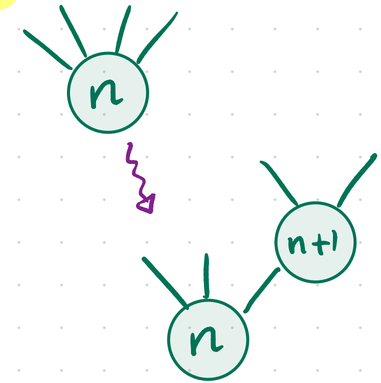
Corollary

Suppose $c = (c_1, \ldots, c_n)$ has $n_1 = 0$

Let $n_0 = \#\{i \in [n] : c_i = 0\}$ ← # of leaves

Let $T \in_u \mathcal{T}_c$ and let T' be a uniform binary tree with n_0 leaves. Then $\text{height}(T) \preccurlyeq_{st} \text{height}(T')$.

Binary trees are the tallest*



Theorem (A-B, Donderwinkel 2022)

If $c \succcurlyeq c'$ and $T \in_u \mathcal{T}_c$, $T' \in_u \mathcal{T}_{c'}$ then

$$\text{height}(T) \preccurlyeq_{st} \text{height}(T')$$

Corollary

Suppose $c = (c_1, \ldots, c_n)$ has $n_1 = 0$

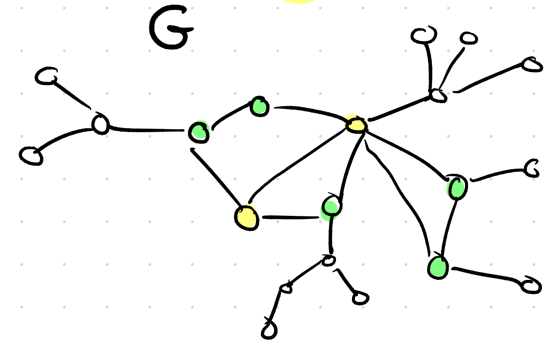
Let $n_0 = \#\{i \in [n] : c_i = 0\}$ ← # of leaves

Let $T \in_u \mathcal{T}_c$ and let T' be a uniform binary tree with n_0 leaves. Then $\text{height}(T) \preccurlyeq_{st} \text{height}(T')$.

First theorem then follows without too much effort.

A tool for Theorem 2: the core-kernel decomposition

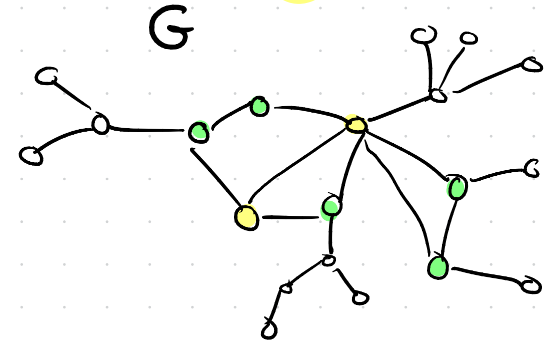
For a graph G:



A tool for Theorem 2: the core-kernel decomposition

For a graph G:

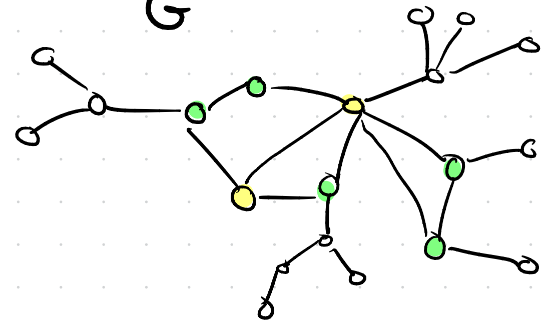
$\mathrm{core}(G) =$ 



A tool for Theorem 2: the core-kernel decomposition

For a graph G:

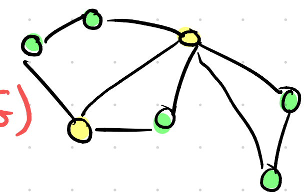
G



core(G) = 

= maximum subgraph of minimum degree 2

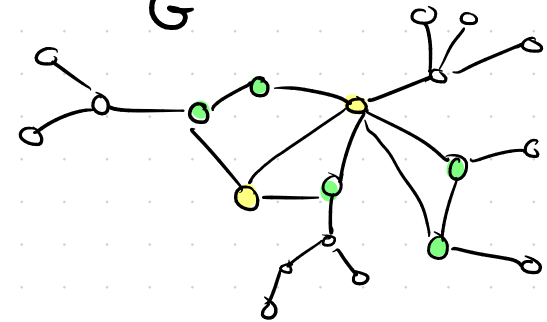
core(G)



A tool for Theorem 2: the core-kernel decomposition

For a graph G:

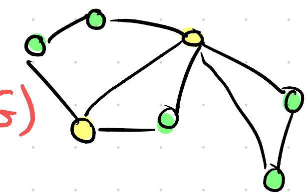
G



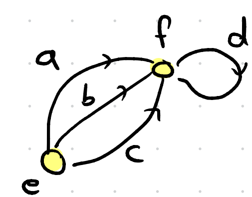
core(G) = 

= maximum subgraph of minimum degree 2

core(G)

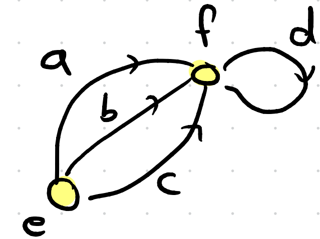


kernel(G) = "replace paths in core(G) by edges"



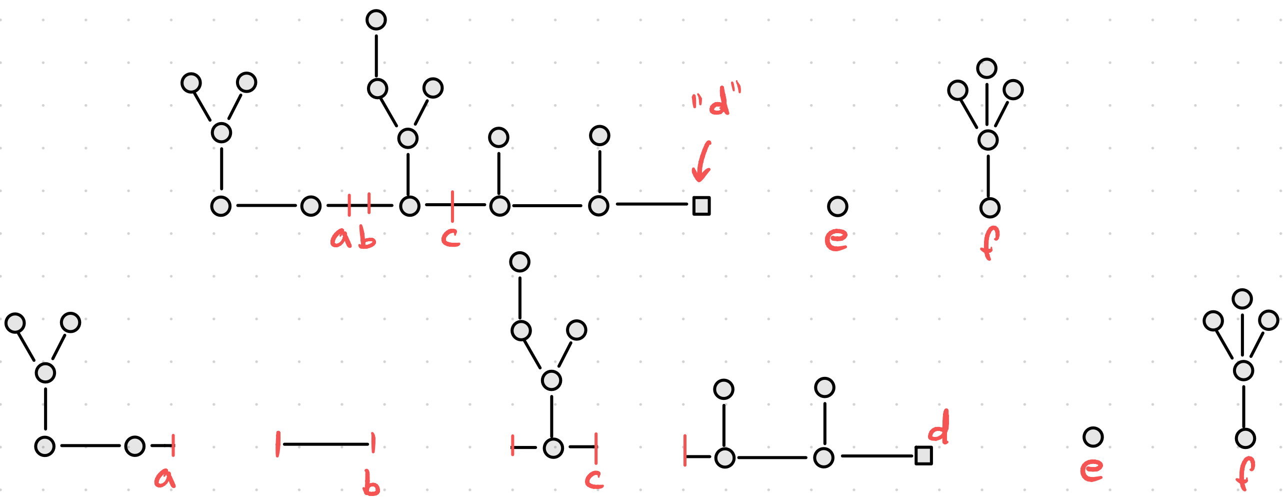
Recovering a graph from its core

Data to reconstruct G from kernel(G)



A) A tree with |E(kernel(G))|-1 marks along the first root-to-leaf path.

B) A forest with root set V(kernel(G))



Diameter bounds for random graphs

[Goal: $G \in_u \mathcal{G}_d$ $n_2(d) \leq (1-\varepsilon)n$ then $\mathbb{E}\,\mathrm{diam}(G) = O_\varepsilon(\sqrt{n})$]

Using the core-kernel decomposition, diameter bounds for $G \in_u \mathcal{G}_d$ follow from

a) bounds on heights in random trees / forests

(similar to above)

b) bounds on the diameter of kernel(G).

(kernel(G) has minimum degree 3 so by our third theorem has logarithmic diameter)

A conjecture

Theorem (A-B, Crudele 2025) Fix a degree sequence $d=(d_1,\dots,d_n)$ s.t. a) $\mathcal{G}_d \neq \emptyset$; b) $d_i \geq 3$ for all i. Let $G \in_u \mathcal{G}_d$.
Then $\mathbb{P}(G \text{ connected}) = 1 - O(\frac{1}{n})$; $\mathbb{P}(\mathrm{diam}(G) > 62 \log n) = O(\frac{1}{n})$

Conjecture (3-regular graphs maximize diameter when $\delta \geq 3$)

Fix a degree sequence $d=(d_1,\dots,d_n)$ with n even and s.t.

a) $\mathcal{G}_d \neq \emptyset$; b) $d_i \geq 3$ for all i.

Let $d' = \underbrace{(3,3,\dots,3)}_{n \text{ terms.}}$. Let $G \in_u \mathcal{G}_d$ and $G' \in_u \mathcal{G}_{d'}$.

Then $\mathrm{diam}(G) \preceq_{st} \mathrm{diam}(G)$.