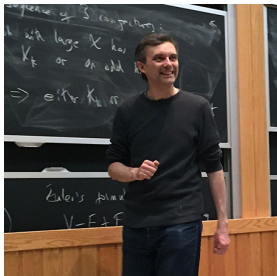


Counting graphic sequences via integrated random walks

Carla Groenland
TU Delft

Joint work with Paul Balister, Serte Donderwinkel,
Tom Johnston and Alex Scott

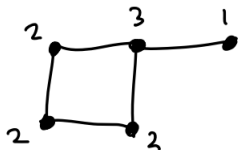
JCB 2024



Started March 2022, Oxford; <https://arxiv.org/pdf/2301.07022.pdf>

Talk overview

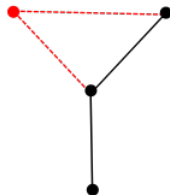
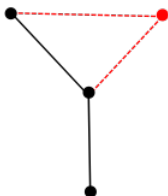
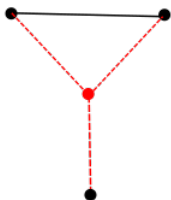
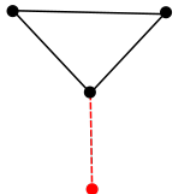
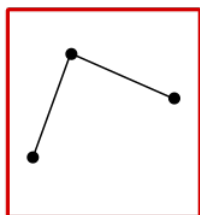
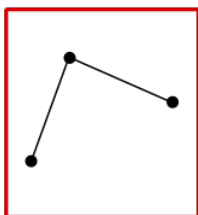
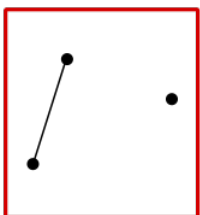
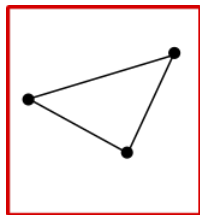
Graphic sequence: integer sequence $n - 1 \geq d_1 \geq \dots \geq d_n \geq 0$ that is the degree sequence of n -vertex graph.



$(3, 2, 2, 2, 1)$

- Why may you want to count these?
- Reformulation to problem about integrated random walks.
- Main result and proof sketch.

Wine glasses are reconstructible



Graph reconstruction conjecture

G is *reconstructible* if every graph with the same multiset of strict induced subgraphs is isomorphic to it.

Reconstruction conjecture (Kelly and Ulam)

Every graph G on $n \geq 3$ vertices is reconstructible.

P. Kelly. *On isometric transformations*. PhD thesis, University of Wisconsin, 1942.

S. Ulam. *A collection of mathematical problems*, volume 8. Interscience Publishers, 1960.

Degree reconstruction from small cards

ℓ -deck: multiset of induced subgraphs on ℓ vertices.

Problem

For which ℓ can the degree sequence of any n -vertex graph G be reconstructed from ℓ -deck?

Want: same ℓ -deck implies same degree sequence.

number of ℓ -decks \geq number of degree sequences for n -vertex graphs

$$\binom{n}{\ell}^{2^{\ell^2}} \geq$$

number of ℓ -decks \geq number of degree sequences for n -vertex graphs
 $\geq 2^n$.

Each ℓ -vertex graph appears at most $\binom{n}{\ell}$ times in the deck.

$$\ell = \Omega(\sqrt{\log n})$$

Simple lower bound

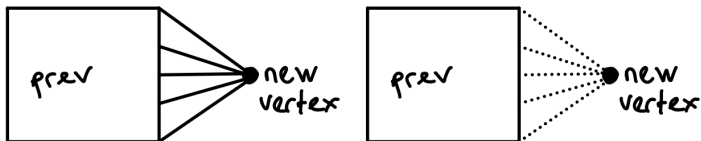
$G(n)$ = number of graphic sequences of length n
= number of degree sequences for n -vertex graphs
 $\geq 2^n$.

Simple lower bound

$G(n)$ = number of graphic sequences of length n
= number of degree sequences for n -vertex graphs
 $\geq 2^n$.

Induction on n , two new sequences from (d_1, \dots, d_{n-1}) :

$(n-1, d_1+1, \dots, d_{n-1}+1)$ new vertex connected to all
 $(d_1, \dots, d_{n-1}, 0)$ new vertex connected to none



Simple upper bound

$G(n)$ = number of graphic sequences of length n
= number of degree sequences for n -vertex graphs
 $\leq 4^n$.

Simple upper bound

$G(n)$ = number of graphic sequences of length n
= number of degree sequences for n -vertex graphs
 $\leq 4^n$.

Stars-and-bars argument: number of integer sequences

$$n - 1 \geq d_1 \geq \cdots \geq d_n \geq 0$$

is at most $\binom{2n}{n} \leq 4^n$.

Counting graphic sequences

$G(n)$ = number of graphic sequences of length n .

Burns (2007): for some constants $c_1, c_2 > 0$, all $n \in \mathbb{N}$

$$c_1 \frac{4^n}{n} \leq G(n) \leq \frac{4^n}{\sqrt{n} \log^{c_2} n}.$$

Counting graphic sequences

$G(n)$ = number of graphic sequences of length n .

Burns (2007): for some constants $c_1, c_2 > 0$, all $n \in \mathbb{N}$

$$c_1 \frac{4^n}{n} \leq G(n) \leq \frac{4^n}{\sqrt{n \log^{c_2} n}}.$$

Balister, Donderwinkel, G., Johnston, Scott, 2023+

$$G(n) = (c_{\text{deg}} + o(1)) \frac{4^n}{n^{3/4}}.$$

Related work: score sequences

$T(n)$ = number of out-degree sequences for n -vertex tournaments.

\approx 1970: Moser, Erdős-Moser and Kleitman investigate + conjecture.
Winston-Kleitman (lower bound, 1983) and Kim-Pittel (upper bound, 2000) show

$$T(n) = \Theta\left(\frac{4^n}{n^{5/2}}\right).$$

Kolesnik (2022) shows there is a constant $c \approx 0.392$ such that

$$T(n) = (c + o(1))\frac{4^n}{n^{5/2}}.$$

0 1 3 6 2 7
: 13
: 20
23 IS 12
10 22 11 21

THE ON-LINE ENCYCLOPEDIA OF INTEGER SEQUENCES[®]

founded in 1964 by N. J. A. Sloane

[Hints](#)

(Greetings from [The On-Line Encyclopedia of Integer Sequences!](#))

A004251 Number of graphical partitions (degree-vectors for simple graphs with n vertices, or possible ordered row-sum vectors for a symmetric 0-1 matrix with diagonal values 0).²⁴
(Formerly M1250)

1, 1, 2, 4, 11, 31, 102, 342, 1213, 4361, 16016, 59348, 222117, 836315, 3166852, 12042620,
45967479, 176005709, 675759564, 2600672458, 10029832754, 38753710486, 149990133774, 581393603996,
2256710139346, 8770547818956, 34125389919850, 132919443189544, 518232001761434, 2022337118015338,
7898574056034636, 30873421455729728 ([list](#): [graph](#): [refs](#): [listen](#): [history](#): [text](#): [internal format](#))

OFFSET 0,3

COMMENTS In other words, $a(n)$ is the number of graphic sequences of length n , where a graphic sequence is a sequence of numbers which can be the degree sequence of some graph.

Previously known: $n \leq 290$. We compute for $n \leq 1651$.

Erdős-Gallai (1960)

$n - 1 \geq d_1 \geq \dots \geq d_n \geq 0$ graphic sequence \iff

- Parity condition: $d_1 + \dots + d_n$ is even.
- Dominating condition: for all k ,
'highest k degrees can feasibly be obtained', that is,

$$d_1 + \dots + d_k \leq k(k - 1) + \min(d_{k+1}, k) + \dots + \min(d_n, k)$$

Erdős-Gallai (1960)

$n - 1 \geq d_1 \geq \dots \geq d_n \geq 0$ graphic sequence \iff

- Parity condition: $d_1 + \dots + d_n$ is even.
- Dominating condition: for all k ,
'highest k degrees can feasibly be obtained', that is,

$$d_1 + \dots + d_k \leq k(k - 1) + \min(d_{k+1}, k) + \dots + \min(d_n, k)$$

Counting \rightarrow probability: what is the probability that a uniformly random $n - 1 \geq d_1 \geq \dots \geq d_n \geq 0$ satisfies both conditions?

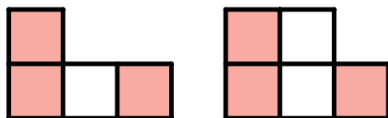
Probability of being even

Lemma

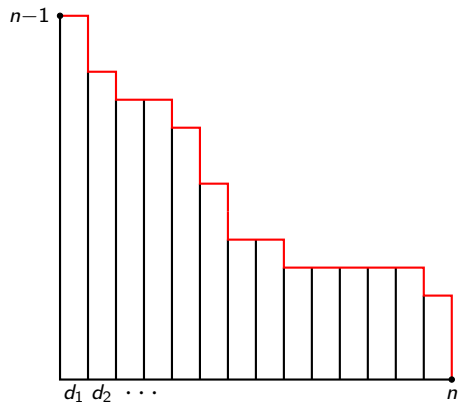
A uniformly random sequence $n - 1 \geq d_1 \geq \dots \geq d_n \geq 0$ which satisfies the dominating condition, has probability $\frac{1}{2} + o(1)$ of satisfying the parity condition (that $\sum_{i=1}^n d_i$ is even).

Idea: 'match sequences with odd and even sums'.

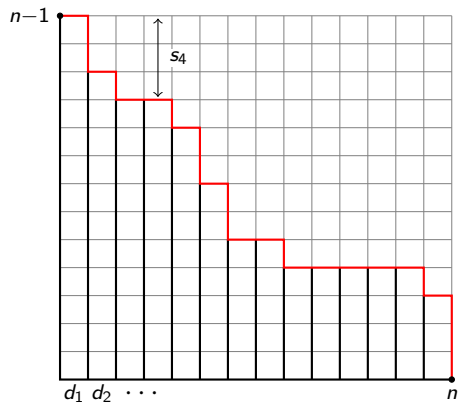
Unexpected experimental observation: $(\# \text{ even} - \# \text{ odd}) \approx 4^n / n^{O(1)}$.



Reformulation as lattice path

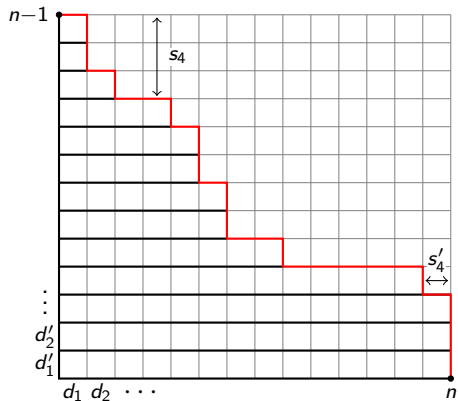


Reformulation as lattice path



$$s_i = (n - 1) - d_i$$

Reformulation as lattice path

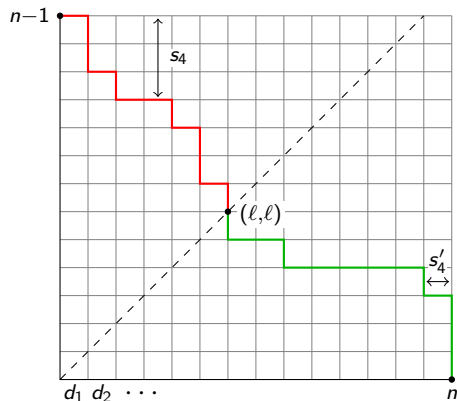


$$s_i = (n - 1) - d_i$$

$$s'_i = n - d'_i \text{ where}$$

$d'_i =$ number of vertices of degree $\geq i$.

Reformulation as lattice path



$$s_i = (n - 1) - d_i$$

$$s'_i = n - d'_i \text{ where}$$

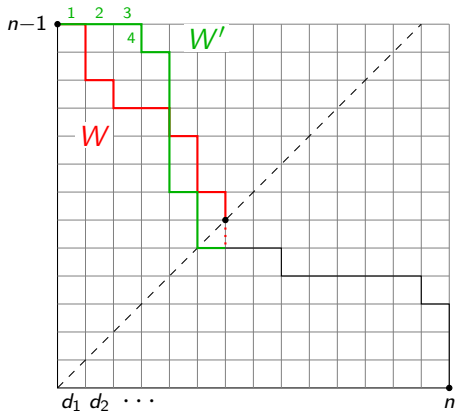
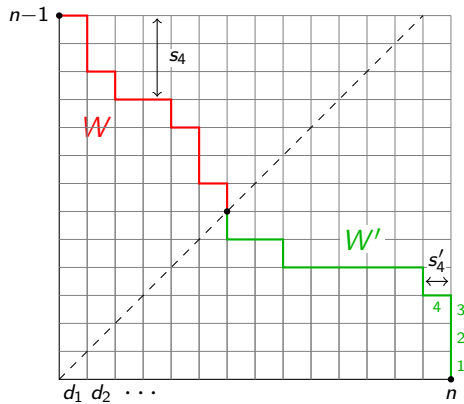
$d'_i =$ number of vertices of degree $\geq i$.

Dominating condition \iff

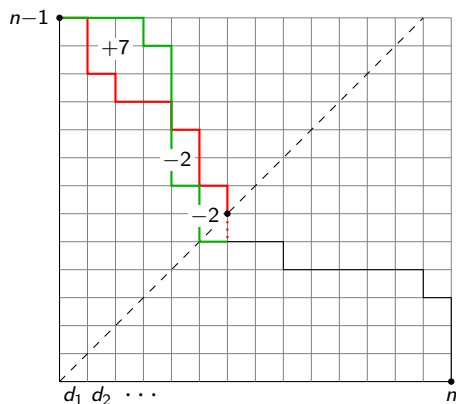
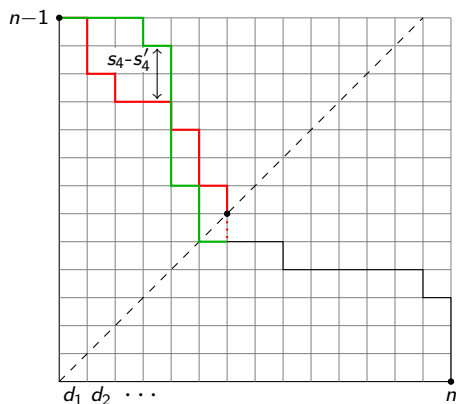
$$(s_1 - s'_1) + \dots + (s_k - s'_k) \geq 0$$

for all $k \leq \ell$

Split in half and reflect path

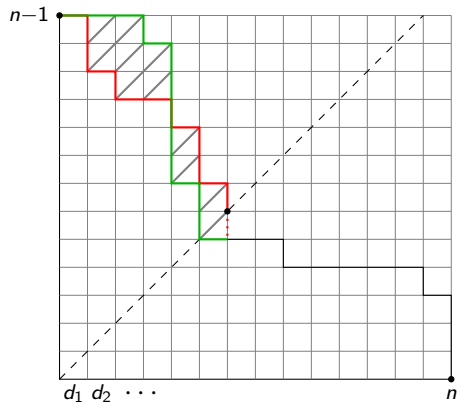
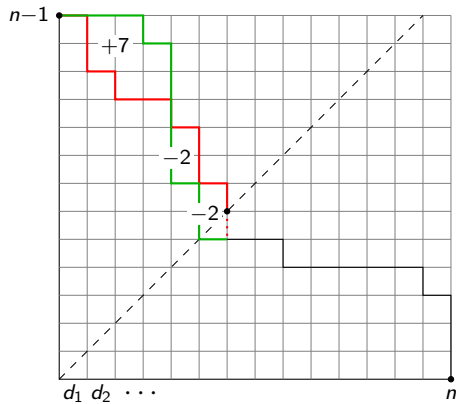


Interpretation of dominating condition

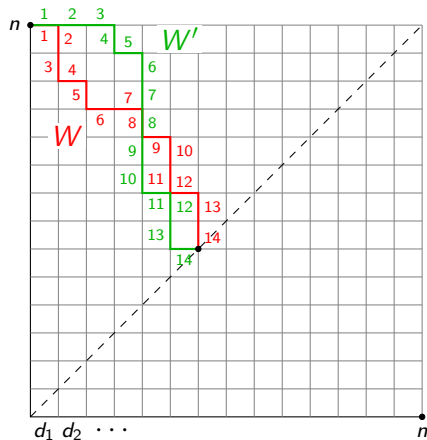
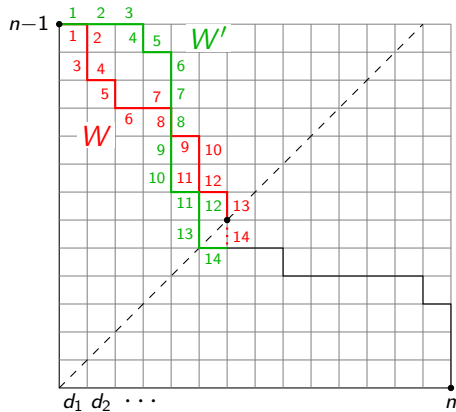


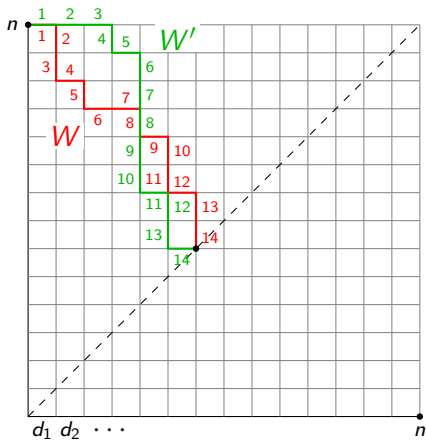
$(s_1 - s'_1) + (s_2 - s'_2) + \dots + (s_k - s'_k)$ equals sum of signed areas.

'Double count': compute area diagonally



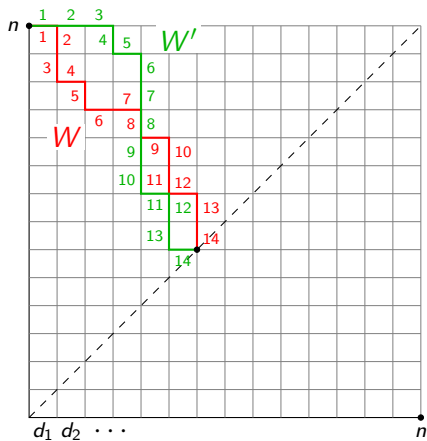
Cheat: pretend $(n, 0)$ to $(0, n)$





$W, W': \ell$ times $\rightarrow, n - \ell$ times \downarrow

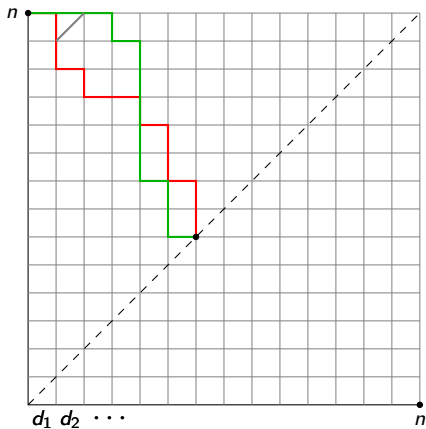
Total of n steps.



W, W' : ℓ times \rightarrow , $n - \ell$ times \downarrow

Total of n steps.

Y_i = difference in 'height' between W' and W after 'step' i



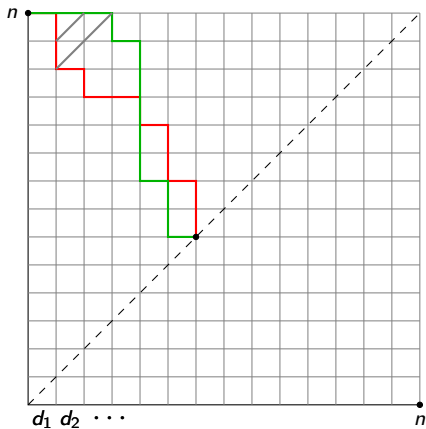
W, W' : ℓ times \rightarrow , $n - \ell$ times \downarrow

Total of n steps.

Y_i = difference in 'height' between W' and W after 'step' i

$$Y_0 = Y_1 = 0$$

$$Y_2 = 1$$



W, W' : ℓ times \rightarrow , $n - \ell$ times \downarrow

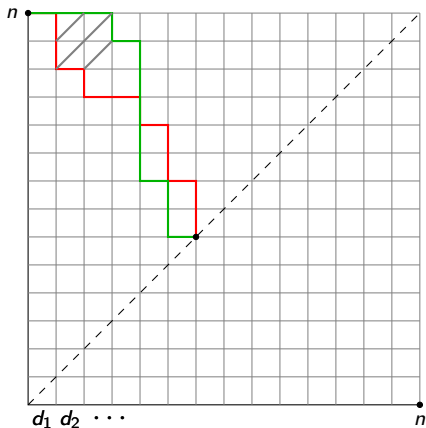
Total of n steps.

Y_i = difference in 'height' between W' and W after 'step' i

$$Y_0 = Y_1 = 0$$

$$Y_2 = 1$$

$$Y_3 = 2$$



W, W' : ℓ times \rightarrow , $n - \ell$ times \downarrow

Total of n steps.

Y_i = difference in 'height' between W' and W after 'step' i

$$Y_0 = Y_1 = 0$$

$$Y_2 = 1$$

$$Y_3 = 2$$

$$Y_4 = 1$$

Symmetric simple random walk (SSRW)

- Steps: $X_1, X_2, \dots \sim \text{Unif}(\{-1, 1\})$ independent.
- SSRW: $S_k = X_1 + \dots + X_k$.

Symmetric simple random walk (SSRW)

- Steps: $X_1, X_2, \dots \sim \text{Unif}(\{-1, 1\})$ independent.
- SSRW: $S_k = X_1 + \dots + X_k$.
- $Z_i = 1$ if W goes \downarrow at step i and -1 otherwise.
- $Z'_i = 1$ if W' goes \rightarrow at step i and -1 otherwise.
- So $L_i = \frac{1}{2}(Z_i + Z'_i) \in \{-1, 0, 0, 1\}$.

- **Lazy** steps: $L_1, L_2, \dots \sim \text{Unif}(\{-1, 0, 0, 1\})$ independent.
- Lazy SSRW: $Y_k = L_1 + \dots + L_k$.

Lazy, bridge, integrated

- **Lazy** steps: $L_1, L_2, \dots \sim \text{Unif}(\{-1, 0, 0, 1\})$ independent.
- Lazy SSRW: $Y_k = L_1 + \dots + L_k$.
- Lazy SSRW **bridge**: condition $(Y_k)_{k=0}^n$ on $Y_n = 0$.
- **Integrated** walk: $A_n = Y_1 + \dots + Y_n$.

Lazy, bridge, integrated

- **Lazy** steps: $L_1, L_2, \dots \sim \text{Unif}(\{-1, 0, 0, 1\})$ independent.
- Lazy SSRW: $Y_k = L_1 + \dots + L_k$.
- Lazy SSRW **bridge**: condition $(Y_k)_{k=0}^n$ on $Y_n = 0$.
- **Integrated** walk: $A_n = Y_1 + \dots + Y_n$.
- Dominating condition: sum of areas is always positive.
Equivalent: $A_1 \geq 0, A_2 \geq 0, \dots, A_n \geq 0$.

Quick asymptotics using theorem of Vysotsky

For $(Y_i)_{i=1}^n$ the lazy SSRW and $A_k = \sum_{i=1}^k Y_i$ its area process,

$$\mathbb{P}(A_1, \dots, A_n \geq 0 \mid Y_n = 0) = \Theta(n^{-1/4}).$$

Theorem of [Vysotsky](#) (2014) applies to wide range of random processes.

Quick asymptotics using theorem of Vysotsky

For $(Y_i)_{i=1}^n$ the lazy SSRW and $A_k = \sum_{i=1}^k Y_i$ its area process,

$$\mathbb{P}(A_1, \dots, A_n \geq 0 \mid Y_n = 0) = \Theta(n^{-1/4}).$$

Theorem of [Vysotsky \(2014\)](#) applies to wide range of random processes.

Adding this to our reformulation ('exercise'),

$$G(n) = \Theta(n^{-1/4})\Theta(4^n/\sqrt{n}) = \Theta(4^n/n^{3/4}).$$

To get correct constant, we strengthen the probabilistic result above.

Persistence probability for integrated (lazy) SSRW

Persistence probability: probability that a random process does not go negative.

Balister, Donderwinkel, G., Johnston, Scott, 2023+

$$n^{1/4} \mathbb{P}(A_1, \dots, A_n \geq 0 \mid Y_n = 0) \rightarrow \frac{\Gamma(3/4)}{\sqrt{2\pi(1-\rho)}}$$

as $n \rightarrow \infty$.

(Definition of ρ later, Γ is extension of factorial.)

Main result: complete statement

Balister, Donderwinkel, G., Johnston, Scott, 2023+

The number of degree sequences for n -vertex graphs

$$G(n) = (c_{\text{deg}} + o(1)) \frac{4^n}{n^{3/4}}$$

where

$$c_{\text{deg}} = \frac{\Gamma(3/4)}{4\pi\sqrt{2(1-\rho)}} \approx 0.099094.$$

$$G(n) \binom{2n-1}{n}^{-1} = 2(1 + o(1)) \mathbb{P}(A_1, \dots, A_n \geq 0 \mid Y_n = 0)$$

Definition of ρ

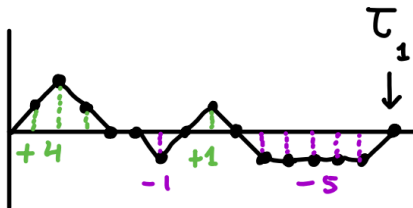
For $(Y_i)_{i=1}^n$ the lazy SSRW and $A_k = \sum_{i=1}^k Y_i$ its area process,

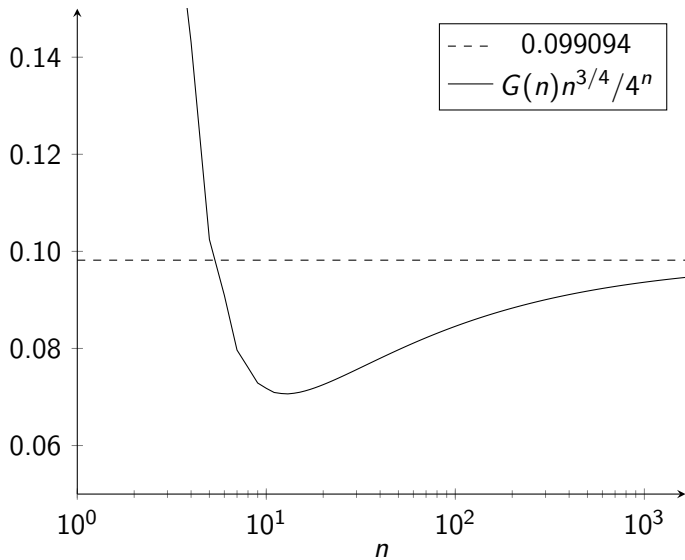
$$\rho = \mathbb{P}(A_{\tau_1} = 0)$$

with

$$\tau_1 = \inf\{k \geq 1 : Y_k = 0, A_k \leq 0\}$$

first visit of Y to 0 at which (A_k) hits non-positive number.





Heuristic for $n^{-1/4}$

- Lazy SSRW bridge: $Y_i = L_1 + \dots + L_i$
- **Excursion:** $Y_i = 0, Y_{i+1}, \dots, Y_{i+k} \neq 0, Y_{i+k+1} = 0$.
- $N = \Theta(\sqrt{n})$ excursions with areas ζ_1, \dots, ζ_N .

- Lazy SSRW bridge: $Y_i = L_1 + \dots + L_i$
- **Excursion:** $Y_i = 0, Y_{i+1}, \dots, Y_{i+k} \neq 0, Y_{i+k+1} = 0$.
- $N = \Theta(\sqrt{n})$ excursions with areas ζ_1, \dots, ζ_N .

- Dominating condition: $E_k = \sum_{i=1}^k \zeta_i \geq 0$ for all k .
- $(E_k)_{k=1}^N$ is almost symmetric random walk;
SRW with N steps remains positive with probability $\Theta(\sqrt{N})$.

Exchangeable and symmetric excursions

- $N = \Theta(\sqrt{n})$ excursions with areas ζ_1, \dots, ζ_N .
- $E_k = \sum_{i=1}^k \zeta_i$.
- Steps are not independent!

Exchangeable and symmetric excursions

- $N = \Theta(\sqrt{n})$ excursions with areas ζ_1, \dots, ζ_N .
- $E_k = \sum_{i=1}^k \zeta_i$.
- Steps are not independent!
- **Exchangeable:** $(\zeta_1, \dots, \zeta_N) \stackrel{d}{=} (\zeta_{\sigma(1)}, \dots, \zeta_{\sigma(N)})$ for any permutation $\sigma \in \Sigma_N$.
- + symmetric: for any permutation σ and $s_1, \dots, s_N \in \{1, -1\}$,

$$(\zeta_1, \dots, \zeta_N) \stackrel{d}{=} (s_1 \zeta_{\sigma(1)}, \dots, s_N \zeta_{\sigma(N)}).$$

Exploiting symmetry and exchangeability

$$\mathbb{P} \left(\sum_{i=1}^k \zeta_i \geq 0 \text{ for all } k \in [M] \right) = \mathbb{P} \left(\sum_{i=1}^k s_i \zeta_{\sigma(i)} \geq 0 \text{ for all } k \in [M] \right)$$

Exploiting symmetry and exchangeability

$$\begin{aligned} \mathbb{P} \left(\sum_{i=1}^k \zeta_i \geq 0 \text{ for all } k \in [M] \right) &= \mathbb{P} \left(\sum_{i=1}^k s_i \zeta_{\sigma(i)} \geq 0 \text{ for all } k \in [M] \right) \\ &= \sum_{x \in \mathbb{R}^n} \mathbb{P}(\zeta_1 = x_1, \dots, \zeta_N = x_N) \mathbb{P} \left(\sum_{i=1}^k s_i x_{\sigma(i)} \geq 0 \text{ for all } k \in [M] \right). \end{aligned}$$

Useful (combinatorial) lemma

For $x = (x_1, \dots, x_N) \in \mathbb{R}^N$, permutation $\sigma \in \Sigma_N$ and signs $s \in \{-1, 1\}^N$ chosen uniformly at random (independently),

$$\mathbb{P} \left(\sum_{i=1}^k s_i x_{\sigma(i)} \geq 0 \text{ for all } k \in [N] \right) = ???$$

Useful (combinatorial) lemma

For $x = (x_1, \dots, x_N) \in \mathbb{R}^N$, permutation $\sigma \in \Sigma_N$ and signs $s \in \{-1, 1\}^N$ chosen uniformly at random (independently),

$$\mathbb{P} \left(\sum_{i=1}^k s_i x_{\sigma(i)} \geq 0 \text{ for all } k \in [N] \right) = \frac{(2N-1)!!}{2^N N!}$$

does not depend on x if x is 'generic': for all distinct $A, A' \subseteq [N]$

$$\sum_{i \in A} x_i \neq \sum_{i \in A'} x_i.$$

Making excursions 'generic'

If the steps L_1, L_2, \dots, L_n of the lazy SSRW bridge are 'perturbed' by

$$\epsilon_1, \dots, \epsilon_n \sim \text{Unif} \left(\left[-\frac{1}{2n}, \frac{1}{2n} \right] \right)$$

slightly, then 'perturbed' $\tilde{\zeta}_1, \dots, \tilde{\zeta}_N$ would be 'generic'!

Where does Γ come from?

If ζ_1, \dots, ζ_N would be 'generic', probability of interest would be

$$\mathbb{E} \left[\frac{(2N-1)!!}{2^N N!} \right] \sim n^{-1/4} \Gamma(3/4) / \sqrt{2\pi}$$

where for N the number of excursions,

$$n^{-1/2} N \rightarrow 2\sqrt{E}, \quad E \sim \text{Exp}(1)$$

Where does ρ come from?

When the area process is > 0 or < 0 , the perturbed one is also > 0 or < 0 respectively.

Still need to account for when it equals zero.

Where does ρ come from?

When the area process is > 0 or < 0 , the perturbed one is also > 0 or < 0 respectively.

Still need to account for when it equals zero.

For i with $\zeta_1 + \dots + \zeta_i = 0$, do we have

$$\epsilon_1 + \dots + \epsilon_i \geq 0?$$

M = number of i such that $\zeta_1 + \dots + \zeta_i = 0$.

M conditional on the dominating condition converges in distribution to $\text{Geom}(\rho)$.

Open problem: upper bound in useful lemma

For $(s, \sigma) \sim \text{Unif}(\{-1, 1\}^N \times \Sigma_N)$ when is

$$\mathbb{P} \left(\sum_{i=1}^k s_i x_{\sigma(i)} \geq 0 \text{ for all } k \in [N] \right)$$

maximal? Is this for $x_1 = \dots = x_N$?

Open problem: fixed number of edges

What if you fix the number of edges instead of the number of vertices?

Let $p(N)$ the fraction of partitions of an integer N that are graphical.

Answering a conjecture of [Wilf \(1982\)](#), [Pittel \(1999\)](#) showed $p(N) \rightarrow 0$ as $N \rightarrow \infty$.

Known: $p(N) = N^{-\Theta(1)}$.

Can you show $\Theta(N^{-\alpha})$ or $(c + o(1))N^{-\alpha}$?

Open problem: other variations

We computed the number of graphic sequences for n -vertex graphs.

What about bipartite graphs, digraphs, uniform hypergraphs?

Open problem: other variations

We computed the number of graphic sequences for n -vertex graphs.

What about bipartite graphs, digraphs, uniform hypergraphs?

We computed the persistence probability for the integrated (lazy) SSRW bridge up to $(1 + o(1))$.

Can you do this for other integrated random processes?

Thank you for your attention!

Balister, Donderwinkel, G., Johnston, Scott, 2023+

For $(Y_i)_{i=1}^n$ lazy SSRW and $(A_i)_{i=1}^n$ its area process,

$$\mathbb{P}(A_1, \dots, A_n \geq 0 \mid Y_n = 0) = (1 + o(1)) \frac{\Gamma(3/4)}{\sqrt{2\pi(1-\rho)}} \frac{1}{n^{1/4}}.$$

Balister, Donderwinkel, G., Johnston, Scott, 2023+

$$G(n) = (1 + o(1)) \frac{\Gamma(3/4)}{4\pi\sqrt{2(1-\rho)}} \frac{4^n}{n^{3/4}}.$$