On the Longest Path and The Diameter in Random Apollonian Networks

Abbas Mehrabian
amehrabi@uwaterloo.ca

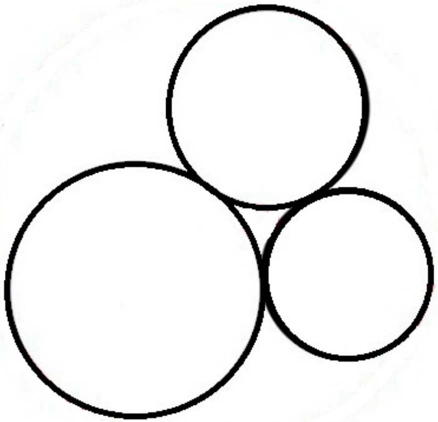
University of Waterloo

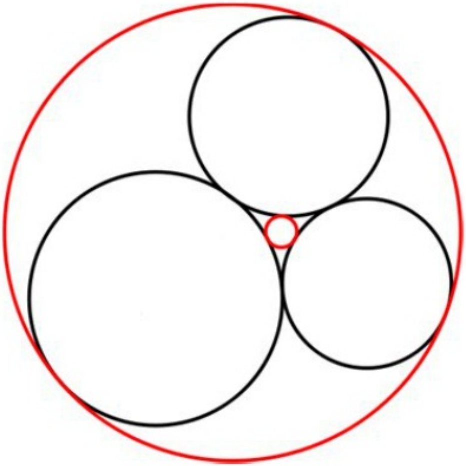
Frontiers in Mathematical Sciences
Honouring Siavash Shahshahani
Sharif University of Technology, Tehran
December 26th, 2012

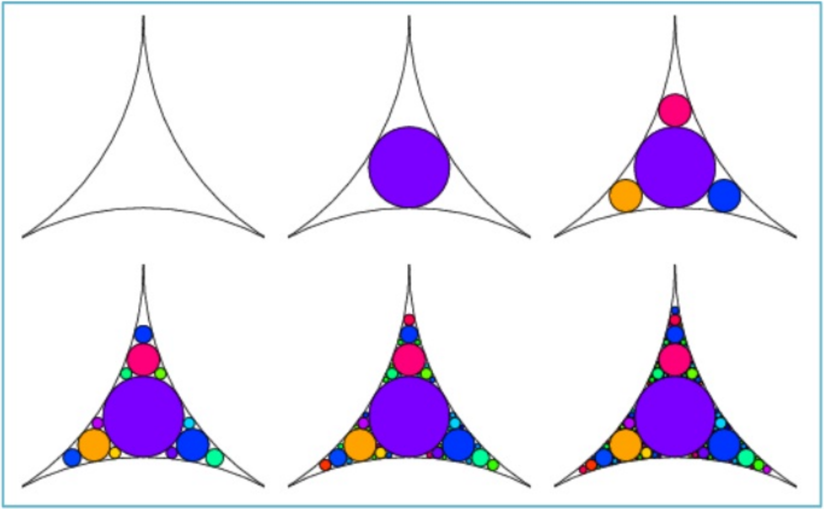
co-authors

- Ehsan Ebrahimzadeh
- Linda Farczadi
- Jane Gao
- Cristiane Sato
- Nick Wormald
- Jonathan Zung

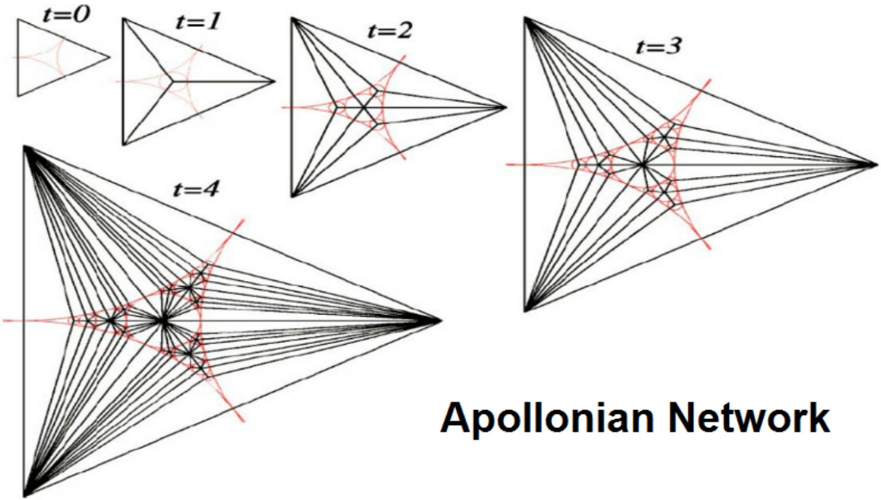
Pictures: Charalampos (Babis) Tsourakakis



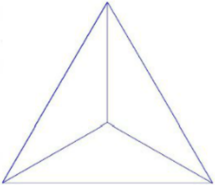




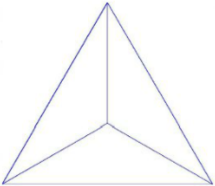
Apollonian Gasket



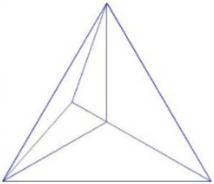
Apollonian Network



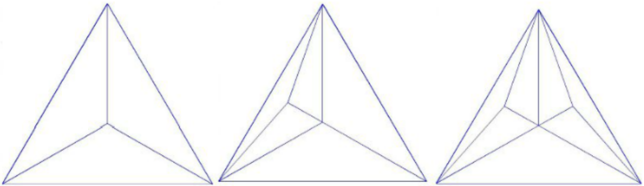
$t = 1$



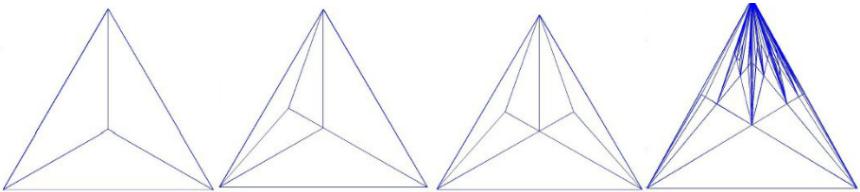
$t = 1$



$t = 2$



$t = 1$ $t = 2$ $t = 3$



$t = 1$ $t = 2$ $t = 3$ $t = 100$

Random Apollonian Network

After t steps,

- a random triangulated plane graph
- $t + 3$ vertices
- $3t + 3$ edges
- $2t + 1$ faces

called a Random Apollonian Network (RAN).
Zhou, Yan, Wang, Physical Review (2005)

Motivation

Modelling real-world networks:

- The Web graph
- Social networks
- Brain neurons
- Protein interactions

Motivation

Modelling real-world networks:

- The Web graph
- Social networks
- Brain neurons
- Protein interactions

Known models include:

- Erdös-Renyi random graphs
- preferential attachment model
- Kronecker graphs
- Cooper-Frieze model
- Aiello-Chung-Lu model
- protean graphs
- Fabrikant-Koutsoupias-Papadimitriou model

Motivation

RANs are an interesting model for generating random planar graphs.

Properties of Real-World Networks

1. Power-law degree distribution:

$$\mathbb{P}\left[\mathrm{deg}(\mathrm{a\ random\ vertex}) = k\right] = Ck^{-\beta}$$

2. Small-world phenomenon (six degrees of separation) :
There is a short path connecting every two vertices.

Known Results

Degree sequence

$n :=$ number of vertices
$Z_{k,n} :=$ number of vertices with degree k.

Theorem (Frieze and Tsourakakis 2012)

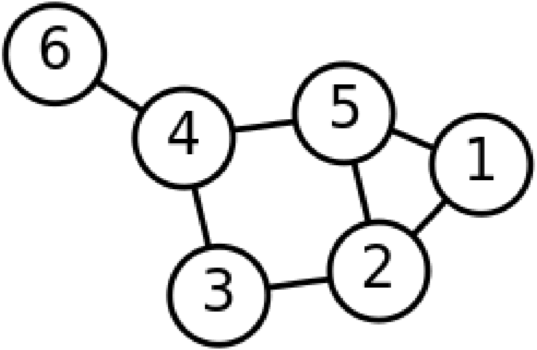
For every $k = k(n) \geq 3$,

$$\mathbb{E}\left[Z_{k,n}\right] = \Theta(nk^{-3}),$$

and,

$$\mathbb{P}\left[\left|Z_{k,n} - \mathbb{E}\left[Z_{k,n}\right]\right| > 10\sqrt{n \log n}\right] \to 0$$

The Diameter of a Graph



Diameter = 3

Known Results

The diameter

Theorem (Albenque and Marckert 2008)

Distance between two random vertices $\to 0.55 \log n$.

Theorem (Frieze and Tsourakakis 2012)

$$\mathbb{P}\left[\text{diameter} > 7.1 \log n\right] \to 0$$

Known Results

The diameter

Theorem (Albenque and Marckert 2008)

Distance between two random vertices $\to 0.55 \log n$.

Theorem (Frieze and Tsourakakis 2012)

$$\mathbb{P}\left[\text{diameter} > 7.1 \log n\right] \to 0$$

Theorem (EFGMSWZ'12+)

$$\frac{\text{diameter}}{\log n} \to c \approx 1.668 \qquad \text{in probability}$$

Our Result on the Diameter

Theorem (EFGMSWZ'12+)

$$f(x) := \frac{12x^3}{1-2x} - \frac{6x^3}{1-x},$$

$y :=$ unique solution to

$$x(x-1)f'(x) = f(x)\log f(x), \quad x \in (0, 1/2),$$

$$c := (1 - y^{-1})/\log f(y) \approx 1.668$$

Then for every fixed $\varepsilon > 0$,

$$\mathbb{P}\left[(1-\varepsilon)c\log n \le \text{diameter} \le (1+\varepsilon)c\log n\right] \to 1$$

The Longest Path

The conjecture

Conjecture (Frieze and Tsourakakis 2012)

$$\mathbb{P}\left[\exists \text{ a path containing a positive fraction of vertices}\right] \longrightarrow 1$$

The Longest Path

The conjecture

Conjecture (Frieze and Tsourakakis 2012)

$$\mathbb{P}\left[\exists \text{ a path containing a positive fraction of vertices}\right] \longrightarrow 1$$

Not true!

The Longest Path

Our results

$m :=$ number of faces $= 2n - 5$

$L_m :=$ length of the longest path

Theorem (EFGMSWZ'12+)

$\exists \theta > 0$ *such that*

$$\mathbb{P}\left[L_m < n/(\log n)^{\theta}\right] \to 1$$

The Longest Path

Our results

$m :=$ number of faces $= 2n - 5$

$L_m :=$ length of the longest path

Theorem (EFGMSWZ'12+)

$\exists\theta > 0$ *such that*

$$\mathbb{P}\left[L_m < n/(\log n)^{\theta}\right] \to 1$$

Theorem (EFGMSWZ'12+)

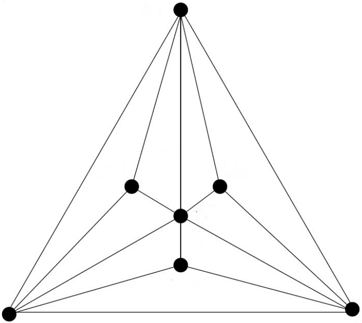
$$L_m > m^{0.63}$$

and

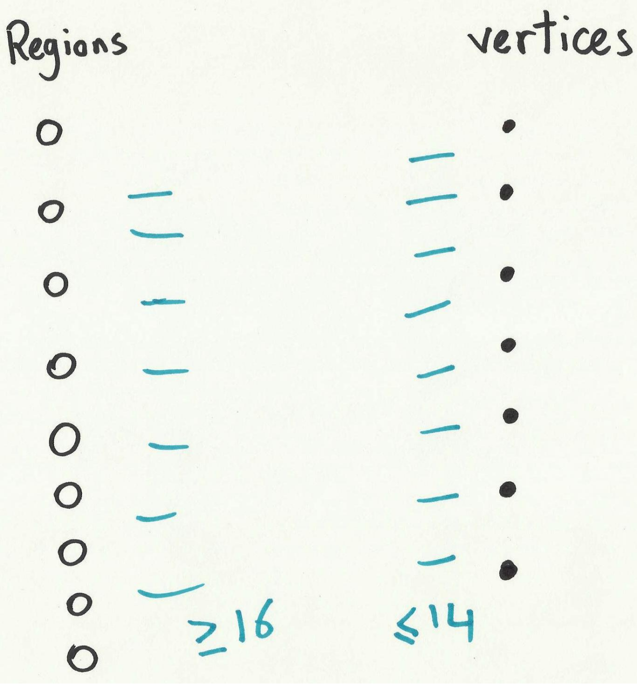
$$\mathbb{E}\left[L_m\right] = \Omega\left(m^{0.88}\right)$$

Upper Bound for Longest Path

The Main Idea

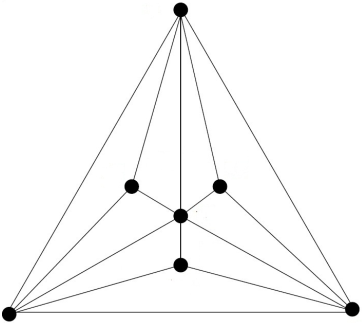


Claim: A simple path cannot contain internal vertices of all 9 regions.

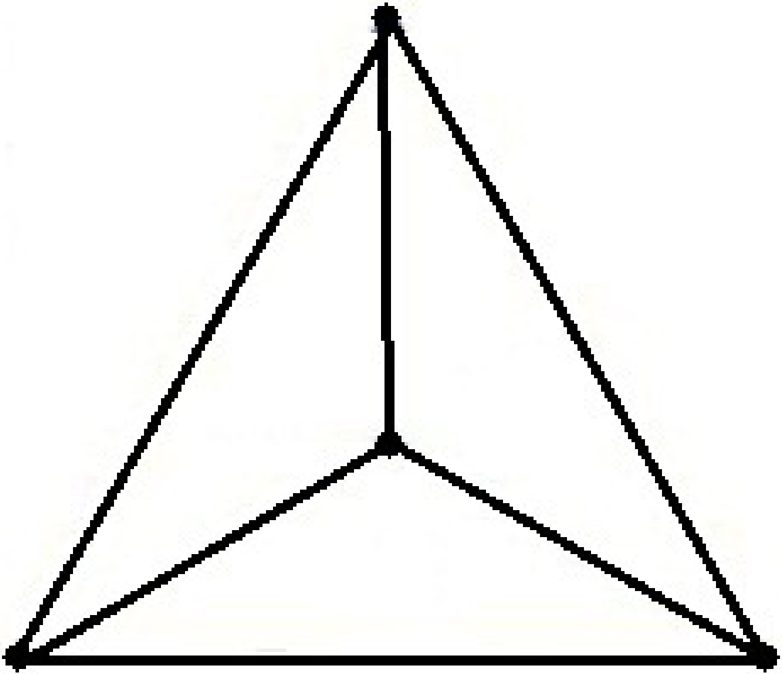

Regions
vertices
≥16
≤14

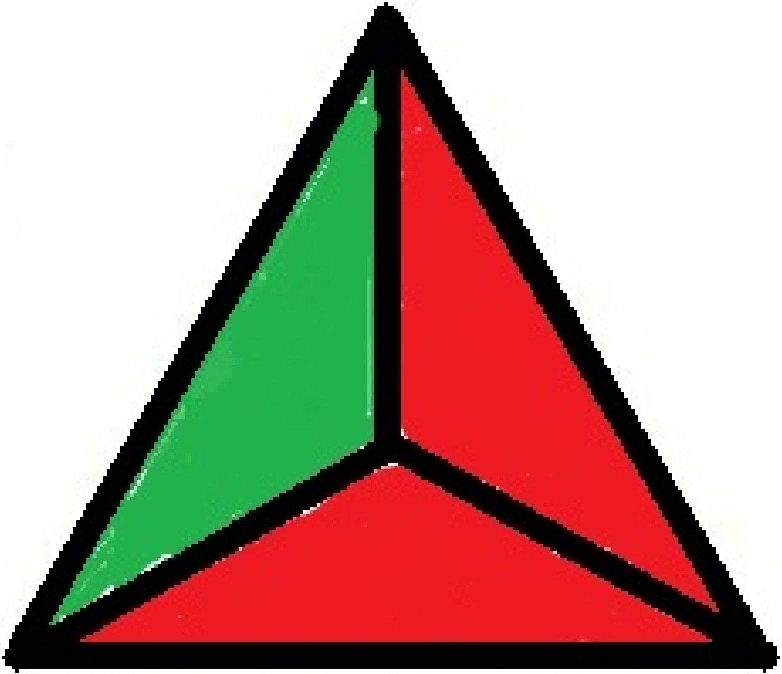
Upper Bound for Longest Path

The Main Idea



Claim: A simple path cannot contain internal vertices of all 9 regions.





Eggenberger-Pólya Urn

Theorem (Eggenberger and Pólya 1923)

Start: g green, r red balls.

In each step:

- *pick a random ball and return it to the urn;*
- *add s balls of the same colour.*

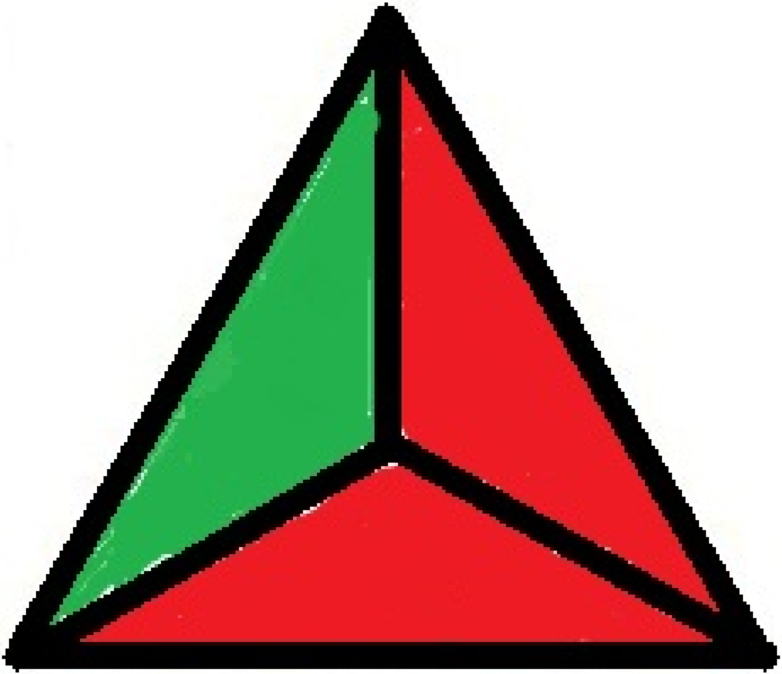
After n draws:

g_n: green balls

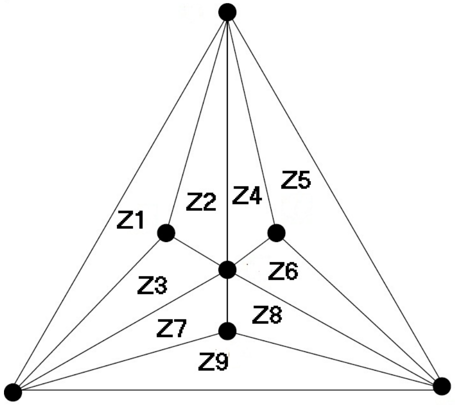
t_n: number of balls

For any $\alpha \in [0,1]$

$$\lim_{n\to\infty} \mathbb{P}\left[\frac{g_n}{t_n} < \alpha\right] = \frac{\Gamma((g+r)/s)}{\Gamma(g/s)\Gamma(r/s)} \int_0^{\alpha} x^{\frac{g}{s}-1}(1-x)^{\frac{r}{s}-1}\,\mathrm{d}x$$
$$= \mathbb{P}\left[Beta(g/s, r/s) < \alpha\right]$$



Upper Bound for Longest Path



Corollary

$$\mathbb{P}\left[\frac{\min\{Z_1,\cdots,Z_9\}}{n} < \epsilon\right] < 13\sqrt[4]{\epsilon}\,.$$

Upper Bound for Longest Path

Fix a small ϵ. We lose

$$n\left[\varepsilon + (1-\varepsilon)\varepsilon + (1-\varepsilon)^2\varepsilon + \cdots + (1-\varepsilon)^k\varepsilon\right] = n\left[1-(1-\varepsilon)^{k+1}\right]$$

vertices in any simple path.

Upper Bound for Longest Path

Fix a small ϵ. We lose

$$n\left[\varepsilon + (1-\varepsilon)\varepsilon + (1-\varepsilon)^2\varepsilon + \cdots + (1-\varepsilon)^k\varepsilon\right] = n\left[1-(1-\varepsilon)^{k+1}\right]$$

vertices in any simple path.

Theorem (EFGMSWZ'12+)

$\exists\theta > 0$ *such that*

$$\mathbb{P}\left[L_m < n/(\log n)^{\theta}\right] \to 1$$

Lower Bounds for Longest Path

$m :=$ number of faces $= 2n - 5$
$L_m :=$ length of the longest path
$\eta := \log 2 / \log 3 \approx 0.63$

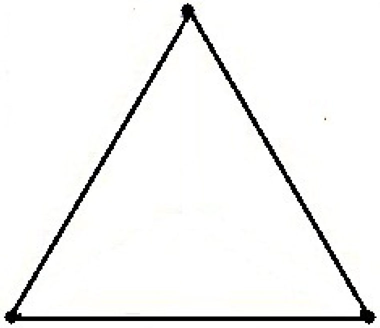
Theorem (EFGMSWZ'12+)

$$L_m > m^\eta$$

and

$$\mathbb{E}[L_m] = \Omega\left(m^{0.88}\right)$$

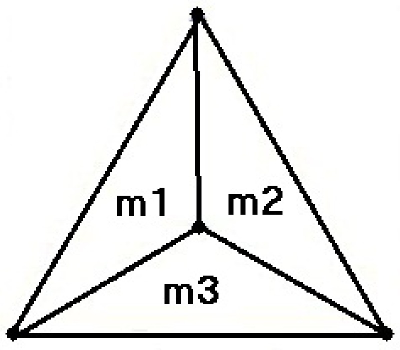
Lower Bounds for Longest Path



Lemma

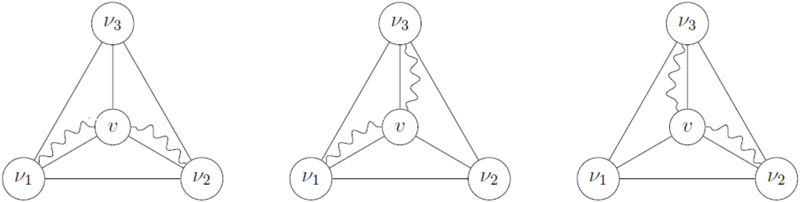
For any two boundary vertices,
$\exists$ *a path of length* $> m^{\eta}$ *connecting them*
not containing the third boundary vertex.

Lower Bounds for Longest Path



Assume that $m_1 \geq m_2 \geq m_3$

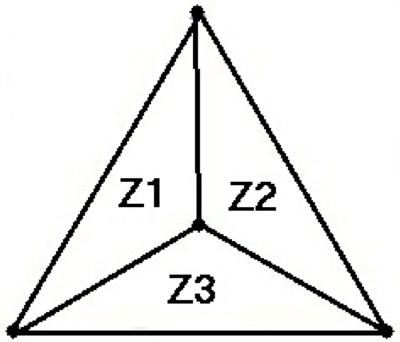
Lower Bounds for Longest Path



$$L_m \geq m_1^{\eta} + m_2^{\eta} \geq 2\left(\frac{m}{3}\right)^{\eta} = m^{\eta}$$

since $m_1 \geq m_2 \geq m_3$ and $m_1 + m_2 + m_3 = m$

Lower Bounds for Longest Path



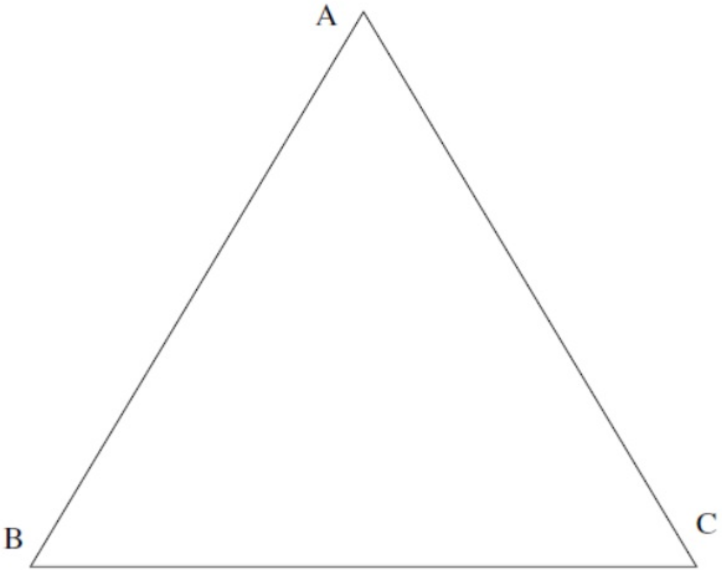
$$\mathbb{E}\left[L_m\right] \geq \mathbb{E}\left[Z_1^{0.88} + Z_2^{0.88} \middle| Z_1 \geq Z_2 \geq Z_3\right] \geq m^{0.88}$$

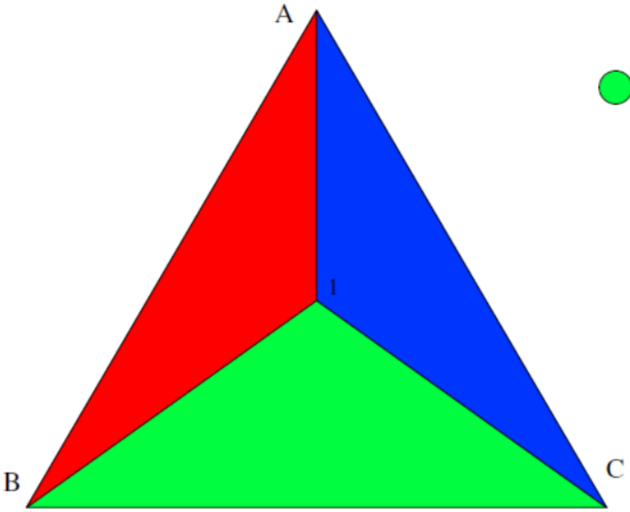
Radius

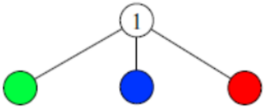
Radius : max distance between a vertex and the boundary

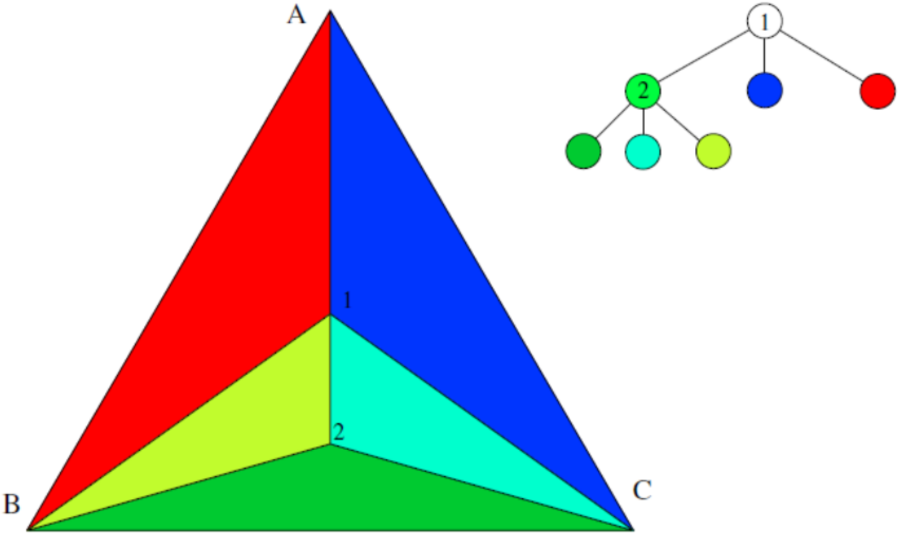
Lemma

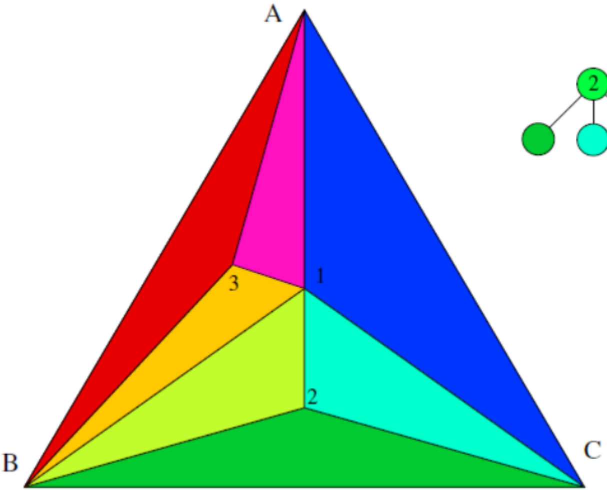
If radius $/\log n \longrightarrow c/2$ *in probability,*
then diameter$/\log n \longrightarrow c$ *in probability.*

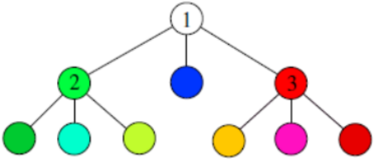

A
B
C

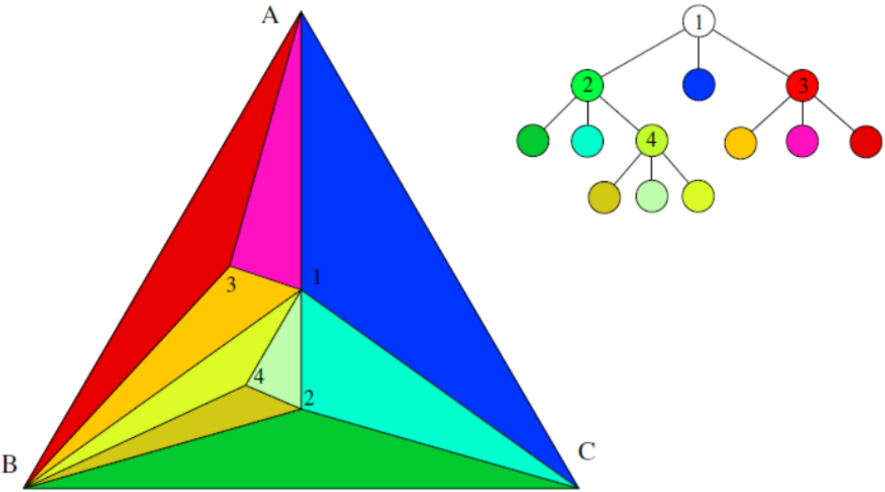

A
1
B
C

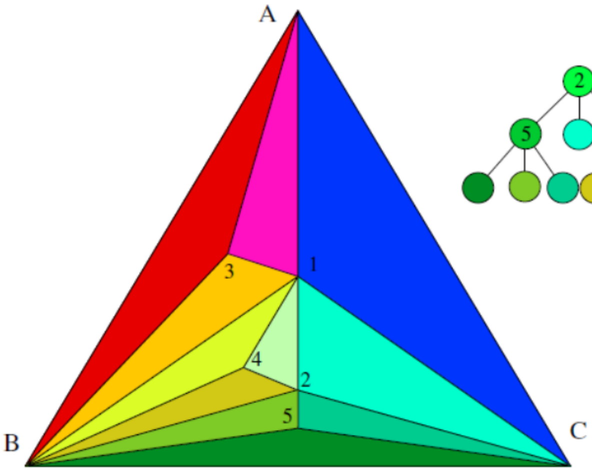

1

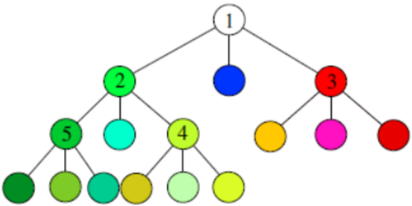

A
1
2
B
C
1
2

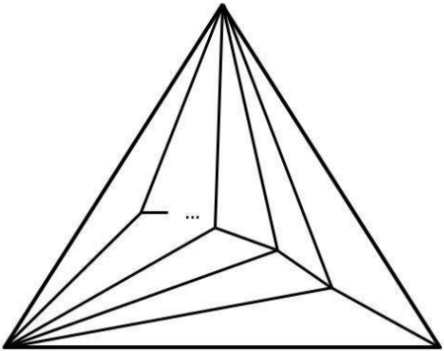

A
B
C
1
2
3

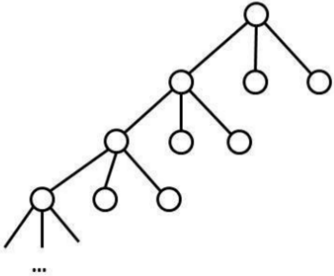

1
2
3

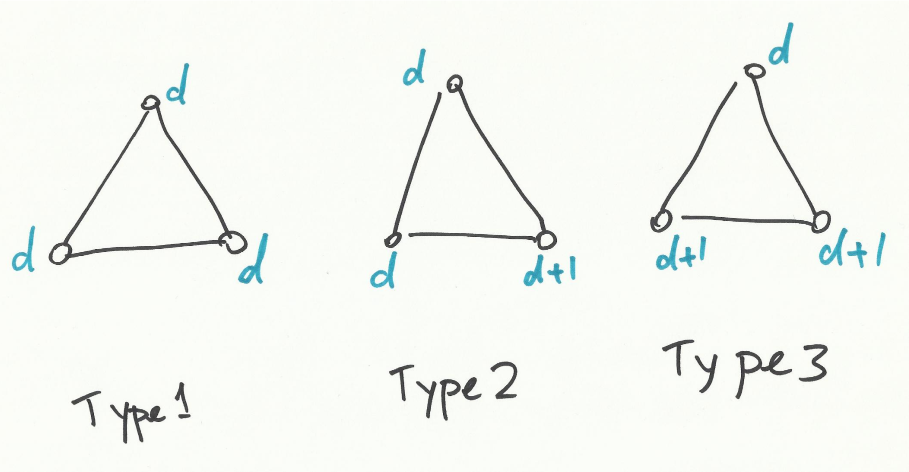

A
B
C
1
2
3
4

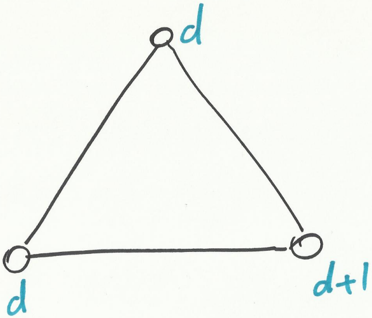

A
B
C
1
2
3
4
5


1
2
3
4
5

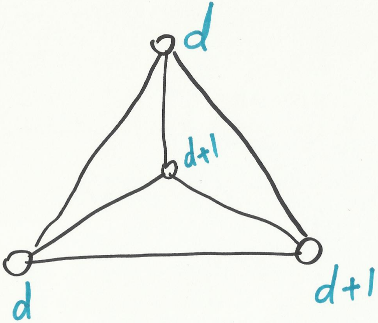




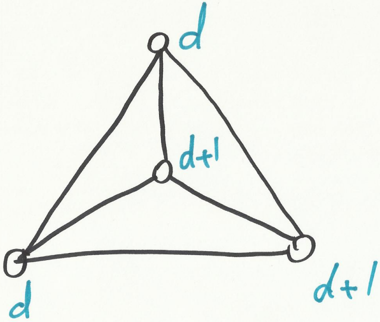

d
d
d
Type 1
d
d
d+1
Type 2
d
d+1
d+1
Type 3



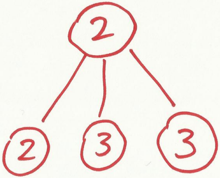
Type 2

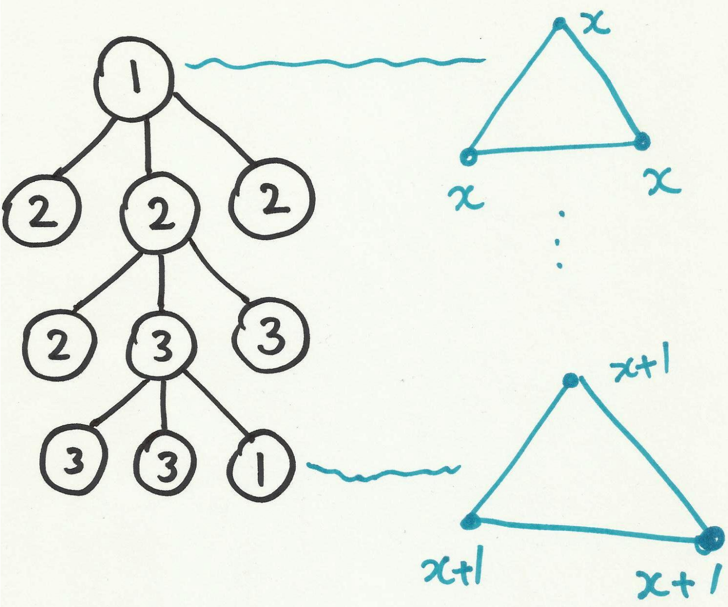


Type 2

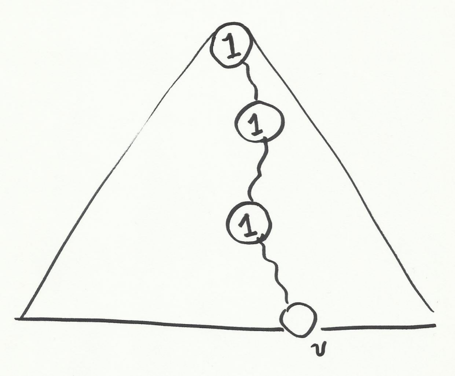


Type2

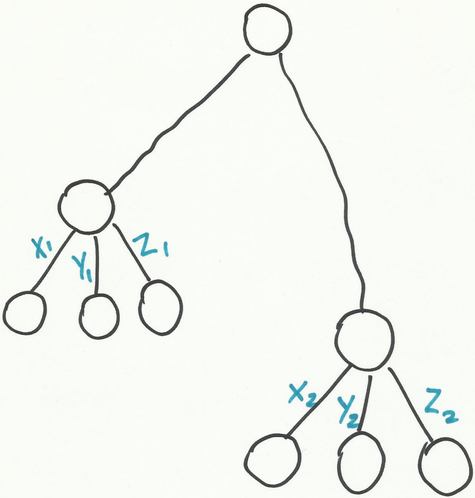



1
2
2
2
2
3
3
3
3
1
x
x
x
x+1
x+1
x+1

Crucial Observation



Distance of a vertex to the boundary
equals number of type-1 nodes
on path of the corresponding node to the root



$$X_1, Y_1, Z_1, X_2, Y_2, Z_2 \overset{d}{\sim} E$$

$$(X_1, Y_1, Z_1) \perp (X_2, Y_2, Z_2)$$

Broutin-Devroye's Theorem

Theorem (Broutin and Devroye 2006)

$E :=$ a positive random variable

$b :=$ a positive integer

$T_\infty :=$ an infinite b-ary tree.

Label the edges of T_∞ randomly,

1. *The label of every edge is distributed like E.*
2. *For vertices u and v, edges going down from u and v are independent.*

$H_t :=$ height of the subtree containing nodes whose sum of labels on their path to root $\leq t$.

Then $\frac{H_t}{t} \to \rho$ in probability

$\rho :=$ unique solution to

$$\sup\{\lambda/\rho - \log(\mathbb{E}[\exp(\lambda E)]) : \lambda \leq 0\} = \log b\,.$$

Two Difficulties

1. Branches are not independent.
2. We do not want the height!

Our Result on the Diameter

Theorem (EFGMSWZ'12+)

$$f(x) := \frac{12x^3}{1-2x} - \frac{6x^3}{1-x},$$

$y :=$ unique solution to

$$x(x-1)f'(x) = f(x)\log f(x), \quad x \in (0, 1/2),$$

$$c := (1 - y^{-1})/\log f(y) \approx 1.668$$

Then for every fixed $\varepsilon > 0$,

$$\mathbb{P}\left[(1-\varepsilon)c\log n \le \text{diameter} \le (1+\varepsilon)c\log n\right] \to 1$$

Open Problems

Concentration of Diameter

We showed

$$\mathbb{P}\left[1.667\log n \le \text{diameter} \le 1.669\log n\right] \longrightarrow 1$$

How much can the diameter deviate from its expected value?

Open Problems

The Longest Path

We showed $\exists \theta > 0$ such that

$$\mathbb{P}\left[L_m < m/(\log m)^{\theta}\right] \to 1$$

and

$$L_m > m^{0.63}$$

and

$$\mathbb{E}\left[L_m\right] = \Omega\left(m^{0.88}\right)$$

All these bounds can perhaps be improved.
Concentration of L_m around its expected value?

Open Problems

Cheeger constant

Definition (Cheeger constant)

$$\min\left\{\frac{|E(S, V\setminus S)|}{|S|} : |S| \le \frac{n}{2}\right\}$$

Open Problems

Cheeger constant

Definition (Cheeger constant)

$$\min\left\{\frac{|E(S, V\setminus S)|}{|S|} : |S| \leq \frac{n}{2}\right\}$$

Frieze and Tsourakakis: maximum degree is $O(\sqrt{n})$.

$$\text{Cheeger constant } \leq \frac{O(\sqrt{n})}{n/6} = O\left(\frac{1}{\sqrt{n}}\right)$$

Is this bound tight?

Thanks for your attention!