A Geometric Preferential Attachment Model of Networks II

Abraham D. Flaxman,
Microsoft Research

Alan M. Frieze,
Carnegie Mellon University

Juan Vera
University of Waterloo

December 11, 2007

Outline

- Introduction
  - Preferential Attachment and its relatives
- Model
  - Geometric Preferential Attachment I
  - Geometric Preferential Attachment II
- Results
  - Theorems
  - Proof techniques
- Conclusion

The Preferential Attachment Graph

- Build a graph dynamically. At time t have $G_t = (V_t, E_t)$.

The Preferential Attachment Graph

- Build a graph dynamically. At time t have $G_t = (V_t, E_t)$.
- At time t, add vertex v_t, and connect it randomly to m neighbors

The Preferential Attachment Graph

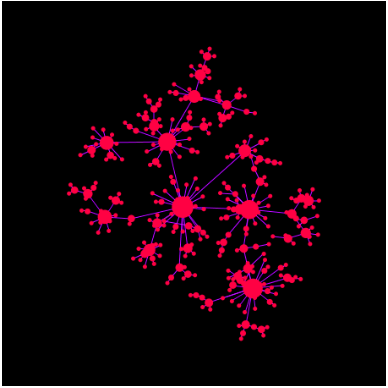
- Build a graph dynamically. At time t have $G_t = (V_t, E_t)$.
- At time t, add vertex v_t, and connect it randomly to m neighbors, with probability given by:

$$\Pr[v_t \to w] = \frac{1}{Z} \deg_t(w).$$

The Preferential Attachment Graph

- Build a graph dynamically. At time t have $G_t = (V_t, E_t)$.
- At time t, add vertex v_t, and connect it randomly to m neighbors, with probability given by:

$$\Pr[v_t \to w] = \frac{1}{Z} \deg_t(w).$$

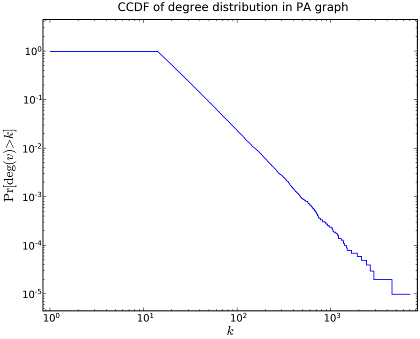


Powerlaw degree distribution

PA graph has a “scale-free” degree distribution:

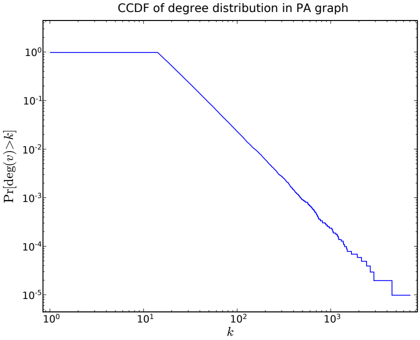
Powerlaw degree distribution

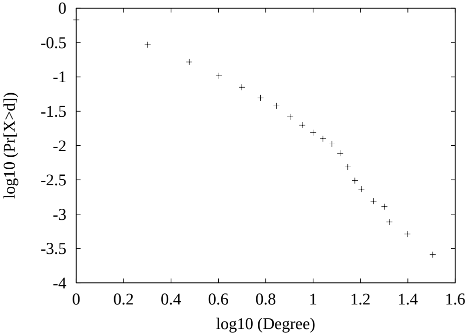
PA graph has a “scale-free” degree distribution:



Powerlaw degree distribution

PA graph has a "scale-free" degree distribution:





(a) Pansiot-Grad

Modifications

It’s fun to analyze, it looks like some graphs from the real-world.
Let’s consider the many possible modifications:

Modifications

It's fun to analyze, it looks like some graphs from the real-world.
Let's consider the many possible modifications:

New concept or mechanism	Limits of γ	Reference
Linear growth, linear pref. attachment	$\gamma=3$	Barabási and Albert, 1999
Nonlinear preferential attachment $\Pi(k_i)\sim k_i^{\alpha}$	no scaling for $\alpha\neq 1$	Krapivsky, Redner, and Leyvraz, 2000
Asymptotically linear pref. attachment $\Pi(k_i)\sim a_{\infty}k_i$ as $k_i\to\infty$	$\gamma\to 2$ if $a_{\infty}\to\infty$ $\gamma\to\infty$ if $a_{\infty}\to 0$	Krapivsky, Redner, and Leyvraz, 2000
Initial attractiveness $\Pi(k_i)\sim A+k_i$	$\gamma=2$ if $A=0$ $\gamma\to\infty$ if $A\to\infty$	Dorogovtsev, Mendes, and Samukhin, 2000a, 2000b
Accelerating growth $\langle k\rangle\sim t^{\theta}$ constant initial attractiveness	$\gamma=1.5$ if $\theta\to 1$ $\gamma\to 2$ if $\theta\to 0$	Dorogovtsev and Mendes, 2001a
Accelerating growth $\langle k\rangle=at+2b$	$\gamma=1.5$ for $k\ll k_c(t)$ $\gamma=3$ for $k\gg k_c(t)$	Barabási *et al.*, 2001 Dorogovtsev and Mendes, 2001c
Internal edges with probab. p	$\gamma=2$ if $q=\frac{1-p+m}{1+2m}$	
Rewiring of edges with probab. q	$\gamma\to\infty$ if $p,q,m\to 0$	Albert and Barabási, 2000
c internal edges or removal of c edges	$\gamma\to 2$ if $c\to\infty$ $\gamma\to\infty$ if $c\to -1$	Dorogovtsev and Mendes, 2000c
Gradual aging $\Pi(k_i)\sim k_i(t-t_i)^{-\nu}$	$\gamma\to 2$ if $\nu\to-\infty$ $\gamma\to\infty$ if $\nu\to 1$	Dorogovtsev and Mendes, 2000b
Multiplicative node fitness $\Pi_i\sim\eta_i k_i$	$P(k)\sim\frac{k^{-1-C}}{\ln(k)}$	Bianconi and Barabási, 2001a
Additive-multiplicative fitness $\Pi_i\sim\eta_i(k_i-1)+\zeta_i$	$P(k)\sim\frac{k^{-1-m}}{\ln(k)}$ $1\leq m\leq 2$	Ergün and Rodgers, 2001
Edge inheritance	$P(k_{in})=\frac{d}{k_{in}^{\sqrt{2}}}\ln(ak_{in})$	Dorogovtsev, Mendes, and Samukhin, 2000c
Copying with probab. p	$\gamma=(2-p)/(1-p)$	Kumar *et al.*, 2000a, 2000b
Redirection with probab. r	$\gamma=1+1/r$	Krapivsky and Redner, 2001
Walking with probab. p	$\gamma\simeq 2$ for $p>p_c$	Vázquez, 2000
Attaching to edges	$\gamma=3$	Dorogovtsev, Mendes, and Samukhin, 2001a
p directed internal edges $\Pi(k_i,k_j)\propto(k_i^{in}+\lambda)(k_j^{out}+\mu)$	$\gamma_{in}=2+p\lambda$ $\gamma_{out}=1+(1-p)^{-1}+\mu p/(1-p)$	Krapivsky, Rodgers, and Redner, 2001
$1-p$ directed internal edges Shifted linear pref. activity	$\gamma_{in}=2+p$ $\gamma_{out}\simeq 2+3p$	Tadić, 2001a

Modifications

It's fun to analyze, it looks like some graphs from the real-world.
Let's consider the many possible modifications:

New concept or mechanism	Limits of γ	Reference
Linear growth, linear pref. attachment	$\gamma=3$	Barabási and Albert, 1999
Nonlinear preferential attachment $\Pi(k_i)\sim k_i^{\alpha}$	no scaling for $\alpha\neq 1$	Krapivsky, Redner, and Leyvraz, 2000
Asymptotically linear pref. attachment $\Pi(k_i)\sim a_{\infty}k_i$ as $k_i\to\infty$	$\gamma\to 2$ if $a_{\infty}\to\infty$ $\gamma\to\infty$ if $a_{\infty}\to 0$	Krapivsky, Redner, and Leyvraz, 2000
Initial attractiveness $\Pi(k_i)\sim A+k_i$	$\gamma=2$ if $A=0$ $\gamma\to\infty$ if $A\to\infty$	Dorogovtsev, Mendes, and Samukhin, 2000a, 2000b
Accelerating growth $\langle k\rangle\sim t^{\theta}$ constant initial attractiveness	$\gamma=1.5$ if $\theta\to 1$ $\gamma\to 2$ if $\theta\to 0$	Dorogovtsev and Mendes, 2001a
Accelerating growth $\langle k\rangle=at+2b$	$\gamma=1.5$ for $k\ll k_c(t)$ $\gamma=3$ for $k\gg k_c(t)$	Barabási *et al.*, 2001 Dorogovtsev and Mendes, 2001c
Internal edges with probab. p	$\gamma=2$ if $q=\frac{1-p+m}{1+2m}$	
Rewiring of edges with probab. q	$\gamma\to\infty$ if $p,q,m\to 0$	Albert and Barabási, 2000
c internal edges or removal of c edges	$\gamma\to 2$ if $c\to\infty$ $\gamma\to\infty$ if $c\to -1$	Dorogovtsev and Mendes, 2000c
Gradual aging $\Pi(k_i)\sim k_i(t-t_i)^{-\nu}$	$\gamma\to 2$ if $\nu\to -\infty$ $\gamma\to\infty$ if $\nu\to 1$	Dorogovtsev and Mendes, 2000b
Multiplicative node fitness $\Pi_i\sim\eta_i k_i$	$P(k)\sim\frac{k^{-1-C}}{\ln(k)}$	Bianconi and Barabási, 2001a
Additive-multiplicative fitness $\Pi_i\sim\eta_i(k_i-1)+\zeta_i$	$P(k)\sim\frac{k^{-1-m}}{\ln(k)}$ $1\leq m\leq 2$	Ergün and Rodgers, 2001
Edge inheritance	$P(k_{in})=\frac{d}{k_{in}^{\sqrt{2}}}\ln(ak_{in})$	Dorogovtsev, Mendes, and Samukhin, 2000c
Copying with probab. p	$\gamma=(2-p)/(1-p)$	Kumar *et al.*, 2000a, 2000b
Redirection with probab. r	$\gamma=1+1/r$	Krapivsky and Redner, 2001
Walking with probab. p	$\gamma\simeq 2$ for $p>p_c$	Vázquez, 2000
Attaching to edges	$\gamma=3$	Dorogovtsev, Mendes, and Samukhin, 2001a
p directed internal edges $\Pi(k_i,k_j)\propto(k_i^{in}+\lambda)(k_j^{out}+\mu)$	$\gamma_{in}=2+p\lambda$ $\gamma_{out}=1+(1-p)^{-1}+\mu p/(1-p)$	Krapivsky, Rodgers, and Redner, 2001
$1-p$ directed internal edges Shifted linear pref. activity	$\gamma_{in}=2+p$ $\gamma_{out}\simeq 2+3p$	Tadić, 2001a

[Barabási, A.-L., and R. Albert, Statistical mechanics of complex networks, Reviews of Modern Physics, Vol 74, page 47-97, 2002.]

One modification that's missing from list

One modification that's missing from list

Underlying geometry of vertices

One modification that’s missing from list

Underlying geometry of vertices:

- A feature nodes have in many real-world networks.

One modification that's missing from list

Underlying geometry of vertices:

- A feature nodes have in many real-world networks.
- Often a reasonable hypothesis even when the nodes do not explicitly live in a metric space.

Central Question in this talk

How does underlying geometric structure affect preferential attachment?

Geometric PA I

Old setup (Geo-PA-I):

- Every vertex v is a uniformly random point on the surface of a 3-dimensional sphere.

Geometric PA I

Old setup (Geo-PA-I):

- Every vertex v is a uniformly random point on the surface of a 3-dimensional sphere.
- At time t, add vertex v_t, and connect it randomly to m neighbors

Geometric PA I

Old setup (Geo-PA-I):

- Every vertex v is a uniformly random point on the surface of a 3-dimensional sphere.
- At time t, add vertex v_t, and connect it randomly to m neighbors, from *only* neighbors within critical radius r

Geometric PA I

Old setup (Geo-PA-I):

- Every vertex v is a uniformly random point on the surface of a 3-dimensional sphere.
- At time t, add vertex v_t, and connect it randomly to m neighbors, from *only* neighbors within critical radius r, with probability given by:

$$\Pr[v_t \to w] = \begin{cases} \frac{1}{Z}\deg_t(w) & \text{if } \|v_t - w\| \leq r; \\ 0 & \text{otherwise.} \end{cases}$$

Geometric PA I

Old setup (Geo-PA-I):

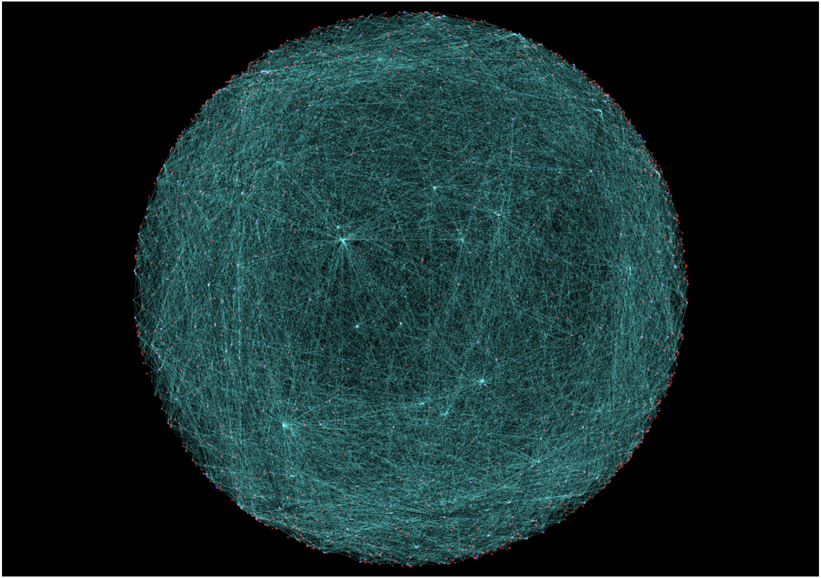
- Every vertex v is a uniformly random point on the surface of a 3-dimensional sphere.
- At time t, add vertex v_t, and connect it randomly to m neighbors, from *only* neighbors within critical radius r, with probability given by:

$$\Pr[v_t \to w] = \begin{cases} \frac{1}{Z}\deg_t(w) & \text{if } \|v_t - w\| \le r; \\ 0 & \text{otherwise.} \end{cases}$$

- We would like to take normalization Z to be

$$T_t(v_t) = \sum_{w:\|v_t - w\| \le r} \deg_t(w).$$

Geometric PA I Image



Geometric PA II

Introduce *affinity function* $F : \mathbb{R}_+ \to \mathbb{R}_+$.

Geometric PA II

Introduce *affinity function* $F : \mathbb{R}_+ \to \mathbb{R}_+$.

- At time t, add vertex v_t, and connect it randomly to m neighbors, with probability given by

$$\Pr[v_t \to w] = \frac{1}{Z}\deg_t(w) \times F\big(\|v_t - w\|\big)$$

Geometric PA II

Introduce *affinity function* $F : \mathbb{R}_+ \to \mathbb{R}_+$.

- At time t, add vertex v_t, and connect it randomly to m neighbors, with probability given by

$$\Pr[v_t \to w] = \frac{1}{Z} \deg_t(w) \times F\big(\|v_t - w\|\big),$$

where $Z = \max\{T_t(v_t), \alpha m t I\}$, with

Geometric PA II

Introduce *affinity function* $F : \mathbb{R}_+ \to \mathbb{R}_+$.

- At time t, add vertex v_t, and connect it randomly to m neighbors, with probability given by

$$\Pr[v_t \to w] = \frac{1}{Z} \deg_t(w) \times F\big(\|v_t - w\|\big),$$

  where $Z = \max\{T_t(v_t), \alpha m t I\}$, with
  - $T_t(v_t) = \sum_{w \in V_t} \deg_t(w) F\big(\|v_t - w\|\big)$,

Geometric PA II

Introduce *affinity function* $F : \mathbb{R}_+ \to \mathbb{R}_+$.

- At time t, add vertex v_t, and connect it randomly to m neighbors, with probability given by

$$\Pr[v_t \to w] = \frac{1}{Z} \deg_t(w) \times F\left(\|v_t - w\|\right),$$

where $Z = \max\left\{T_t(v_t), \alpha m t I\right\}$, with
  - $T_t(v_t) = \sum_{w \in V_t} \deg_t(w) F\left(\|v_t - w\|\right)$,
  - $I = \int_{S^2} F\left(\|w - v_t\|\right) dw$,

Geometric PA II

Introduce *affinity function* $F : \mathbb{R}_+ \to \mathbb{R}_+$.

- At time t, add vertex v_t, and connect it randomly to m neighbors, with probability given by

$$\Pr[v_t \to w] = \frac{1}{Z} \deg_t(w) \times F\big(\|v_t - w\|\big),$$

  where $Z = \max\{T_t(v_t), \alpha m t I\}$, with
  - $T_t(v_t) = \sum_{w \in V_t} \deg_t(w) F\big(\|v_t - w\|\big)$,
  - $I = \int_{S^2} F\big(\|w - v_t\|\big) dw$,
  - α is bias towards self loops.

Geometric PA II

Introduce *affinity function* $F : \mathbb{R}_+ \to \mathbb{R}_+$.

- At time t, add vertex v_t, and connect it randomly to m neighbors, with probability given by

$$\Pr[v_t \to w] = \frac{1}{Z} \deg_t(w) \times F\big(\|v_t - w\|\big),$$

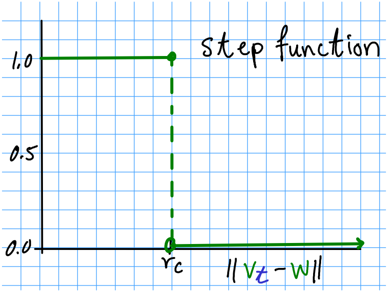
where $Z = \max\{T_t(v_t), \alpha m t I\}$, with

  - $T_t(v_t) = \sum_{w \in V_t} \deg_t(w) F\big(\|v_t - w\|\big)$,
  - $I = \int_{S^2} F\big(\|w - v_t\|\big) dw$,
  - α is bias towards self loops.

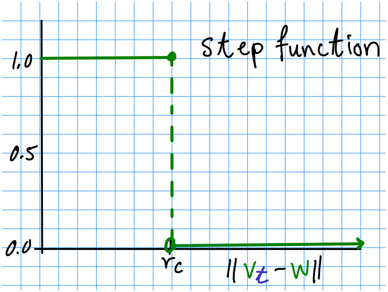
Restrictions on F: I must exist, $0 < I < \infty$.

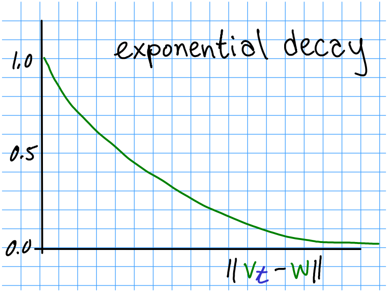
Prototypical affinity functions:

Prototypical affinity functions:

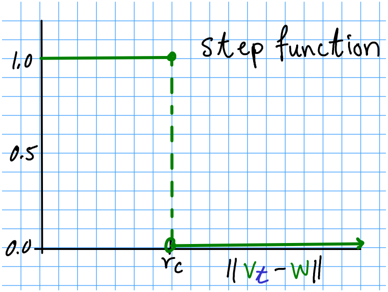


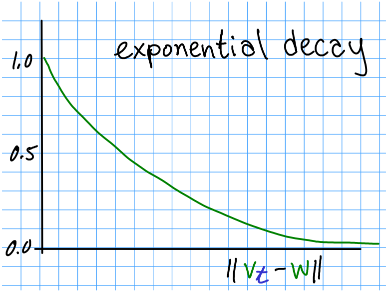
Prototypical affinity functions:

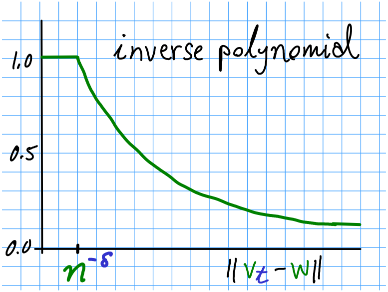




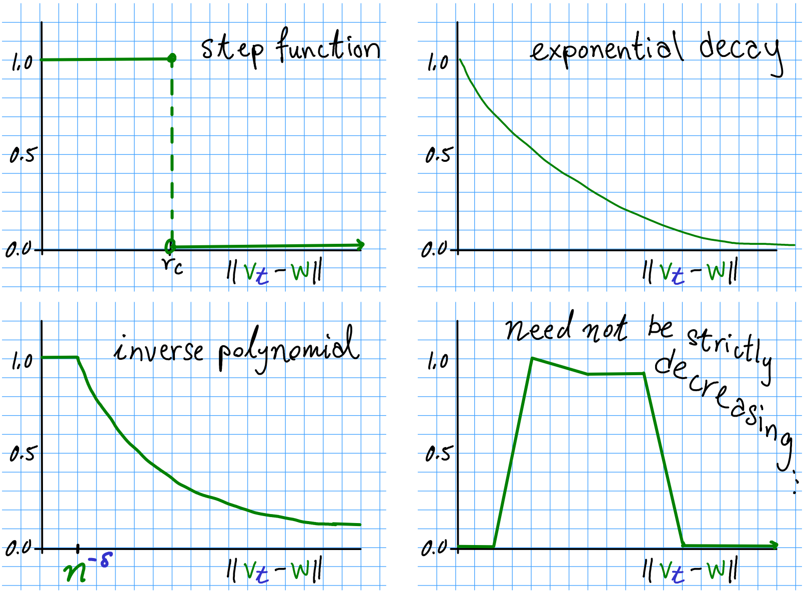
Prototypical affinity functions:







Prototypical affinity functions:



What happens?

In the Geo-PA-II model, what do you think happens to:

What happens?

In the Geo-PA-II model, what do you think happens to:

- The degree distribution?

What happens?

In the Geo-PA-II model, what do you think happens to:

- The degree distribution?
- The conductance/sparsest cut?

What happens?

In the Geo-PA-II model, what do you think happens to:

- The degree distribution?
- The conductance/sparsest cut?
- The diameter?

Degree distribution

Theorem

Degree distribution

Theorem

For $\alpha > 2$ *and*

$$\int_0^{\pi} F(x)^2 \sin x dx = \mathcal{O}\left(t^{1-\epsilon} l^2\right),$$

Degree distribution

Theorem

For $\alpha > 2$ *and*

$$\int_0^{\pi} F(x)^2 \sin x dx = \mathcal{O}\left(t^{1-\epsilon} l^2\right),$$

we have

$$\mathrm{E}\left[\#\{w : \deg_t(w) = k\}\right] = C_k(m, \alpha)\left(\frac{m}{k}\right)^{1+\alpha} t + \mathcal{O}(t^{1-\delta}),$$

Degree distribution

Theorem

For $\alpha > 2$ *and*

$$\int_0^{\pi} F(x)^2 \sin x dx = \mathcal{O}\left(t^{1-\epsilon} l^2\right),$$

we have

$$\mathrm{E}\left[\#\{w : \deg_t(w) = k\}\right] = C_k(m, \alpha)\left(\frac{m}{k}\right)^{1+\alpha} t + \mathcal{O}(t^{1-\delta}),$$

where

$$C_k(m, \alpha) \to C_\infty(m, \alpha) \textit{ as } k \to \infty.$$

Degree distribution

Theorem

For $\alpha > 2$ *and*

$$\int_0^\pi F(x)^2 \sin x dx = \mathcal{O}\left(t^{1-\epsilon} l^2\right),$$

we have

$$\mathrm{E}\left[\#\{w : \deg_t(w) = k\}\right] = C_k(m,\alpha)\left(\frac{m}{k}\right)^{1+\alpha} t + \mathcal{O}(t^{1-\delta}),$$

where

$$C_k(m,\alpha) \to C_\infty(m,\alpha) \text{ as } k \to \infty.$$

(We also have a concentration result.)

Conductance/Sparsest cut

Theorem

For $\alpha > 0$ and m a sufficiently large constant, if there exist ϕ and η with

$$\frac{1}{n} \ll \phi \ll 1 \text{ and } \eta \ll 1$$

such that

$$\frac{1}{2} \int_{\eta}^{\pi} F(x) \sin x \, dx \leq \phi I$$

then the cut induced by a great circle of the sphere contains $\mathcal{O}((\eta + \phi)mn)$ edges **whp**.

Conductance/Sparsest cut

Example:

$$F(x) = \min\left\{ n^{\delta\beta}, \frac{1}{x^{\beta}} \right\}.$$

Conductance/Sparsest cut

Example:

$$F(x) = \min\left\{ n^{\delta\beta}, \frac{1}{x^{\beta}} \right\}.$$

For $\beta > 2$, get

$$e(S, \bar{S})/|S| = \mathcal{O}\left(mn^{-\delta(\beta-1)} \right).$$

Conductance/Sparsest cut

Example:

$$F(x) = \min \left\{ n^{\delta\beta}, \frac{1}{x^{\beta}} \right\}.$$

For $\beta > 2$, get

$$e(S, \bar{S})/|S| = \mathcal{O}\left(mn^{-\delta(\beta-1)} \right).$$

For $\beta = 2$, get

$$e(S, \bar{S})/|S| = \mathcal{O}\left(\frac{m \log \log n}{\log n} \right).$$

Conductance/Sparsest cut

Example:

$$F(x) = \min\left\{n^{\delta\beta}, \frac{1}{x^{\beta}}\right\}.$$

For $\beta > 2$, get

$$e(S, \bar{S})/|S| = \mathcal{O}\left(mn^{-\delta(\beta-1)}\right).$$

For $\beta = 2$, get

$$e(S, \bar{S})/|S| = \mathcal{O}\left(\frac{m \log\log n}{\log n}\right).$$

For $\beta < 2$, G is an expander.

Expander Criteria

Call F *tame* if exist constants C_1, C_2 such that

- $F(x) \geq C_1$ for $0 \leq x \leq \pi$,
- $I \leq C_2$.

Theorem

If $\alpha > 2$, F is tame, and $m \geq K \log n$ for sufficiently large K, then **whp**

- *G_n has conductance bounded below by a constant.*
- *G_n is connected.*
- *G_n has diameter $\mathcal{O}(\log n / \log m)$.*

Diameter

We also have some results for diameter when affinity function is not tame.

Lemma 1: a simple expectation

Lemma

For u chosen u.a.r. in S^2 and $t > 0$, we have

$$\mathrm{E}[T_t(u)] = 2lmt.$$

Proof

$$\begin{aligned}
\mathrm{E}[T_t(u)] &= \mathrm{E}\left[\sum_{w \in V_t} \deg_t(w) F(\|u - w\|)\right] \\
&= \sum_{w \in V_t} \deg_t(w) \int_{S^2} F(\|u - w\|) dw \\
&= \sum_{w \in V_t} \deg_t(w) I = 2Imt.
\end{aligned}$$

□

Lemma 2: a not-so-simple concentration inequality

Lemma

For any $t > 0$ *and for* u *chosen u.a.r. in* S^2*,*

$$\Pr\left[\left|T_t(u) - 2lmt\right| \geq ml(t^{2/\alpha} + t^{1/2}\ln t)\ln n\right] = \mathcal{O}\left(n^{-2}\right).$$

Proof by Azuma-Hoeffding, using a coupling argument.

Summary

Geo-PA-II: choose your own affinity function $F(x)$.

- Degree distribution has power $1 + \alpha$.
- Expander/Sparse cuts depend on $F(x)$.
- Diameter does as well.
- Proof uses tight concentration, coupling.

Future work

- Technical work:
  - $\alpha = 2$ (i.e. remove α)
  - non-uniform random points
  - necess. and suff. condition on F for expansion
- Modelling work: The sparse cuts are “wrong”.

Future work: getting sparse cuts right