Kronecker graphs

Jure Leskovec (jure@cs.stanford.edu)
Computer Science Department
Cornell University / Stanford University

Joint work with: Deepay Chakrabarty, Jon Kleinberg and Christos Faloutsos

Mechanistic vs. Statistical

- Mechanistic/algorithmic type models for macro-scale:
  - E.g.: Preferential attachment, Small-world
  - Good for predicting macro scale properties
  - Bad at predicting individual edges
  - Analytically tractable
- Statistical models to model micro-scale:
  - E.g.: Exponential random graphs
  - Good for modeling diads, reciprocity, etc.
  - Troubles with macro scale (e.g., power-law degree distributions)
  - Computationally expensive to fit
  - Not analytically tractable

Can we have both?

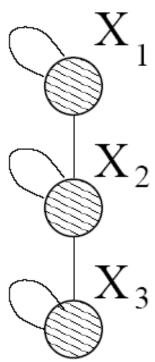
- A model that brings the two worlds together
  - A model that can model heterogeneous nodes while having the macro-scale properties naturally emerge
- Would like to fit the model to large networks

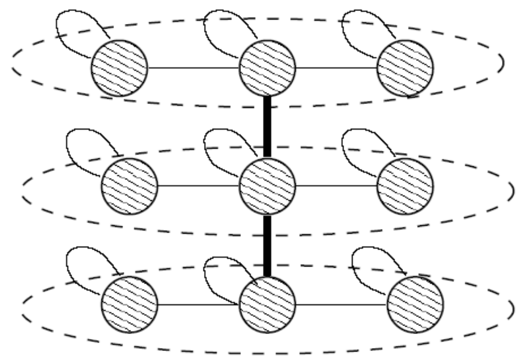
- → **Kronecker graphs:**
  - Can be analytically analyzed
  - Can be estimated from large networks

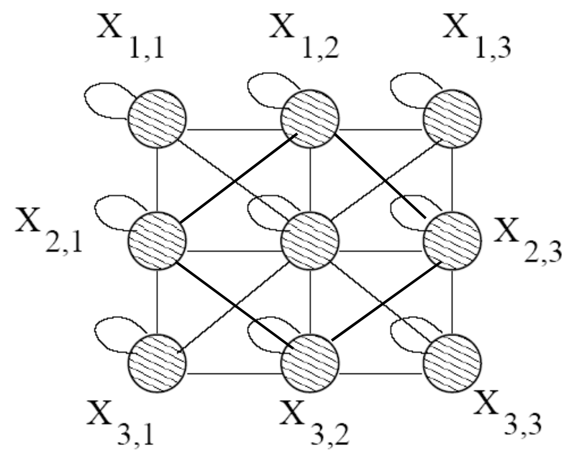
- **Idea:** Generate graphs recursively

[w/ Chakrabarti-Kleinberg-Faloutsos, PKDD '05]

Kronecker product: Graph







Intermediate stage

1	1	0
1	1	1
0	1	1

(3x3)

$$G_1$$

Adjacency matrix

G_1	G_1	0
G_1	G_1	G_1
0	G_1	G_1

(9x9)

$$G_2 = G_1 \otimes G_1$$

Adjacency matrix

Kronecker product: Definition

- Kronecker product of matrices A and B is given by

$$\underset{N\,x\,M}{\mathbf{C} = \mathbf{A}} \otimes \underset{K\,x\,L}{\mathbf{B}} \doteq \underset{N{*}K\,x\,M{*}L}{\begin{pmatrix} a_{1,1}\mathbf{B} & a_{1,2}\mathbf{B} & \dots & a_{1,m}\mathbf{B} \\ a_{2,1}\mathbf{B} & a_{2,2}\mathbf{B} & \dots & a_{2,m}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1}\mathbf{B} & a_{n,2}\mathbf{B} & \dots & a_{n,m}\mathbf{B} \end{pmatrix}}$$

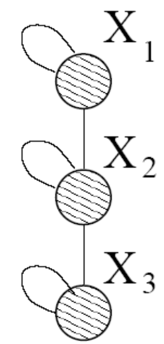
- We define a Kronecker product of two graphs as a Kronecker product of their adjacency matrices

[w/ Chakrabarti-Kleinberg-Faloutsos, PKDD '05]

Kronecker graphs

- **Kronecker graph:** a growing sequence of graphs by iterating the Kronecker product

$$G_k = \underbrace{G_1 \otimes G_1 \otimes \ldots G_1}_{k\ times}$$



G_1

- Each Kronecker multiplication exponentially increases the size of the graph

- One can easily use multiple initiator matrices (G_1', G_1'', G_1''') that can be of different sizes

Kronecker product: Graph

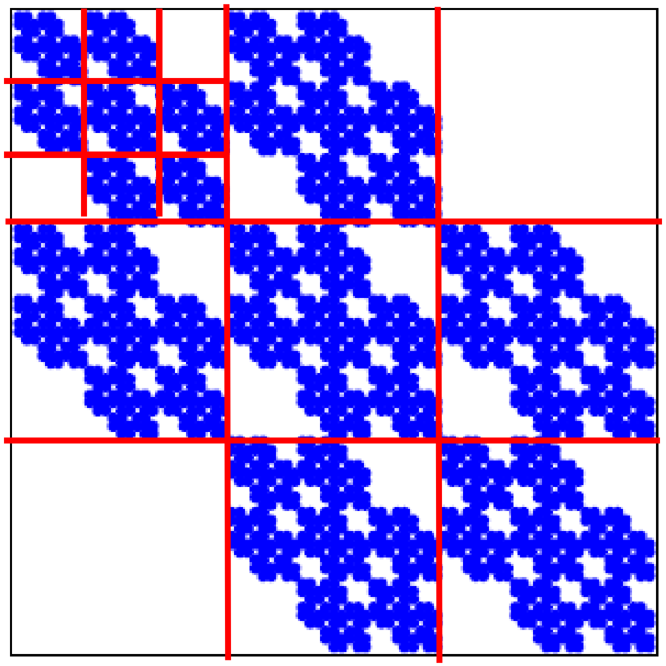
- Continuing multypling with G_1 we obtain G_4 and so on ...

1	1	0
1	1	1
0	1	1

G_1

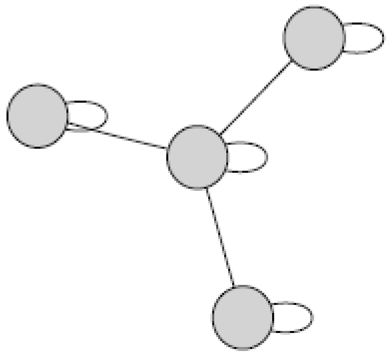
G_1	G_1	0
G_1	G_1	G_1
0	G_1	G_1

$G_2 = G_1 \otimes G_1$

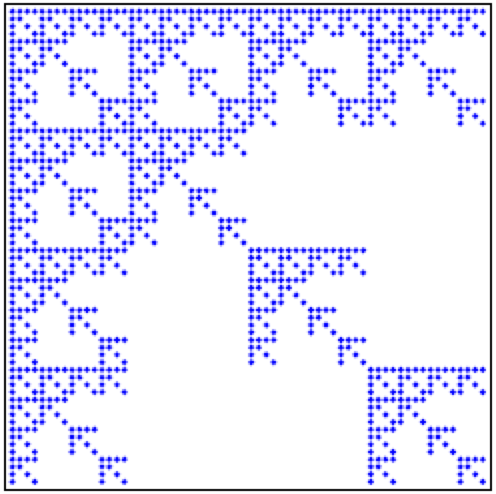


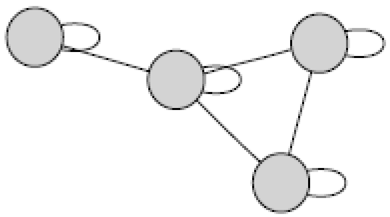
G_4 adjacency matrix

Various initiator matrices

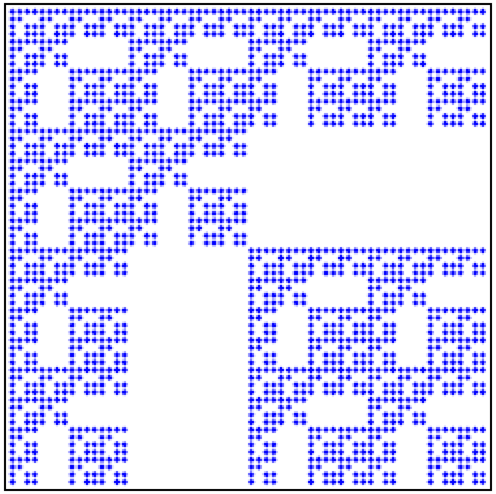


1	1	1	1
1	1	0	0
1	0	1	0
1	0	0	1





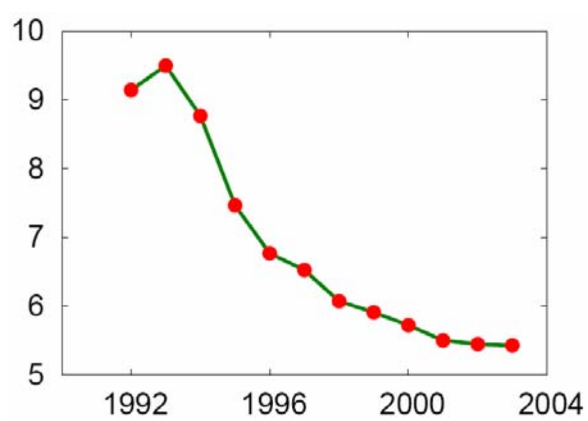
1	1	1	1
1	1	0	0
1	0	1	1
1	0	1	1



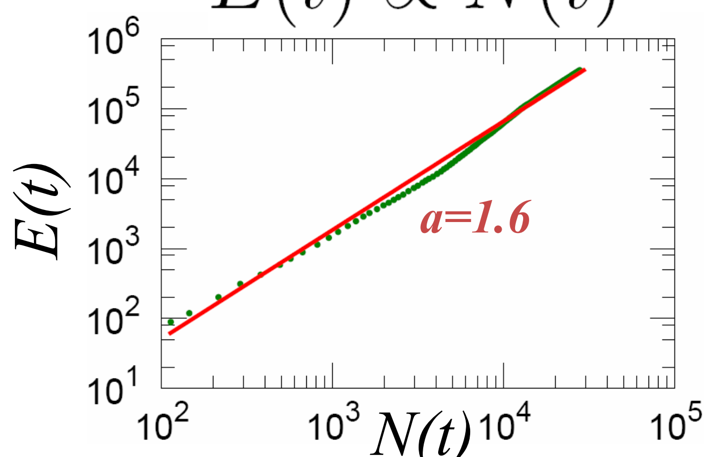
Initiator K_1 | K_1 adjacency matrix | K_3 adjacency matrix

Properties of Kronecker graphs

- We prove Kronecker graphs have:
  - Multinomial Degree Distribution
  - Multinomial eigenvalue and eigenvector distribution
  - Small Diameter
  - Densification Power Law
  - Shrinking/Stabilizing Diameter



$$E(t) \propto N(t)^a$$



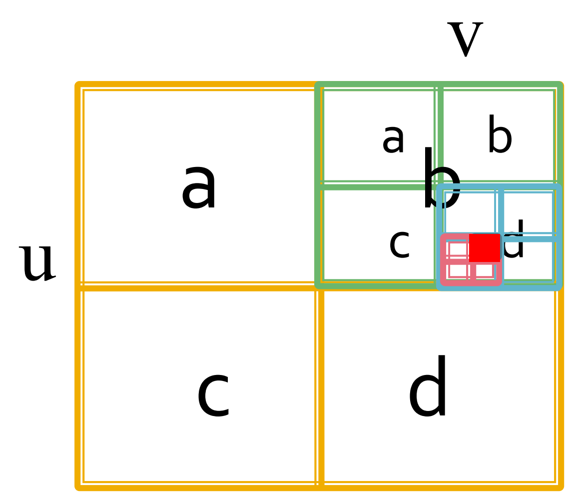
Kronecker graphs: Interpretation

- Initiator matrix is a similarity matrix
- Node u is described with k binary attributes $u_1, u_2, \ldots, u_k$
- Probability of a link between nodes u, v:

$$P(u,v) = \prod G_1[u_i, v_i]$$

$$G_1 = \begin{array}{cc} 0 & 1 \\ \begin{array}{|c|c|} \hline a & b \\ \hline c & d \\ \hline \end{array} & \begin{array}{c} 0 \\ 1 \end{array} \end{array}$$

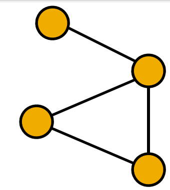
$$\mathrm{u} = (0,1,1,0)$$
$$\mathrm{v} = (1,1,0,1)$$
$$\mathrm{P(u,v)} = \mathrm{b \cdot d \cdot c \cdot b}$$



[w/ Faloutsos, ICML '07]

Estimating the Kronecker model

- Given a graph G and Kronecker matrix Θ we calculate probability that Θ generated G $P(G|\Theta)$



0.5	0.2
0.1	0.3

Θ

Prob. of edge (i,j)

0.25	0.10	0.10	0.04
0.05	0.15	0.02	0.06
0.05	0.02	0.15	0.06
0.01	0.03	0.03	0.09

$\Theta_k = \Theta_1 \otimes \Theta_1$

1	0	1	1
0	1	0	1
1	0	1	1
1	1	1	1

G

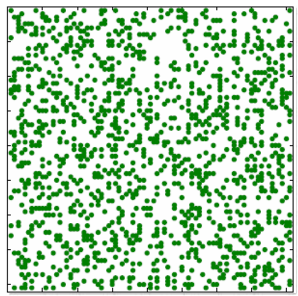
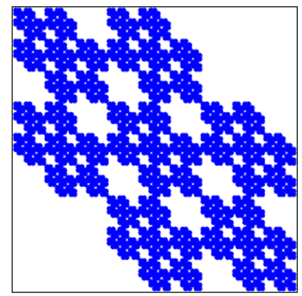
$\boxed{P(G|\Theta)}$

$$P(G|\Theta) = \prod_{(u,v)\in G} \Theta_k[u,v] \prod_{(u,v)\notin G} (1-\Theta_k[u,v])$$

[w/ Faloutsos, ICML '07]

Kronecker graphs: Estimation

- Maximum likelihood estimation

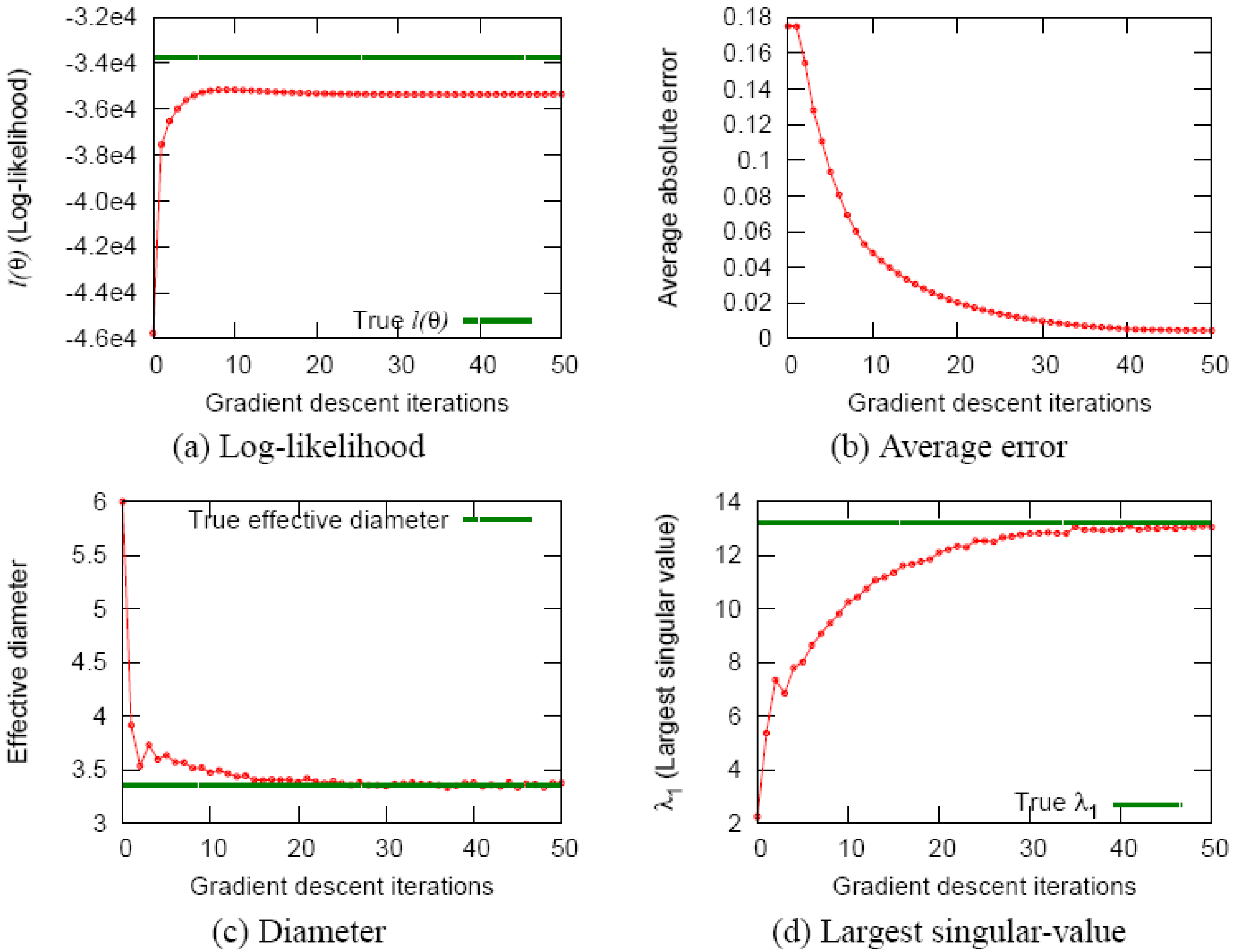
$\arg\max_{G_1} P\Big($  $\Big|$  $\xleftarrow{\text{Kronecker}} G_1 \Big)$

- Naïve estimation takes *O(N!N²)*:
  - *N!* for different node labelings:
    - **Our solution:** Metropolis sampling: *N!* → *(big) const*
  - N^2 for traversing graph adjacency matrix
    - **Our solution:** Kronecker product ($E << N^2$): $N^2 \rightarrow E$
- Do stochastic gradient descent

$G_1 = \begin{array}{|c|c|} \hline a & b \\ \hline c & d \\ \hline \end{array}$

We estimate the model in O(E)

Stochastic gradient descent



(a) Log-likelihood

(b) Average error

(c) Diameter

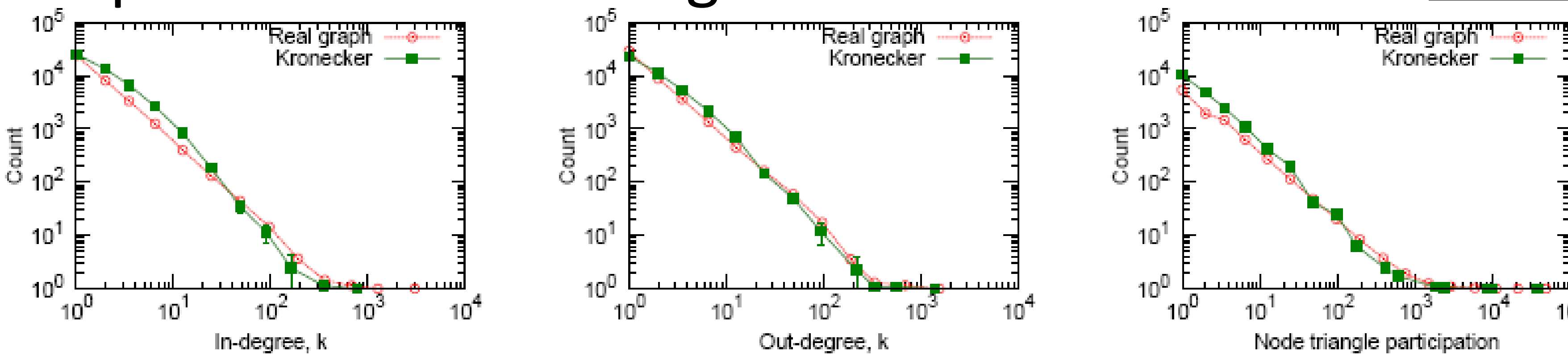
(d) Largest singular-value

Generate a Kronecker graph, start a gradient descent from a random point

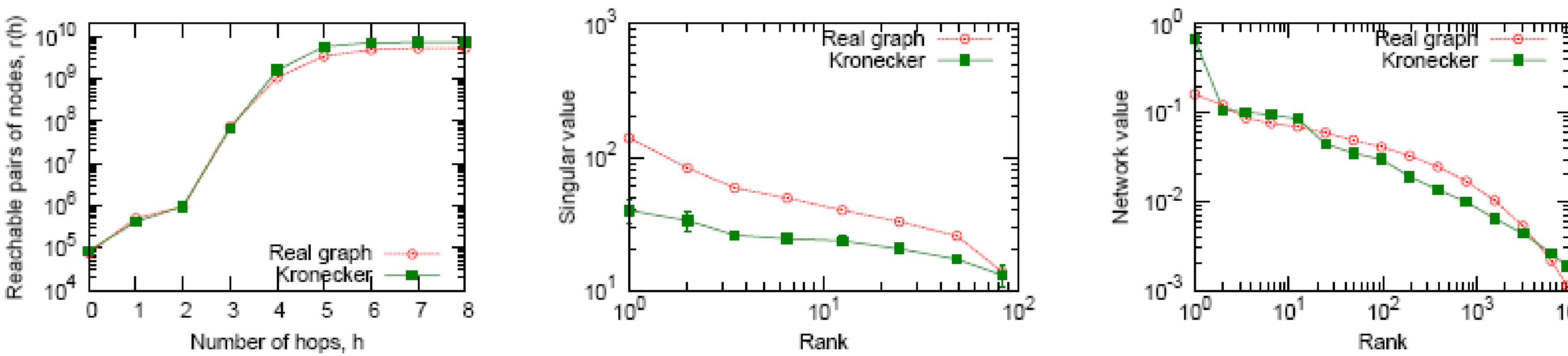
Estimation: Epinions (N=76k, E=510k)

- We search the space of $\sim 10^{1,000,000}$ permutations. Fitting takes 2 hours

$$G_1 = \begin{array}{|c|c|} \hline 0.99 & 0.54 \\ \hline 0.49 & 0.13 \\ \hline \end{array}$$



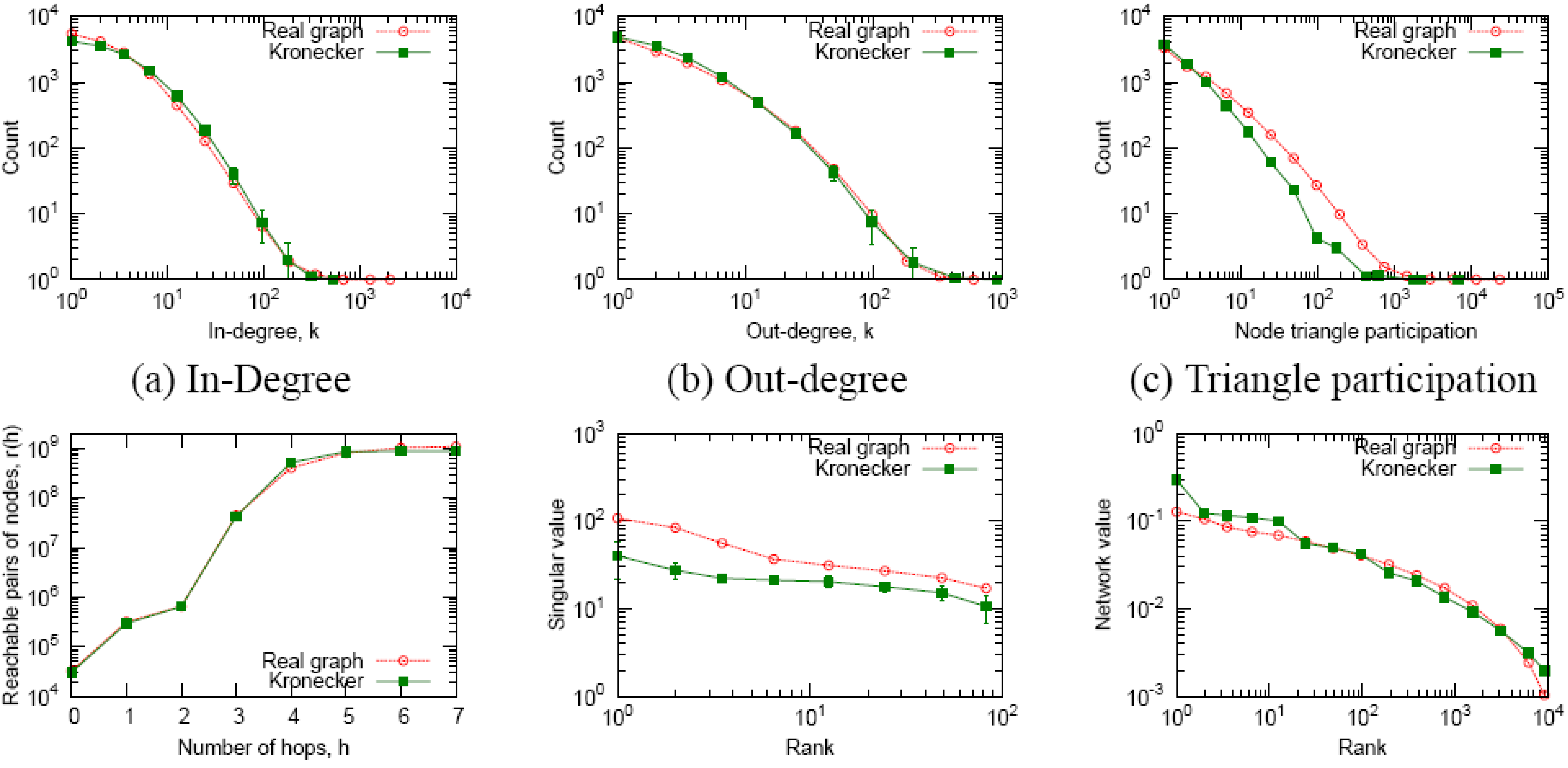
(a) In-Degree (b) Out-degree (c) Triangle participation



(d) Hop plot (e) Scree plot (f) "Network" value

Estimation: Blogs (N=32k, E=318k)

$$G_1 = \begin{array}{|c|c|} \hline 0.99 & 0.58 \\ \hline 0.51 & 0.22 \\ \hline \end{array}$$



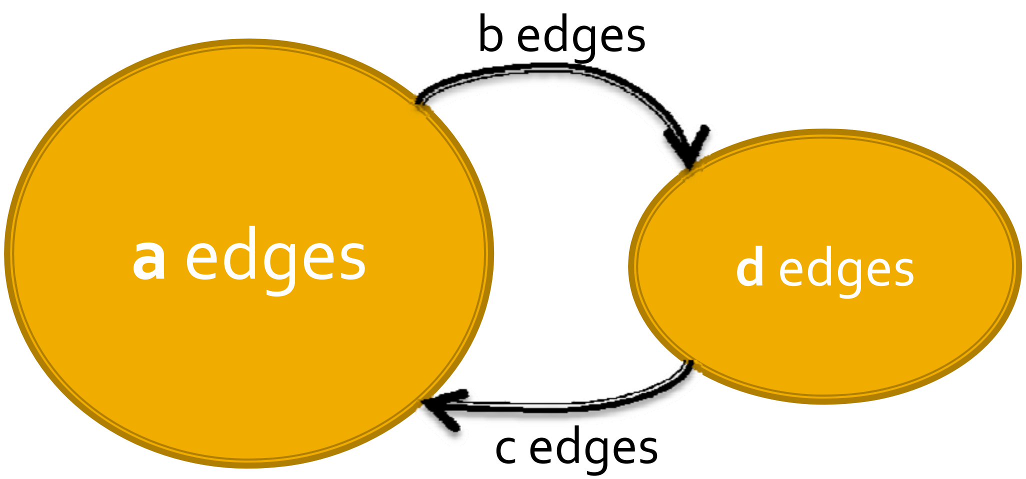
(a) In-Degree (b) Out-degree (c) Triangle participation

(d) Hop plot (e) Scree plot (f) "Network" value

Kronecker & Network structure

- What do estimated parameters tell us about the network structure?

$$G_1 = \begin{array}{|c|c|} \hline a & b \\ \hline c & d \\ \hline \end{array}$$



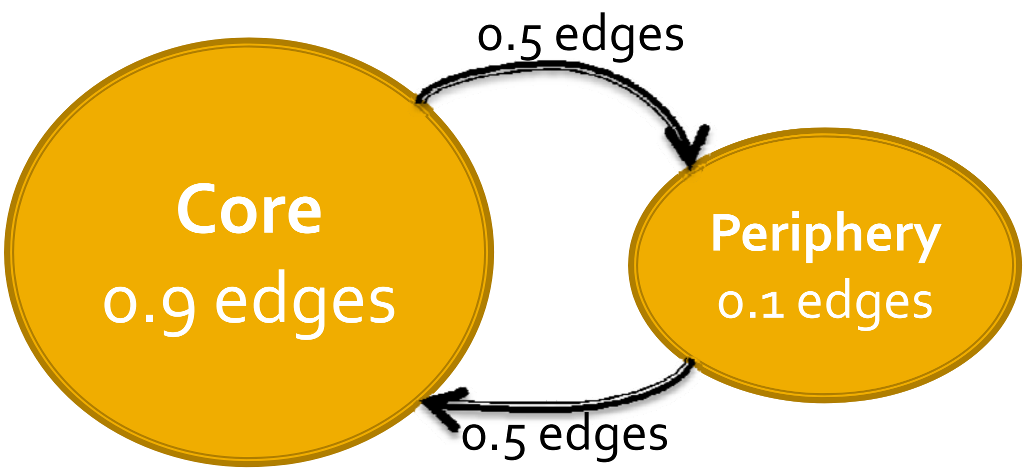
and each sub community has a similar structure

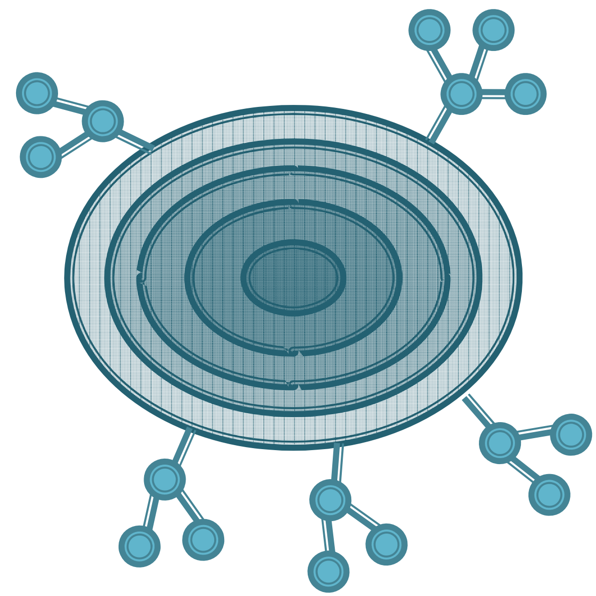
[w/ Dasgupta-Lang-Mahoney, WWW ’08]

Kronecker & Network structure

- What do estimated parameters tell us about the network structure?

$$G_1 = \begin{array}{|c|c|}\hline 0.9 & 0.5 \\ \hline 0.5 & 0.1 \\ \hline \end{array}$$

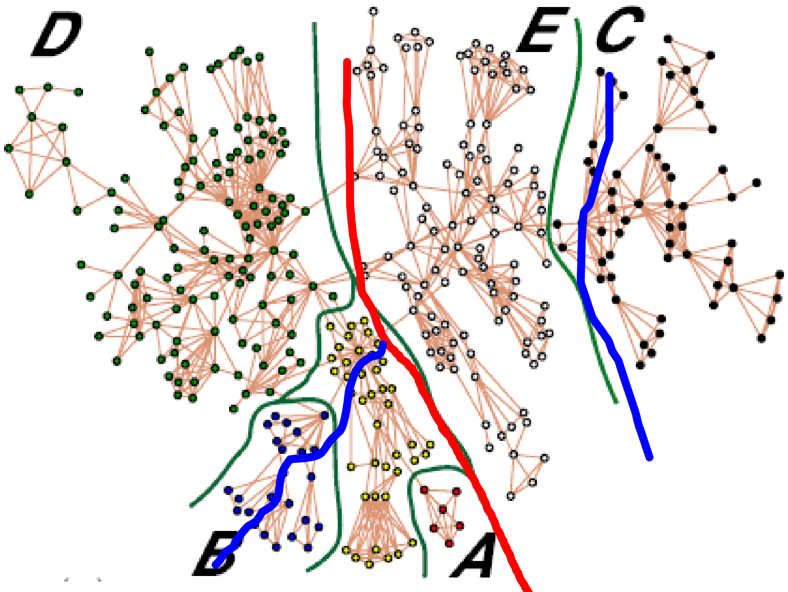




Recursive core-periphery
(jellyfish, octopus)

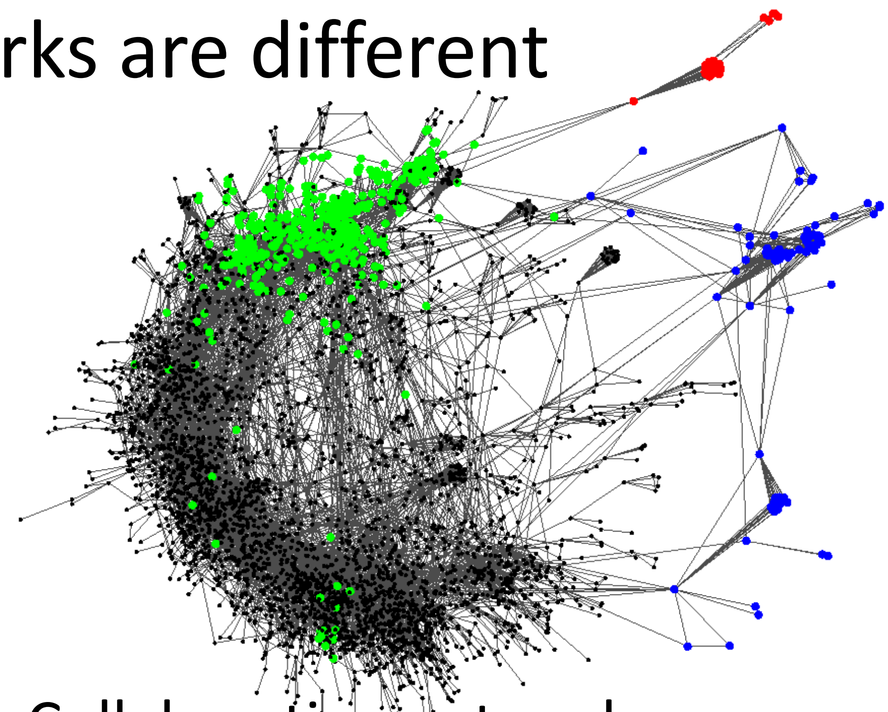
Small vs. Large networks

- Small and large networks are different



Scientific collaborations
(N=397, E=914)

$$G_1 = \begin{bmatrix} 0.99 & 0.17 \\ 0.17 & 0.82 \end{bmatrix}$$



Collaboration network
(N=4,158, E=13,422)

$$G_1 = \begin{bmatrix} 0.99 & 0.54 \\ 0.49 & 0.13 \end{bmatrix}$$

Conclusion

- Kronecker Graph model has
  - provable properties
  - small number of parameters
- Scalable algorithms for fitting Kronecker Graphs
- Efficiently search large space ($\sim 10^{1,000,000}$) of permutations
- Kronecker graphs fit well real networks using few parameters
- Can easily be extended to estimate node attributes or treat them as given
- Kronecker graphs match graph properties without a priori deciding on which ones to fit

References

- Graphs over Time: Densification Laws, Shrinking Diameters and Possible Explanations, by Jure Leskovec, Jon Kleinberg, Christos Faloutsos, ACM KDD 2005
- Realistic, Mathematically Tractable Graph Generation and Evolution, Using Kronecker Multiplication, by Jure Leskovec, Deepay Chakrabarti, Jon Kleinberg and Christos Faloutsos, PKDD 2005
- Scalable Modeling of Real Graphs using Kronecker Multiplication, by Jure Leskovec and Christos Faloutsos, ICML 2007
- Graph Evolution: Densification and Shrinking Diameters, by Jure Leskovec, Jon Kleinberg and Christos Faloutsos, ACM TKDD 2007