Analytic Theory of Power Law Graphs

Jeremy Kepner

MIT Lincoln Laboratory

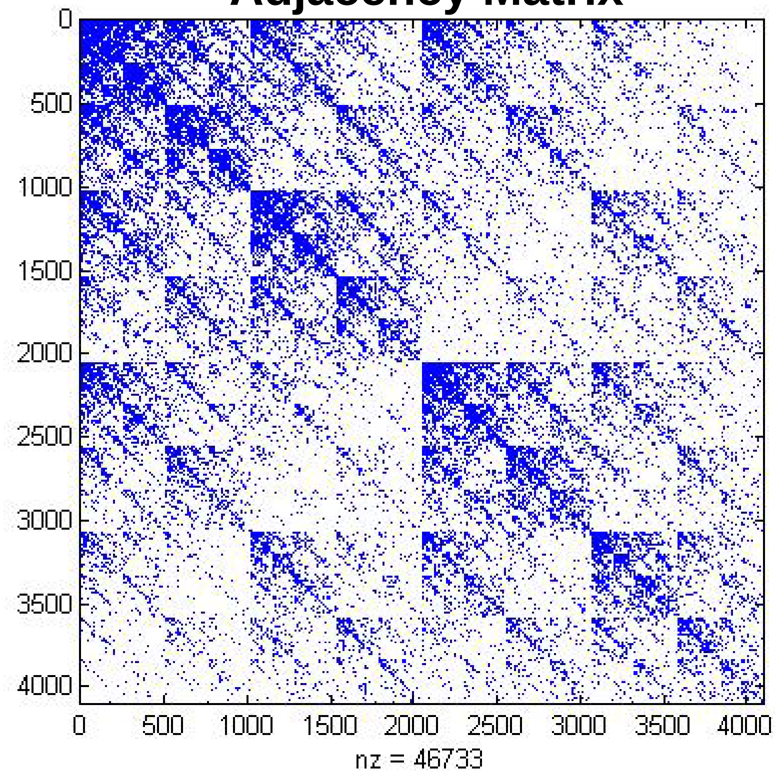
This work is sponsored by the Department of Defense under Air Force Contract FA8721-05-C-0002.
Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the United States Government.

Outline

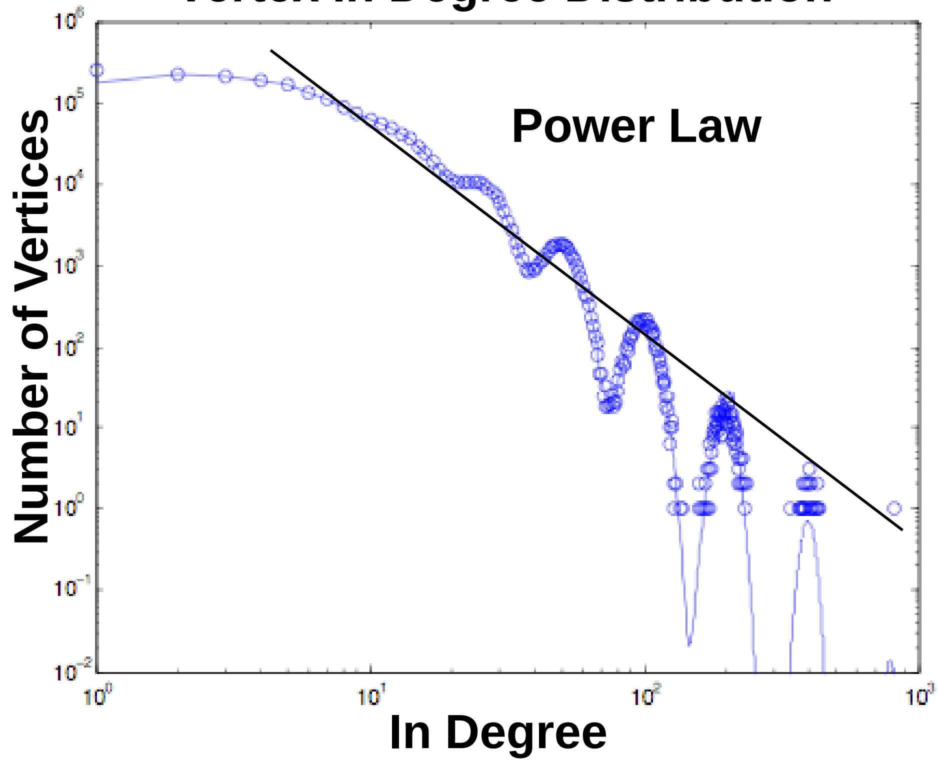
- **Introduction**
  - ***Kronecker Graphs***
  - ***Graphs as Matrices***
  - ***Algorithm Comparison***
- **$B^{\otimes K}$ Graphs**
- **$(B+I)^{\otimes K}$ Graphs**
- **Summary**

Power Law Modeling of Kronecker Graphs

Adjacency Matrix

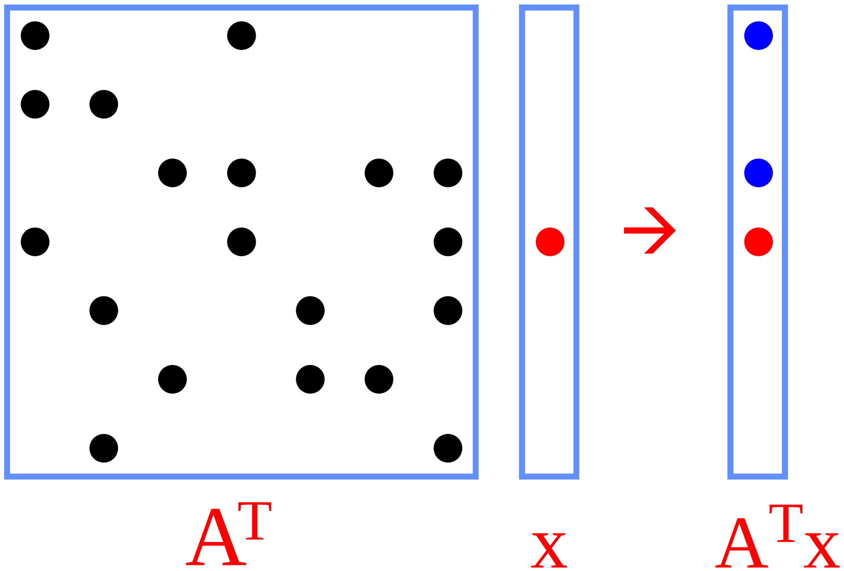


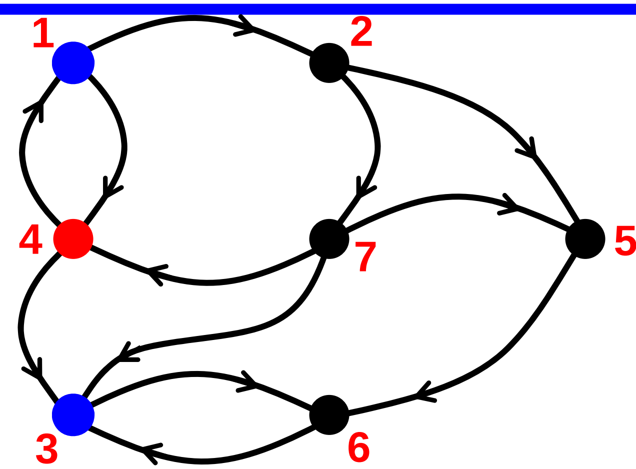
Vertex In Degree Distribution



- **Real world data (internet, social networks, …) has connections on all scales (i.e power law)**
- **Can be modeled with Kronecker Graphs: $G^{\otimes k} = G^{\otimes k-1} \otimes G$**
  - **Where "⊗"denotes the Kronecker product of two matrices**

Graphs as Matrices





- **Graphs can be represented as a sparse matrices**
  - **Multiply by adjacency matrix → step to neighbor vertices**
  - **Work-efficient implementation from sparse data structures**
- **Most algorithms reduce to products on semi-rings: C = A “+”.“x” B**
  - **“x” : associative, distributes over “+”**
  - **□“+” : associative, commutative**
  - **Examples: +.* min.+ or.and**

Algorithm Comparison

Algorithm (Problem)	Canonical Complexity	Array-Based Complexity	Critical Path (for array)
Bellman-Ford (SSSP)	$\Theta(mn)$	$\Theta(mn)$	$\Theta(n)$
Generalized B-F (APSP)	NA	$\Theta(n^3 \log n)$	$\Theta(\log n)$
Floyd-Warshall (APSP)	$\Theta(n^3)$	$\Theta(n^3)$	$\Theta(n)$
Prim (MST)	$\Theta(m+n \log n)$	$\Theta(n^2)$	$\Theta(n)$
Borůvka (MST)	$\Theta(m \log n)$	$\Theta(m \log n)$	$\Theta(\log^2 n)$
Edmonds-Karp (Max Flow)	$\Theta(m^2n)$	$\Theta(m^2n)$	$\Theta(mn)$
Push-Relabel (Max Flow)	$\Theta(mn^2)$ (or $\Theta(n^3)$)	$O(mn^2)$	?
Greedy MIS (MIS)	$\Theta(m+n \log n)$	$\Theta(mn+n^2)$	$\Theta(n)$
Luby (MIS)	$\Theta(m+n \log n)$	$\Theta(m \log n)$	$\Theta(\log n)$

Majority of selected algorithms can be represented with array-based constructs with equivalent complexity.

($n = |V|$ and $m = |E|$.)

Outline

- Introduction
- $B^{\otimes K}$ Graphs
  - *Definitions*
  - *Bipartite Graphs*
  - *Degree Distribution*
- $(B+I)^{\otimes K}$ Graphs
- Summary

Kronecker Products and Graph

Kronecker Product

- **Let B be a N_BxN_B matrix**
- **Let C be a N_CxN_C matrix**
- **Then the Kronecker product of B and C will produce a N_BN_CxN_BN_C matrix A:**

$$A = B \otimes C = \begin{pmatrix} b_{1,1}C & b_{1,2}C & \dots & b_{1,M_B}C \\ b_{2,1}C & b_{2,2}C & \dots & b_{2,M_B}C \\ \vdots & \vdots & & \vdots \\ b_{N_B,1}C & b_{N_B,2}C & \dots & b_{N_B,M_B}C \end{pmatrix}$$

Kronecker Graph (Leskovec 2005 & Chakrabati 2004)

- **Let G be a NxN adjacency matrix**
- **Kronecker exponent to the power k is:**

$$G^{\otimes k} = G^{\otimes k-1} \otimes G$$

Types of Kronecker Graphs

<u>Explicit</u>

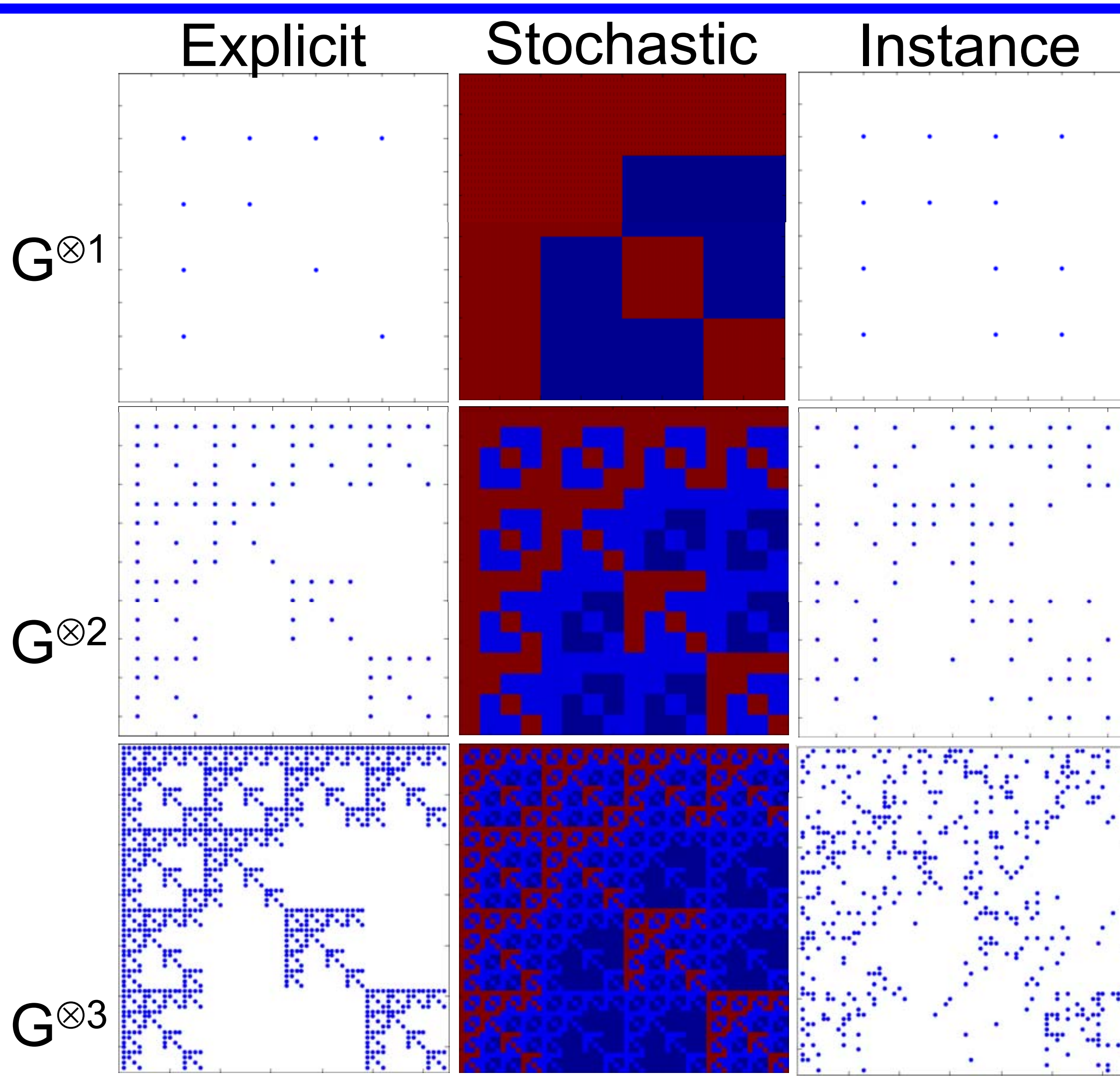
- **G only 1 and 0s**

<u>Stochastic</u>

- **G contains probabilities**

<u>Instance</u>

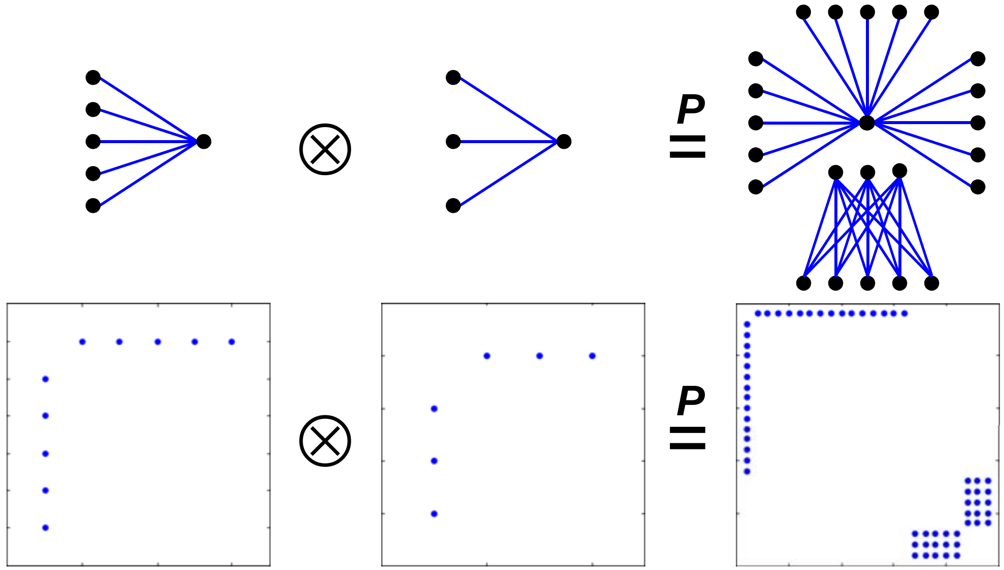
- **A set of M points (edges) drawn from a stochastic**



Kronecker Product of a Bipartite Graph

$\overset{P}{=}$

Equal with the right permutation



$B(5,1) \otimes B(3,1) \overset{P}{=} B(15,1) \cup B(3,5)$

- **Fundamental result [Weischel 1962] is that the Kronecker product of two complete bipartite graphs is two complete bipartite graphs**
- **More generally**

$$B(n_1,m_1) \otimes B(n_2,m_2) \overset{P}{=} B(n_1 n_2, m_1 m_2) \cup B(n_2 m_1, n_1 m_2)$$

Degree Distribution of Bipartite Kronecker Graphs

- **Kronecker exponent of a bipartite graph produces many independent bipartite graphs**

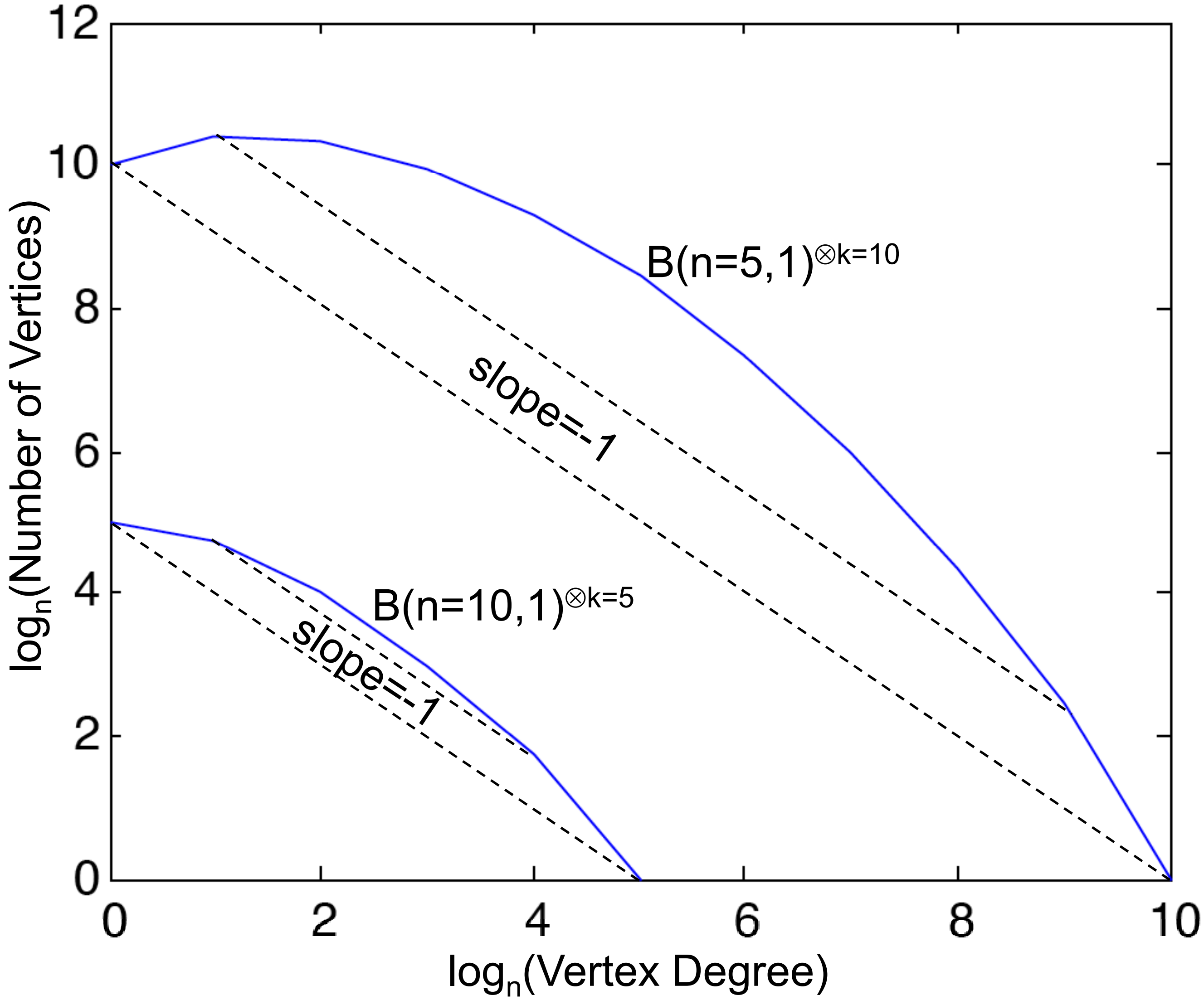
$$B(n,m)^{\otimes k} \stackrel{P}{=} \bigcup_{r=0}^{k-1} \bigcup^{\binom{k-1}{r}} B(n^{k-r}m^r, n^r m^{k-r})$$

- **Only k+1 different kinds of nodes in this graph, with degree distribution**

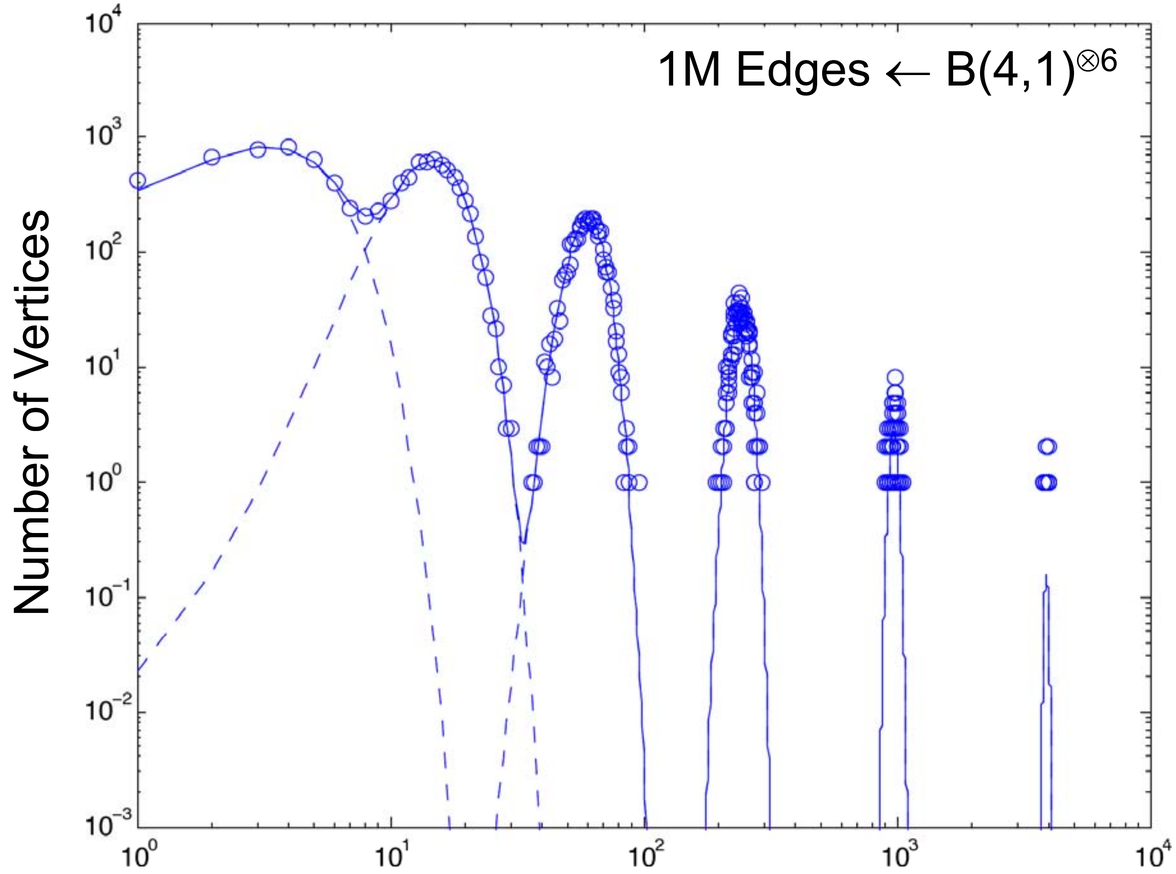
$$Count[Deg = n^r m^{k-r}] = \binom{k}{r} n^{k-r} m^r$$

Explicit Degree Distribution

- **Kronecker exponent of bipartite graph naturally produces exponential distribution**



Instance Degree Distribution



- **An instance graph drawn from a stochastic bipartite graph is just the sum of Poisson distributions taken from the explicit bipartite graph**

Outline

- Introduction
- $B^{\otimes K}$ Graphs
- **$(B+I)^{\otimes K}$ Graphs**
  - *Bipartite + Identity Graphs*
  - *Permutations and substructure*
  - *Degree Distribution*
  - *Iso Parametric Ratio*
- Summary

Theory

- **Bipartite Kronecker graphs highlight the fundamental structures in a Kronecker graph, but**
  - **Are not connected (i.e. many independent bipartite graphs)**

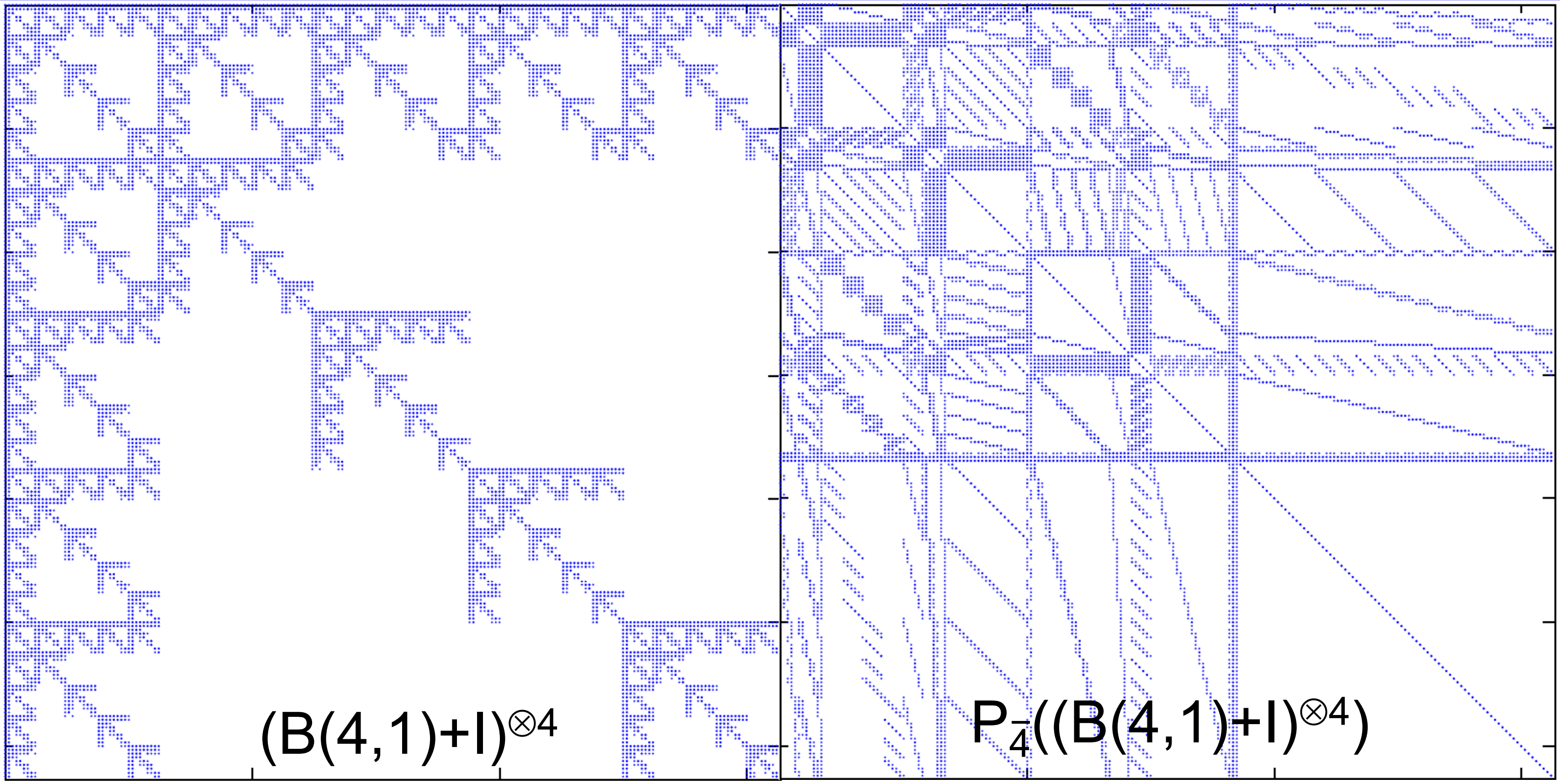
- **Adding identity matrix creates connections on all scales**
  - **Resulting explicit graph has diameter = 2**
  - **Sub-structures in the graph are given by**

$$(B+I)^{\otimes k} \stackrel{P}{=} \sum_{r=1}^{k} \text{``}\binom{k}{r}\text{''} \bigcup^{N^{k-1}} B^{\otimes k}$$

  - **Where “” indicates permutations are required to add the matrices**

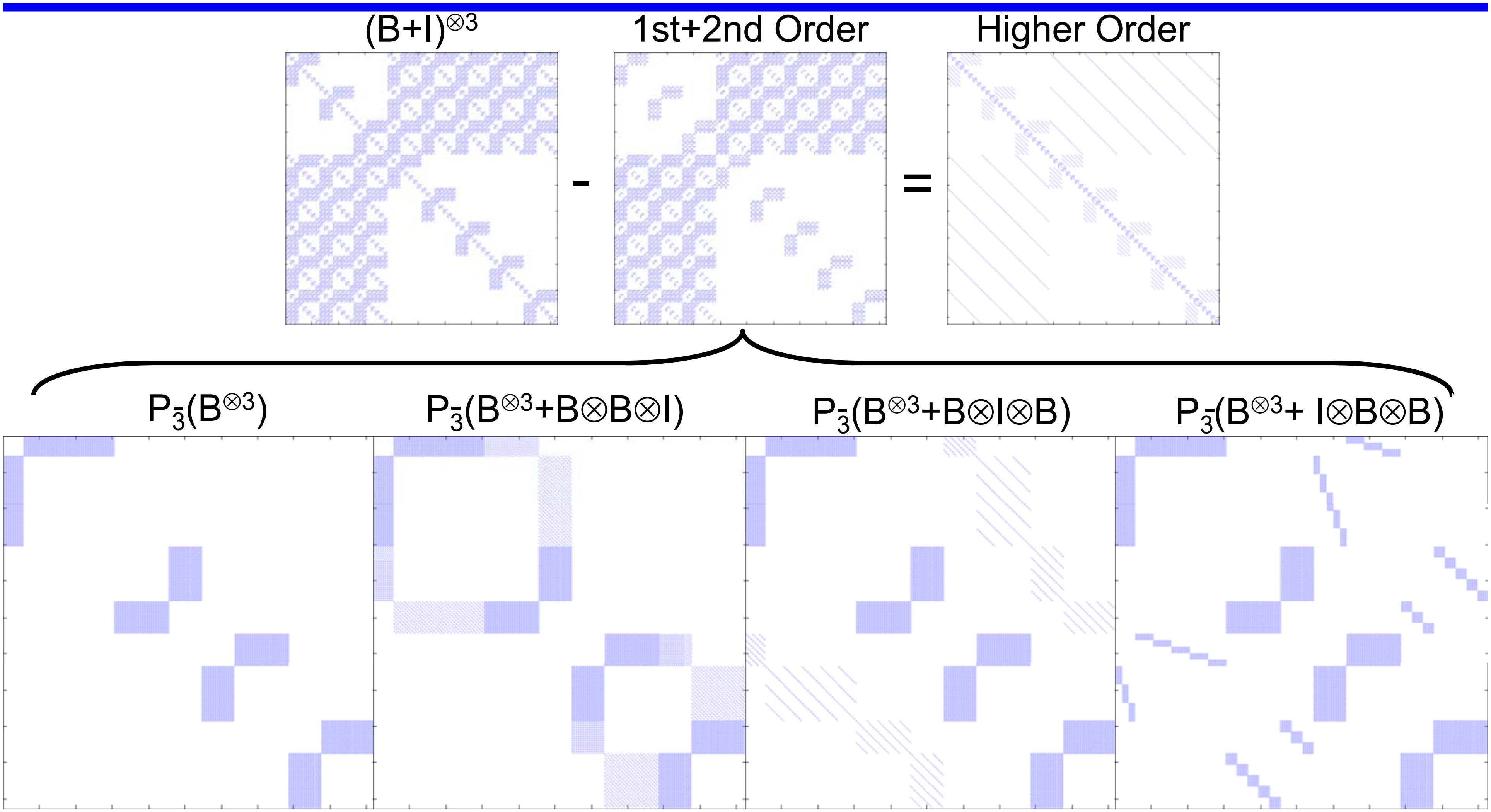
- **Sub-structure can be revealed by applying permutation that “groups” vertices by their bipartite sub-graph**

Bipartite Permutation



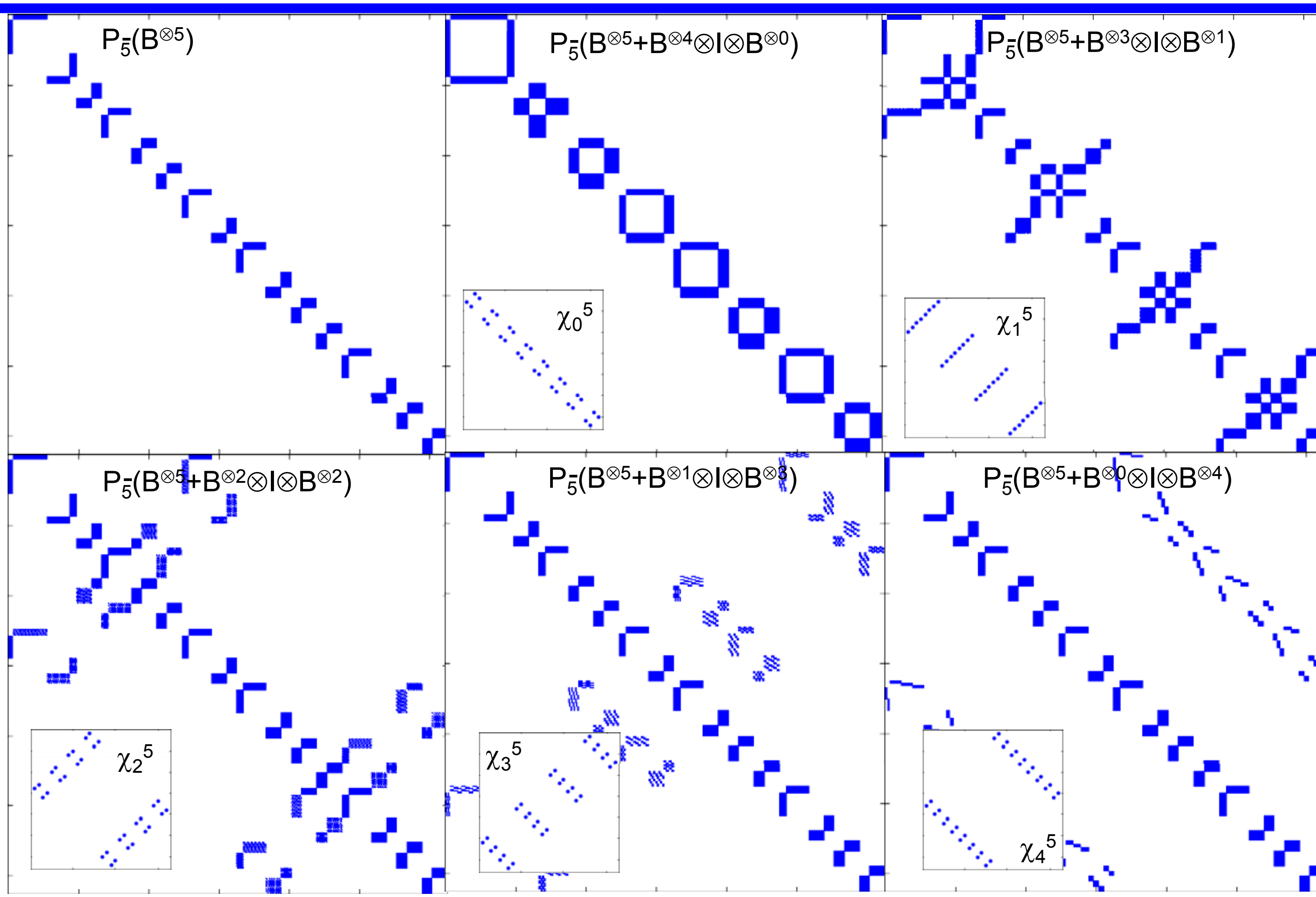
- **Left: unpermuted $(B+I)^{\otimes 4}$ kronecker graph**
- **Right: permuted $(B+I)^{\otimes 4}$ kronecker graph**

Identifying Substructure



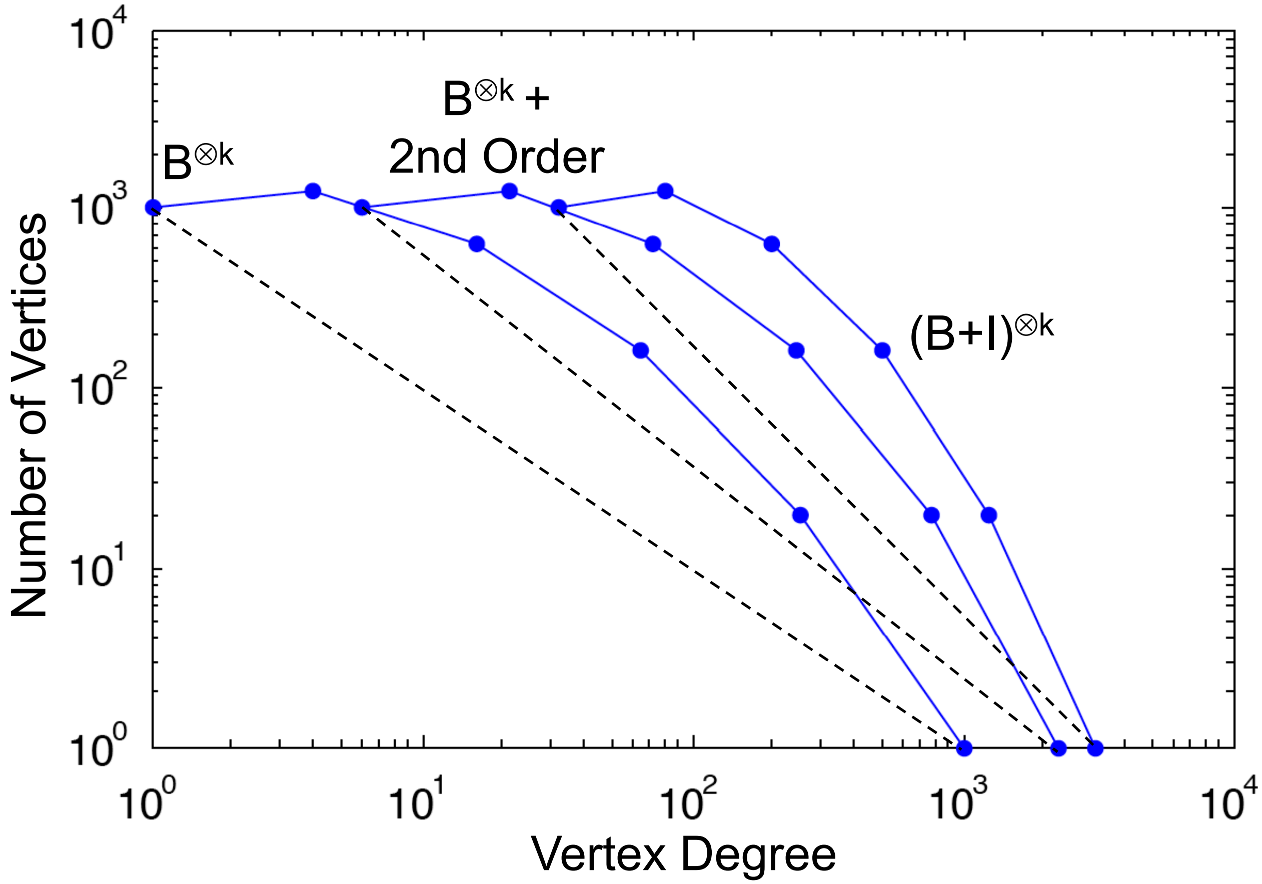
- **Permuting specific terms shows their contributions to the graph**

Quantifying Substructure



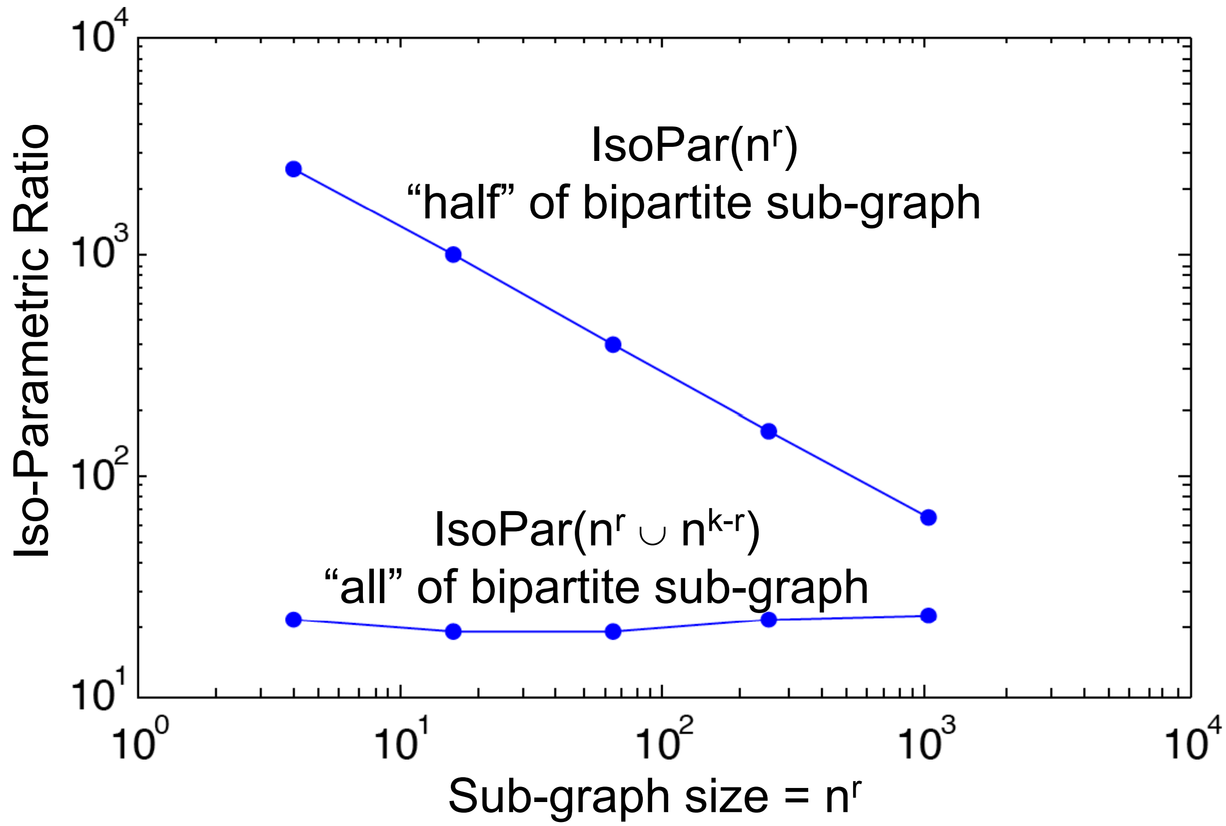
- **Connections between bipartite subgraphs are the Kronecker product of corresponding 2x2 matrices, e.g. $B(1,1)^{\otimes 4}\otimes I(2)$**

Substructure Degree Distribution



- **Only k+1 different kinds of nodes in this graph, with same degree distribution, only differing values of vertex degree**
- **$(B+I)^{\otimes k}$ is steeper than $B^{\otimes k}$**

Example Result: Iso-Parametric Ratio



- **Iso-parametric ratios measure the "surface" to "volume" of a sub-graph**
- **Can analytically compute for a Kronecker graph: $(B+I)^{\otimes k}$**
- **Shows large effect of including "half" or "all" of bipartite sub-graph**

Kronecker Graph Theory
-Summary of Current Results-

Quantity	Graph: $B(n,m)^{\otimes k}$	Graph: $(B+I)^{\otimes k}$
Degree Distribution	$Count[Deg = n^r m^{k-r}] = \binom{k}{r} n^{k-r} m^r$	$Count[Deg = (n+1)^r (m+1)^{k-r}] = \binom{k}{r} n^{k-r} m^r$
Betweenness Centrality	$Count[C_b = (n/m)^{2r-k}(n^{k-r}m^r - 1)] = \binom{k}{r} n^{k-r} m^r$	
Diameter	$Diam(B^{\otimes k}) = \infty$	$Diam((B+I)^{\otimes k}) = 2$
Eigenvalues	$eig(B(n,m)^{\otimes k}) = \{\overbrace{(nm)^{k/2}, ..., (nm)^{k/2}}^{2^{k-1}}, \overbrace{-(nm)^{k/2}, ..., -(nm)^{k/2}}^{2^{k-1}}\}$	$eig((B+I)^{\otimes k}) = \{((nm)^{1/2}+1)^k, ((nm)^{1/2}+1)^{k-1}, ((nm)^{1/2}-1)^2((nm)^{1/2}+1)^{k-2}, ...\}$
Iso-parametric Ratio "half"	$IsoPar(n_k(i)) = \infty$	$IsoPar(n_k(i)) = 2(n+1)^{k-r}(m+1)^r - 2$
Iso-parametric Ratio "all"	$IsoPar(n_k(i) \cup m_k(i)) = 0$	$IsoPar(n_k(i) \cup m_k(i)) = 2\frac{n^r m^{k-r}(n+1)^{k-r}(m+1)^r + n^{k-r}m^r(n+1)^r(m+1)^{k-r}}{2n^k m^k + n^r m^{k-r} + n^{k-r}m^r + [\chi \text{ terms}]} - 2$

Reference

- **Book: “Graph Algorithms in the Language of Linear Algebra”**
- **Editors: Kepner (MIT-LL) and Gilbert (UCSB)**
- **Contributors**
  - **Bader (Ga Tech)**
  - **Chakrabart (CMU)**
  - **Dunlavy (Sandia)**
  - **Faloutsos (CMU)**
  - **Fineman (MIT-LL & MIT)**
  - **Gilbert (UCSB)**
  - **Kahn (MIT-LL & Brown)**
  - **Kegelmeyer (Sandia)**
  - **Kepner (MIT-LL)**
  - **Kleinberg (Cornell)**
  - **Kolda (Sandia)**
  - **Leskovec (CMU)**
  - **Madduri (Ga Tech)**
  - **Robinson (MIT-LL & NEU), Shah (UCSB)**