Spectrum and Algebraic Connectivity of semi-regular random graphs

Theodore Kolokolnikov

Algebraic connectivity (A.C.) is a measure of how fast the information can diffuse through a graph. It corresponds to the smallest non-zero eigenvalue of the graph laplacian:

$$L = D - A$$

Laplacian — degree matrix (diagonal) — adjacency matrix

- E.g. consensus model of information flow:

$$\frac{dx_j}{dt} = \underbrace{\sum_{i \sim j} (x_i - x_j)}_{-(L\vec{x})_j}$$

sol'n: $\vec{x} = \sum_i c_i \vec{v}_i e^{-\mu_i t}$ where $\mu_i \in \text{spectrum}(L)$, $\mu_1 = 0$, $\mu_2 > 0$ if G is connected

Large t: $\vec{x} \sim x_{avg} + c_2 \vec{v}_2 e^{-\mu_2 t}$ as $t \to \infty$

$\Rightarrow \mu_2$ is the rate of convergence to consensus.

Question: Among graphs of n vertices and m edges, which graph maximizes A.C.? [i.e. has best "diffusion" property?]

- d-regular graphs $\left(m = \frac{d}{2} n\right)$ are better than just Erdos-Reney graphs
- For a random d-regular graph, it is known that $\mu_2 \sim d - 2\sqrt{d-1}$ as $n \to \infty$

[Alon 86, Nilli 2004, Friedman 91, Broder-Shamir 87]

- Can we do better if we relax the "regularity" assumption? Suppose that the average degree is d, but graph is not necessary regular...
  - Yes if $d \in [3, 8]$!
  - Semi-regular random bipartite graph can have higher A.C...

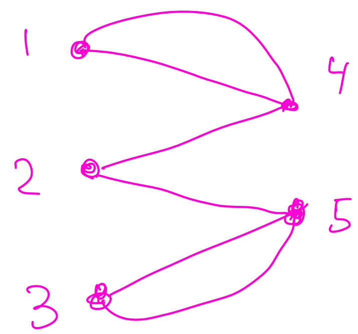
Random Semi-Regular Bipartite (RSRB) graph model:

- Put d_1 copies of vertices labelled $1 \dots n_1$ into bag 1
- Put d_2 copies of vertices labelled $n_1+1 \dots n_1+n_2$ into bag 2
- Draw vertices at random, one from each bag, to create edges.
- Note that we must have $n_1 d_1 = n_2 d_2 \Rightarrow \boxed{d = \frac{2 d_1 d_2}{d_1 + d_2}}$

avg. deg.

Example: $n_1 = 3,\ d_1 = 2,\quad n_2 = 2,\ d_2 = 3$

B_1:	2	3	1	3	1	2
	|	|	|	|	|	|
B_2:	4	5	4	5	4	5



Main Result 1.1. *Consider a (d_1, d_2) RSRB graph. In the limit $n \to \infty$, its spectrum density asymptotes to*

$$\rho(x) = \begin{cases} \frac{1}{\pi}\frac{d_1 d_2}{d_1+d_2}\frac{\sqrt{(x^2-r_-^2)(r_+^2-x^2)}}{(d_1 d_2 - x^2)|x|}, & |x| \in (r_-, r_+) \\ \frac{|d_2-d_1|}{d_1+d_2}\delta(x), & |x| < r_- \\ 0, & |x| > r_+ \end{cases} \tag{1.3}$$

where δ is the Dirac-delta function and

$$r_\pm = \left(d_1 + d_2 - 2 \pm \sqrt{(d_1+d_2-2)^2 - (d_2-d_1)^2}\right)^{1/2}. \tag{1.4}$$

In other words, the number of eigenvalues inside any interval (a, b) asymptotes to $\int_a^b \rho(x)dx$ as $n \to \infty$.
Moreover, its algebraic connectivity asymptotes to

$$\mu \sim \frac{d_1+d_2}{2} - \left(\left(\frac{d_2-d_1}{2}\right)^2 + r_+^2\right)^{1/2}, \quad n \gg 1 \tag{1.5}$$

• Note: if $d_1 = d_2 \Rightarrow \mu \sim d - 2\sqrt{d-1}$ [Alon, 86...]

$$\rho \sim \frac{d}{2\pi}\frac{\sqrt{4(d-1) - x^2}}{d^2 - x^2}, \quad |x| < 2\sqrt{d-1}$$

[Mackay, 81]

So we recover d-reg. graphs as special case...

Ex:

$d_1 = 2, d_2 = 3$

$d = 3$:

- 3-reg : $\mu \sim 3 - 2\sqrt{2} = 0.1716$
- $d_1 = 2, d_2 = 6$: $d = 3$, $\mu \sim 0.1997$

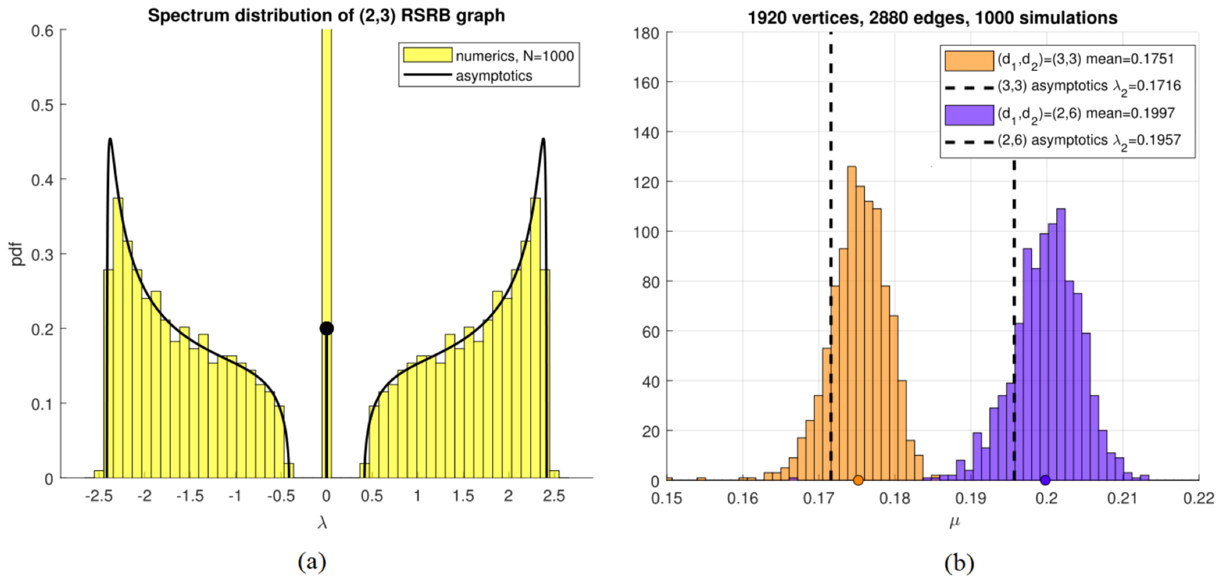


FIG. 2. (a) Full spectrum of random semi-regular bipartite graph with $(d_1, d_2) = (2, 3)$. Numerics correspond to the histograph of eigenvalues of a single such graph with 1000 vertices, computed numerically using Matlab. Asymptotics corresponds to the formula (1.3). The height of the lollypop corresponds to the weight delta function at the origin. (b) Comparison of algebraic connectivity between $(3, 3)$ regular bipartite, $(2, 6)$ semi-regiular bipartite graphs, and the asymptotic theory. The two classes have the same number of vertices and edges, and $(2, 6)$ is 15% better than (3,3) (both for asymptotics and numerics).

Conclusion : (2, 6) semiregular is 15% "better" than 3-reg [while having same # vertices & edges]

All RSRB graphs with integer average degree $d = 3, \ldots, 8$													
d	3		4		5		6		7		8		
d_1	3	2	4	3	5	3	6	4	7	4	8	6	5
d_2	3	6	4	6	5	15	6	12	7	28	8	12	20
μ_{asympt}	0.1715	**0.1957**	0.5358	**0.5535**	1	**1.0890**	1.5278	**1.5587**	2.1010	**2.1435**	**2.7084**	2.6887	2.6671
μ_{numerics}	0.178	0.205	0.553	0.572	1.027	1.122	1.565	1.596	2.150	2.205	2.766	2.745	2.729
std	0.006	0.006	0.011	0.010	0.015	0.017	0.018	0.018	0.021	0.020	0.026	0.022	0.022
diff %	3.8%	4.7%	3.1%	3.2%	2.7%	3.0%	2.4%	2.4%	2.3%	2.87%	2.1%	2.1%	2.2%

Spectrum: **Trace method:** $\sum_i^n \lambda_i^s = \text{trace } A^s$;

Let $\varphi_s = \frac{1}{n} \text{trace}(A^s)$;

Let $\rho(x)$ be the spectrum density.

Formally: $\rho(x) = \frac{1}{n}\sum_{j=1}^{n} \delta(x - \lambda_j)$; **then**

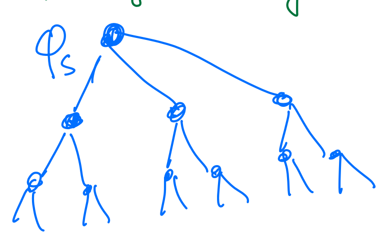
$$\int x^s \rho(x)\, dx = \varphi_s \qquad (*)$$

Step 1: compute φ_s. **Step 2:** Invert $(*)$.

Computing φ_s: $(A^s)_{ii}$ is the # of closed walks of length s that start and end at vertex i.

Key insight: **Locally**, semiregular graph looks like a tree. So it is enough to compute # of walks on a tree.

3-reg locally looks like tree:

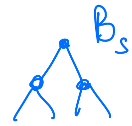


Warm-up: 3-reg graph:

Then $\varphi_s \equiv$ # of closed walks of length s on this tree:

$$\varphi_0 = 1,$$
$$\varphi_{odd} = 0$$
$$\varphi_2 = 3$$
$$\varphi_4 = \underbrace{\downarrow\downarrow\uparrow\uparrow}_{3\times 2} + \underbrace{\downarrow\uparrow\downarrow\uparrow}_{3\times 3} = 15$$

$$\varphi_6 = 3\underset{\downarrow\uparrow}{\varphi_4} + 3\underset{\downarrow}{B_2}\underset{\uparrow}{\varphi_2} + 3\underset{\downarrow}{B_4}\,\uparrow$$



where B_s: # closed walks of length s on a binary tree:

$B_2 = 2$, $\quad B_4 = 2\cdot 2 + 2\cdot 2 = 8 \Rightarrow \varphi_6 = 3[15 + 2\cdot 3 + 8] = 87$

$$\varphi_8 = 3B_6 + 3B_4\varphi_2 + 3B_2\varphi_4 + 3\varphi_6$$

$$B_6 = 2B_4 + 2B_2B_2 + 2B_4$$

General:
$$\begin{cases} \varphi_s = \sum_{j=0}^{s-2} 3\varphi_j B_{s-2-j} \\ B_s = \sum_{j=0}^{s-2} 2B_j B_{s-2-j} \end{cases}$$

Generating fcn:

$$B(x) = \sum_{s=0}^{\infty} B_s x^s; \quad \varphi(x) = \sum \varphi_s x^s \Rightarrow \begin{cases} B = 1 + 2x^2B^2 \\ \varphi = 1 + 3x^2\varphi B \end{cases}$$

$$\Rightarrow \begin{cases} B = \frac{1}{4x^2}\left(1 - \sqrt{1-8x^2}\right) \\ \varphi = \frac{1}{2}\left(\frac{3\sqrt{1-8x^2} - 1}{1-9x^2}\right) = 1 + 3x^2 + 15x^4 + 87x^6 + \dots \end{cases}$$

- Note singularity of $\varphi(x)$ at $x = \frac{1}{\sqrt{8}}$

$$\Rightarrow \boxed{\varphi_s \sim c\left(\sqrt{8}\right)^s \quad \text{as } s \to \infty}$$

$$\Rightarrow \lambda_{max} \sim \sqrt{8} = 2\sqrt{2}$$ (for adjacency matrix A)

Now $L = 3I - A$

$$\Rightarrow \boxed{\mu_2 \sim 3 - 2\sqrt{2}}$$ [Recovering Alon asymptotics]

Full spectrum:

[Mackay, 81]:

$$\begin{cases} \varphi = \frac{1}{2}\left(\frac{3\sqrt{1-8x^2}-1}{1-9x^2}\right) = \sum \varphi_s x^s; \\ \int \rho\, x^s\, ds = \varphi_s \end{cases}$$

Cauchy formula:

[Stieltjes Inversion]

$$\varphi_s = \oint_{|z|=\varepsilon} z^{-s-1}\varphi(z)\frac{dz}{2\pi i}$$

$$= \frac{i}{\pi}\int_{-\frac{1}{\varepsilon}}^{\frac{1}{\varepsilon}} x^s\,\frac{1}{x}\,\varphi\left(\frac{1}{x}\right)dx$$

$$\Rightarrow \rho(x) = -\frac{1}{\pi}\,\mathrm{Im}\left(\frac{1}{x}\,\varphi\left(\frac{1}{x}\right)\right)$$

For 3-regular graphs: $\rho(x) = \frac{1}{\pi}\,\frac{3}{2}\,\frac{\sqrt{8-x^2}}{9-x^2}$ [Mackay, 81]

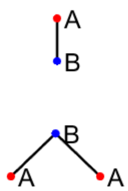
Full plan:

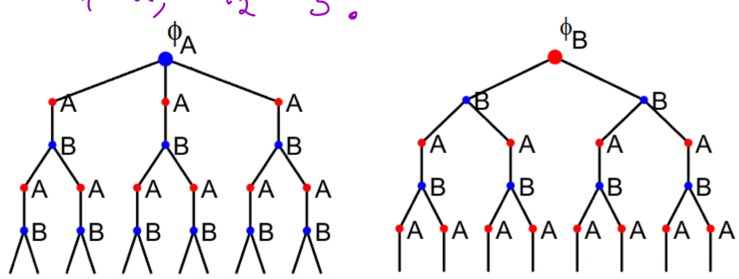
$$\phi_s = \frac{d_1}{d_1+d_2}\phi_{A,s} + \frac{d_2}{d_1+d_2}\phi_{B,s},$$

$$\phi_{A,s} = d_2\sum_{j=0}^{s-2}\phi_{A,j}A_{s-2-j}, \qquad \phi_{B,s} = d_1\sum_{j=0}^{s-2}\phi_{B,j}B_{s-2-j},$$

$$A_s = (d_1-1)\sum_{j=0}^{s-2}A_jB_{s-2-j}, \qquad B_s = (d_2-1)\sum_{j=0}^{s-2}B_jA_{s-2-j},$$

Ex: $d_1 = 2$, $d_2 = 3$:





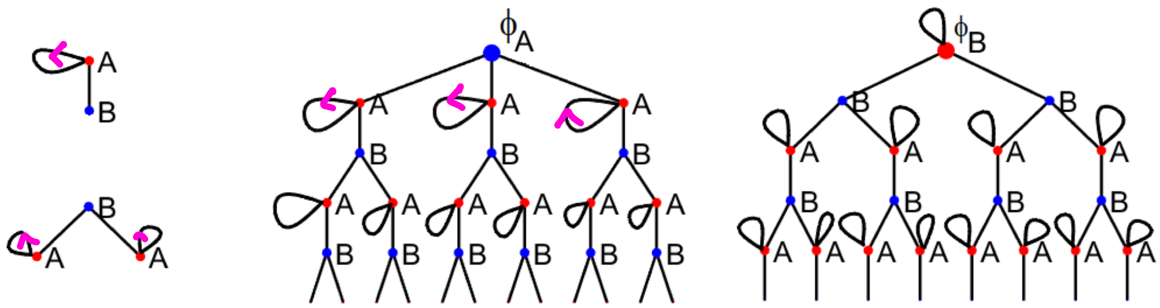
Ex: $B_6 = 2(A_0B_4 + A_2B_2 + A_4B_0)$

$$\Rightarrow \quad \phi(x) = \frac{d_1 d_2}{d_1+d_2} \frac{\frac{d_1+d_2}{d_1 d_2} - \sqrt{(d_2-d_1)^2 x^4 + (4-2d_1-2d_2)x^2+1} - 1}{d_1 d_2 x^2 - 1}.$$

$$\Rightarrow \quad \rho(x) = -\frac{1}{\pi} \operatorname{Im}\left(\frac{1}{x}\phi\left(\frac{1}{x}\right)\right) = \frac{1}{\pi}\frac{1}{|x|}\frac{d_1 d_2}{d_1+d_2}\frac{\sqrt{(\gamma_+^2-x^2)(x^2-\gamma_-^2)}}{d_1 d_2 - x^2}$$

AC: $L = D - A$; regularize:

e.x. $d_1 = 2$, $d_2 = 3$; add directed loop to vertices of degree 2; then each vertex will have "degree" 3:



- Then $L = D - A$ where $D = d_2 I$

$$\begin{aligned}
A &= 1 + (d_2 - d_1) x A + (d_1 - 1) x^2 AB, \\
B &= 1 + (d_2 - 1) x^2 BA, \\
\phi_A &= 1 + d_2 x^2 \phi_A A, \\
\phi_B &= 1 + x(d_2 - d_1)\phi_B + d_1 x^2 \phi_B B, \\
\phi &= \frac{d_1}{d_1+d_2}\phi_A + \frac{d_2}{d_1+d_2}\phi_B
\end{aligned}$$

- get a quartic for μ which yields (1.4).

Random Semiregular Graphs (RSR): given d_1, d_2, p, n construct a random graph of n vertices, where each vertex has degree d_1 with prob. $1-p$, and degree d_2 with prob p.

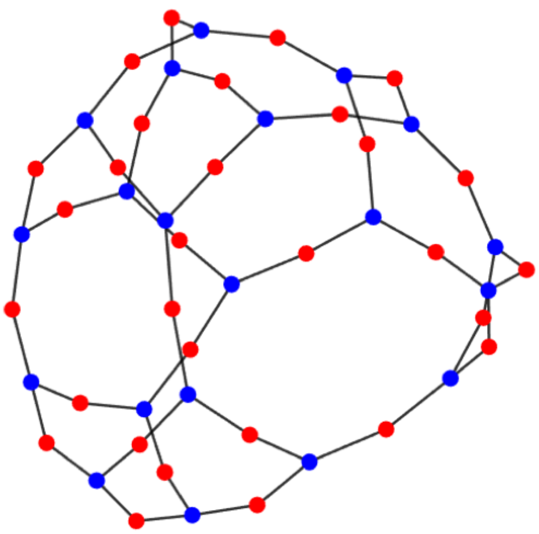
Ex: $d = 2.4$:

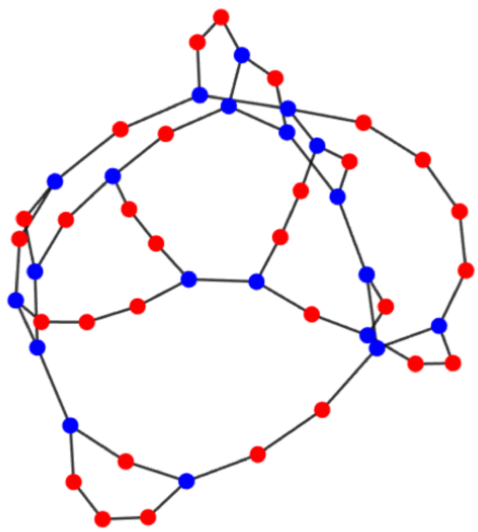
RSRB

```
d1=2; d2=3; n1=30; n2=20;
bag1=mod([0:n1*d1-1], n1)+1;
bag2=mod([0:n2*d2-1], n2)+1+n1;
bag2=bag2(randperm(numel(bag2)));
G=graph(bag1, bag2);
plot(G);
```

RSR

```
p=0.4; d1=2; d2=3; n=50;
n1=(1-p)*n; n2=p*n;
v1=mod([0:n1*d1-1], n1)+1;
v2=mod([0:n2*d2-1], n2)+1+n1;
bag=[v1, v2];
bag=bag(randperm(numel(bag)));
G=graph(bag(1:end/2), bag(end/2+1:end));
plot(G);
```





Main Result 1.2. *Consider a* $(p,\ d_1, d_2)$ *random semi-regular graph. Let*

$$F(R,x) = x\,(d_2-d_1)\,(1-Rx)\,p + \left(Rx^2(d_2-1)-1\right)\left(R^2x^2\,(d_1-1)+Rx\,(d_2-d_1)-R+1\right). \tag{1.7}$$

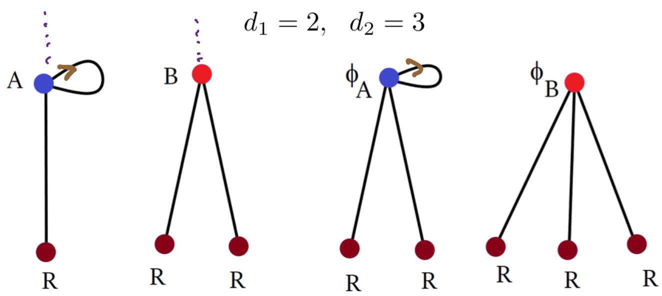
Let x *be the smallest root of the system* $F = 0 = \partial F/\partial R$*. Then in the limit* $n \to \infty$*, the AC is given by* $\mu = d_2 - 1/x$*.*

In general, eliminating R from the system $F = 0 = \partial F/\partial R$ is a straightforward computer algebra computation using a resultant, and yields in a 6th degree polynomial for x. It is too ugly to write down here for general d_1, d_2 – see Appendix A for Maple code. In the case $d_1 = 2, d_2 = 3$, RSR graph has average degree $2+p$, and μ is the smallest root of

$$0 = \mu\,(\mu-4)\left(\mu^2-4\mu-1\right) + 2\mu\left(3\mu^3-33\mu^2+89\mu-19\right)p + \left(-15\mu^2-30\mu+1\right)p^2 + 8p^3. \tag{1.8}$$

Figure 3 compares μ given by (1.8) with numerical computations of μ for randomly chosen $(p, 2, 3)$ RSR graphs. Note that the numerical result approaches the asymptotic value of μ as the number of edges n is increased.

Sketch of proof:



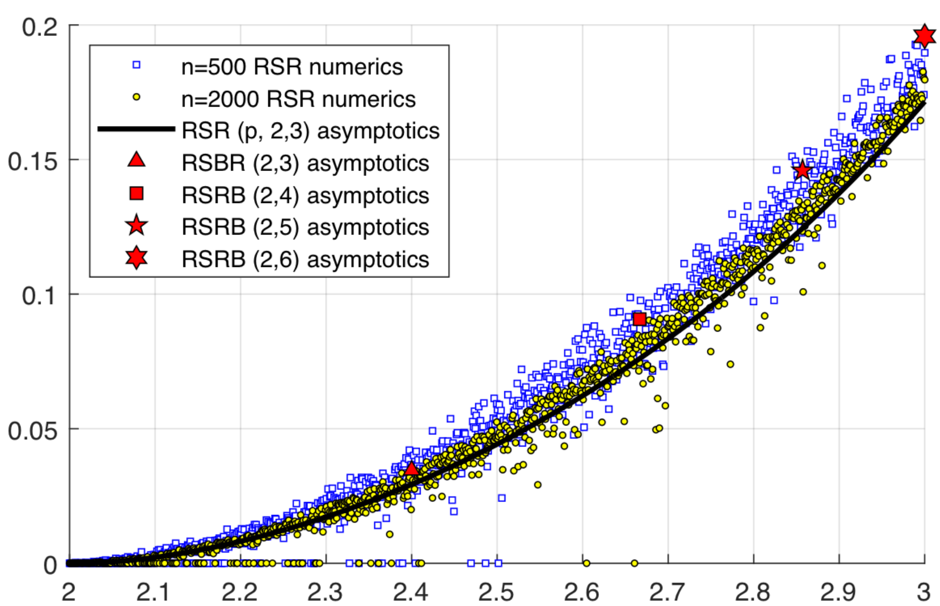
$R = (1-p)\,A + pB$

The corresponding generating function $\phi(x) = \sum \phi_s x^s$ solves the equations

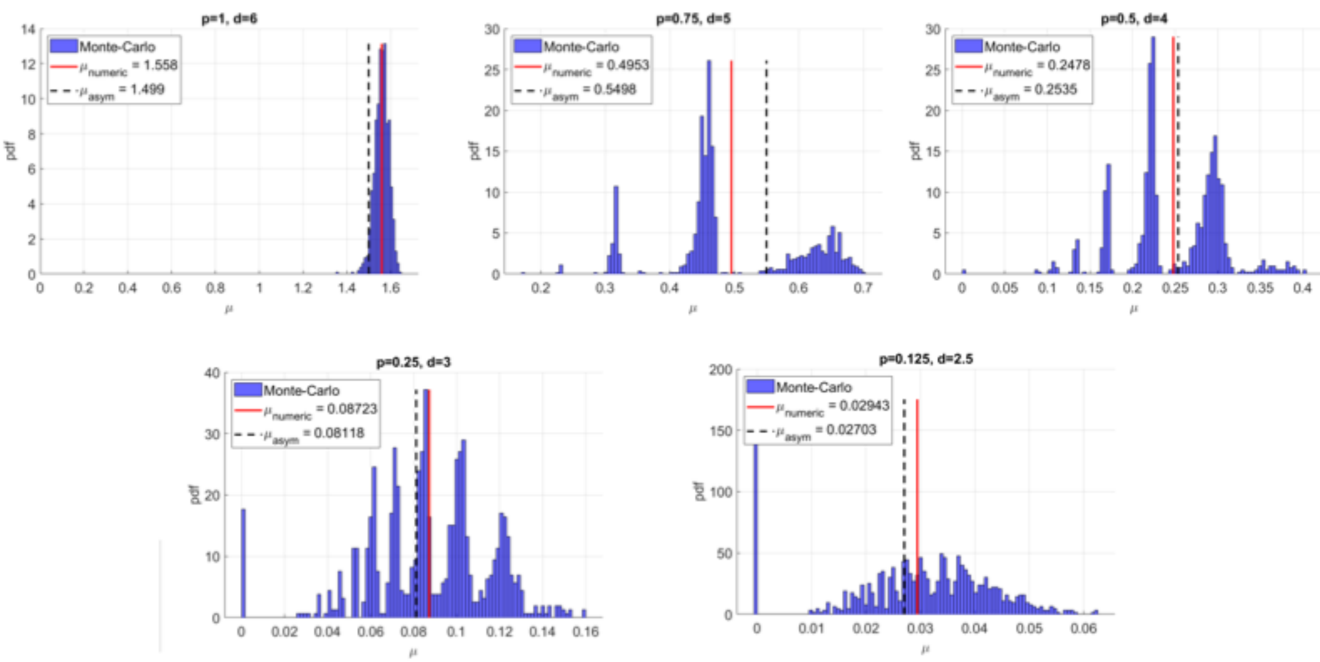
$$\begin{aligned} A &= 1 + x\,(d_2-d_1)\,A + (d_1-1)\,x^2 AR \\ B &= 1 + (d_2-1)\,BR \\ R &= (1-p)\,A + pB \\ \phi_A &= 1 + x\,(d_2-d_1)\,\phi_A + d_1 x^2 \phi_A R \\ \phi_B &= 1 + d_2 x^2 \phi_B R \\ \phi &= (1-p)\phi_A + p\phi_B. \end{aligned} \tag{2.17}$$

Eliminating A and B yields a cubic $F(R;x) = 0$ given by (1.7). The AC is then given by $\mu = d_2 - 1/x$, where x is the singularity of $R(x)$ that is closest to the origin. By implicit function theorem, this happens when $F_R = 0$. In other words, x satisfies $F = 0 = \partial F/\partial R$.■

RSR model with $d=4$						
d_1	4	3	3	2	2	2
d_2	4	5	6	5	6	7
p		0.5	1/3	2/3	0.5	0.4
μ_{asympt}	**0.5359**	0.44261	0.39162	0.3333	0.25352	0.20748
μ_{numerics}	0.551	0.488	0.451	0.286	0.217	0.174
std	0.010	0.020	0.022	0.062	0.051	0.045
diff %	2.8%	10%	15%	-14%	-14%	-19%



$d_1=2,\ d_2=6$:



Open Questions:

Challenge 1. *Describe the full distribution of AC, particularly for RSR graphs. Explain why it can be multi-peaked when* $d_1 \neq d_2$.

Challenge 2. *Find a family of random graphs which has a higher algebraic connectivity than d-regular random graphs when average degree* $d \geq 10$. *Explore if more complex degree distribution (e.g. tri-regular) can be better than semi-regular for say,* $d = 3$.

Challenge 3. *For fixed average degree* d *and fixed number of vertices* n, *find graphs (not necessarily d-regular) with highest possible girth.*

Preprint:

It is better to be semi-regular when you have a low degree

T. Kolokolnikov

- available on arxiv or my website.

Thank you!