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

T. Kolokolnikov<sup>1</sup>

<sup>1</sup>Department of Mathematics and Statistics, Dalhousie University Halifax, Nova Scotia, B3H3J5, Canada\*

We study the algebraic connectivity for several classes of random semi-regular graphs. For large random semi-regular bipartite graphs, we explicitly compute both their algebraic connectivity and as well as the full spectrum distribution. For an integer  $d \in [3, 7]$ , we find families of random semi-regular graphs that have higher algebraic connectivity than a random *d*-regular graphs with the same number of vertices and edges. On the other hand, we show that regular graphs beat semiregular graphs when  $d \geq 8$ . More generally, we study random semi-regular graphs whose average degree is *d*, not necessary an integer. This provides a natural generalization of a *d*-regular graph in the case of a non-integer *d*. We characterise their algebraic connectivity in terms of a root of a certain 6th-degree polynomial. Finally, we construct a small-world-type network of average degree 2.5 with a relatively high algebraic connectivity. We also propose some related open problems and conjectures.

#### 1. INTRODUCTION

Algebraic connectivity (AC; also called the spectral gap) of a graph is a fundamental property that measures how fast information diffuses throughout the graph [1, 2]. It corresponds to the second smallest eigenvalue of the graph Laplacian matrix (the smallest eigenvalue is always zero). In many applications, it is desirable to maximize the algebraic connectivity (i.e. speed of diffusion) subject to certain constraints; this and related problems have a long history [3–9]. For example in communications, the "cost" of a network increases with the number of links. It is therefore desirable to have as few edges as possible, while at the same time maximizing the algebraic connectivity. This leads to a natural question: For a fixed average degree d, what is the graph that maximizes algebraic connectivity as the number of vertices  $n \to \infty$ ?

Numerous papers address various aspects of this question, see e.g. [6–9]. This is one of those situations where the best answer is elusive [10], but a decent answer can be found relatively quickly [3, 4]. Graphs with high algebraic connectivity are related to expander graphs, and are important in many applications [11, 12]. In this paper, we study the algebraic connectivity of *sparse* random graphs in the case when the number of edges m scales linearly with the number of vertices n, i.e. m = O(n). In other words, we fix the average degree d = 2m/n = O(1) while letting  $n \to \infty$ .

It is well known that a random Erdos-Renyi graph (where m edges are taken at random) needs  $O(n \log n)$  edges to be fully connected [13, 14], and as such, they are poor candidates for maximizing algebraic connectivity in this sparse regime (since AC of a disconnected graph is zero). A good candidate are d-regular graphs for which every



FIG. 1. Types of graphs considered in this work. (a) Random semi-regular bipartite graph with  $(d_1, d_2) = (2, 3)$ . (b) Random semi-regular graph with  $(p, d_1, d_2) = (0.4, 2, 3)$ . Both graphs have the same average degree of d = 2.4. (c) Small-world network consisting of a ring and with edges added at random between the odd-numbered vertices. It has an average degree of d = 2.5.

<sup>\*</sup> tkolokol@gmail.com

vertex has degree d. It is well known that random regular d-graphs have an algebraic connectivity that asymptotes to  $\mu \sim d - 2\sqrt{d-1}$  as  $n \to \infty$  with fixed d [15–19]. For integer  $d \ge 3$ , this is quantity is bounded away from zero, which assures these graphs have good expander properties [20, 21]. A natural question is whether one can do better than regular graph, but without "too much work" (in, say, O(n) time). In this paper we give an affirmative answer when  $d \le 7$ : we introduce a class of semi-regular random graphs (whose vertices have degree either  $d_1$  or  $d_2 \ne d_1$ ) which are as easy to construct as random d-regular graphs, but which have better AC.

Another question is whether "expander-type" graphs (which we define to be graphs with AC bounded away from 0 as  $n \to \infty$ ) are possible when the average degree d is less than 3, with  $n \to \infty$ . For 2 < d < 3, the answer is yes, and it is provided by semi-regular graphs of degrees 2 and 3, whose average degree is 2 + p (with  $0 ). We show that AC of such graphs asymptotes to <math>\mu \sim p^2/4$  in the limit of small p and large n, with  $n \to \infty$  independent of p. (The answer is no when  $d \leq 2$  since any graph needs at least n - 1 edges to be connected).

Our first result is on random semi-regular *bipartite* graphs, where each vertex has degree either  $d_1$  or  $d_2$ , and we compute their asymptotic AC in the limit  $n \to \infty$ . In particular, we will exhibit a family of semi-random bipartite graphs having the same average degree as random *d*-regular graphs, but which have higher algebraic connectivity when  $d \leq 7$ .

Before stating our result, let us define what we mean by such graphs. Consider a bipartite graph with one part having  $n_1$  vertices of degree  $d_1$ , and the second part having  $n_2$  vertices of degree  $d_2$ , such that every edge is between these two parts, and no edges are within each part. Then  $n_1d_1 = n_2d_2$ . Such graph has  $n = n_1 + n_2$  vertices so that

$$n_1 = \frac{d_2}{d_1 + d_2}n, \quad n_2 = \frac{d_1}{d_1 + d_2}n \tag{1.1}$$

and therefore the average degree is

$$d = \frac{2d_1d_2}{d_1 + d_2}.\tag{1.2}$$

We call such a graph a  $(d_1, d_2)$  semi-regular bipartite graph. We now introduce the following random model.

**Random semi-regular bipartite (RSRB) graph model**: Take two bags. In the first bag, put  $d_1$  copies of  $n_1$  vertices labelled  $1 \dots n_1$ . In the second bag, put  $d_2$  copies of  $n_2$  vertices labelled  $n_1 + 1 \dots n_1 + n_2$ . Start with an empty graph of  $n_1 + n_2$  vertices. Then randomly pick two vertices without replacement – one from each bag – and add an edge between them to the graph. Repeat until the bags are empty. Refer to Figure 1(a) and Matlab code in Appendix A.

We remark that this model (and the theory below) generally allows for multiple edges. If desired, they can be eliminated through a random rewiring postprocessing step<sup>1</sup>. These cases are sufficiently rare that the postprocessing step does not effect the asymptotic results in the large n limit. We now state our main results for RSRB graphs.

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

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

where  $\delta$  is the Dirac-delta function and

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

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

<sup>&</sup>lt;sup>1</sup> Given a multiple edge (a, b) or a self loop (a = b), choose another edge (c, d) at random. Then replace edges (a, b) and (c, d) by edges (a, c) and (b, d). Repeat until all multiple edges/loops are eliminated. This operation preserves degree distribution and edge count.



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-regular 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).

Figure 2(a) shows the shape of the distribution (1.3) for the case  $d_1 = 2$ ,  $d_2 = 3$  (having an average degree of d = 2.4), and compares it to an numerical histogram of eigenvalues of a 1000-vertex graph for this case. Very good agreement is observed.

Note that in the case  $d_1 = d_2$ , formula (1.3) reduces to the well-known Mackay distribution of the spectrum of a regular graph [15]:

$$\rho(x) = \begin{cases} \frac{1}{\pi} \frac{d}{2} \frac{\sqrt{4(d-1)-x^2}}{d^2-x^2}, & |x| < 2\sqrt{d-1} \\ 0, & |x| > 2\sqrt{d-1}. \end{cases}$$
(1.6)

Moreover, formula (1.5) simplifies to the classical result  $\mu \sim d - 2\sqrt{d-1}$  for the algebraic connectivity of a random regular graph of d vertices <sup>2</sup>.

Consider a random cubic graph  $(d_1 = d_2 = 3)$  and contrast it with an RSRB graph with  $d_1 = 2$ ,  $d_2 = 6$ . Both cases have the same average degree of d = 3. Formula (1.5) gives the asymptotic values of  $\mu \sim 0.17157$  for the former, and  $\mu \sim 0.19577$ , for the latter. Thus, (2,6) RSRB graph is about 15% better than a random cubic graph with respect to its expander properties, while having the same number of vertices and edges. Figure 2(b) shows a histogram for  $\mu$  of 1000 randomly constructed such graphs, comparing these two cases with n = 2000. Very good agreement between numerics and asymptotics is observed in both cases.

More generally, the following table shows all possible combinations of  $d_1, d_2$  such that  $d = \frac{2d_1d_2}{d_1+d_2}$  is an integer between 3 and 8, and the corresponding value of  $\mu$  (1.5).

 $<sup>^2</sup>$  This also shows that a random regular graph and random bipartite regular graphs have the same algebraic connectivity, at least to leading order in n.

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
$\mu_{\rm asympt}$	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
$\mu_{\text{numerics}}$	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%

The row  $\mu_{\text{asympt}}$  is the asymptotic formula given by (1.5). Row  $\mu_{\text{numerics}}$  corresponds to Monte Carlo simulations of  $\mu$ . It shows the average  $\mu$  for 200 randomly chosen RSRB graphs with n = 1000 edges. The row "diff %" is  $\frac{\mu_{\text{numerics}} - \mu_{\text{asympt}}}{\mu_{\text{asympt}}} \times 100$ . Uniformly good agreement with between asymptotics and numerics is observed.

Parameters with higher AC are in shown in bold. For  $d \leq 7$ , RSRB with  $d_1 \neq d_2$  have higher algebraic connectivity than the *d*-regular graph. On the other hand, for  $d \geq 8$ , *d*-regular graphs win.

It is interesting to note that for any integer  $d \ge 3$ , equation (1.2) always has a solution with integers  $2 \le d_1 < d_2$ . When d is prime, this solution is unique and is given by  $d_1 = (d+1)/2$ ,  $d_2 = d(d+1)/2$ . More generally, the number of such solutions is precisely the number of pythegorean triples of leg d (sequence A046079 in OEIS).

RSRB graphs above have a constraint  $n_1d_1 = n_2d_2$ . In particular the minimum attainable average degree of such graphs is d = 2.4 corresponding to  $(d_1, d_2) = (2, 3)$ . We can remove this constraint by instead introducing the probability of having degree  $d_1$  or  $d_2$  as follows.

**Random semi-regular (RSR) model:** Given  $p, d_1, d_2$ , and n, let  $n_1 = \lfloor (1-p)n \rfloor$  and let  $n_2 = n - n_1$ . In the same bag, put  $d_1$  copies of  $n_1$  vertices labelled  $1 \dots n_1$  and  $d_2$  copies of  $n_2$  vertices labelled  $n_1 + 1 \dots n$ . Create edges by drawing two vertices from the bag at random (without replacement), until only one or zero vertices are left in the bag. See Appendix A for Matlab code.

An example of RSR graph is shown in Figure 1(b). Note that such a graph has an average degree of  $d = (1 - p) d_1 + pd_2$ . We have the following.

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

$$F(R,x) = x \left( d_2 - d_1 \right) \left( 1 - Rx \right) p + \left( Rx^2 (d_2 - 1) - 1 \right) \left( R^2 x^2 (d_1 - 1) + Rx (d_2 - d_1) - R + 1 \right).$$
(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  $\mu$  is the smallest root of

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

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

The following table gives the value of  $\mu$  for several choices of  $p, d_1, d_2$  for which the average degree d = 4, and comparison with numerics:

RSR model with $d = 4$										
$d_1$	4	3	3	2	2	2				
$d_2$	$d_2$ 4		6	5	6	7				
p		0.5	1/3	2/3	0.5	0.4				
$\mu_{\rm asympt}$	0.5359	0.44261	0.39162	0.3333	0.25352	0.20748				
$\mu_{\text{numerics}}$	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%				

Here,  $\mu_{\text{asympt}}$  is as given by Main Result 1.2;  $\mu_{\text{numerics}}$  is the average of 200 simulations with n = 1000. As apparent from this table, RSR model performs worse than RSRB of average degree 4 (including 4–regular graphs). This



FIG. 3. Comparison between RSR, RSRB asymptotics, and RSR numerics, with average degree  $2 < d \leq 3$ . Numerics represent AC of 1000 randomly chosen RSR graphs with  $p \in (0, 1)$  and  $(d_1, d_2) = (2, 3)$ . Here,  $\mu$  is plotted against d = 2 + p. Asymptotics for RSR correspond to roots of (1.8). Asymptotics for RSRB are given by (1.5).

appears to be true for any average degree d. However one advantage of RSR graphs is that they produce graphs with an average degree 2 < d < 3, with  $\mu$  bounded away from zero as  $n \to \infty$ . The smallest average degree that RSRB model can have is 2.4 corresponding to  $(d_1, d_2) = (2, 3)$ . In fact, (1.8) shows that  $\mu \sim p^2/4$  when the average degree is 2 + p with  $d_1 = 2, d_2 = 3$ , and 0 .

### 2. DERIVATION OF MAIN RESULTS

**Derivation of spectral density** (1.3). Following [15, 22] we use the trace method. It is based on the fact that  $\operatorname{trace}(A^s) = \sum_{j=1}^s \lambda_j^s$ . Define  $\phi_s = \frac{1}{n} \operatorname{trace}(A^s)$ . Then  $\phi_s$  represents the average number of closed walks of length s on the graph whose adjacency matrix is A. In the limit  $n \to \infty$ , the eigenvalue distribution therefore satisfies

$$\int x^s \rho(x) dx = \phi_s. \tag{2.9}$$

As in [15, 22], the derivation consists of (a) computing  $\phi_s$ ; and (b) inverting the integral equation (2.9) to determine  $\rho(x)$ .

To compute  $\phi_s$ , we write

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

where  $\phi_{A,s}$  is the number of closed walks of length s starting from a vertex of degree  $d_2$ , and  $\phi_{B,s}$  is that number for vertices of degree  $d_1$ . A key insight [15] that allows the asymptotic determination of  $\phi_{A,s}$  and  $\phi_{B,s}$  is that locally, a large random regular graph looks like a tree because the probability of encountering a cycle of length s is vanishinly small as  $n \to \infty$  (for fixed s). The same is true of semi-regular random graphs. For the RSRB graphs, each successive level alternates between vertices of degree  $d_1$  and  $d_2$ . We therefore decompose these trees as illustrated in the following diagram, for the case  $d_1 = 2$  and  $d_2 = 3$ :



Accordingly,  $\phi_{A,s}$  and  $\phi_{B,s}$  satisfy the recursion

$$\phi_{A,s} = d_2 \sum_{j=0}^{s-2} \phi_{A,j} A_{s-2-j}, \qquad \phi_{B,s} = d_1 \sum_{j=0}^{s-2} \phi_{B,j} B_{s-2-j},$$
$$A_s = (d_1 - 1) \sum_{j=0}^{s-2} A_j B_{s-2-j}, \qquad B_s = (d_2 - 1) \sum_{j=0}^{s-2} B_j A_{s-2-j}, \qquad (2.10)$$

with  $A_0 = B_0 = \phi_{A,0} = \phi_{B,0} = 1$ . Let  $A(x) = \sum A_s x^s$  be the generating function for  $A_s$  and similarly for  $B, \phi_A, \phi_B, \phi$ . From recursion relations (2.10), the corresponding generating functions satisfy

$$A = 1 + (d_1 - 1) x^2 AB,$$
  

$$B = 1 + (d_2 - 1) x^2 BA,$$
  

$$\phi_A = 1 + d_2 x^2 \phi_A A,$$
  

$$\phi_B = 1 + d_1 x^2 \phi_B B,$$
  

$$\phi = \frac{d_1}{d_1 + d_2} \phi_A + \frac{d_2}{d_1 + d_2} \phi_B$$
  
(2.11)

Solving (2.11), while keeping in mind that  $\phi(0) = 1$ , we find that

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

In principle,  $\phi_s$  can be computed from  $\phi(x)$  by Taylor-expanding near the origin; the first few terms are  $\phi_0 = 1$ ,  $\phi_2 = 2\frac{d_1d_2}{d_1+d_2}$ ,  $\phi_4 = 2\frac{d_1d_2}{d_1+d_2}$  ( $d_1 + d_2 - 1$ ). In [15] explicit computation of  $\phi_s$  was combined with Tchebychev polynomials to compute the McKay distribution for *d*-regular random graphs. Here, we use a simpler and more powerful technique using complex variables introduced in [23].

Write  $\phi_s$  using Cauchy's integral formula as

$$\phi_s = \int z^{-s-1} \phi(z) \frac{dz}{2\pi i}$$

where the integration is around a circle  $|z| = \varepsilon$ , with  $\varepsilon$  sufficiently small to avoid any branch cuts of  $\phi$ . Using the fact that  $\phi(z)$  is even, we then have

$$\phi_s = \frac{1}{\pi} \int_0^{\pi} \varepsilon^{-s} \phi\left(\varepsilon e^{i\theta}\right) e^{-si\theta} d\theta$$
$$= \frac{i}{\pi} \int_{-1/\varepsilon}^{1/\varepsilon} \frac{1}{x} \phi\left(\frac{1}{x}\right) x^s dx$$

Taking  $\varepsilon \to 0$  and recalling (2.9) we obtain

$$\int \rho(x) x^s dx = \frac{i}{\pi} \int_{-\infty}^{\infty} \frac{1}{x} \phi\left(\frac{1}{x}\right) x^s dx.$$

This yields the formula for  $\rho$ , namely

$$\rho(x) = -\frac{1}{\pi} \operatorname{Im} \left\{ \frac{1}{x} \phi\left(\frac{1}{x}\right) \right\}.$$
(2.12)

This formula was derived in [23] using an equivalent technique (Stieltjes inversion). Next we compute

$$\frac{1}{x}\phi\left(\frac{1}{x}\right) = \frac{1}{x}\frac{d_1d_2}{d_1+d_2}\frac{\frac{d_1+d_2}{d_1d_2}x^2 - \sqrt{(x^2-r_-^2)(x^2-r_+^2) - x^2}}{d_1d_2 - x^2},$$
(2.13)

where  $r_{\pm}$  are given by (1.4). It follows that  $\operatorname{Im}\left\{\frac{1}{x}\phi\left(\frac{1}{x}\right)\right\} = 0$  for  $|x| \notin (r_{-}, r_{+})$  and  $x \neq 0$ . On the other hand when  $|x| \in (r_{-}, r_{+})$  we obtain

$$\rho(x) = \frac{1}{\pi} \frac{1}{|x|} \frac{d_1 d_2}{d_1 + d_2} \frac{\sqrt{(r_+^2 - x^2)(x^2 - r_-^2)}}{d_1 d_2 - x^2}, \quad |x| \in (r_-, r_+).$$

$$(2.14)$$

This recovers the formula (1.3) except for the delta mass at the center which is due to the singularity there. To compute the weight of the delta mass, one can integrate the overall density and impose that  $\int \rho = 1$ . While the exact integration of (2.14) is possible, it is easier and more instructive to compute the number of zero eigenvalues directly as follows.

Suppose that  $d_2 > d_1$ . Label the components of the eigenvector  $x_1 \ldots x_{n_1}$  and  $y_1 \ldots y_{n_2}$ , corresponding to vertices of degree  $d_1$  and  $d_2$ , respectively. A zero eigenvalue satisfies  $\sum_{x_j \in nbr(y_k)} x_j = 0$ ,  $k = 1 \ldots n_2$  and  $\sum_{y_j \in nbr(x_k)} y_j = 0$ ,  $k = 1 \ldots n_1$ . Look for solutions of the form where  $y_j = 0$ ,  $j = 1 \ldots n_2$ . This corresponds to solving  $n_2$  linear equations  $\sum_{x_j \in nbr(y_k)} x_j = 0$ , for the  $n_1$  unknows  $x_1 \ldots x_{n_1}$ . This is an under-determined system, since  $n_2 < n_1$ . Generically, it has  $n_1 - n_2$  independent solutions. Therefore the weight of the delta function at zero should be  $\frac{n_1 - n_2}{n}$ . Using (1.2) yields  $\frac{n_1 - n_2}{n} = \frac{d_2 - d_1}{d_1 + d_2}$ , which recovers the weight of the delta function in (1.3). We also verified using computer algebra that this indeed agrees with the full integration of (2.14), namely, that  $\int_{r_-}^{r_+} \rho(x) dx = \frac{d_1}{d_1 + d_2}$ ,  $d_2 > d_1$ , so that indeed  $\int_{-\infty}^{\infty} \rho = 1$ . This completes the derivation of the spectral density (1.3).

Finally, another derivation of the delta weight, as pointed out by James Mingo, is to use Proposition 3.8 in [24]. It says that if  $\frac{1}{x}\phi\left(\frac{1}{x}\right)$  has a singularity at x = a, then the associated measure  $\rho(x)$  has a delta mass at x = a whose weight is given by  $\lim_{x\to a} \left| (x-a) \frac{1}{x}\phi\left(\frac{1}{x}\right) \right|$ . Here, a = 0 and the mass weight is then given by  $\lim_{x\to 0} \left|\phi\left(\frac{1}{x}\right)\right| = \frac{d_1d_2}{d_1+d_2} \frac{\sqrt{r_-^2 r_+^2}}{d_1d_2} = \frac{|d_2-d_1|}{d_1+d_2}$  in agreement with direct computation.

**Derivation of AC formula for RSRB** (1.5) Next we derive the formula for algebraic connectivity (1.5). For *d*-regular graphs, the Laplacian graph is given by dI - A where *I* is the identity and *A* is the adjacency matrix. This allows to characterize the AC in terms of second-largest eigenvalue  $\lambda$  of *A*:  $\mu = d - \lambda$ . Of course this is not true if the graph is not regular. The trick is to **regularize the graph by adding enough self-loops to vertices until all vertices have the same degree**. Each self-loops counts as a single additional edge (i.e. adding a loop to vertex *j* corresponds to adding "1" to the *j*-th diagonal entry of the associated adjacency matrix) and crucially, adding self-loops does not change the Laplacian of the graph.

For a  $(d_1, d_2)$  semi-regular graph, assume that  $d_1 < d_2$  and add  $d_2 - d_1$  loops to all vertices of degree  $d_1$ . This results in a  $d_2$ -regular graph with loops. As before let  $\phi_s$  be the number of closed walks of size s. Let  $\lambda$  be the *local expansion rate*, that is, the rate at which  $\phi_s$  grows:  $\lambda = \lim_{s \to \infty} \phi_{s+1}/\phi_s$ . Then by analogy to regular graphs, the AC is given by  $\mu = d_2 - \lambda$ . To compute  $\lambda$ , we decompose closed loops similarly to (2.11). The decomposition now has loops as illustrated below:



Correspondingly, the associated generating functions satisfy

$$A = 1 + (d_2 - d_1)xA + (d_1 - 1)x^2AB,$$
  

$$B = 1 + (d_2 - 1)x^2BA,$$
  

$$\phi_A = 1 + d_2x^2\phi_AA,$$
  

$$\phi_B = 1 + x(d_2 - d_1)\phi_B + d_1x^2\phi_BB,$$
  

$$\phi = \frac{d_1}{d_1 + d_2}\phi_A + \frac{d_2}{d_1 + d_2}\phi_B$$
  
(2.15)

To determine AC, it suffices to determine the growth rate of  $\phi_s$ . This growth rate is in turn determined by the locations of singularities in the associated generating functions [25]. In particular, if  $\sum \phi_s z^s = g(z) + h(z)(z-r)^p$ , where p is non-integer and g, h analytic, then  $\phi_s$  grows with the rate  $r^{-1}$ : in other wors,  $\phi_{s+1}/\phi_s \to r^{-1}$  as  $s \to \infty$ . The location of the singularity can be found without solving the full system, and corresponds to the smallest root of

the discriminant of the system. To illustrate this, consider the Catalan numbers  $c_n = \frac{1}{n+1} \begin{pmatrix} 2n \\ n \end{pmatrix}$  whose generating

function satisfies  $c = 1 + xc^2$ . Explicitly,  $c(x) = \frac{1-\sqrt{1-4x}}{2x}$ , and has a singularity at x = 1/4, corresponding to the zero of the discriminant of the quadratic  $xc^2 - c + 1 = 0$ . It follows that  $c_{n+1}/c_n \to 4$  as  $n \to \infty$  (indeed more precise asymptotics  $c_n \sim \frac{1}{\sqrt{\pi}} 4^n n^{-3/2}$  can also be derived from its generating function, although here, we only need the growth rate).

Let us now get back to the system (2.15). Eliminating B, we find that A satisfies a quadratic,

$$(d_2 - 1) (x (d_2 - d_1) - 1) x^2 A^2 + [x^2 (d_2 - d_1) - x (d_2 - d_1) + 1] A - 1 = 0$$

Setting its discriminant to zero, we obtain a quartic polynomial for x, namely

$$(d_1 - d_2)^2 x^4 + 2(d_1 + d_2 - 2)(d_2 - d_1)x^3 + ((d_1 - d_2)^2 - 2(d_1 + d_2 - 2))x^2 + 2(d_1 - d_2)x - 1 = 0.$$
(2.16)

The expansion rate is then  $\lambda = 1/x$ , so that  $\mu = d_2 - 1/x$ . Substituting  $x = 1/(\mu - d_2)$  into (2.16), one obtains a 4th degree polynomial for  $\mu$ ,

$$0 = \mu^{4} - 2(d_{1} + d_{2})\mu^{3} + \left((d_{1} + d_{2})^{2} + 2d_{1}d_{2} - 2d_{1} - 2d_{2} + 4\right)\mu^{2} + 2(d_{1} + d_{2})(d_{1} + d_{2} - 2 - d_{1}d_{2})\mu + (d_{1}d_{2} - d_{1} - d_{2})^{2}$$

Shift this polynomial to eliminate the  $\mu^3$  term by substituting  $\mu = y + \frac{d_1+d_2}{2}$ . Magically, it simultaneously elimintes the  $\mu^1$  term resulting in a quadratic for  $y^2$ :

$$0 = y^{4} + \left(4 - 2\left(d_{1} + d_{2}\right) - \left(d_{2} - d_{1}\right)^{2}/2\right)y^{2} + \frac{\left(d_{2} - d_{1}\right)^{2}\left(d_{1} + d_{2}\right)}{2} + \frac{\left(d_{2} - d_{1}\right)^{4}}{16}$$

Solving for y using quadratic formula yields the solution (1.5).

**Derivation of Result 1.2.** Again, we compute the average number of closed walks of length s,  $\phi_s$ . As before, each vertex looks like a tree locally. Each child in this tree has probability (1 - p) chance of having degree  $d_1$ , and

probability p of having degree  $d_2$ . As in the derivation of (1.5), assume that  $d_1 \leq d_2$ , and add  $d_2 - d_1$  self-loops to each vertex of degree  $d_1$ . The walks on these trees then decompose according to the following diagram (here given in the case  $d_1 = 2$ ,  $d_2 = 3$ ):



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

$$A = 1 + x (d_2 - d_1) A + (d_1 - 1) x^2 A R$$
  

$$B = 1 + (d_2 - 1) B R$$
  

$$R = (1 - p) A + p B$$
  

$$\phi_A = 1 + x (d_2 - d_1) \phi_A + d_1 x^2 \phi_A R$$
  

$$\phi_B = 1 + d_2 x^2 \phi_B R$$
  

$$\phi = (1 - p) \phi_A + p \phi_B.$$
  
(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  $F = 0 = \partial F/\partial R$ .

#### 3. DISCUSSION AND OPEN PROBLEMS

We computed asymptotics of AC for two models of semiregular random graphs: either RSR or RSRB models. While RSR model is shown to have smaller AC than a random *d*-regular graph with the same average degree, we found that RSRB model has *higher* AC than *d*-regular when  $d \leq 7$ .

Ramunajan expander graphs are defined [11, 12] as being d-regular graphs whose algebraic connectivity is at least  $d - 2\sqrt{d-1}$ . As illustrated in Figure 2(b) (see also Figure 9 in [11] and related table there), a random d-regular graph has a decent chance of being Ramunajan graph (around 66% when d = 3 as  $n \to \infty$  according to simulations in [11]). We can extend the definition of Ramunajan exander graphs as being any graph of *average* degree d whose AC is at least  $d - 2\sqrt{d-1}$ . When d = 3, an RSRB (2, 6) graph is actually Ramunajan graph with very high probability as  $n \to \infty$ . This is illustrated in Figure 2(b) where  $n \approx 2000$ . Out of 1000 randomly chosen such graphs, only one had AC less than  $d - 2\sqrt{d-1} = 0.1716$ .

We relied on careful but formal power series computations. While the results were validated using numerics, a rigorous analysis of the results in this paper is an open problem. Related to this, it is an open question to estimate the accuracy of asymptotics as a function of n. In particular, the accuracy appears to be significantly degraded for the RSR graphs when  $d_1$  and  $d_2$  far from each other. Take for example the case  $(d_1, d_2) = (2, 6)$ . Intriguingly, the distribution for AC appears to have multiple peaks and does not necessary concentrate around the mean, as illustrated in Figure 4 (note that this does not appear to be the case for RSRB graphs as figure 2(b) illustrates). We pose this as an open problem.

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

Another class of graphs having high algebraic connectivity are the complete bipartite graphs  $K_{b,n-b}$ , which have AC  $\mu = b$  (with b < n/2). The average degree of such a graph asymptotes to d = 2b as  $n \to \infty$ , so that  $\mu \sim d/2$ . Contrast this to d-regular random graphs ( $\mu \sim d - 2\sqrt{d-1}$ ). As was noted in [7],  $K_{d/2,n-d/2}$  have higher AC than the d-regular graphs (for asymptotically same number of edges) provided that d < 15 (since  $d/2 > d - 2\sqrt{d-1}$  in that



FIG. 4. Distribution of AC for RSR graphs with  $(d_1, d_2) = (2, 6)$  and with p as indicated. Average degree d = 2 + 4p is also indicated. We used n = 500 nodes, and 1000 simulations.  $\mu_{\text{numeric}}$  is the average of the distribution whereas  $\mu_{\text{asym}}$  is the asymptotics according to Main Result 1.2. Note the multi-peaked shape of the distribution, and the fact that the distribution does not concentrate around the mean when  $d \leq 5$ .

case) while the converse is true when  $d \ge 15$ . While complete bipartite graphs have a relatively high AC when d < 15, they are also fragile, in the sense that removing any single edge causes AC to decrease by one. Random graphs are more robust with respect to edge deletions as they seem to have some redundancy built-in.

Consider the limit of large  $d_2$  for the RSRB model. From (1.2) and (1.5), one obtains that  $d \sim 2d_1$  and  $\mu \sim \frac{d}{2} - 1$  as  $d_2 \to \infty$ . This is one less than the AC of complete bipartite graph  $K_{d/2,n-d/2}$  (which also has average degree d as  $n \to \infty$ ). More generally, let  $\mu_{d_1,d_2}$  be as in (1.5) and let  $\mu_d = d - 2\sqrt{d-1}$ , where  $d = 2d_1d_2/(d_1+d_2)$  is the average degree. It can be shown that  $\mu_{d_1,d_2} < \mu_d$  for any  $d_1, d_2$  when  $d \ge 10$ . This discussion suggests the following question.

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

There are several notions of graph connectivity, with AC being just one. Another notion is using the average of reliability polynomial [26, 27] to compute the effect of edge deletions on graph connectivity. While the full reliability polynomial requires exponential time to compute, we performed the following simple experiment to quickly measure graph robustness. Start with a random 6-regular graph on 500 vertices (i.e. containing 1500 edges). Then delete edges at random one by one until the graph becomes disconnected. Over 100 simulations, it took on average 460 edge deletions until the graph became disconnected (and in every instance, disconnectivity first occurred when a single vertex lost all of its edges). We then repeated this experiment with  $(d_1, d_2) = (4, 12)$  RSRB graph with 500 vertices (which also contains 1500 edges but has higher AC than 6-regular random graph). It took on average 325 edge deletions until the disconnection was achieved. These preliminary experiments indicate that *d*-regular graphs are more reliable with respect to edge deletions than semi-regular bipartite graphs, even though the latter might have a higher AC. We state this as a conjecture.

**Conjecture 1.** For a given average degree d, the most reliable graph is a d-regular graph.

Generally speaking, cubic graphs of high girth are good candidates for high-AC graphs [7]. Since RSRB can have higher AC, they are also natural candidates for searching high-girth graphs. We pose this as a challenge.

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

A *d*-regular graph with the smallest possible *n* for a given girth *g* is called a cage. There is an extensive literature for looking for high-girth graphs; see [28, 29] for an overview. There are powerful methods to do computer searches for high-girth *d*-regular graphs [30, 31]. As an example, consider (2, 6) RSRB graphs, whose average degree is d = 3. Do (2, 6) semi-regular graphs give better cages than cubic graphs with the same number of vertices? The answer is,

it depends. Note that that (2, 6) graphs can be obtained from 6-regular graphs by inserting a vertex in the middle of each edge; conversely, any (2, 6) graph yields a 6-regular graph by vertex contraction: just delete all 2-degree vertices. According to [28, 32], a 6-regular 5-cage graph has 40 vertices. By vertex insertion, this yields a girth 10 (2, 6) semi-regular graph with 80 vertices. On the other hand, there are three cubic graphs having girth 10 and only 70 vertices [28]. So cubic is better for girth 10.

Similarly, a 6-regular cage of girth 6 has 62 vertices [28], which induces a (2,6) graph of girth 12 and 124 vertices. On the other hand, a cubic graph of girth 12 has at least 126 vertices [28, 33]. So (2,6) semi-regular graph wins for girth 12.

The techniques in this paper are rather general, and can be used to derive algebraic connectivity for many other random graph families. As an example, consider the "small-world"-type graph illustrated in Figure 1(c). Start with a cycle of *n* vertices. Then connect all odd-numbered vertices to each other at random. See Appendix A for Matlab code that generates such a graph. The resulting graph has an average degree of 2.5. To compute its AC, as in the derivation of (1.5), we add a self-loop to even-numbered vertices so that each vertex has degree 3. We leave it as an excercise to the reader that the average walk-counting sequence  $\phi_s$  has the generating function which satisfies

$$A = 1 + x^{2}(A\hat{A} + AB)$$

$$\hat{A} = 1 + 2x^{2}\hat{A}B$$

$$B = 1 + xB + x^{2}AB$$

$$\phi_{o} = 1 + x^{2}\phi_{o}\left(\hat{A} + 2B\right)$$

$$\phi_{e} = 1 + x\phi_{e} + 2x^{2}\phi_{e}A$$

$$\phi = \frac{1}{2}\phi_{o} + \frac{1}{2}\phi_{e}$$
(3.18)

Looking at the radius of convergence of the generating functions as before, we find that in the limit  $n \to \infty$  its AC asymptotes to the smallest root of the polynomial

$$\mu^9 - 24\mu^8 + 249\mu^7 - 1454\mu^6 + 5184\mu^5 - 11400\mu^4 + 14848\mu^3 - 10368\mu^2 + 3108\mu - 136 = 0,$$

explicitly given by  $\mu = 0.0521926$ . This is validated using Monte Carlo simulations: the average AC of 1000 random such graphs with n = 1000 vertices is 0.0557 (std=0.0029), the difference of about 7%. This is significantly higher than a RSR graph with  $(p, d_1, d_2) = (0.5, 2, 3)$ , which also has average degree of 2.5, but whose AC is  $\mu \sim 0.044241$ .

### APPENDIX A: MATLAB AND MAPLE CODE

## RSRB model (Figure 1(a)):

```
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 model (Figure 1(b)):

```
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);
```

# Small-world-like model (Figure 1(c)):

```
n=52;
v1=[1:n];
v2=[2:n,1];
bag=[1:2:n];
bag=bag(randperm(numel(bag)));
v1=[v1,bag(1:end/2)];
v2=[v2,bag(end/2+1:end)];
G=graph(v1,v2);
t=(1:n)/n*2*pi;
plot(G, 'XData', cos(t), 'YData', sin(t));
```

Maple code to compute AC of RSR graphs:

$$\begin{bmatrix} > [eq := A=1+x*(d2-d1)*A+(d1-1)*x^2*A*R, B=1+(d2-1)*x^2*2B*R, R=(1-p)*A+p*B; eq := A=1+x(d2-d1)A+(d1-1)x^2AR, B=1+(d2-1)x^2BR, R=(1-p)A+pB \\ > eliminate({eq}, (A, B))[2][1]: F:=collect(%, p, factor); F:=-x(-d2+d1)(Rx-1)p-(Rd2x^2-x^2R-1)(R^2d1x^2-R^2x^2-Rd1x+Rd2x-R+1) \\ > eliminate({F, diff(F,B)}, (R))[2][1]: pol := collect(%, p, factor); pol := 4x6(d2-1)(d1-1)(-d2+d1)3p3+x4(-d2+d1)2(8d12d22x2+20d1d23x2-d24x2-16d12d2x2 - 64d1d22x2-16d23x2-28d12d2x+8d12x2-30d1d22x+68d1d2x2+2d23x+50d22x2+8d12x + 76d1d2x-24d1x2+24d22x-48d2x2-d12+10d1d2-46xd1-d22-62xd2+15x2-8d1-8d2 + 36x+8)p2+2x2(-d2+d1)(2d13d23x4-6d12d2x4+6d1d25x4-2d26x4-6d13d22x4+8d12d33x4 - 7d1d24x4+8d13d2x3-2d13x4-14d12d22x3-24d12d2x4+16d1d25x4-2d26x4-6d13d22x4+8d12d33x4 - 7d1d24x4+8d13d2x3-2d13x4-14d12d22x3-24d12d2x4+30d1d23x3-18d1d24x4+7d55x3 - 12d24x4+8d13d2x3-2d13x4-14d12d22x2-17d12d2x3+10d12x4+22d1d233x2-32d1d22x3 + 34d1d2x4-9d24x2+17d23x3-8d22x4-4d13x2+15d12d2x2+16d12x3-37d1d22x2+46d1d2x3 + 5d23x3 - 15d1x4+13d23x2-8d22x3-9d2x4+4d13x+6d12d2x-114d2x2+9x3-3d1d2x2+46d1d2x3 + 5d23x - 10d2x3+6x4-9d1x2x+16d1d2x-16d1x2-d22x-114d2x2+9x3-d12+4d1d2+xd1 - d22-7xd2+10x2-2d1-2d2+2x+2)p-(d12x2-2d1d2x2+d2x2+4d1x2+2xd1-2xd2 + 4x2+1)(-d22x2+d1d2x-d22x+2d2x-xd1+xd2-x2-d1+d2)2 > d1 := 2; d2 := 3; d1 := 2 d2 := 3 > pol;-8x6p3+x4(224x2-120x+15)p2-2x2(96x4+52x3-46x2-3x+3)p-(-3x2-2x+1)(-4x2-2x + +1)2 > collect(numer(eval(%, x = 1/(3 - mu))), p, factor);-8p3+(15p2+30p-1)p2-2x2(96x4+52x3-46x2-3x+3)p-(-3x2-2x+1)(-4x2-2x + +1)2 > collect(numer(eval(%, x = 1/(3 - mu))), p, factor);$$

- [1] M. Fiedler, Algebraic connectivity of graphs, Czechoslovak Mathematical Journal 23 (2) (1973) 298–305.
- [2] N. M. M. de Abreu, Old and new results on algebraic connectivity of graphs, Linear algebra and its applications 423 (1) (2007) 53–73.
- [3] A. Ghosh, S. Boyd, Growing well-connected graphs, 2006 45th IEEE Conference on Decision and Control (2006) 6605–6611.
- [4] H. Wang, P. Van Mieghem, Algebraic connectivity optimization via link addition, Proceedings of the 3rd International Conference on Bio-Inspired Models of Network, Information and Computing Systems (2008) 22.
- [5] A. Sydney, C. Scoglio, D. Gruenbacher, Optimizing algebraic connectivity by edge rewiring, Applied Mathematics and computation 219 (10) (2013) 5465-5479.
- [6] R. Olfati-Saber, Ultrafast consensus in small-world networks, American Control Conference, 2005. Proceedings of the 2005 (2005) 2371–2378.
- [7] T. Kolokolnikov, Maximizing algebraic connectivity for certain families of graphs, Linear Algebra and its Applications 471 (2015) 122–140.
- [8] K. Ogiwara, T. Fukami, N. Takahashi, Maximizing algebraic connectivity in the space of graphs with a fixed number of vertices and edges, IEEE Transactions on Control of Network Systems 4 (2) (2015) 359–368.
- [9] K. Shahbaz, M. N. Belur, A. Ganesh, Algebraic connectivity: local and global maximizer graphs, arXiv preprint arXiv:2110.01918 (2021).
- [10] D. Mosk-Aoyama, Maximum algebraic connectivity augmentation is NP-hard, Operations Research Letters 36 (6) (2008) 677–679.
- [11] S. Hoory, N. Linial, A. Wigderson, Expander graphs and their applications, Bulletin of the American Mathematical Society 43 (4) (2006) 439–561.
- [12] A. Lubotzky, Expander graphs in pure and applied mathematics, Bulletin of the American Mathematical Society 49 (1) (2012) 113–162.
- [13] P. Erdos, A. Renyi, On Random Graphs. I, Publicationes Mathematicae 6 (1959) 290–297.

- [14] N. Alon, J. H. Spencer, The probabilistic method, Wiley, 2004.
- [15] B. D. McKay, The expected eigenvalue distribution of a large regular graph, Linear Algebra and its Applications 40 (1981) 203–216.
- [16] N. Alon, Eigenvalues and expanders, Combinatorica 6 (2) (1986) 83–96.
- [17] A. Broder, E. Shamir, On the second eigenvalue of random regular graphs, Foundations of Computer Science, 1987., 28th Annual Symposium on (1987) 286–294.
- [18] J. Friedman, On the second eigenvalue and random walks in randomd-regular graphs, Combinatorica 11 (4) (1991) 331–362.
- [19] A. Nilli, Tight estimates for eigenvalues of regular graphs, Electron. J. Combin 11 (9) (2004) 1–4.
- [20] N. C. Wormald, Models of random regular graphs, London Mathematical Society Lecture Note Series (1999) 239–298.
- [21] M. R. Murty, Ramanujan graphs, Journal-Ramanujan Mathematical Society 18 (1) (2003) 33–52.
- [22] E. P. Wigner, Characteristic vectors of bordered matrices with infinite dimensions, Annals of Mathematics 62 (3) (1955) 548–564.
- [23] I. S. A. Longoria, J. A. Mingo, Freely independent coin tosses, standard Young tableaux, and the Kesten-McKay law, arXiv preprint arXiv:2009.11950 (2020).
- [24] J. A. Mingo, R. Speicher, Free probability and random matrices, Vol. 35, Springer, 2017.
- [25] H. S. Wilf, generatingfunctionology, CRC press, 2005.
- [26] J. I. Brown, D. Cox, R. Ehrenborg, The average reliability of a graph, Discrete Applied Mathematics 177 (2014) 19–33.
- [27] H. Pérez-Rosés, Sixty years of network reliability, Mathematics in Computer Science 12 (3) (2018) 275–293.
- [28] G. Exoo, R. Jajcay, Dynamic cage survey, The electronic journal of combinatorics (2012) DS16–July.
- [29] N. Biggs, Constructions for cubic graphs with large girth, the electronic journal of combinatorics (1998) A1–A1.
- [30] B. McKay, W. Myrvold, J. Nadon, Fast backtracking principles applied to find new cages, in: Proceedings of the ninth annual ACM-SIAM symposium on Discrete algorithms, 1998, pp. 188–191.
- [31] G. Exoo, B. D. McKay, W. Myrvold, J. Nadon, Computational determination of (3, 11) and (4, 7) cages, Journal of Discrete Algorithms 9 (2) (2011) 166–169.
- [32] M. O'Keefe, P.-K. Wong, A smallest graph of girth 5 and valency 6, Journal of Combinatorial Theory, Series B 26 (2) (1979) 145–149.
- [33] C. T. Benson, Minimal regular graphs of girths eight and twelve, Canadian Journal of Mathematics 18 (1966) 1091–1094.