ON THE SPECTRAL DISTRIBUTION OF LARGE WEIGHTED RANDOM REGULAR GRAPHS

LEO GOLDMAKHER, CAP KHOURY, STEVEN J. MILLER, AND KESINEE NINSUWAN

ABSTRACT. McKay proved the limiting spectral measures of the ensembles of d-regular graphs with N vertices converge to Kesten's measure as $N \to \infty$. Given a large d-regular graph we assign random weights, drawn from some distribution $\mathcal{W}$, to its edges. We study the relationship between $\mathcal{W}$ and the associated limiting spectral distribution obtained by averaging over the weighted graphs. We establish the existence of a unique 'eigendistribution' (a weight distribution $\mathcal{W}$ such that the associated limiting spectral distribution is a rescaling of $\mathcal{W}$). Initial investigations suggested that the eigendistribution was the semi-circle distribution, which by Wigner's Law is the limiting spectral measure for real symmetric matrices. We prove this is not the case, though the deviation between the eigendistribution and the semi-circular density is small (the first seven moments agree, and the difference in each higher moment is $O(1/d^2)$). Our analysis uses combinatorial results about closed acyclic walks in large trees, which may be of independent interest.

1. INTRODUCTION

The eigenvalues of adjacency matrices associated to graphs encode a wealth of information, and are thus a natural and important object to study and understand. We consider d-regular graphs below. Thus d is always an eigenvalue of the adjacency matrix, and is the largest eigenvalue in absolute value. The simplest application of the eigenvalues is to determine if a graph is connected, which happens if and only if d is a simple eigenvalue. Our next application depends on the difference between the second largest (in absolute value) eigenvalue and d; this is called the spectral gap. A large spectral gap implies many desirable properties. Such graphs are well-connected, meaning the graph has very few edges but all vertices are able to communicate with each other very quickly. These graphs arise in communication network theory, allowing the construction of superconcentrators and non-blocking networks [Bien, Pi], in coding theory [SS] and in cryptography [GILVZ]. Alon [Al] conjectured that as $N \to \infty$, for $d \geq 3$ and any $\epsilon > 0$, "most" d-regular graphs on N vertices have their second largest (in absolute value) eigenvalue at most $2\sqrt{d-1} + \epsilon$; it is known that the $2\sqrt{d-1}$ cannot be improved upon. Friedman [Fr1, Fr2, Fr3] proved this, though the finer behavior around this critical threshold is still open (see [MNS] for numerics and conjectures). For some basics of graph theory and constructions of families of expanders (graphs with a large spectral gap and thus good connectivity properties), see [DSV, LPS, Mar, Sar1, Sar2].

After investigating the largest two eigenvalues and their consequences, it is natural to study the rest of the spectrum. Thirty years ago, McKay [McK] investigated the distribution of eigenvalues

2010 *Mathematics Subject Classification.* 15B52, 05C80, 60F05 (primary), 05C22, 05C38 (secondary).

Key words and phrases. Random graphs, spectral distribution, weighted graphs, regular graphs.

Portions of this work were completed at REUs at AIM and Williams College; we thank our colleagues there for helpful conversations on earlier drafts, and the referee for a very careful reading of the manuscript and many suggestions which improved the exposition. The first named author was partially supported by an NSERC Discovery Grant. The third named author was partially supported by NSF grants DMS0600848, DMS0970067 and DMS1265673. The fourth named author was partially supported by NSF grant DMS0850577, Brown University and Williams College.

of large, random d-regular graphs; ***we always assume our graphs do not contain any self-loops or multiple edges***. When the number of cycles is small relative to the size of the graph (which is true for most d-regular graphs as the number of vertices grows), he proved the existence of a limiting spectral distribution ν_d depending only on d, and gave an explicit formula for ν_d. Recent work [DP, TVW] proves that if we renormalize ν_d so that its associated density function has support $[-1,1]$, then the sequence of renormalized measures converges to Wigner's semicircle measure so long as $d \to \infty$ with the number of vertices. We explore the more complicated situation for randomly weighted regular graphs (see [KSV] for a related problem, where the graphs are not regular). We weigh the graphs by attaching weights to each edge. There is an extensive literature on properties of weighted graphs (where we may weight either the edges or the graphs in the family) and weighted ensembles; see [ALHM, AL, BLMST, BM, Bo1, Bo2, ES, Ga, McD1, McD2, Po, So, Ven] and the references therein for some results and applications.

More precisely, suppose $\mathcal{W}$ is a random variable with finite moments on $\mathbb{R}$ and density $p_{\mathcal{W}}$, and $G \in \mathcal{R}_{N,d}$, the set of simple d-regular graphs on N vertices with no self-loops. We weigh each edge by independent identically distributed random variables (iidrv's) drawn from $\mathcal{W}$. In other words, we replace all nonzero entries in the adjacency matrix of G by iidrv's drawn from $\mathcal{W}$; this is the same as taking the Hadamard product of a real symmetric weight matrix with the graph's adjacency matrix. Denote the spectrum of the weighted graph G by $\{\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_N\}$, and consider the uniform measure $\nu_{d,G,\mathcal{W}}$ on this spectrum:

$$\nu_{d,G,\mathcal{W}}(A) \ = \ \frac{1}{N}\#\big\{j \leqslant N : \lambda_j = A\big\}. \tag{1.1}$$

As indicated by the subscripts, this measure depends on d, G, and $\mathcal{W}$. We are interested in the limiting behavior, so rather than focusing on any particular graph G we take a sequence of graphs of increasing size. We first set some notation.

- $\mathcal{R}_{N,d}$: The set of simple d-regular graphs on N vertices without self-loops.
- $|G|$, a_{ij}, $n_{\rm cyl}(k;G)$: $|G|$ is the number of vertices, $a_{ij} = 1$ if vertices i and j are connected by an edge and 0 otherwise, and $n_{\rm cyl}(k;G)$ is the number of cycles of length k.
- c_m: We set c_m to be the $m^{\rm th}$ moment of the semi-circle distribution, normalized to have variance 1/4; $c_{2k+1} = 0$ and $c_{2k} = \frac{1}{4^k(k+1)}\binom{2k}{k}$ (with $\frac{1}{k+1}\binom{2k}{k}$ the $k^{\rm th}$ Catalan number).
- $\mu_{\mathcal{X}}(k)$, $p_{\mathcal{X}}$, $\mathbf{x}$: For $\mathcal{X}$ a random variable whose density has finite moments, $\mu_{\mathcal{X}}(k)$ is its $k^{\rm th}$ moment and $p_{\mathcal{X}}$ is the density associated to $\mathcal{X}$. Finally, $\mathbf{x}$ is an $N(N+1)/2$ vector (or, equivalently, an $N \times N$ real symmetric matrix) of independent random variables x_{ij} drawn from X. We typically take X to be our weight random variable $\mathcal{W}$.
- $G_{\mathbf{w}}$, $\mu_{d,\mathcal{W}}(k;G)$, $\mu_{d,\mathcal{W}}(k)$: For a fixed d, weight distribution $\mathcal{W}$ and graph G, $G_{\mathbf{w}}$ denotes the graph obtained by weighting the edges of G by $\mathbf{w}$, $\mu_{d,\mathcal{W}}(k;G)$ is the average (over weights drawn from $\mathcal{W}$) $k^{\rm th}$ moment of the associated spectral measure $\nu_{d,G_{\mathbf{w}}}$, while $\mu_{d,\mathcal{W}}(k)$ is the average of $\mu_{d,\mathcal{W}}(k;G)$ over $G \in \mathcal{R}_{N,d}$.

The following result is the starting point of our investigations. The unweighted case is due to McKay [McK]; the existence proof in the general case proceeds similarly.

Theorem 1.1. *For any sequence of d-regular graphs $\{G_i\}$ with $|G_i| \to \infty$ and $n_{\rm cyl}(k;G_i) = o(|G_i|)$ for every $k \geqslant 3$, the limiting distribution $\nu_{d,\mathcal{W}}(x) := \lim_{i\to\infty}\nu_{d,G_i,\mathcal{W}}(x)$ exists and depends only on d and $\mathcal{W}$. In the unweighted case (i.e., each weight is 1) the density is given by Kesten's*

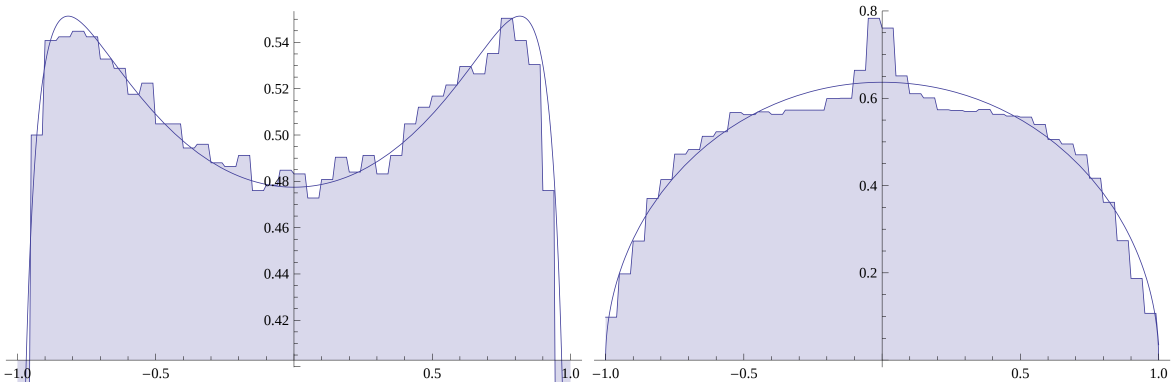

FIGURE 1. Evidence 'supporting' the semi-circular conjecture: 100 4-regular 200×200 matrices. (Left) Unweighted eigenvalue density vs Kesten's measure. (Right) Semi-circular weights vs the semi-circular distribution.

measure:

$$f(x) \;=\; \begin{cases} \frac{d}{2\pi(d^2-x^2)}\sqrt{4(d-1)-x^2}, & |x| \le 2\sqrt{d-1} \\ 0 & \text{otherwise.} \end{cases} \tag{1.2}$$

Note that as $d \to \infty$, Kesten's measure tends to the semi-circle distribution.

The difficulty is deriving good, closed-form expressions when the weights are non-trivial. To this end, we study the one-parameter family of maps

$$T_d : \mathcal{W} \longmapsto \nu_{d,\mathcal{W}} \tag{1.3}$$

(see Theorem 1.1 for a definition of $\nu_{d,\mathcal{W}}$). To understand the behavior of T_d, we investigate its eigendistributions, a concept we now explain. Recall that any measure ν can be rescaled by a real $\lambda > 0$ to form a new measure $\nu^{(\lambda)}$ by setting

$$\nu^{(\lambda)}(A) \;:=\; \nu(\lambda A) \qquad (\text{for all } A \subseteq \mathbb{R}). \tag{1.4}$$

If a distribution $\mathcal{W}$ satisfies

$$T_d\mathcal{W} = \mathcal{W}^{(\lambda)} \tag{1.5}$$

for some $\lambda > 0$, we say $\mathcal{W}$ is an eigendistribution of T_d with eigenvalue λ. We prove in §3 that for each d the map T_d has a unique eigendistribution, up to rescaling; this existence proof is a straightforward application of standard techniques.

Thus the natural question is to determine the eigendistribution for each d. Explicit formulas exist for the moments, but quickly become very involved. Brute force computations show that the first seven moments of the eigendistribution agree with the moments of a semi-circular distribution, suggesting that the semi-circle is the answer. If true this is quite interesting, as the semi-circle is the limiting spectral measure for real symmetric matrices (Wigner's law); moreover, as $d \to \infty$ the limiting spectral measure of the unweighted ensemble of d-regular graphs converges to the semi-circle. In fact, the motivation for this research was the following question: What weights must be introduced so that the weighted ensemble has the semi-circle as its density?

While a determination of the first few moments and numerical investigations (see Figure 1) seemed to support the semi-circle as the eigendistribution, this conjecture is false, though the two distributions are close and agree as $d \to \infty$. For another ensemble where numerical data and heuristic arguments suggested a specific limiting spectral measure which was close to but not equal to the answer, see the work on real symmetric Toeplitz matrices [BDJ, HM].

To state our results precisely, we switch to the language of moments. In §2 we define our notation, which relates the moments of $\mathcal{W}$ and $T_d\mathcal{W}$ in terms of *closed acyclic path patterns*, a combinatorial notion we develop in §2.1. From this we deduce our main result.

Theorem 1.2. *There is a unique eigendistribution of T_d which has second moment equal to $1/4$, which we denote $\mathcal{W}_d$. Let $\mu_{\mathcal{W}_d}(k)$ denote the $k^{\rm th}$ moment of $\mathcal{W}_d$. Then for all non-negative integer k we have $\mu_{\mathcal{W}_d}(2k+1) = 0$ and*

$$\mu_{\mathcal{W}_d}(2k) \ = \ c_{2k} + O\left(1/d^2\right), \tag{1.6}$$

where c_m is the $m^{\rm th}$ moment of the semi-circle distribution normalized to have second moment 1/4. We have $\mu_{\mathcal{W}_d}(2) = 1/4$, $\mu_{\mathcal{W}_d}(4) = 1/8$, $\mu_{\mathcal{W}_d}(6) = 5/64$ (all agreeing with the normalized semi-circular density), but instead of 7/128 (the eight moment of the semi-circle) we find

$$\mu_{\mathcal{W}_d}(8) \ = \ \frac{7}{128} + \frac{1}{128(d^2+d+1)}. \tag{1.7}$$

The eighth moment shows our error term is optimal. The fact that the error decays like $1/d^2$, and not $1/d$, is the consequence of a beautiful combinatorial alignment (see Lemma 2.6).

We concentrate on deriving results about the eigendistribution $\mathcal{W}_d$ and not on the convergence of the individual weighted spectral measures to the average, as the techniques from [McK] and standard arguments (see for example [Bai, Bi, HM, Ta]) suffice to prove such convergence. We only quoted part of Theorem 1.1 of [McK]; the rest of it refers to convergence of the corresponding cumulative distribution functions for graphs satisfying the two conditions in the theorem, and his argument applies with trivial modifications in our setting.

One could also investigate the distribution of gaps between adjacent, normalized eigenvalues. This was studied in [JMRR] for d-regular graphs. Their numerics support a GOE spacing law, which also governs the behavior for the ensemble of real symmetric matrices, but we are far from having a proof in this setting. The distribution of gaps is significantly harder than the density of eigenvalues, and it was only recently (see [ERSY, ESY, TV1, TV2]) where these spacing measures were determined for non-Gaussian random matrix ensembles. There is now a large body of work on the density of eigenvalues and the gaps between them for different structured random matrix ensembles; see [FM, Fo, Meh] for a partial history and the general theory, and [BLMST, BCG, BHS1, BHS2, HM, KKMSX] and their references for some results on structured ensembles.

2. Combinatorial Preliminaries

Below we expand upon the ideas in the introduction, and develop needed combinatorial notions. In particular, we introduce closed acyclic path patterns, which play a crucial role in our work.

We begin by formalizing the notion of a randomly weighted graph. Suppose as before that $G \in \mathcal{R}_{N,d}$ has adjacency matrix $A = \left(a_{ij}\right)$, and let $\mathcal{W}$ be a random variable whose probability density has finite moments. Let $\mathbf{w} = \left\{w_{ij} : 1 \leqslant i \leqslant j \leqslant N\right\}$ denote a set of independent random variables drawn from $\mathcal{W}$, and form an $N \times N$ matrix $A_{\mathbf{w}} = \left(b_{ij}\right)$, where

$$b_{ij} \ = \ \begin{cases} w_{ij}a_{ij} & \text{if } i \leqslant j \\ w_{ji}a_{ji} & \text{otherwise.} \end{cases} \tag{2.1}$$

Observe that $A_{\mathbf{w}}$ is a real symmetric matrix, and that $b_{nn} = 0$ for all n. We may therefore interpret $A_{\mathbf{w}}$ as the adjacency matrix of a weighted graph $G_{\mathbf{w}}$ whose edges are weighted by the random variables $\mathbf{w}$; equivalently, $G_{\mathbf{w}}$ is the Hadamard product of our weight matrix and G's adjacency matrix. We also note that at most dN of the entries b_{ij} are nonzero.

We are interested in the relationship between the distribution $\mathcal{W}$ and the corresponding spectral distribution. Denote the eigenvalues of $A_{\mathbf{w}}$ by $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_N$, and let $\nu_{d,G_{\mathbf{w}}}$ be the uniform measure on this spectrum, as in (1.1); its density is thus

$$\mathrm{d}\nu_{d,G_{\mathbf{w}}}(x) \ = \ \frac{1}{N}\sum_{n=1}^{N}\delta(x-\lambda_n)\mathrm{d}x, \tag{2.2}$$

where $\delta(u)$ is the Dirac delta functional.[1] While we do not need the subscript d as $G_{\mathbf{w}}$ implicitly encodes the degree of regularity d, we prefer to be explicit and highlight the role of this important parameter. By definition and the eigenvalue trace formula, the k^{th} moment $\mu_{\nu_{d,G_{\mathbf{w}}}}(k)$ is

$$\mu_{\nu_{d,G_{\mathbf{w}}}}(k) \ = \ \int_{-\infty}^{\infty} x^k \mathrm{d}\nu_{d,G_{\mathbf{w}}}(x) \ = \ \frac{1}{N}\sum_{n=1}^{N}\lambda_n^k \ = \ \frac{1}{N}\mathrm{Tr}(A_{\mathbf{w}}^k); \tag{2.3}$$

we write $\mu_{\nu_{d,G_{\mathbf{w}}}}(k)$ to emphasize that d is fixed and we are studying a specific weighted graph $G_{\mathbf{w}}$.

The following approach is standard and allows us to convert information on the matrix elements of $A_{\mathbf{w}}$ (which we know) to information on the eigenvalues (which we desire). We have

$$\mathrm{Tr}(A_{\mathbf{w}}^k) \ = \ \sum_{i_1=1}^{N}\sum_{i_2=1}^{N}\cdots\sum_{i_k=1}^{N} b_{i_1i_2}b_{i_2i_3}\cdots b_{i_ki_1}. \tag{2.4}$$

Thus we see that the k^{th} moment of the spectral distribution associated to $G_{\mathbf{w}}$ is the average weight of a closed walk of length k in G (where by the weight of a walk we mean the product of the weights of all edges traversed, counted with multiplicity).

Since we are interested in the dependence on the distribution $\mathcal{W}$, not on the specific values of the $N(N+1)/2$ random variables $\mathbf{w} = (w_{ij})_{1\le i,j\le N}$, we average over $\mathbf{w}$ drawn from $\mathcal{W}$'s density $p_{\mathcal{W}}$ to obtain the 'typical' k^{th} moment $\mu_{d,\mathcal{W}}(k;G)$ of the weighted spectral distributions:

$$\mu_{d,\mathcal{W}}(k;G) \ := \ \int_{-\infty}^{\infty}\mu_{\nu_{d,G_{\mathbf{w}}}}(k)\mathrm{d}\mathbf{w} \ = \ \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\mu_{\nu_{d,G_{\mathbf{w}}}}(k)\prod_{1\le i\le j\le N}p_{\mathcal{W}}(w_{ij})\mathrm{d}w_{ij}, \tag{2.5}$$

where $p_{\mathcal{W}}$ is the density function corresponding to distribution $\mathcal{W}$.

To build intuition for the later calculations, we calculate the first and second moments.

Lemma 2.1. *Fix d, $G \in \mathcal{R}_{N,d}$ and $\mathcal{W}$. We have $\mu_{d,\mathcal{W}}(1;G) = 0$ and $\mu_{d,\mathcal{W}}(2;G) = d\mu_{\mathcal{W}}(2)$, where $\mu_{\mathcal{W}}(k)$ is the k^{th} moment of $\mathcal{W}$. Thus $\mu_{d,\mathcal{W}}(1) = 0$ and $\mu_{d,\mathcal{W}}(2) = d\mu_{\mathcal{W}}(2)$.*

Proof. Since $b_{nn} = 0$ for all n, we see that

$$\mu_{d,\mathcal{W}}(1;G) \ = \ \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\mu_{\nu_{d,G_{\mathbf{w}}}}(1)\mathrm{d}\mathbf{w} \ = \ \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\frac{1}{N}\sum_{n=1}^{N}b_{nn}\prod_{1\le i\le j\le N}p_{\mathcal{W}}(w_{ij})\mathrm{d}w_{ij} \ = \ 0. \tag{2.6}$$

[1]We write d for the degree of regularity, and d for differentials.

For the second moment, we use G is d-regular, $b_{ij} = a_{ij}w_{ij}$, and $b_{nn} = 0$ and $b_{ij} = b_{ji}$. The number of non-zero a_{ij} is $dN/2$ (each vertex has d edges emanating from it, and each edge is doubly counted), and recall $\mu_{\mathcal{W}}(2)$ denotes the second moment of the weight distribution $\mathcal{W}$. Thus

$$\begin{aligned}
\mu_{d,\mathcal{W}}(2;G) &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\mu_{\nu_{d,G_{\mathbf{w}}}}(2)\mathrm{d}\mathbf{w} \ = \ \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\frac{2}{N}\sum_{1\leqslant i<j\leqslant N} b_{ij}^2 p_{\mathcal{W}}(w_{ij})\mathrm{d}w_{ij} \\
&= \frac{2}{N}\sum_{\substack{1\leqslant i<j\leqslant N\\ a_{ij}=1}}\int_{-\infty}^{\infty} w_{ij}^2 p_{\mathcal{W}}(w_{ij})\mathrm{d}w_{ij} \ = \ \frac{2}{N}\sum_{\substack{1\leqslant i<j\leqslant N\\ a_{ij}=1}}\mu_{\mathcal{W}}(2) \\
&= \frac{2}{N}\cdot\frac{dN}{2}\mu_{\mathcal{W}}(2) \ = \ d\mu_w(2). \qquad (2.7)
\end{aligned}$$

□

The first two moments are independent of G; however, this is not the case for higher moments (for example, in the third moment we have the possibility of a loop). For these higher moments, we need to perform an averaging over G as well, and study

$$\mu_{d,\mathcal{W}}(k) \ = \ \frac{1}{|\mathcal{R}_{N,d}|}\sum_{G\in\mathcal{R}_{N,d}}\mu_{d,\mathcal{W}}(k;G). \qquad (2.8)$$

While we can compute any $\mu_{d,\mathcal{W}}(k)$, the calculations quickly become very involved, and indicate the need for a unified approach if we desire a tractable closed form expression. For example, the average (over weights drawn from a fixed $\mathcal{W}$ and $G \in \mathcal{R}_{N,d}$) for the next two even moments are

$$\begin{aligned}
\mu_{d,\mathcal{W}}(4) &= d\mu_{\mathcal{W}}(4) + 2d(d-1)\mu_{\mathcal{W}}(2)^2 \\
\mu_{d,\mathcal{W}}(6) &= d\mu_{\mathcal{W}}(6) + 6d(d-1)\mu_{\mathcal{W}}(4)\mu_{\mathcal{W}}(2) + [3d(d-1)^2 + 2d(d-1)(d-2)]\mu_{\mathcal{W}}(2)^3,
\end{aligned} \qquad (2.9)$$

where as always $\mu_{\mathcal{W}}(k)$ denotes the k^{th} moment of the weight distribution $\mathcal{W}$ (the odd moments are easily shown to vanish). We prove these formulas in Lemma 2.4.

Recall that our goal is to find a distribution $\mathcal{W}$ so that $T_d\mathcal{W} = \mathcal{W}^{(\lambda)}$ for some λ, normalized to have second moment equal to 1/4 (the second moment of the semi-circle). Our second moment calculation in Lemma 2.1 suggests that $\lambda = \sqrt{d}$. If the semi-circle is a fixed eigendistribution, then we must have $\mu_{d,\mathcal{W}}(4) = d^2/8$ and $\mu_{d,\mathcal{W}}(6) = 5d^3/64$. From (2.9), we see that if we choose $\mathcal{W}$ so that the fourth moment is 1/8 then we do get $\mu_{d,\mathcal{W}}(4) = d^2/8$, and if the sixth moment of $\mathcal{W}$ is also $5/64$ then $\mu_{d,\mathcal{W}}(6) = 5d^3/64$. These results suggest that we can inductively show that the semi-circle is a fixed eigendistribution, but a more involved calculation (see Lemma 2.4) shows this breaks down at the eighth moment:

$$\begin{aligned}
\mu_{d,\mathcal{W}}(8) &= d\mu_{\mathcal{W}}(8) + 8d(d-1)\mu_{\mathcal{W}}(6)\mu_{\mathcal{W}}(2) + 6d(d-1)\mu_{\mathcal{W}}(4)^2 \\
&\quad + 16d(d-1)^2\mu_{\mathcal{W}}(4)\mu_{\mathcal{W}}(2)^2 + 12d(d-1)(d-2)\mu_{\mathcal{W}}(4)\mu_{\mathcal{W}}(2)^2 \\
&\quad + 4d(d-1)^3\mu_{\mathcal{W}}(2)^4 + 8d(d-1)^2(d-2)\mu_{\mathcal{W}}(2)^4 \\
&\quad + 2d(d-1)(d-2)(d-3)\mu_{\mathcal{W}}(2)^4. \qquad (2.10)
\end{aligned}$$

If $\mathcal{W}$ is an eigendistribution of T_d with $\lambda = \sqrt{d}$ then $\mu_{d,\mathcal{W}}(8)$ must equal $d^4\mu_{\mathcal{W}}(8)$, which implies

$$\mu_{\mathcal{W}}(8) \ = \ \frac{7}{128} + \frac{1}{128(d^2+d+1)}. \qquad (2.11)$$

This is almost, but not quite, 7/128, the eighth moment of the normalized semi-circle.

To unify the derivation of (2.9) and (2.10), as well as the higher moments, we introduce some notation. This allows us to give a compact, tractable closed form expression for these moments, and helps us prove that there is a unique eigendistribution (and determine its moments).

2.1. **Closed acyclic path patterns.** From (2.3) and (2.4), it is clear that moments of the spectral distribution are closely related to the set of closed walks in G. Moreover, we shall demonstrate below that it suffices to restrict our attention to walks containing no cycles, as all the walks with at least one closed cycle contribute a negligible amount to (2.4). We now introduce a combinatorial object which will keep track of all closed walks on a large tree.

Definition 2.2. *A* closed acyclic path pattern *(*CAPP*) is a string of symbols such that*

1. *every symbol which appears at all appears an even number of times; and*
2. *in the substring of symbols between any two consecutive instances of the same symbol, every symbol which appears at all appears an even number of times.*

We call two CAPPs *equivalent* if they differ only by a relabeling of the symbols. The following is the *raison d'être* for our definition.

Lemma 2.3 (Classification of closed walks)**.** *The closed acyclic path patterns classify the closed walks beginning at a given vertex in a large tree.*

Proof. There is a natural map from the set of paths (closed or not) in a large tree to the set of sequences, where we treat the edges as symbols and just record the edges used in order. It is evident that this map is "injective" (the relevant equivalence relations on paths and sequences coincide). There are two issues. We must show

1. every closed path corresponds to a sequence which is a CAPP; and
2. every CAPP is realizable as the edge sequence of some path.

These are not hard to see. Removing any edge from a tree disconnects the tree into two connected components, so we can ask if two vertices are on the "same side" of an edge or on "opposite sides". Furthermore two vertices are on the same side of *every* edge if and only if they are the same vertex. If we follow a path in a tree, then the start and end points are on the same side of an edge if and only if we traverse that edge an even number of times. By a straightforward induction on the length of the path/sequence, a sequence corresponds to an actual path in a tree if and only if the second condition in the definition of a CAPP holds. Likewise, a path is closed if and only if the corresponding sequence satisfies the first condition in the definition of a CAPP holds. □

We can now define the terms that will appear in Lemma 2.4, our closed form expression for the moments $\mu_{d,\mathcal{W}}(k)$. Given a CAPP π, let $e_1, e_2, e_3, \dots, e_r$ denote all the distinct symbols appearing in π, in order of appearance. Equivalently, the e_i denote the edges composing the walk represented by π, ordered by first traversal. We need the following definitions.

- We denote the set of (equivalence classes of) CAPPs of length k by $\mathcal{P}_k$. Note $\mathcal{P}_k$ is empty for k odd. For $\pi \in P_{2k}$, we define the **diagram** of π to be the minimal ordered, rooted tree which is traversed by the path described by the pattern, with edges repeated according to how often the edge is traversed in each direction.
- The **multiplicity** of a CAPP π is $m_\pi(d)$, where m_π is the polynomial

$$m_\pi(x) \;=\; \prod_{j=1}^{r}(x-\alpha_j), \tag{2.12}$$

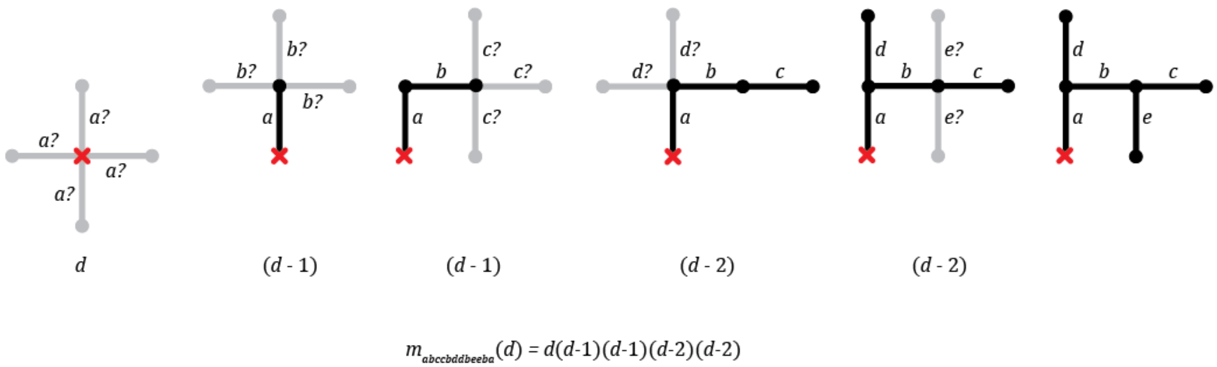

FIGURE 2. A realization of a particular CAPP in a d-regular graph with $d = 4$; this illustrates the multiplicity formula as an instance of standard counting principles.

where $\alpha_j := \#\{i < j : e_i \text{ is adjacent to } e_j\}$; we call α_j the multiplicity of edge e_j. Note that $(d - \alpha_j)/d$ is the proportion of edges emanating from vertex j that are not yet used in π when vertex j is first visited. This is used in calculating contributions to the moments, as $d - \alpha_j$ represents the number of possibilities available in choosing the next distinct edge. We measure adjacency by looking at the edges on the tree, *not* by the ordering of the edges in our symbol. Thus if $\pi = abccbddbeeba$ the multiplicity of a is 0, that of b and of c is 1, and that of d and of e is 2. Figure 2 illustrates (in the case of a 4-regular graph) how the number of choices at each stage depends on the shape of the path so far.

- The **signature** of π is
$$\sigma(\pi) \ := \ (n_1, n_2, \dots, n_r), \tag{2.13}$$
where n_i denotes the number of times the symbol e_i appears in π. Thus each n_i is a positive integer. If $\pi \in \mathcal{P}_k$ then the sum of the entries of its signature is k.
  - ◇: $\mathcal{P}_k^{(2)}$ is the set of all CAPPs in $\mathcal{P}_k$ with signature $(2, 2, \dots, 2)$.
  - ◇: $\mathcal{P}_k^{(4)}$ is the set of all CAPPs in $\mathcal{P}_k$ with signature $(4, 2, \dots, 2)$.
  - ◇: $\mathcal{P}_k^{\circ}$ is the set of all CAPPs in $\mathcal{P}_k$ excluding the pattern with signature (k).
- Given a signature $\sigma(\pi) = (n_1, n_2, \dots, n_r)$ and a random variable $\mathcal{W}$, the **moment contribution associated to** π **with respect to** $\mathcal{W}$ is
$$\mu_{\mathcal{W}}(\sigma(\pi)) \ := \ \mu_{\mathcal{W}}(n_1)\mu_{\mathcal{W}}(n_2)\cdots\mu_{\mathcal{W}}(n_r). \tag{2.14}$$

We can now give a complete description of the moments of the limiting spectral distribution (averaging over weights drawn from a fixed $\mathcal{W}$ and averaging over $G \in \mathcal{R}_{N,d}$ with $N \to \infty$). Our answer is in terms of the moments of the weight distribution $\mathcal{W}$ and some combinatorial data.

Lemma 2.4 (Moment Expansion)**.** *Fix a weight $\mathcal{W}$ and a degree of regularity d. Let $\mu_{\mathcal{W}}(k)$ be the k^{th} moment of $\mathcal{W}$, $\mu_{d,\mathcal{W}}(k)$ the average over $G \in \mathcal{R}_{N,d}$ and over weights w_{ij} drawn from $\mathcal{W}$ of the k^{th} moments of the measures ν_{d,G_w}, and $\mathcal{P}_k$ the collection of all* CAPP*s of length k. Then*
$$\mu_{d,\mathcal{W}}(k) \ = \ \sum_{\pi \in \mathcal{P}_k} m_\pi(d)\mu_{\mathcal{W}}(\sigma(\pi)), \tag{2.15}$$
where $m_\pi(d)$, $\sigma(\pi)$ and $\mu_{\mathcal{W}}(\sigma(\pi))$ are defined in (2.12) *through* (2.14)*, and as Carleman's condition is satisfied the moments uniquely determine a probability distribution.*

As $\mathcal{P}_k$ is trivially empty for k odd, Lemma 2.4 implies all odd moments vanish in the limit.

2.2. **Proof of Lemma 2.4.** Before proving Lemma 2.4, we show it is reasonable by deriving its prediction for the moment expansions of (2.9) (we leave the eighth moment, (2.10), to the reader). For the fourth moment, we need all CAPPs of length 4. We have

$$\mathcal{P}_4 \ = \ \{\pi_{4;1} = e_1e_1e_1e_1,\ \pi_{4;2} = e_1e_2e_2e_1,\ \pi_{4;3} = e_1e_1e_2e_2\}\,. \tag{2.16}$$

The signatures are $\sigma(\pi_{4;1}) = (4)$, $\sigma(\pi_{4;2}) = (2,2)$, $\sigma(\pi_{4;3}) = (2,2)$. Recall the multiplicity α_j of π is the number of $i < j$ such that e_i is adjacent to e_j. We have

$$m_{\pi_{4,1}}(d) \ = \ d-0, \quad m_{\pi_{4;2}}(d) \ = \ (d-0)(d-1), \quad m_{\pi_{4;3}}(d) \ = \ (d-0)(d-1). \tag{2.17}$$

Thus

$$\sum_{\pi\in\mathcal{P}_4} m_\pi(d)\mu_{\mathcal{W}}(\sigma(\pi)) \ = \ d\mu_{\mathcal{W}}(4) + 2d(d-1)\mu_{\mathcal{W}}(2)\mu_{\mathcal{W}}(2), \tag{2.18}$$

in agreement with the first part of (2.9).

The calculation of the sixth moment is more involved, as we need to carefully determine the multiplicities. There are three cases. Note the sum of the entries of the signatures must equal 6, so there are only three possibilities: (6), $(4,2)$, and $(2,2,2)$.

- Signature of (6): The only π that gives this is $e_1e_1e_1e_1e_1e_1$. The multiplicity is $d-0$, and the contribution is $d\mu_{\mathcal{W}}(6)$.
- Signature of $(4,2)$: There are six possibilities: $e_1e_1e_1e_1e_2e_2$, $e_1e_1e_1e_2e_2e_1$, $e_1e_1e_2e_2e_1e_1$, $e_1e_2e_2e_1e_1e_1$, $e_1e_2e_2e_2e_2e_1$ and $e_1e_1e_2e_2e_2e_2$. We always have e_1 occur first, but note that either e_1 or e_2 could be the most frequently occurring symbol. The other possibilities violate some of the conditions; for example, both $e_1e_1e_2e_1e_1e_2$ and $e_1e_2e_1e_1e_2e_1$ violate the second condition in Definition 2.2 as each has the consecutive string $e_1e_2e_1$. Each of the six valid choices has signature $d(d-1)$, and the total contribution is thus $6d(d-1)\mu_{\mathcal{W}}(4)\mu_{\mathcal{W}}(2)$.
- Signature of $(2,2,2)$: This is the first non-trivial case, as we have to carefully look and see where we are in our walk to determine the multiplicity. There are five terms. Three have multiplicity $d(d-1)^2$; they are $e_1e_1e_2e_3e_3e_2$, $e_1e_2e_3e_3e_2e_1$ and $e_1e_2e_2e_1e_3e_3$. Two have multiplicity $d(d-1)(d-2)$; they are $e_1e_1e_2e_2e_3e_3$ and $e_1e_2e_2e_3e_3e_1$. For example, for the last one we start at vertex 0 and move to vertex 1 by e_1, then to vertex 2 by e_2, then back to vertex 1 by e_2, then to vertex 3 by e_3, back to vertex 1 by e_3 and then return to vertex 0 by e_1. As all edges include vertex 1, they are all adjacent, thus $\alpha_1 = 0$, $\alpha_2 = 1$ and $\alpha_3 = 2$. The contribution from these five terms is $3d(d-1)^2\mu_{\mathcal{W}}(2)^3 + 2d(d-1)(d-2)\mu_{\mathcal{W}}(2)^3$.

We now turn to the proof of the Moment Expansion Lemma. We start with an informal discussion of the issues. We know that we can write the n^{th} spectral moment of a d-regular graph (not worrying yet about a limit along a sequence of graphs or averaging over the weights) as a sum of terms, where each term corresponds to a closed path in the graph of length n. On the other hand, the summation in Lemma 2.4 can also be thought of as a sum of similar terms if we interpret the summand $m_\pi(d)\mu_{\mathcal{W}}(\sigma(\pi))$ as $m_\pi(d)$ separate summands $\mu_{\mathcal{W}}(\sigma(\pi))$, one for each of the paths starting and ending at a given vertex with pattern π. While these summations are similar, they are *not* identical, since G is not a tree but rather a specific d-regular graph which may or may not contain cycles. There are qualitatively different types of discrepancy here, both caused by small cycles:

- Paths which actually include a non-trivial cycle have no corresponding summand in our formula, since there is no path through a d-ary tree involving a cycle.
- Paths which go partway around a cycle in both directions may have a corresponding summand in our formula, but their weights do not match. For example, suppose there is a triangle with vertices u, v, w, where u is the root. Then the length 8 path u, v, w, v, u, w, v, w, u uses edge uv twice, edge uw twice, and edge vw four times, so this path contributes $\mu_{\mathcal{W}}(2)^2\mu_{\mathcal{W}}(4)$ to the summation. This gives the "pattern" $abbacbbc$, which is not a CAPP because of the substring $bacb$. The closest analogous path in an *acyclic* graph would involve four different edges, since the two edges which are at distance 2 from the root in opposite directions cannot be the same edge. This path *does* correspond to a term in our formula, $abbacddc$, but the signature is different. That path contributes $\mu_{\mathcal{W}}(2)^4$ to our formula (2.15).

The idea of the proof is to determine the contribution from a tree, and bound the average deviation of our d-regular graphs from being a tree. Although (2.15) does not give the correct spectral moments for individual graphs, it can give the correct limiting spectral moments for a sequence of graphs. The technical condition that the number of small cycles in the graphs is growing slowly is precisely what is needed to guarantee that these discrepancies vanish in the limit. Fortunately there exist good bounds on the numbers of such small cycles in the family $\mathcal{R}_{N,d}$.

Proof of Lemma 2.4. We first recall some notation. Given a d-regular graph G on N vertices (so $G \in \mathcal{R}_{N,d}$) and a probability distribution $\mathcal{W}$, we form the weighted graph $G_{\mathbf{w}}$ whose edges are weighed by iidrv's drawn from $\mathcal{W}$. We denote average (with respect to the weights w_{ij} being drawn from $\mathcal{W}$) of the k^{th} moment of the associated spectral distributions $\nu_{d,G_{\mathbf{w}}}$ by $\mu_{d,\mathcal{W}}(k;G)$. From (2.3) and (2.4) we know that $\mu_{d,\mathcal{W}}(k;G)$ is the average weight of a closed walk of length k in G. The first step in the proof is to show that only acyclic walks contribute significantly to this average; i.e., all walks which contain cycles contribute a negligible amount.

We thus consider a closed path of length k, denoting the vertices by $i_1, i_2, \dots, i_k$. Let

$$C_{d,G_{\mathbf{w}}}(k) \ := \sum_{\substack{\langle i_1,i_2,\dots,i_k,i_1\rangle \\ \text{contains a cycle}}} b_{i_1i_2}b_{i_2i_3}\cdots b_{i_ki_1} \tag{2.19}$$

denote the contribution to the k^{th} moment of $\nu_{d,G_{\mathbf{w}}}$ from paths containing a cycle, and

$$A_{d,G_{\mathbf{w}}}(k) \ := \sum_{\substack{\langle i_1,i_2,\dots,i_k,i_1\rangle \\ \text{contains no cycles}}} b_{i_1i_2}b_{i_2i_3}\cdots b_{i_ki_1} \tag{2.20}$$

the contribution from the acyclic closed paths. We may thus rewrite equations (2.3) and (2.4) as

$$\mu_{d,G_{\mathbf{w}}}(k) \ = \ \frac{1}{N}C_{d,G_{\mathbf{w}}}(k) + \frac{1}{N}A_{d,G_{\mathbf{w}}}(k). \tag{2.21}$$

We prove that the first term tends to 0 as $N \to \infty$, which implies that $\mu_{d,\mathcal{W}}(k)$ only depends on paths with no cycles (i.e., CAPPs). From Lemma 2.1 we may assume $k \geq 3$. Fix a $G \in \mathcal{R}_{N,d}$ and a weight vector **w** with components independently drawn from $\mathcal{W}$. We may take all but $Nd/2$ of the entries of **w** to be 0 without affecting the weighted adjacency matrix; for notational convenience we label those weights which aren't necessarily 0 by $\{w_1, w_2, \dots, w_s\}$ (where $s = Nd/2$).

Choose some k-cycle in G; as it can only traverse these weighted edges, its contribution is $w_1^{r_1}w_2^{r_2}\cdots w_s^{r_s}$, where $r_i \geqslant 0$ and $\sum r_i = k$. Here r_i represents the total number of times our

k-cycle traverses the edge with weight w_i. Averaging over $\mathcal{W}$ and using the independence of the weights, we have that the expected contribution of a k-cycle is

$$\begin{aligned}\mathbb{E}[w_1^{r_1}\cdots w_s^{r_s}] &= \mathbb{E}[w_1^{r_1}]\cdots\mathbb{E}[w_s^{r_s}] \\ &= \mu_{\mathcal{W}}(r_1)\mu_{\mathcal{W}}(r_2)\cdots\mu_{\mathcal{W}}(r_s) \\ &= \mu_{\mathcal{W}}(1)^{\beta_1}\mu_{\mathcal{W}}(2)^{\beta_2}\cdots\mu_{\mathcal{W}}(s)^{\beta_s}\end{aligned} \tag{2.22}$$

for some non-negative integers β_i satisfying $\sum_{i=1}^{s} i\beta_i = k$. This immediately implies $\alpha_{k+1} = \alpha_{k+2} = \cdots = \alpha_s = 0$, whence

$$\mathbb{E}[w_1^{r_1}\cdots w_s^{r_s}] = \mu_{\mathcal{W}}(1)^{\alpha_1}\cdots\mu_{\mathcal{W}}(k)^{\alpha_k}. \tag{2.23}$$

Let

$$M = \max\left\{\mu_{\mathcal{W}}(1)^{\alpha_1}\mu_{\mathcal{W}}(2)^{\alpha_2}\cdots\mu_{\mathcal{W}}(k)^{\alpha_k} : \alpha_i \geqslant 0, \sum_{i=1}^{k} i\alpha_i = k\right\}. \tag{2.24}$$

Note that M depends only $\{\mu_{\mathcal{W}}(i)\}_{i=1}^{k}$ (the first k moments of $\mathcal{W}$) and k; in particular, it is bounded independent of N. We highlight this fact by writing $M = M(\mathcal{W}, k)$.

Let $C_{G,i}$ be the number of i-cycles in G, and let $C_{i,N,d}$ be the average number of of cycles of length i in the ensemble of d-regular graphs on N vertices. For a fixed weight distribution $\mathcal{W}$, the contribution of paths with cycles to $\mu_{d,\mathcal{W}}(k)$ from averaging over weights drawn from $\mathcal{W}$ and graphs $G \in \mathcal{R}_{N,d}$ is

$$\begin{aligned}&\frac{1}{|\mathcal{R}_{N,d}|}\sum_{G\in\mathcal{R}_{N,d}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\frac{1}{N}C_{d,G_{\mathbf{w}}}(k)\prod_{1\le i\le j\le N}p_{\mathcal{W}}(w_{ij})\mathrm{d}w_{ij} \\ &\le \frac{1}{|\mathcal{R}_{N,d}|}\sum_{G\in\mathcal{R}_{N,d}}\left(\frac{1}{N}\sum_{i=3}^{k}C_{G,i}M(\mathcal{W},k)\right) \\ &= \frac{M(\mathcal{W},k)}{N}\cdot\sum_{i=3}^{k}\frac{1}{|\mathcal{R}_{N,d}|}\sum_{G\in\mathcal{R}_{N,d}}C_{G,i} = \frac{M(\mathcal{W},k)}{N}\cdot\sum_{i=3}^{k}C_{i,N,d},\end{aligned} \tag{2.25}$$

Here $i \geqslant 3$, as otherwise G has no cycles. By Lemma 4.1 of [McK], for $i \ge 3$ we have

$$\lim_{N\to\infty} C_{i,N,d} = \frac{(d-1)^i}{2i}. \tag{2.26}$$

Combining this with the above, we deduce that the contribution from the paths with a cycle to $\mu_{d,\mathcal{W}}(k)$ is $O(1/N)$, and thus negligible as $N \to \infty$. In particular, this implies

$$\mu_{d,\mathcal{W}}(k) = \lim_{N\to\infty}\frac{1}{|\mathcal{R}_{N,d}|}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\frac{1}{N}A_{d,G_{\mathbf{w}}}(k)\prod_{1\le i\le j\le N}p_{\mathcal{W}}(w_{ij})\mathrm{d}w_{ij}. \tag{2.27}$$

The proof is completed by noting that this is equivalent to $\sum_{\pi\in\mathcal{P}_k} m_\pi(d)\mu_{\mathcal{W}}(\sigma(\pi))$. This follows from the definition of CAPPs, multiplicities and signatures, and similar arguments as in [McK]. The factor $\mu_{\mathcal{W}}(\sigma(\pi))$ comes from how often each weight occurs and averaging over the weights.

The factor $m_\pi(d)$ requires a bit more work. As we take the limit as $N \to \infty$, there is no loss in assuming we have a tree. As we need a closed path to have a contribution, each k must be even and each edge must be traversed an even number of times (as we have a tree, there are no cycles). By Lemma 2.3 there is a one-to-one correspondence between CAPPs and legal walks along edges. Each time we hit a vertex and go off along a new edge, the number of choices we have equals d

(the regularity degree) minus the number of edges we have already taken from the vertex. This is why the multiplicity of edge e_j is d minus the number of edges e_i adjacent to e_j with $i < j$, and adjacency is measured relative to the tree and *not* the string of edges. This completes the proof. □

While Lemma 2.4 gives a closed form expression for the limiting moments, it is not immediately apparent that it is a *useful* expansion. We need a way of computing the sum over π of $m_\pi(d)\mu_{\mathcal{W}}(\sigma(\pi))$, which is our next subject.

2.3. **Counting walks by signature.** We conclude this section by showing how to count walks with certain simple signatures. We use these results to prove our theorems on eigendistributions in §3. We first recall some notation.

- $\mathcal{P}_k^{(2)}$ is the set of all CAPPs in $\mathcal{P}_k$ with signature $(2, 2, 2, \ldots, 2)$.
- $\mathcal{P}_k^{(4)}$ is the set of all CAPPs in $\mathcal{P}_k$ with signature $(4, 2, 2, \ldots, 2)$.
- $\mathcal{T}_k$ is the set of all triples (π, x, y), where $\pi \in \mathcal{P}_k^{(2)}$ and x, y are symbols corresponding to distinguished edges in the diagram, which must be adjacent and first traversed in that order.
- $\mathcal{P}_k^{\circ}$ is the set of all CAPPs in $\mathcal{P}_k$ excluding the pattern with signature (k).

Lemma 2.5 (Counting Walks without Repeated Edges)**.** *There are exactly* $\frac{1}{k+1}\binom{2k}{k}$ CAPPs *of length* $2k$ *and signature* $(2, 2, 2, \ldots, 2)$*. That is,* $|\mathcal{P}_{2k}^{(2)}| = \frac{1}{k+1}\binom{2k}{k}$.

Proof. Walks of signature $(2, 2, 2, \ldots, 2)$ use each edge exactly twice. Such a walk is determined by its diagram, regarded as an ordered tree (in the sense that the children of each vertex "remember" in which order they were visited). It is well-known that the Catalan numbers count such trees, and appear throughout random matrix theory (see for example [AGZ]). □

The following lemma plays a key role in computing the lower order term to the moments in Theorem 1.2, and allows us to improve our error from $O(1/d)$ to $O(1/d^2)$.

Lemma 2.6 (Serendipitous Correspondence)**.** *There is a two-to-one correspondence between length-*$2k$ CAPPs *whose signature is* $(2, 2, 2, \ldots, 2)$ *with a distinguished pair of adjacent edges, and length-*$2k$ CAPPs *with signature* $(4, 2, 2, \ldots, 2)$*. That is,* $|\mathcal{T}_{2k}| = 2|\mathcal{P}_{2k}^{(4)}|$.

Proof. In this proof, we always use the symbols x, y (in that order) as the distinguished symbols for an object in $\mathcal{T}_{2k}$. These symbols will occur either in the order $xyyx$ or $xxyy$ (the order cannot be $xyxy$ or the CAPP condition would be violated). Consider sequences $AxByCyDxE$ (case 1) or $AxBxCyDyE$ (case 2), where capital letters denote substrings, where every symbol occurring in ABCDE does so exactly twice in total. In order for this to be a genuine pattern, for each non-distinguished symbol that occurs, one of the following must be true.

(1) Both occurrences are in the same substring (A, B, C, D, or E).
(2) One occurrence is in A and the other is in E.
(3) One occurrence is in B and the other is in D (case 1 only).
(4) One occurrence is in C and the other is in A or E (case 2 only).

Since x and y are adjacent edges, the last two possibilities are ruled out. But now it is not hard to see that elements of $\mathcal{P}_{2k}^{(4)}$ have the form $AzBzCzDzE$, with precisely the same conditions on A, B, C, D, E. Then we correspond the patterns $AxByCyDxE$ and $AxBxCyDyE$ to the pattern $AzBzCzDzE$, giving the desired two-to-one correspondence. □

3. THE EIGENDISTRIBUTION

Our goal is to find the eigendistributions $\mathcal{W}^{(\lambda)}$ of the maps T_d from (1.3). Recall that T_d maps a given weight distribution $\mathcal{W}$ to a spectral distribution. In this section we prove that for each d there exists a unique (up to rescaling) eigendistribution of T_d. To do this, we first apply Lemma 2.4 to obtain a recursive identity on the moments of any eigendistribution; it will then be seen that there exists a distribution possessing these moments. Moreover, we show that after appropriate rescaling, the moments grow very similarly to those of the semicircle distribution. The two distributions are not exactly the same, and we quantify the extent to which they differ.

We first prove that for any fixed d, T_d has at most one eigenvalue. Given any distribution $\mathcal{W}$ with density $p_{\mathcal{W}}$ and any $\lambda > 0$, let $\mu_{\mathcal{W}}(k)$ denote the k^{th} moment of $\mathcal{W}$ and $\mu_{\mathcal{W}^{(\lambda)}}(k)$ denote the k^{th} moment of the rescaled distribution $\mathcal{W}^{(\lambda)}$ (see (1.4) for the effect of scaling by λ). We have

$$\mu_{\mathcal{W}^{(\lambda)}}(k) \;=\; \int_{-\infty}^{\infty} x^k\, d\mathcal{W}^{(\lambda)}(x) \;=\; \int_{-\infty}^{\infty} x^k \lambda\, p_{\mathcal{W}}(\lambda x)\,\mathrm{d}x \;=\; \int_{-\infty}^{\infty} \left(\frac{x}{\lambda}\right)^k p_{\mathcal{W}}(x)\,\mathrm{d}x \;=\; \frac{1}{\lambda^k}\mu_{\mathcal{W}}(k). \tag{3.1}$$

In particular, if $\mathcal{W}$ is an eigendistribution with eigenvalue λ, and $\mu_{d,\mathcal{W}}(k)$ denotes the moments of the spectral distribution $T_d\mathcal{W} = \mathcal{W}^{(\lambda)}$, then $\mu_{d,\mathcal{W}}(2) = \lambda^{-2}\mu_{\mathcal{W}}(2)$. On the other hand, from Lemma 2.1 we know that $d^{-1}\mu_{d,\mathcal{W}}(2) = \mu_{\mathcal{W}}(2)$, whence $\lambda = d^{-1/2}$. We thus obtain a relation for the even moments of an eigendistribution:

$$\mu_{d,\mathcal{W}}(2k) \;=\; d^k \mu_{\mathcal{W}}(2k). \tag{3.2}$$

Substituting this into Lemma 2.4 and simplifying yields the following formula.

Lemma 3.1 (Eigenmoment Formulas)**.** *Suppose $\mathcal{W}_d$ is an eigendistribution of T_d, i.e., $T_d\mathcal{W}_d = \mathcal{W}_d^{(\lambda)}$ for some $\lambda > 0$. Denote the moments of $\mathcal{W}_d$ by $\mu_{\mathcal{W}_d}(k)$. We may assume (without loss of generality) that $\mathcal{W}_d$ is scaled so that $\mu_{\mathcal{W}_d}(2) = 1/4$ (the second moment of the normalized semi-circle distribution). Then $\mu_{\mathcal{W}_d}(k) = 0$ for all odd k, and*

$$\mu_{\mathcal{W}_d}(2k) \;=\; \frac{1}{d^k - d} \sum_{\pi \in \mathcal{P}_{2k}^{\circ}} m_\pi(d)\mu_{\mathcal{W}_d}(\sigma(\pi)). \tag{3.3}$$

We can now prove Theorem 1.2, namely that there exists a unique eigendistribution, as well as determine properties of its moments.

Proof of Theorem 1.2. As the signature $\sigma(\pi)$ involves numbers strictly smaller than $2k$ for all $\pi \in \mathcal{P}_{2k}^{\circ}$, (3.3) gives a recursive formula for the moments. Thus, if an eigendistribution $\mathcal{W}_d$ exists, then its moments are uniquely specified. The even moments are easily bounded above by 1 and below by the moments of the normalized semi-circular distribution. Thus Carleman's condition is satisfied ($\sum_{k=1}^{\infty} \mu_{\mathcal{W}_d}(2k)^{-1/2k} = \infty$), and the moments uniquely determine a distribution (see [Bi, Ta]).

From Lemma 3.1 we easily find $\mu_{\mathcal{W}_d}(2) = 1/4$, $\mu_{\mathcal{W}_d}(4) = 1/8$, $\mu_{\mathcal{W}_d}(6) = 5/64$, and

$$\mu_{\mathcal{W}_d}(8) \;=\; \frac{7}{128} + \frac{1}{128(d^2+d+1)}. \tag{3.4}$$

From this data it seems safe to guess that the main term of $\mu_{\mathcal{W}_d}(2k)$ is

$$c_{2k} \;:=\; \frac{1}{4^k(k+1)}\binom{2k}{k} \tag{3.5}$$

(the $2k^{\text{th}}$ moment of the normalized semi-circular distribution), which we now prove. We first show

$$\mu_{\mathcal{W}_d}(k) \;=\; c_k + O(1/d), \tag{3.6}$$

and then with a bit more work improve the error to $O(1/d^2)$.

For odd k there is nothing to prove, since both $\mu_{\mathcal{W}_d}(k)$ and c_k vanish. We thus restrict our attention to even k, and proceed by induction; to emphasize that we have even subscripts we write $2k$ for k. For $2k \leq 8$, we have already verified the conjecture. The only role of the inductive hypothesis is to ensure that, when computing $\mu_{\mathcal{W}_d}(2k)$, we can treat all lower eigenmoments as $O(1)$. The recursion formula (3.3) gives

$$(d^k - d)\mu_{\mathcal{W}_d}(2k) \;=\; \sum_{\pi \in \mathcal{P}^{\circ}_{2k}} m_\pi(d)\mu_{\mathcal{W}_d}(\sigma(\pi)). \tag{3.7}$$

The total contribution from those π which involve fewer than k symbols is $O(d^{k-1})$. Thus, the main term must come from the patterns involving k edges, i.e., π whose signature $\sigma(\pi) = (2, 2, \dots, 2)$. Recall $\mathcal{P}_k^{(2)}$ is the set of all CAPPs of length k which possess a signature of this form. Note $\mu_{\mathcal{W}_d}(2) = 1/4$ and $\mu_{\mathcal{W}_d}(4) = 1/8$, so if $\pi \in \mathcal{P}_{2k}^{(2)}$ then $\mu_{\mathcal{W}_d}(\sigma(\pi)) = (1/4)^k$, while if $\pi \in \mathcal{P}_{2k}^{(4)}$ then $\mu_{\mathcal{W}_d}(\sigma(\pi)) = (1/8)(1/4)^{k-1}$. We have

$$\begin{aligned}
(d^k - d)\mu_{\mathcal{W}_d}(2k) \;&=\; \sum_{\pi \in \mathcal{P}_{2k}^{(2)}} m_\pi(d)\mu_{\mathcal{W}_d}(\sigma(\pi)) + O(d^{k-1}) \\
&=\; \left(\sum_{\pi \in \mathcal{P}_{2k}^{(2)}} \mu_{\mathcal{W}_d}(\sigma(\pi))\right) d^k + O(d^{k-1}) \;=\; |\mathcal{P}_{2k}^{(2)}| 2^{-2k} d^k + O(d^{k-1}),
\end{aligned} \tag{3.8}$$

where we used for $\pi \in \mathcal{P}_{2k}^{(2)}$

$$m_\pi(d) \;=\; \prod_{i=1}^{k}(d - \alpha_i) \;=\; d^k + O(d^{k-1}) \tag{3.9}$$

(where α_i is the number of edges prior to the i^{th} which are adjacent to the i^{th} as in (2.12)). Lemma 2.5 yields the desired conclusion.

The serendipitous correspondence from Lemma 2.6 allows us to improve the error to

$$\mu_{\mathcal{W}_d}(2k) \;=\; c_{2k} + O(1/d^2). \tag{3.10}$$

As above, we have already verified the theorem (with no error term) in the cases when k is odd or at most 8. Henceforth we assume that $2k > 8$. We analyze the contribution from patterns with at least $k - 1$ distinct symbols (in other words, we allow at most one repetition), and trivially bound the contribution from the remaining by $O(d^{k-2})$. We have

$$\begin{aligned}
(d^k-d)\mu_{\mathcal{W}_d}(2k) &= \sum_{\pi\in\mathcal{P}_{2k}^\circ} m_\pi(d)\mu_{\mathcal{W}_d}(\sigma(\pi)) \\
&= \left(\sum_{\pi\in(\mathcal{P}_{2k}^{(2)}\cup\mathcal{P}_{2k}^{(4)})} m_\pi(d)\mu_{\mathcal{W}_d}(\sigma(\pi))\right) + O(d^{k-2}) \\
&= (1/4)^k\left(\sum_{\pi\in\mathcal{P}_{2k}^{(2)}} m_\pi(d)\right) + (1/8)(1/4)^{k-1}\left(\sum_{\pi\in\mathcal{P}_{2k}^{(4)}} m_\pi(d)\right) + O(d^{k-2}) \\
&= (1/4)^k\left(\sum_{\pi\in\mathcal{P}_{2k}^{(2)}} m_\pi(d) + 2\sum_{\pi\in\mathcal{P}_{2k}^{(4)}} m_\pi(d)\right) + O(d^{k-2}). \qquad (3.11)
\end{aligned}$$

The strategy is to compute the secondary terms of $m_\pi(d)$, multiply them by the correct factor and then substitute back into (3.11). If $\pi\in\mathcal{P}_{2k}^{(2)}$, then

$$m_\pi(d) \;=\; \prod_{i=1}^k (d-\alpha_i) \;=\; d^k - \left(\sum_{i=1}^k \alpha_i\right) d^{k-1} + O(d^{k-2}), \qquad (3.12)$$

where α_i is the number of edges prior to the i^{th} which are adjacent to the i^{th} as in (2.12). Summing over i gives the number of pairs of adjacent edges in the diagram. Summing over $\mathcal{P}_{2k}^{(2)}$, we obtain

$$\sum_{\pi\in\mathcal{P}_{2k}^{(2)}} m_\pi(d) \;=\; |\mathcal{P}_{2k}^{(2)}|d^k - |\mathcal{T}_{2k}|d^{k-1} + O(d^{k-2}). \qquad (3.13)$$

All of these terms have the same value for $\mu_{\mathcal{W}_d}(\sigma(\pi))$, namely $(1/4)^k$.

For the other summation we need only the dominant term:

$$\sum_{\pi\in\mathcal{P}_{2k}^{(4)}} m_\pi(d) \;=\; |\mathcal{P}_{2k}^{(4)}|d^{k-1} + O(d^{k-2}). \qquad (3.14)$$

All of these terms have the same value for $\mu_{\mathcal{W}_d}(\sigma(\pi))$, namely $(1/8)(1/4)^{k-1} = (1/2)(1/4)^k$.

Using the above, we find that the contribution from $\pi\in\mathcal{P}_{2k}^{(2)}\cup\mathcal{P}_{2k}^{(4)}$ to (3.11) is

$$|\mathcal{P}_{2k}^{(2)}|\frac{d^k}{4^k} - |\mathcal{T}_{2k}|\frac{d^{k-1}}{4^k} + |\mathcal{P}_{2k}^{(4)}|\frac{d^{k-1}}{2\cdot 4^k} + O\left((d/4)^{k-2}\right); \qquad (3.15)$$

however, by Lemma 2.6 $|\mathcal{P}_{2k}^{(4)}| = 2|\mathcal{T}_{2k}|$, and thus the order d^{k-1} terms above cancel, yielding

$$\mu_{\mathcal{W}_d}(2k) \;=\; |\mathcal{P}_{2k}^{(2)}|\frac{d^k}{4^k(d^k-d)} + O\left(\frac{d^{k-2}}{4^k(d^k-d)}\right). \qquad (3.16)$$

From Lemma 2.5 we have $|\mathcal{P}_{2k}^{(2)}| = \frac{1}{k+1}\binom{2k}{k} = 4^k c_{2k}$ (where c_{2k} is the $2k^{\text{th}}$ moment of the semi-circle distribution normalized to have variance 1/4), and thus

$$\mu_{\mathcal{W}_d}(2k) \;=\; c_{2k} + O\left(\frac{c_{2k}}{d^{k-1}}\right) + O\left(\frac{1}{4^k d^2}\right); \qquad (3.17)$$

as $k > 3$ the second error term dominates. We conclude that the error is $O(1/d^2)$, as claimed. $\square$

REFERENCES

[ALHM] D. Allen, T.-C. Lu, D. Huber and H. Moon, *Hierarchical Random Graphs for Networks with Weighted Edges and Multiple Edge Attributes*, Universidad Nacional de General Sarmiento, May 2011.

[Al] N. Alon, *Eigenvalues and expanders*, Combinatorica **6** (1986), no. 2, 83–96.

[AL] H. Amini and M. Lelargey, *The Diameter Of Weighted Random Graphs*, to appear in the Annals of Applied Probability. http://arxiv.org/abs/1112.6330.

[AGZ] G. W. Anderson, A. Guionnet and O. Zeitouni, *An Introduction to Random Matrices*, Cambridge Studies in Advanced Mathematics **118**, Cambridge University Press, 2010.

[Bai] Z. Bai, *Methodologies in spectral analysis of large-dimensional random matrices, a review*, Statist. Sinica **9** (1999), no. 3, 611–677.

[BLMST] O. Beckwith, V. Luo, S. J. Miller, K. Shen and N. Triantafillou, *Distribution of Eigenvalues of Weighted, Structured Matrix Ensembles*, preprint, 2014. http://arxiv.org/abs/1112.3719

[Bien] F. Bien, *Constructions of telephone networks by group representations*, Notices of the AMS **36** (1989), no. 1, 5–22.

[Bi] P. Billingsley, *Probability and Measure* (third edition), Wiley, New York, 1995.

[Bo1] M. Bolla, *Distribution of the eigenvalues of random block-matrices*, Lin. Alg. Appl. **377** (2004), 219–240.

[Bo2] M. Bolla, *Noisy random graphs and their Laplacians*, Discrete Mathematics **308** (2008) 4221–4230.

[BCG] A. Bose, S. Chatterjee, and S. Gangopadhyay, *Limiting spectral distributions of large dimensional random matrices*, J. Indian Statist. Assoc. **41** (2003), 221–259.

[BHS1] A. Bose, R. S. Hazra, and K. Saha, *Patterned random matrices and notions of independence*, Technical report R3/2010 (2010), Stat-Math Unit, Kolkata. Available online at http://www.isical.ac.in/∼statmath.

[BHS2] A. Bose, R. S. Hazra, and K. Saha, *Patterned random matrices and method of moments*, in Proceedings of the International Congress of Mathematicians Hyderabad, India, 2010, 2203–2230. (Invited article). World Scientific, Singapore and Imperial College Press, UK.

[BDJ] W. Bryc, A. Dembo, T. Jiang, *Spectral measure of large random Hankel, Markov and Toeplitz matrices*, Ann. Probab. **34** (2006), no. 1, 1–38.

[BM] A. Bose and S. S. Mukherjee, *Bulk behaviour of Schur-Hadamard products of symmetric random matrices*, to appear in Random Matrices: Theory and Applications, 2014.

[DSV] G. Davidoff, P. Sarnak, and A. Valette, *Elementary Number Theory, Group Theory, and Ramanujan Graphs*, London Mathematical Society, Student Texts, Vol. 55, Cambridge University Press, Cambridge 2003.

[DP] I. Dumitriu and S. Pal, *Sparse regular random graphs: spectral density and eigenvectors*, Ann. Probab. **40** (2012), no. 5, 2197–2235.

[ES] Y. Emek and Y. Shavitt, *Approximating the Statistics of various Properties in Randomly Weighted Graphs*, Proceeding SODA '11 Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, 2011, Pages 1455–1467.

[ERSY] L. Erdős, J. A. Ramirez, B. Schlein, and H.-T. Yau, *Bulk Universality for Wigner Matrices*, Comm. Pure Appl. Math. **63** (2010), no. 7, 895–925.

[ESY] L. Erdős, B. Schlein, and H.-T. Yau, *Wegner estimate and level repulsion for Wigner random matrices*, Int. Math. Res. Not. IMRN 2010, no. 3, 436–479.

[FM] F. W. K. Firk and S. J. Miller, *Nuclei, Primes and the Random Matrix Connection*, Symmetry **1** (2009), 64–105; doi:10.3390/sym1010064.

[Fo] P. J. Forrester, *Log-Gases and Random Matrices*, London Mathematical Society Monographs **34**, Princeton University Press, Princeton, NJ 2010.

[Fr1] J. Friedman, *Some geometric aspects of graphs and their eigenfunctions*, Duke Math. J. **69** (1993), no. 3, 487–525.

[Fr2] J. Friedman, *A proof of Alon's second eigenvalue conjecture*, Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing, 720–724 (electronic), ACM, New York, 2003.

[Fr3] J. Friedman, *A proof of Alon's second eigenvalue conjecture and related problems*, Mem. Amer. Math. Soc. **195** (2008), no. 910, viii+100 pp.

[Ga] D. Garlaschelli, *The weighted random graph model*, New J. Phys. **11** (2009) 073005, doi:10.1088/1367-2630/11/7/073005.

[GILVZ] O. Goldreich, R. Impagliazzo, L. Levin, R. Venkatesan, and D. Zuckerman, *Security preserving amplification of hardness*. In 31st Annual Symposium on Foundations of Computer Science, Vol. I, II (St. Louis, MO, 1990), 318–326, IEEE Comput. Soc. Press, Los Alamitos, CA, 1990.

[HM] C. Hammond and S. J. Miller, *Eigenvalue spacing distribution for the ensemble of real symmetric Toeplitz matrices*, Journal of Theoretical Probability **18** (2005), no. 3, 537–566.

[JMRR] D. Jakobson, S. D. Miller, I. Rivin, and Z. Rudnick, *Eigenvalue spacings for regular graphs*. Pages 317–327 in *Emerging Applications of Number Theory (Minneapolis, 1996)*, The IMA Volumes in Mathematics and its Applications, Vol. 109, Springer, New York, 1999.

[KSV] O. Khorunzhy, M. Shcherbina1 and V. Vengerovsky, *Eigenvalue distribution of large weighted random graphs*, Journal of Mathematical Physics **45** (2004), no. 4, 1968–1672.

[KKMSX] G. S. Kopp, M. Koloğlu, S. J. Miller, F. Strauch and W. Xiong, *The Limiting Spectral Measure for Ensembles of Symmetric Block Circulant Matrices*, Journal of Theoretical Probability **26** (2013), no. 4, 1020–1060.

[LPS] A. Lubotzky, R. Phillips, and P. Sarnak, *Ramanujan graphs*, Combinatorica **8** (1988), no. 3, 261–277.

[Mar] G. A. Margulis, *Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators (Russian)*, Problemy Peredachi Informatsii **24** (1988), no. 1, 51–60; translation in Problems Inform. Transmission **24** (1988), no. 1, 39–46.

[McD1] C. McDiarmid, *Random graphs from a weighted minor-closed class*, The Electronic Journal of Combinatorics **20** (2013), no. 2, #P52

[McD2] C. McDiarmid, *Connectivity for random graphs from a weighted bridge-addable class*, The Electronic Journal of Combinatorics **19** (2012), no. 4, #P53.

[McK] B. McKay, *The expected eigenvalue distribution of a large regular graph*, Linear Algebra Appl. **40** (1981), 203–216.

[McW] B. McKay and N. Wormald, *The degree sequence of a random graph. I. The models*, Random Structures & Algorithms **11** (1997), no. 2, 97–117.

[Meh] M. Mehta, *Random Matrices*, 2nd edition, Academic Press, Boston, 1991.

[MNS] S. J. Miller, T. Novikoff and A. Sabelli, *The distribution of the second largest eigenvalue in families of random regular graphs*, Experimental Mathematics **17** (2008), no. 2, 231–244.

[Pi] Pippenger, *Super concentrators*, SIAM Journal Comp. **6** (1977), 298–304.

[Po] M. Porfiri, *Synchronization in Random Weighted Directed Networks*, Circuits and Systems I: Regular Papers, IEEE Transactions on (Volume:55 , Issue: 10), 2008, 3170–3177.

[Sar1] P. Sarnak, *Some applications of modular forms*, Cambridge Trusts in Mathemetics, Vol. 99, Cambridge University Press, Cambridge, 1990.

[Sar2] P. Sarnak, *What is an Expander?*, Notices of the AMS **51** (2004), no. 7, 762–763.

[So] S. Sodin, *An estimate for the average spectral measure of random band matrices*, J. Stat. Phys. **144** (2011), no. 1, 46–59.

[SS] M. Sipser and D. A. Spielman, *Expander codes*, IEEE Trans. Inform. Theory **42** (1996), no. 6, part 1, 1710–1722.

[Ta] L. Takacs, *A Moment Convergence Theorem*, The American Mathematical Monthly **98** (Oct., 1991), no. 8, 742–746.

[TV1] T. Tao and V. H. Vu, *From the Littlewood-Offord problem to the Circular Law: universality of the spectral distribution of random matrices*, Bull. Amer. Math. Soc. **46** (2009), 377–396.

[TV2] T. Tao and V. H. Vu, *Random matrices: universality of local eigenvalue statistics up to the edge*, Comm. Math. Phys. **298** (2010), no. 2, 549–572.

[TVW] L. V. Tran, V. H. Vu and K. Wang, *Sparse random graphs: eigenvalues and eigenvectors*, Random Structures Algorithms **42** (2013), no. 1, 110–134.

[Ven] V. Vengerovsky, *Eigenvalue distribution of large weighted bipartite random graphs*, Zh. Mat. Fiz. Anal. Geom. **10** (2014), no. 2, 240–255.

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA, USA
E-mail address: `Leo.Goldmakher@williams.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MI, USA
E-mail address: `prof.cap.khoury@gmail.com`

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA, USA
E-mail address: `sjm1@williams.edu, Steven.Miller.MC.96@aya.yale.edu`

INSTITUTE FOR COMPUTATIONAL AND MATHEMATICAL ENGINEERING, STANFORD UNIVERSITY, STANFORD, CA, USA
E-mail address: `eveve@stanford.edu`