Spectral partitioning in random regular blockmodels

Paolo Barucca
University of Zurich[*]
(Dated: October 24, 2016)

Graph partitioning problems emerge in a wide variety of complex systems, ranging from biology to finance, but can be rigorously analyzed and solved only for a few graph ensembles. Here, an ensemble of random graphs with regular block structure is defined, for which analytical results can be obtained. In particular, the spectral density of such random regular blockmodel is computed exactly for a modular, bipartite and core-periphery structure. McKay's law for random regular graphs is found analytically to apply also for regular modular and bipartite structures when blocks are homogeneous. In core-periphery structures, where blocks are intrinsically heterogeneous, a new law is found to apply for the spectral density. Exact solution to the inference problem is provided for the models discussed. All analytical results show perfect agreement with numerical experiments. Final discussion summarizes results and outlines the relevance of the results for the solution of graph partitioning problems in other graph ensembles, in particular for the study of detectability thresholds and resolution limits.

I. INTRODUCTION

The recent developments of network theory driven by the increasing number of applications in biology, ecology, social systems, economics and finance [1, 2], have pushed forward theoretical research in graph theory. In particular the need to establish the statistical significance of various network metrics in real systems has ignited new results in random graph theory: on statistical inference [3], spectral theory of random graphs [4–8], ensembles of exponential random graphs [9, 10].
In this paper the cavity method, a standard tool in statistical physics [11], is applied to characterize the spectral properties of random graphs with regular block structure. The graph ensembles defined in this work represent the regular, i.e. where intra- and inter- blocks degrees are fixed, counterpart of the long-studied stochastic block models (SBM) [18, 39], for such reason I will refer to these models as random regular blockmodels (rrBM). This class of random graph models has first been analyzed in [17] under the name of equitable random graphs. From community detection to noise reduction in cross-correlation matrices, spectral theory of random matrices and random graphs is crucial to disentangle noise and signal in systems with large number of components. Following the derivation in [4] a finite set of non-linear equations for the spectral density is found and the solutions are provided for all possible two-blocks undirected structures: modular, bipartite and core-periphery.
The paper is organized as follows: in Section II the random regular blockmodel is defined and its isolated eigenvalues, including the highest eigenvalue, are explicitly computed.
In Section III a brief introduction of the cavity approach to the spectral density is provided and the general expression for the cavity variances of random regular blockmodel graphs is derived. In Section IV the expression is solved for all symmetric two block structures and its consequences on spectral clustering are outlined, and a general methodology for the block structure inference in rrBM graphs based on eigenvectors's localization is introduced, to overcome the limitation of naive spectral clustering. Numerical evidence of all results is presented. Finally in Section V the relevance of the results with respect to open questions in theory of random graphs is discussed and possible directions of research both for analytical results in spectral theory and for spectral methodologies for graph partitioning are outlined.

II. RANDOM REGULAR BLOCKMODELS

A random regular blockmodel is a graph ensemble defined by a set of vertices V, a partition $B = \{B_a\}_{a=1}^{m}$ dividing V in m non-overlapping sets of vertices, also called blocks, and a connectivity matrix $\mathbf{c}$, a $m \times m$ matrix of non-negative integer numbers [52]. For the sake of simplicity in the following we will refer to block B_a with its corresponding integer index a. For later use, we also introduce for all nodes the assignments g_i, such that for each node i holds $i \in B_{g_i}$.
Each graph $G = (V, E)$ of a random regular blockmodel must satisfy the constraints:

$$\forall a, b \in B \, \forall i \in a \;\; |\{(i,j) \in E \,|\, j \in b\}| = c_{ab}, \qquad (1)$$

i.e. the total number of edges of node i in block a with a vertex in b equals c_{ab}, for every vertex i and every pair of blocks a and b. In the case of blocks of different sizes, $|B_a| = N_a$ such that $\sum_{a=1}^{m} N_a = |V|$, then, for the system to have solution the connectivity matrix $\mathbf{c}$ and block sizes must obey the relations

$$\forall a, b \in B \; N_a c_{ab} = N_b c_{ba}, \qquad (2)$$

i.e. the total number of edges between blocks a and b must be uniquely defined.

[*] Also at London Institute for Mathematical Sciences and INET Oxford

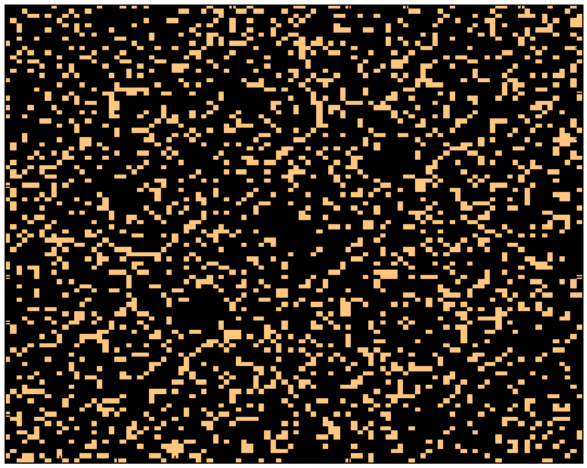
All graphs satisfying (1) have equal probability in the ensemble.
If we introduce the block degrees $k_{i\to a} = |\{(i,j) \in E | j \in a\}|$, (1) can be reformulated as follows: the vector of block degrees of each node in a given block equals the row of the connectivity matrix corresponding to the block index, i.e. $\forall i \in a\ k_{i\to b} = c_{ab}$.
Both stochastic blockmodels and random regular blockmodels are based on an analogous set of parameters, i.e. block assignments and connectivity matrix, nevertheless there is no trivial mapping between stochastic blockmodels, which are defined via link probabilities and their regular counterpart, which are defined via (2). A useful analogy to make sense of their relationship is the following: stochastic blockmodels [18] correspond to random regular blockmodels as the Erdos-Renyi random graphs correspond to the k-regular random graphs, in the sense that in both cases the randomness which is eliminated from the ensemble is the one given by the (block) degree distribution.
Random regular blockmodels are a block structured generalization of k-regular random graphs in the sense that when $c_{ab} = c$ for all pairs (a, b) the corresponding blockmodel ensemble is the set of k-regular random graphs with $k = cm$.
Moreover, the form of the constraints in (2) allow edges to be drawn independently for each pair of blocks, and in particular, for the case of blocks of the same-size, it is possible to sample rrBM graphs simply by assembling regular graphs: between each pair of blocks the edges are drawn according to a k-regular graph, where the value of k equals the corresponding element of the connectivity matrix, then the total set of edges is given by the union of the sets for each of the m^2 regular graphs.

The inference problem

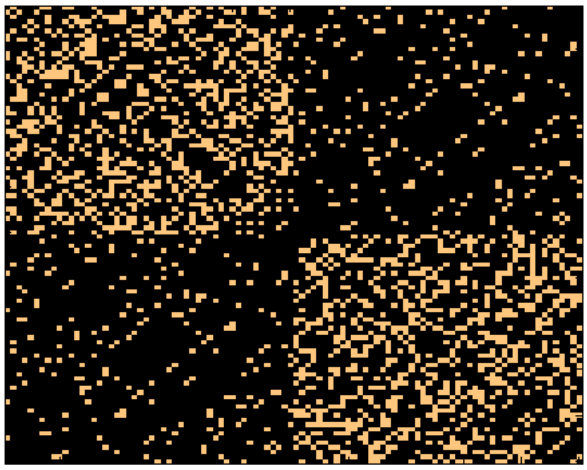
Given a random regular blockmodel graph G, the inference problem consists in reconstructing the parameters, i.e. the partition B and the connectivity matrix $\mathbf{c}$, that generated the graph.
In the latter the focus will be entirely on the representation of G in terms of its adjacency matrix $A = (a_{ij})_{i,j=1}^{N}$ where,

$$a_{ij} = \begin{cases} 1, & \text{if } (i,j) \in E \\ 0, & \text{otherwise} \end{cases}$$

which allows to compute graph properties in algebraic form and can also be used to visualize the inference problem associated with the graph (Figs.1a-1b): when parameters are unknown, there is no a-priori criterion to sort indices and the non-zero elements of the adjacency matrix do not display any specific block structure; once the parameters are known, rows and columns can be sorted according to the block indices, and the structure arises in a clear manner.



(a)



(b)

FIG. 1: (a) Adjacency matrix of a random regular blockmodel graph with a community structure where no block structure seems to be present even though the connectivity matrix reads, $\mathbf{c} = [16,\ 4;\ 4,\ 16]$.
(b) Same graph but rows and columns of the adjacency matrix are sorted according to the block structure, which becomes evident.

III. SPECTRAL THEORY

In this section, the spectral properties of random regular blockmodel graphs with blocks of same size are in-

vestigated. Analytical results are presented, both for the discrete and the continuous part of the spectrum.

A. Discrete part: the signal

Here is shown how isolated eigenvectors of the adjacency matrix entail exact informations on the block structure of rrBMs graphs. Starting from the secular equation

$$\sum_{j=1}^{N} a_{ij}u_j = \lambda u_i, \tag{3}$$

an ansatz of *block-simmetry* can be made such that nodes in the same block share the same eigencomponent, i.e. for all i is hypothesized that $u_i = u_{g_i}$. Since the number of neighbors between different groups is fixed it follows that:

$$\sum_{b=1}^{m} c_{ab}u_b = \lambda u_a, \tag{4}$$

which yields the useful conclusion that each block-symmetric eigenvector of the adjacency matrix corresponds to an eigenvector of the connectivity matrix $\mathbf{c}$, and viceversa. These eigenvectors correspond to a finite set of non-densely distributed, at most degenerate eigenvalues. Generally they can be positioned everywhere in the spectrum but when the block structure is particularly evident they will lie far from the bulk of the spectrum and are therefore referred to as outliers.

B. Continuous part: the noise

In this paragraph statistical physics techniques are used to compute the bulk of the spectrum of the adjacency matrices in a rrBM. Given an ensemble of $N \times N$ symmetric matrices the set of eigenvalues of a given adjacency matrix A is denoted by $\{\lambda_i^A\}_{i=1}^N$. The corresponding empirical spectral density is defined as:

$$\rho(\lambda; A) = \frac{1}{N}\sum_{i=1}^{N} \delta(\lambda - \lambda_i^A), \tag{5}$$

which satisfies the identity[13]:

$$\rho(\lambda; A) = \frac{2}{\pi}\lim_{\epsilon\to 0^+}\frac{1}{N}\Im\left[\frac{\partial}{\partial z}\log \mathcal{Z}(z; A)\right]_{z=\lambda-i\epsilon} \tag{6}$$

where $\log \mathcal{Z}(z; A)$ is obtained via Gaussian integrals as in[13], i.e.:

$$\mathcal{Z}(z; A) = \int\left[\Pi_{i=1}^{N}\frac{dx_i}{\sqrt{(}2\pi)}e^{-H(x;z,A)}\right] \tag{7}$$

with $H(x; z, A) = \frac{z}{2}\sum_i^N x_i^2 - \frac{1}{2}\sum_{i,j}^N A_{ij}x_i x_j$. Such formulation yields an expression for the spectral density of any graph of the ensemble in terms of the variances of the Gaussian variables introduced in (7),

$$\rho(\lambda; A) = \frac{1}{\pi}\lim_{\epsilon\to 0^+}\frac{1}{N}\Im\left[\sum_i^N \langle x_i^2\rangle_z\right]_{z=\lambda-i\epsilon}. \tag{8}$$

In principle, computing variances in (8) is not easier than diagonalizing the adjacency matrix but for sparse graphs an approximation method has been proposed that holds exactly in the large N limit, the cavity method [11].
In the cavity method, conditional probability distributions are introduced for each node and are parametrized by specific variables, i.e. the cavity variances $\Delta_i^{(j)}$, each representing the variance of x_i if its neighbor j is not taken into account. With such approximation the following set of self-consistent equations can be derived [4]:

$$\Delta_i^{(j)}(z) = \frac{1}{z - \sum\limits_{l\in\partial i\setminus j}^{N} A_{il}^2\Delta_l^{(i)}(z)}, \tag{9}$$

where ∂i is the set of neighbor of node i, i.e. $\partial i = \{e \in E | i \in e\}$. From cavity variances it is possible to compute node variances via the equations

$$\Delta_i(z) = \frac{1}{z - \sum\limits_{l\in\partial i}^{N} A_{il}^2\Delta_l^{(i)}(z)}, \tag{10}$$

which lead to compute the spectral density $\rho(\lambda; A)$.
In the case of rrBMs the ansatz of *block-simmetry* can be made for the cavity variances:

$$\Delta_i^{(j)}(z) = \Delta_{g_i}^{(g_j)}(z). \tag{11}$$

This ansatz reduces the set of equations for cavity variances from a size of order N (in the sparse case) to the following set of m^2 equations:

$$\Delta_a^{(b)}(z) = \frac{1}{z - \sum\limits_{c}^{m}(c_{ac} - \delta_{bc})^+\Delta_c^{(a)}(z)}, \tag{12}$$

where $(x)^+ = \max(x, 0)$. Block variances can then be computed,

$$\Delta_a(z) = \frac{1}{z - \sum\limits_{c}^{m} c_{ac}\Delta_c^{(a)}(z)}, \tag{13}$$

and finally the spectral density,

$$\rho(\lambda) = \frac{1}{\pi m}\sum_{a=1}^{m}\Im[\Delta_a(z)]_{z=\lambda-i\epsilon}. \tag{14}$$

IV. RESULTS

In this section spectral theory is applied to characterize the three main block structures: modular, bipartite and

core-periphery. Such categorization is fundamental because any structured pair of blocks, perhaps embedded in more complicated block structures, must fit it. For all these categories both the bulk and the outliers of the spectrum are computed analytically and numerically and the inference problem is solved.
In particular the orderings relations between the highest and smallest eigenvalue of the bulk and the outliers are considered. A new spectral density function is found in the core-periphery case.

A. Modular structures

Graph partitioning, and in particular spectral bisection, is a long-standing problem in graph theory[14, 27, 28, 33]. Here for such modular case, two homogenous blocks are analyzed: the blocks share the same size and the connectivity matrix reads:

$$\mathbf{c} = \begin{pmatrix} c_{in} & c_{out} \\ c_{out} & c_{in} \end{pmatrix} \tag{15}$$

where c_{in} and c_{out} are non-negative integers such that $c_{in} > c_{out}$. For later use I also define $c = c_{out}$ and $r = c_{in}/c_{out}$, $\epsilon = 1/r$ is a quantity that has been used to characterize the strength of the assortative structure [40].
Assortative structures have been widely investigated with various approaches: multilevel schemes [26], spectral methods [21, 45], modularity maximization[35], belief-propagation [40], Markov-Chain Montecarlo methods [48], and other heuristic algorithms [12]. Stochastic blockmodels have been shown to display a detectability transition [19, 40].
In this homogenous case cavity equations can be further simplified, cavity variances associated to the two blocks can be hypothesize to be equal, $\Delta_1^{(b)} = \Delta_2^{(b)}$ for $b = 1, 2$, and given the form of the equations, by inspection it is also possible to look for fully-symmetric solutions such that $\Delta_a^{(b)} = \Delta^{(cav)}$ for all (a, b). This ansatz yields:

$$\Delta^{(cav)}(z) = \frac{1}{z - (c + cr - 1)\Delta^{(cav)}}, \tag{16}$$

Now the equation is identical to the one derived in [4], and analogously carrying out the algebra the spectral density found is the McKay's law [46]:

$$\rho(\lambda) = \frac{k\sqrt{4(k-1) - \lambda^2}}{2\pi(k^2 - \lambda)} \tag{17}$$

where $k = c(1 + r)$ (Fig.3).
(17) yields the maximal eigenvalue in the bulk, $\lambda_{bulk}^{+} = 2\sqrt{c_{in} + c_{out} - 1}$. Outlier eigenvalues can be easily computed via the characteristic polynomial:

$$(rc - \lambda)^2 - c^2 = 0. \tag{18}$$

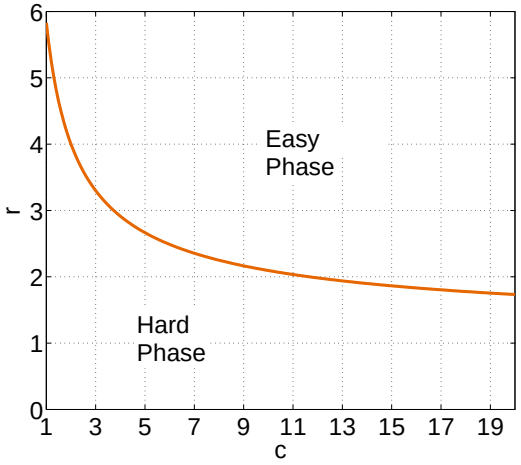


FIG. 2: Critical line in the plane c-r. Above the line, in the easy phase standard spectral bisection works, while below the line, in the hard phase more informations need to be used, e.g. eigenvectors's statistics.

The first outlier λ_{max} equals the total connectivity k and its corresponding eigenvector is constant and uninformative. On the other hand, the second outlier, which we refer to as λ_o, is informative and its relationship with λ_b^+ is crucial for the inference problem: when $\lambda_o > \lambda_b^+$ it is simply the second largest eigenvalue and its corresponding eigenvector can be easily and fast computed, but when $\lambda_o < \lambda_b^+$ then it is no longer the second eigenvalue and its ranking becomes unknown.
Such transition occurs when $\lambda_o = \lambda_b^+$, which corresponds to the critical line in the plane $c - r$ (Fig.2):

$$r_c = \frac{2 + 2\sqrt{2c} + c}{c} \tag{19}$$

Below the critical line, the outlier eigenvector gets lost in the bulk and a criterion is needed to identify the right eigenvector.
The solution to this detectability problem in random regular blockmodels can be found by exploiting the information about the eigenvectors. In fact the distribution of the eigencomponents of the outlier eigenvector and of bulk's eigenvectors turn out to be significantly different: from the *block-symmetry* ansatz, the eigenvector corresponding to λ_o is more extended than the typical eigenvector of the bulk.
By looking at a measure of extendedness, such as the inverse participation ratio (IPR) [36, 49], it is possible to recognize the informative outlier eigenvector associated to λ_o:

$$IPR = \frac{N}{2}\sum_{a}^{2}(u_a)^4 = \frac{1}{N}, \tag{20}$$

independently from r. The inverse participation ratio of the outlier eigenvector is then $1/N$ while the random

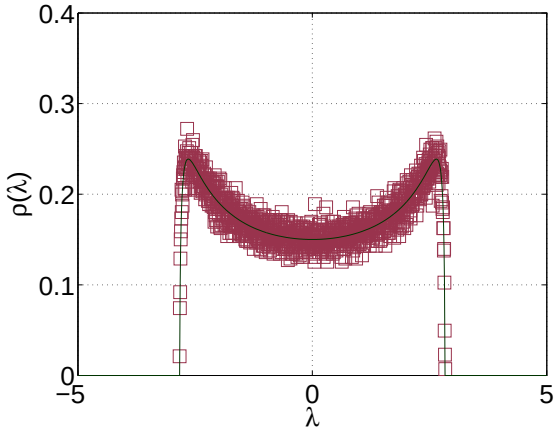


FIG. 3: (Modular case). Spectral density for $c = 1$ and $r = 2$, it corresponds to McKay's law for a k-regular graph with $k = 3$. Squares come from numerical diagonalization of a sample of 100 rrBM graphs of size $N = 1000$.

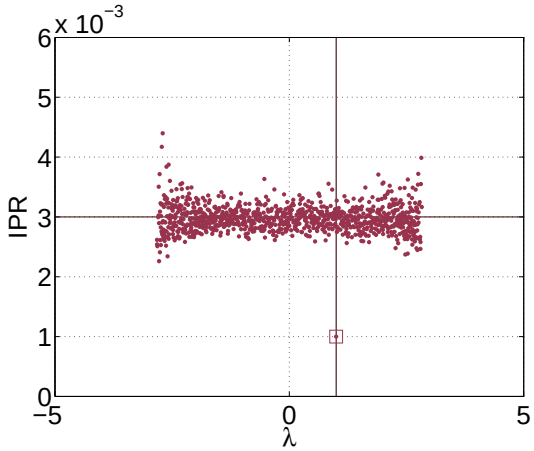


FIG. 4: (Modular case). Inverse participation ratio for each eigenvector in the plane λ-IPR. The eigenvectors of the bulk all share an IPR fluctuating around $3/N$ while the outlier eigenvector (square) has an IPR equal to $1/N$, which allows to solve the inference problem also in the hard phase. Parameters are $c = 1$ and $r = 2$.

eigenvectors in the bulk have an expected IPR of $3/N$ [47] and a standard deviation of order $N^{-3/2}$, so that the signal-to-noise ratio grows with $N^{1/2}$.
Therefore, for large graphs, the inference problem can be solved for all values of $r > 1$, as long as the signal is actually present, also below the critical line by searching for the most extended eigenvector of the adjacency matrix A (Fig. 4), excluding the one associated with the uninformative maximum eigenvalue.

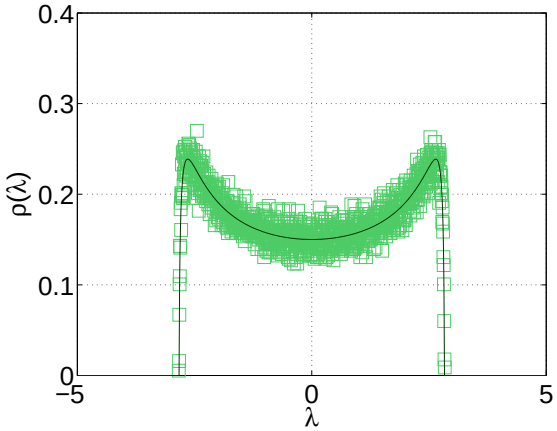


FIG. 5: (Bipartite case). Spectral density for $c_{in} = 1$ and $c_{out} = 2$, it corresponds to McKay's law for a k-regular graph with $k = 3$. Squares come from numerical diagonalization of a sample of 100 rrBM graphs of size $N = 1000$.

B. Bipartite structures

Bipartite structures have been investigated in various contexts [20, 25] for their many applications, from social network analysis to economics [16, 22–24].
In this paragraph I focus on a random bipartition, where edges within a block can be present but are always less than edges towards the other block.
The connectivity matrix c_{ab} considered is the following:

$$\mathbf{c} = \begin{pmatrix} c_{in} & c_{out} \\ c_{out} & c_{in} \end{pmatrix} \tag{21}$$

where c_{in} and c_{out} are non-negative integers such that $c_{out} > c_{in}$. The analysis is entirely analogous to the one put forward for the assortative case, once the appropriate parallels are drawn: for large c_{out}/c_{in} the informative outlier eigenvector corresponds to the lowest eigenvalue, $\lambda_o = c_{in} - c_{out}$, and the critical line is defined by the condition $\lambda_o = \lambda_b^-$. Also in this case the cavity equations admit a fully symmetric solution for the variances that leads to McKay's law for the spectral density (Fig. 5).

C. Core-periphery structures

Core-periphery structure, in which an high-density core is less densely connected to a low-density periphery, has been documented in a large variety of complex systems, such in social networks [37, 43, 44], world trade networks [31], and financial networks[30].
It has been investigated with various metrics, as for instance MINRES [37], non-backtracking centrality [36], probability marginals [42] and simply, but efficiently, measuring degree centrality [15, 50]; recently PageRank

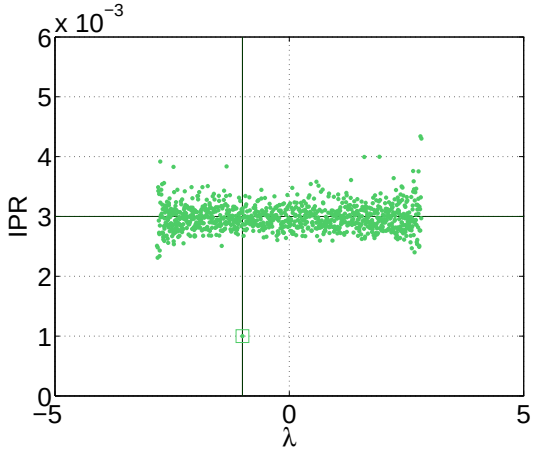


FIG. 6: (Bipartite case) Inverse participation ratio for each eigenvector in the plane λ-IPR. The eigenvectors of the bulk all share an IPR fluctuating around $3/N$ while the outlier eigenvector (square) has an IPR equal to $1/N$, which allows to solve the inference problem also in the hard phase. Parameters are $c = 1$ and $r = 2$.

has been shown to be the most robust measure of coreness in heterogenous degree-corrected SBM graphs [15]. The connectivity matrix $\mathbf{c}$ considered is:

$$c = \begin{pmatrix} c_{cc} & c_{cp} \\ c_{cp} & c_{pp} \end{pmatrix} \tag{22}$$

where $c_{cc} \geq c_{cp} \geq c_{pp}$.The degree regularity in rrBM trivializes the inference problem, since degree alone is sufficient to establish the assignment of a node.
Besides, the intrinsic block heterogeneity of core-periphery structure makes it incompatible with the ansatz of full symmetry that has been used for modular and bipartite structure.
Therefore it is necessary to solve the full system of four equations:

$$\Delta_c^{(c)}(z) = \frac{1}{z - (c_{cc}-1)^+\Delta_c^{(c)}(z) - c_{cp}\Delta_p^{(c)}(z)}$$
$$\Delta_c^{(p)}(z) = \frac{1}{z - c_{cc}\Delta_c^{(c)}(z) - (c_{cp}-1)^+\Delta_p^{(c)}(z)}$$
$$\Delta_p^{(c)}(z) = \frac{1}{z - (c_{cp}-1)^+\Delta_c^{(p)}(z) - c_{pp}\Delta_p^{(p)}(z)}$$
$$\Delta_p^{(p)}(z) = \frac{1}{z - c_{cp}\Delta_c^{(p)}(z) - (c_{pp}-1)^+\Delta_p^{(p)}(z)}.$$

Let us now consider a specific choice of the parameters, i.e. $c_{cc} = k > 1$, $c_{cp} = 1$, $c_{pp} = 0$ for which (23) read:

$$\Delta_c^{(c)}(z) = \frac{1}{z - (k-1)\Delta_c^{(c)}(z) - \Delta_p^{(c)}(z)} \tag{23}$$
$$\Delta_c^{(p)}(z) = \frac{1}{z - k\Delta_c^{(c)}(z)} \tag{24}$$
$$\Delta_p^{(c)}(z) = \frac{1}{z} \tag{25}$$
$$\Delta_p^{(p)}(z) = \frac{1}{z - \Delta_c^{(p)}(z)}. \tag{26}$$

The only equation to be solved in such system turns out to be:

$$\Delta_c^{(c)}(z) = \frac{1}{z - \frac{1}{z} - (k-1)\Delta_c^{(c)}(z)},$$

that yields the solutions:

$$\alpha \equiv \Re[\Delta_c^{(c)}(z)]_{z=\lambda-i0^+} = \begin{cases} \frac{\mu \pm \sqrt{\mu^2 - 4(k-1)}}{2(k-1)}, & \text{if } |\mu| \geq 2\sqrt{k-1} \\ \frac{\mu}{2(k-1)}, & \text{otherwise} \end{cases} \tag{27}$$

$$\beta \equiv \Im[\Delta_c^{(c)}(z)]_{z=\lambda-i0^+} = \begin{cases} 0, & \text{if } |\mu| \geq 2\sqrt{k-1} \\ \frac{\sqrt{4(k-1)-\mu^2}}{2(k-1)}, & \text{otherwise,} \end{cases} \tag{28}$$

where $\mu = \lambda - \frac{1}{\lambda}$. Since a non-zero imaginary part is a necessary condition for a non-zero support of the spectral density, we can immediately find the boundaries of its support from the condition $|\lambda - \frac{1}{\lambda}| = 2\sqrt{k-1}$.
From $\Delta_c^{(c)}(z)$, can then be derived $\Delta_c^{(p)}(z)$, $\Delta_p^{(p)}(z)$, and finally the block variances $\Delta_{c/p}$:

$$[\Delta_c(z)]_{z=\lambda-i0^+} = \frac{1}{\mu - k\alpha - ki\beta} \tag{29}$$
$$[\Delta_p(z)]_{z=\lambda-i0^+} = \frac{1}{\lambda - \frac{1}{\lambda - k\alpha - ki\beta}}, \tag{30}$$

which yield the explicit form of the spectral density:

$$\rho(\lambda) = \frac{1}{2\pi}\left(\frac{k\beta}{\delta} + \frac{k\beta}{\gamma}\frac{1}{\left(\lambda - \frac{\lambda - k\alpha}{\gamma}\right)^2 + \left(\frac{k\beta}{\gamma}\right)^2}\right), \quad \text{with } |\lambda - \frac{1}{\lambda}| < 2\sqrt{k-1}, \tag{31}$$

where $\gamma = (\lambda - k\alpha)^2 + (k\beta)^2$ and $\delta = (\mu - k\alpha)^2 + (k\beta)^2$. As an example, for $k = 2$, the case depicted in Fig.7, the equation (31) reads:

$$\rho(\lambda) = \frac{1}{2\pi}\left(\frac{1}{\sqrt{4-\mu^2}} + \frac{\sqrt{4-\mu^2}}{6-\lambda^2}\frac{1}{\left(\lambda - \frac{1/\lambda}{6-\lambda^2}\right)^2 + \left(\frac{\sqrt{4-\mu^2}}{6-\lambda^2}\right)^2}\right),$$

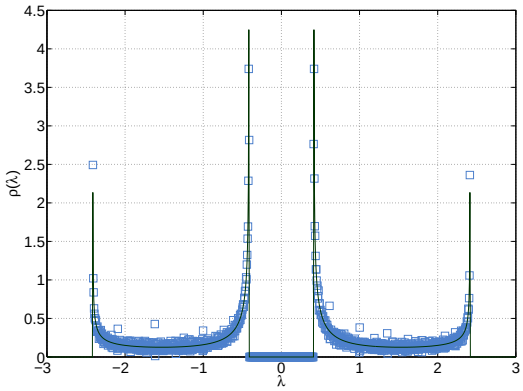


FIG. 7: (Core-periphery case). Spectral density for $c_{cc} = 2$, $c_{cp} = 1$ and $c_{pp} = 0$, corresponding to ((IV C)). Squares come from numerical diagonalization of a sample of 100 rrBM graphs of size $N = 1000$.

with the support defined by the condition $|\lambda - \frac{1}{\lambda}| < 2$.

V. CONCLUSIONS

In this paper the random regular blockmodel, or equivalently equitable random graph ensemble [17], is defined and used for testing graph partitioning algorithms and analytical methods of spectral graph theory.
In particular in the framework of random regular blockmodels the picture of the detectability threshold for naive spectral clustering, i.e. only using the eigenvector associated to the second largest eigenvalue of the adjacency matrix, emerges distinctly, as well as the crucial role of eigenvectors's statistics. Insights from random regular blockmodels could be used to develop new spectral methods based on both eigenvalues and eigenvectors properties. For what concerns analytical methods, it has been shown how the cavity method for spectral theory introduced in [4], can be used for obtaining results on spectra beyond Erdos-Renyi and random regular graphs: McKay's law is found to hold also for homogenous modular and bipartite regular structures, and a new spectral law has been derived for core-periphery structures. Future work will be dedicated to the study of regular heterogenous and multi-modular structures, such as the regular counterpart of planted partition model [51], in relation to the problem of resolution limit [32, 38] in modularity maximization. further studies will deal with the assessment of higher order properties, such as triangles and other motifs, of the ensemble in both the undirected and the directed case.

ACKNOWLEDGEMENT

The author acknowledges support from: FET Project DOLFINS nr. 640772, and FET IP Project MULTIPLEX nr. 317532. The author would like to thank Ton Coolen, Fabrizio Lillo, Elisa Letizia, Piero Mazzarisi and Daniele Tantari for fruitful discussions.

[1] M. E. J. Newman, A. L. Barabasi, and D. J. Watts. *The structure and dynamics of networks.* Princeton University Press, 2006.
[2] Easley, David, and Jon Kleinberg. *Networks, crowds, and markets: Reasoning about a highly connected world.* Cambridge University Press, (2010).
[3] A. Decelle, F. Krzakala, C. Moore, L. Zdeborov *Inference and Phase Transitions in the Detection of Modules in Sparse Networks*, Phys. Rev. Lett. 107, 065701 (2011).
[4] Rogers, Tim, et al. *Cavity approach to the spectral density of sparse symmetric random matrices.* Physical Review E 78.3 (2008): 031116.
[5] Rogers, Tim, and Isaac Pérez Castillo. *Cavity approach to the spectral density of non-Hermitian sparse matrices.* Physical Review E 79.1 (2009): 012101.
[6] F. Chung, L. Lu, V. Vu, *Eigenvalues of random power law graphs* Annals of Combinatorics, (2003).
[7] L. Erdos, A. Knowles, H. Yau, and J. Yin, *Spectral statistics of Erdos-Renyi graphs I: Local semicircle law*, Ann. Probab. Volume 41, Number 3B (2013), 2279-2375.
[8] I. J. Farkas, I. Derenyi, A.L. Barabasi, T. Vicsek, *Spectra of real-world graphs: Beyond the semicircle law.* Phys. Rev. E 64, 026704 (2001).
[9] Robins, Garry, et al. *An introduction to exponential random graph (p*) models for social networks.* Social net-

works 29.2 (2007): 173-191.

[10] Garlaschelli, Diego, and Maria I. Loffredo. *Maximum likelihood: Extracting unbiased information from complex networks.* Physical Review E 78.1 (2008): 015101.

[11] Mézard, Marc, Giorgio Parisi, and Riccardo Zecchina. *Analytic and algorithmic solution of random satisfiability problems.* Science 297.5582 (2002): 812-815.

[12] Blondel, Vincent D., et al. *Fast unfolding of communities in large networks.* Journal of statistical mechanics: theory and experiment 2008.10 (2008): P10008.

[13] Edwards, S. F., and Raymund C. Jones. *The eigenvalue spectrum of a large symmetric random matrix.* Journal of Physics A: Mathematical and General 9.10 (1976): 1595.

[14] Bollobás, Béla, and Alex D. Scott. *Exact bounds for judicious partitions of graphs.* Combinatorica 19.4 (1999): 473-486.

[15] Barucca, Paolo, Daniele Tantari, and Fabrizio Lillo. *Centrality metrics and localization in core-periphery networks." Journal of Statistical Mechanics: Theory and Experiment* 2016.2 (2016): 023401.

[16] Barucca, Paolo, and Fabrizio Lillo. *Disentangling bipartite and core-periphery structure in financial networks.* Chaos, Solitons & Fractals 88 (2016): 244-253.

[17] Newman, M. E. J., and Travis Martin. *Equitable random graphs.* Physical Review E 90.5 (2014): 052824.

[18] Holland, Paul W., Kathryn Blackmond Laskey, and Samuel Leinhardt. *Stochastic blockmodels: First steps.* Social networks 5.2 (1983): 109-137.

[19] Mossel, Elchanan, Joe Neeman, and Allan Sly. *A proof of the block model threshold conjecture.* arXiv preprint arXiv:1311.4115 (2013).

[20] Larremore, Daniel B., Aaron Clauset, and Abigail Z. Jacobs. *Efficiently inferring community structure in bipartite networks.* Physical Review E 90.1 (2014): 012805.

[21] Hendrickson, Bruce, and Robert Leland. *An improved spectral graph partitioning algorithm for mapping parallel computations.* SIAM Journal on Scientific Computing 16.2 (1995): 452-469.

[22] Strogatz, Steven H. *Exploring complex networks.* Nature 410.6825 (2001): 268-276.

[23] Hidalgo, César A., and Ricardo Hausmann. *The building blocks of economic complexity.* proceedings of the national academy of sciences 106.26 (2009): 10570-10575.

[24] Newman, Mark EJ, Duncan J. Watts, and Steven H. Strogatz. *Random graph models of social networks.* Proceedings of the National Academy of Sciences 99.suppl 1 (2002): 2566-2572.

[25] Reed, Bruce, Kaleigh Smith, and Adrian Vetta. *Finding odd cycle transversals.* Operations Research Letters 32.4 (2004): 299-301.

[26] Karypis, George, and Vipin Kumar. *A fast and high quality multilevel scheme for partitioning irregular graphs.* SIAM Journal on scientific Computing 20.1 (1998): 359-392.

[27] Pothen, Alex, Horst D. Simon, and Kang-Pu Liou. *Partitioning sparse matrices with eigenvectors of graphs.* SIAM journal on matrix analysis and applications 11.3 (1990): 430-452.

[28] Boppana, Ravi B. *Eigenvalues and graph bisection: An average-case analysis.* Foundations of Computer Science, (1987), 28th Annual Symposium on. IEEE.

[29] Zdeborová, Lenka, and Stefan Boettcher. *A conjecture on the maximum cut and bisection width in random regular graphs.* Journal of Statistical Mechanics: Theory and Experiment 2010.02 (2010): P02020.

[30] Fricke, Daniel, and Thomas Lux. *Coreperiphery structure in the overnight money market: evidence from the e-mid trading platform.* Computational Economics 45.3 (2015): 359-395.

[31] Smith, David A., and Douglas R. White. *Structure and dynamics of the global economy: network analysis of international trade 19651980.* Social forces 70.4 (1992): 857-893.

[32] Fortunato, Santo, and Marc Barthelemy. *Resolution limit in community detection.* Proceedings of the National Academy of Sciences 104.1 (2007): 36-41.

[33] Fortunato, Santo. *Community detection in graphs.* Physics reports 486.3 (2010): 75-174.

[34] B. D. McKay. *The expected eigenvalue distribution of a large regular graph.* Linear Algebra Appl.,(1981)

[35] M. E. J. Newman *Community detection in networks: Modularity optimization and maximum likelihood are equivalent.* arXiv preprint arXiv:1606.02319 (2016).

[36] T. Martin, X. Zhang, and M. E. J. Newman, *Localization and centrality in networks*, Phys. Rev. E 90, 052808 (2014)

[37] S. P. Boyd, P. John , J. F. William, and R. J. Beck. *Computing core/periphery structures and permutation tests for social relations data.* Social networks 28, no. 2 (2006): 165-178.

[38] R.R. Nadakuditi, M.E.J. Newman, *Graph spectra and the detectability of community structure in networks.* Phys Rev Lett 108(18):188701 (2002).

[39] B. Karrer, M.E.J. Newman, *Stochastic blockmodels and community structure in networks* Physical Review E 83, 016107, (2011).

[40] A. Decelle, F. Krzakala, C. Moore, L. Zdeborov, *Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications*, Physical Review E, 84,066106, (2011).

[41] M.E.J. Newman, S.H. Strogatz, D.J. Watts, *Random graphs with arbitrary degree distributions and their applications*, Physical review E, Phys. Rev. E 64, 026118 (2001).

[42] X. Zhang, T. Martin, M. E. J. Newman. *Identification of core-periphery structure in networks.* Phys. Rev. E 91, 032803 (2015).

[43] S. P. Borgatti, P. Stephen, and M. G. Everett. *Models of core/periphery structures.* Social networks 21, no. 4 (2000): 375-395.

[44] M. G. Everett, G. Martin, and S. P. Borgatti. *The centrality of groups and classes.* The Journal of mathematical sociology 23, no. 3 (1999): 181-201.

[45] F. Krzakala, C. Moore, E. Mossel, J. Neeman, A. Sly, L. Zdeborova, P. Zhang, *Spectral redemption in clustering sparse networks.* Proc. Natl. Acad. Sci. USA 110 20935-20940 (2013).

[46] McKay, Brendan D. *The expected eigenvalue distribution of a large regular graph.* Linear Algebra and its Applications 40 (1981): 203-216.

[47] Dumitriu, Ioana, and Soumik Pal. *Sparse regular random graphs: spectral density and eigenvectors.* The Annals of Probability 40.5 (2012): 2197-2235.

[48] P.T. Peixoto, *Parsimonious module inference in large networks*, Physical review letters 110, 14: 148701 (2013).

[49] Kawamoto, Tatsuro, and Yoshiyuki Kabashima. *Limitations in the spectral method for graph partitioning: Detectability threshold and localization of eigenvectors.*

Physical Review E 91.6 (2015): 062803.

[50] M.P. Rombach, A. Mason, J. H. F. Porter, P. J. Mucha. *Core-periphery structure in networks.* SIAM Journal on Applied mathematics 74, no. 1 (2014): 167-190.

[51] Condon, Anne, and Richard M. Karp. *Algorithms for graph partitioning on the planted partition model.* Random Structures and Algorithms 18.2 (2001): 116-140.

[52] The same model was also studied under the name microcanonical stochastic block model in:
Peixoto, Tiago P. *Nonparametric Bayesian inference of the microcanonical stochastic block model.* arXiv preprint arXiv:1610.02703 (2016).