


Modularity of Erdős-Rényi Random Graphs

Colin McDiarmid

Department of Statistics, University of Oxford, Oxford, UK
cmcd@stats.ox.ac.uk

Fiona Skerman¹

Department of Mathematics, Uppsala University, Uppsala, Sweden
fiona.skerman@math.uu.se

 <https://orcid.org/0000-0003-4141-7059>

Abstract

For a given graph G , modularity gives a score to each vertex partition, with higher values taken to indicate that the partition better captures community structure in G . The modularity $q^*(G)$ (where $0 \leq q^*(G) \leq 1$) of the graph G is defined to be the maximum over all vertex partitions of the modularity value. Given the prominence of modularity in community detection, it is an important graph parameter to understand mathematically.

For the Erdős-Rényi random graph $G_{n,p}$ with n vertices and edge-probability p , the likely modularity has three distinct phases. For $np \leq 1 + o(1)$ the modularity is $1 + o(1)$ with high probability (whp), and for $np \rightarrow \infty$ the modularity is $o(1)$ whp. Between these regions the modularity is non-trivial: for constants $1 < c_0 \leq c_1$ there exists $\delta > 0$ such that when $c_0 \leq np \leq c_1$ we have $\delta < q^*(G) < 1 - \delta$ whp. For this critical region, we show that whp $q^*(G_{n,p})$ has order $(np)^{-1/2}$, in accord with a conjecture by Reichardt and Bornholdt in 2006 (and disproving another conjecture from the physics literature).

2012 ACM Subject Classification Theory of computation \rightarrow Random network models

Keywords and phrases Community detection, Newman-Girvan Modularity, random graphs

Digital Object Identifier 10.4230/LIPIcs.AofA.2018.31

1 Introduction

We start this section with some background and definitions, and then present our results on the modularity of the random graph $G_{n,p}$. After that, we sketch previous work on modularity, and then give a plan of the rest of the paper, which essentially consists of the proofs of the three phases. The remaining proofs will be given in the extended version of this paper.

1.1 Definitions

The large and increasing quantities of network data available in many fields has led to great interest in techniques to discover network structure. We want to be able to identify if a network can be decomposed into communities or highly clustered components.

Modularity was introduced by Newman and Girvan in 2004 [27]. It gives a measure of how well a graph can be divided into communities, and now forms the backbone of the most popular algorithms used to cluster real data [18]. Here a ‘community’ is a collection of nodes which are more densely interconnected than one would expect – see the discussion following the definition of modularity below. There are many applications, including protein discovery, identifying connections between websites, and mapping the onset of schizophrenia on neuron

¹ Supported by the Knut and Alice Wallenberg Foundation and the Swedish Research Council.



clusters in the brain [2]. Its widespread use and empirical success in finding communities in networks makes modularity an important function to understand mathematically. See [11] and [28] for surveys on the use of modularity for community detection in networks.

Given a graph G , modularity gives a score to each vertex partition: the modularity $q^*(G)$ (sometimes called the ‘maximum modularity’) of G is defined to be the maximum of these scores over all vertex partitions. For a set A of vertices, let $e(A)$ be the number of edges within A , and let the *volume* $\text{vol}(A)$ be the sum over the vertices v in A of the degree d_v .

► **Definition 1.1** (Newman & Girvan [27], see also Newman [26]). Let G be a graph with $m \geq 1$ edges. For a vertex partition \mathcal{A} of G , the modularity of \mathcal{A} on G is

$$q_{\mathcal{A}}(G) = \frac{1}{2m} \sum_{A \in \mathcal{A}} \sum_{u,v \in A} \left(\mathbf{1}_{uv \in E} - \frac{d_u d_v}{2m} \right) = \frac{1}{m} \sum_{A \in \mathcal{A}} e(A) - \frac{1}{4m^2} \sum_{A \in \mathcal{A}} \text{vol}(A)^2;$$

and the modularity of G is $q^*(G) = \max_{\mathcal{A}} q_{\mathcal{A}}(G)$, where the maximum is over all partitions \mathcal{A} of the vertices of G .

Isolated vertices are irrelevant. We need to give empty graphs (graphs with no edges) some modularity value. Conventionally we set $q^*(G) = 1$ for each empty graph G [5] (though the value will not be important). The second equation for $q_{\mathcal{A}}(G)$ expresses modularity as the difference of two terms, the *edge contribution* or coverage $q_{\mathcal{A}}^E(G) = \frac{1}{m} \sum_{A \in \mathcal{A}} e(A)$, and the *degree tax* $q_{\mathcal{A}}^D(G) = \frac{1}{4m^2} \sum_{A \in \mathcal{A}} \text{vol}(A)^2$. Since $q_{\mathcal{A}}^E(G) \leq 1$ and $q_{\mathcal{A}}^D(G) > 0$, we have $q_{\mathcal{A}}(G) < 1$ for any non-empty graph G . Also, the trivial partition \mathcal{A}_0 with all vertices in one part has $q_{\mathcal{A}_0}^E(G) = q_{\mathcal{A}_0}^D(G) = 1$, so $q_{\mathcal{A}_0}(G) = 0$. Thus we have

$$0 \leq q^*(G) \leq 1.$$

Suppose that we pick uniformly at random a multigraph with degree sequence (d_1, \dots, d_n) where $\sum_v d_v = 2m$. Then the expected number of edges between vertices u and v is $d_u d_v / (2m - 1)$. This is the original rationale for the definition: whilst rewarding the partition for capturing edges within the parts, we should penalise by (approximately) the expected number of edges.

A differentiation between graphs which are truly modular and those which are not can ... only be made if we gain an understanding of the intrinsic modularity of random graphs. – Reichardt and Bornholdt [30]. In this paper we investigate the likely value of the modularity of an Erdős-Rényi random graph. Let n be a positive integer. Given $0 \leq p \leq 1$, the random graph $G_{n,p}$ has vertex set $[n] := \{1, \dots, n\}$ and the $\binom{n}{2}$ possible edges appear independently with probability p . Given an integer m with $0 \leq m \leq \binom{n}{2}$, the random graph $G_{n,m}$ is sampled uniformly from the m -edge graphs on vertex set $[n]$. These two random graphs are closely related when $m \approx \binom{n}{2} p$: we shall investigate only $q^*(G_{n,p})$ here, but in the extended version of the paper we shall also deduce corresponding results for $q^*(G_{n,m})$.

For a sequence of events A_n we say that A_n holds *with high probability (whp)* if $\mathbb{P}(A_n) \rightarrow 1$ as $n \rightarrow \infty$. For a sequence of random variables X_n and a real number a , we write $X_n \xrightarrow{p} a$ if X_n converges in probability to a as $n \rightarrow \infty$ (that is, if for each $\varepsilon > 0$ we have $|X_n - a| < \varepsilon$ whp).

1.2 Results on the modularity of the random graph $G_{n,p}$

Our first theorem, the Three Phases Theorem, gives the big picture. The three phases correspond to when (a) the expected vertex degree (essentially np) is at most about 1, (b) bigger than 1 but bounded, or (c) tending to infinity.

► **Theorem 1.2.** *Let $p = p(n)$ satisfy $0 \leq p \leq 1$.*

(a) *If $n^2p \rightarrow \infty$ and $np \leq 1 + o(1)$ then $q^*(G_{n,p}) \xrightarrow{p} 1$.*

(b) *Given constants $1 < c_0 \leq c_1$, there exists $\delta = \delta(c_0, c_1) > 0$ such that if $c_0 \leq np \leq c_1$ for n sufficiently large, then whp $\delta < q^*(G_{n,p}) < 1 - \delta$.*

(c) *If $np \rightarrow \infty$ then $q^*(G_{n,p}) \xrightarrow{p} 0$.*

We are able to confirm the $(np)^{-1/2}$ growth rate conjectured to hold for the critical region in [30]. The edge probabilities p correspond to parts (b) or (c) of Theorem 1.2.

► **Theorem 1.3.** *There exists b such that for all $0 < p = p(n) \leq 1$ we have $q^*(G_{n,p}) < \frac{b}{\sqrt{np}}$ whp. Also, given $0 < \varepsilon < 1$, there exists $a = a(\varepsilon) > 0$ such that, if $p = p(n)$ satisfies $np \geq 1$ and $p \leq 1 - \varepsilon$ for n sufficiently large, then $q^*(G_{n,p}) > \frac{a}{\sqrt{np}}$ whp.*

Observe that the upper bound here on $q^*(G_{n,p})$ implies part (c) of Theorem 1.2. As an immediate corollary of Theorem 1.3 we have:

► **Corollary 1.4.** *There exists $0 < a < b$ such that, if $1/n \leq p = p(n) \leq 0.99$ then*

$$\frac{a}{\sqrt{np}} < q^*(G_{n,p}) < \frac{b}{\sqrt{np}} \quad \text{whp.}$$

This result confirms the $\Theta((np)^{-1/2})$ growth rate predicted to hold in this range by Reichardt and Bornholdt [30]: further details of their prediction are given in Section 1.3.

In this extended abstract we give a full proof of the Three Phases Theorem, Theorem 1.2. For Theorem 1.3 we give a proof of the upper bound. We also give a sketch proof of the lower bound, based on an algorithm we call *Swap*, which whp outputs a bipartition achieving the required modularity.

A higher modularity score is taken to indicate a better community division. Thus to determine whether a clustering \mathcal{A} in a graph G shows significant community structure we should compare $q_{\mathcal{A}}(G)$ to the likely (maximum) modularity for an appropriate null model, that is, to the likely value of $q^*(\tilde{G})$ for null model \tilde{G} . It is an interesting question which null model may be most appropriate in a given situation. For example, real networks have been shown to exhibit power law degree behaviour and so null models which can mimic this have been suggested; for example the Chung-Lu model [1] and random hyperbolic graphs [17]. However, a natural minimum requirement is not to consider a community division of a real network as statistically significant unless it has higher modularity than the Erdős-Rényi random graph of the same edge density.

1.3 Previous work on Modularity

The vast majority of papers referencing modularity are papers in which real data, clustered using modularity based algorithms, are analysed. Alongside its use in community detection, many interesting properties of modularity have been documented. A basic observation is that, given a graph G without isolated vertices, in each optimal partition, for each part the corresponding induced subgraph of G must be connected.

Properties and modularity of graph classes

The idea of a *resolution limit* was introduced by Fortunato and Barthélemy [12] in 2007: in particular, if a connected component C in an m -edge graph has strictly fewer than $\sqrt{2m}$ edges, then every optimal partition will cluster the vertices of C together. This is so even if the connected component C consists of two large cliques joined by a single edge. This

property highlights the sensitivity of modularity to noise in the network: if that edge between the cliques, perhaps a mistake in the data, had not been there, then the cliques would be in separate parts in every optimal partition.

The complexity is known. Brandes et al. showed in 2007 that finding the (maximum) modularity of a graph is NP-hard [4]. The reduction required some properties of optimal partitions; for example it was shown that a vertex of degree 1 will be placed in the same part as its neighbour in every optimal partition. Indeed, every part in every optimal partition has size at least 2 or is an isolated vertex, see Lemma 1.6.5 in [31]. The paper [4] also began the rigorous study of the modularity of classes of graphs, in particular of cycles and complete graphs. Later Bagrow [3] and Montgolfier et al. [9] proved that some classes of trees have high modularity, and this was extended in [21] to all trees with maximum degree $o(n)$, and indeed to all graphs where the product of treewidth and maximum degree grows more slowly than the number of edges. There is a growing literature concerning the modularity behaviour of different classes of graphs, see for example [3, 9, 20, 21, 29, 32].

Franke and Wolfe in [13] look at a very different topic, namely the distribution of the modularity of a random partition of a graph or random graph, rather than the modularity of the graph, which is the maximum modularity of a partition. The paper covers some random weighted models where the probability of an edge is proportional to the product of the weights of the end-vertices, including the case of the Erdős-Rényi random graph $G_{n,p}$ for $np \rightarrow \infty$. They show that the modularity of a random partition is asymptotically normally distributed. Their results do not imply anything about the (maximum) modularity $q^*(G_{n,p})$; see also the discussion in the conclusion of [21].

Statistical Physics predictions

In 2004 Guimera et al. [15] observed through simulations that the modularity of random graphs can be surprisingly high. In [15] they conjectured that, for each (large) constant $c > 1$, if $p = c/n$ then whp $q^*(G_{n,p}) \approx c^{-2/3}$. In 2006 Reichardt and Bornholdt [30] made a different conjecture for the modularity in this range. They assumed that an optimal partition will have parts of equal size, then approximated the number of edges between parts using results from [16], where the authors give spin glass predictions for the minimum number of crossing edges in an equipartition of a random graph. For $p = c/n$ their prediction was $q^*(G_{n,c/n}) \approx 0.97 c^{-1/2}(1+o(1))$ whp and we confirm this growth rate. Indeed they predicted $q^*(G_{n,p}) \approx 0.97\sqrt{(1-p)/np}$ which is $\Theta((np)^{-1/2})$ for $1/n \leq p \leq 0.99$. Hence Corollary 1.4 proves that for a large range of p the prediction of Reichardt and Bornholdt [30] is correct up to constant factors (and refutes that of Guimera et al.).

1.4 Plan of the paper

The three phases theorem Theorem 1.2 gave an overview of the behaviour of the modularity $q^*(G_{n,p})$, with the three parts (a), (b) and (c) corresponding to increasing edge-probability p , starting with the sparse case. Theorem 1.3, gave more detailed results for the critical region and confirmed the $(np)^{-1/2}$ growth rate conjectured in the physics community.

Our proofs are organised by starting with the sparse case. In Section 2 we prove Theorem 1.2 part (a), by showing that the partition \mathcal{C} into connected components satisfies $q_{\mathcal{C}}(G_{n,p}) \xrightarrow{p} 1$ in the sparse case. We prove Theorem 1.2 part (b) in Section 3: the lower bound follows quickly from counting isolated edges in $G_{n,p}$, and to prove the upper bound we use expansion properties of the giant component. Section 4 concerns the $a(np)^{-1/2}$ lower bound on $q^*(G_{n,p})$ in Theorem 1.3, and we give a sketch of the proof. In Section 5, we use a robustness result and spectral methods to prove the upper bound $b(np)^{-1/2}$ on $q^*(G_{n,p})$ in Theorem 1.3.

2 The sparse phase: proof of Theorem 1.2 (a)

We can prove that sufficiently sparse random graphs whp have modularity near 1 without developing any extra theory, and we do so here. Lemma 2.2 gives part (a) of the three phases result Theorem 1.2. It is convenient to record first one standard preliminary result on degree tax.

► **Lemma 2.1.** *Let the graph G have $m \geq 1$ edges, and let \mathcal{A} be a k -part vertex partition. If \mathcal{A} has k parts then $q_{\mathcal{A}}^D(G) \geq 1/k$; and if x, y are respectively the largest, second largest volume of a part, then $q_{\mathcal{A}}^D(G) \leq x/2m$ and $q_{\mathcal{A}}^D(G) \leq (x/2m)^2 + y/2m$.*

Proof. All the bounds follow from the convexity of $f(t) = t^2$. Let x_i be the volume of the i th part in \mathcal{A} . For the lower bound, observe that $x_1, \dots, x_k \geq 0$ and $\sum_{i=1}^k x_i = 2m$ together imply that $\sum_{i=1}^k x_i^2 \geq k(2m/k)^2 = 4m^2/k$; and thus $q_{\mathcal{A}}^D(G) = \sum_i x_i^2 / (2m)^2 \geq 1/k$.

For the upper bounds, observe that $0 \leq x_1, \dots, x_k \leq x$ and $\sum_{i=1}^k x_i = 2m$ together imply that $\sum_{i=1}^k x_i^2 \leq (2m/x)x^2 = 2mx$; and so $q_{\mathcal{A}}^D(G) \leq x/2m$. Similarly, supposing that $x_k = x$ and $x_i \leq y$ for $i = 1, \dots, k - 1$, we have $\sum_{i=1}^{k-1} x_i^2 \leq (2m - x)y \leq 2my$; and so $q_{\mathcal{A}}^D(G) \leq (x^2 + 2my) / (2m)^2 = (x/2m)^2 + y/2m$. ◀

► **Lemma 2.2.** *Let $0 < \varepsilon \leq 1/4$, and let $p = p(n)$ satisfy $n^2p \rightarrow \infty$ and $np \leq 1 + \varepsilon$ for n sufficiently large. Then $q^*(G_{n,p}) \geq q_C(G_{n,p}) > 1 - (4\varepsilon)^2$ whp.*

Proof. Let $m = e(G_{n,p})$, and let X be the maximum number of edges in a connected component of $G_{n,p}$. Note that for the connected components partition \mathcal{C} , the edge contribution is 1, and so by the first upper bound on the degree tax in Lemma 2.1, we have $q_C(G_{n,p}) \geq 1 - \frac{X}{m}$. We shall see that when $np \leq 1$ we have $X/m = o(1)$ whp, and so $q_C(G_{n,p}) = 1 - o(1)$ whp. To prove this we break into three ranges of p . The final range, when $1 < np \leq 1 + \varepsilon$ will require a little more care. Observe that since $n^2p \rightarrow \infty$ we have $m \sim n^2p/2$ whp.

Range 1: $n^2p \rightarrow \infty$ and $np \leq n^{-3/4}$. Whp $G_{n,p}$ consists of disjoint edges. This follows by the first moment method, since the expected number of paths on three vertices is $\Theta(n^3p^2)$. Hence whp $X/m = 1/m = o(1)$.

Range 2: $n^{-3/4} \leq np \leq 1/2$. Whp all components are trees or unicyclic and have $O(\log n)$ vertices. Hence whp $X = O(\log n)$ and whp $X/m = O(\log n/n^2p) = o(1)$.

Range 3: $1/2 \leq np \leq 1$. Since $np \leq 1$, whp the maximum number of edges in any component is $o(n)$ (see the next range). But whp $m = \Theta(n)$, and so whp $X/m = o(1)$.

Range 4: $1 < np \leq 1 + \varepsilon/4$. Let $c = 1 + \varepsilon$. Let $x = x(c)$ be the unique root in $(0, 1)$ of $xe^{-x} = ce^{-c}$. Then, for $G_{n,c/n}$, whp $X = (1 + o(1))(1 - x^2/c^2)cn/2$ and each component other than the giant has $O(\log n)$ edges (see for example Theorem 2.14 of [14]). We claim that

$$(1 - x^2/c^2)c < 4\varepsilon/(1 + \varepsilon). \tag{1}$$

To see this, let $f(t) = (1 + t)e^{-(1+t)} - (1 - t)e^{-(1-t)}$ for $t \geq 0$. Then $f(0) = 0$; and for $t > 0$,

$$f'(t) = e^{-(1+t)}(-1 + t + 1) - e^{-(1-t)}((1 - t) - 1) = te^{-1}(e^t - e^{-t}) > 0;$$

31:6 Modularity of Erdős-Rényi Random Graphs

and so $f(t) > 0$ for all $t > 0$. Then $f(\varepsilon) > 0$, that is $(1 - \varepsilon)e^{-(1-\varepsilon)} < (1 + \varepsilon)e^{-(1+\varepsilon)}$, and it follows that $1 - x < \varepsilon$. Hence, $1 - x^2/c^2 < 1 - (1 - \varepsilon)^2/(1 + \varepsilon)^2$. But now

$$(1 - x^2/c^2)c < (1 - (1-\varepsilon)^2/(1+\varepsilon)^2)(1+\varepsilon) = ((1+\varepsilon)^2 - (1-\varepsilon)^2)/(1+\varepsilon) = 4\varepsilon/(1+\varepsilon),$$

and we have proved (1). Hence, for $G_{n,c/n}$, whp $X \leq \frac{4\varepsilon}{1+\varepsilon} \frac{n}{2}$; and so by monotonicity this holds also for $G_{n,p}$ (with $p \leq cn$ as here). Also, $e(G_{n,1/n}) \geq \frac{1+\varepsilon/2}{1+\varepsilon} \frac{n}{2}$ whp, and so by monotonicity this holds also for $G_{n,p}$. Now by the last part of Lemma 2.1, whp

$$q_{\mathcal{C}}(G_{n,p}) \geq 1 - (X/m)^2 - O((\log n)/n) \geq 1 - \left(\frac{4\varepsilon}{1+\varepsilon/2}\right)^2 - O((\log n)/n) > 1 - (4\varepsilon)^2.$$

This completes the proof of the lemma. ◀

3 The middle phase: proof of Theorem 1.2 (b)

It is straightforward to use known results to prove Theorem 1.2 part (b). First we show that the connected components partition \mathcal{C} yields the lower bound. The lower bound will follow also from the lower bound in Theorem 1.3 part (b), but that has quite a long and involved proof, whereas the proof below is only a few lines. As we noted earlier, the upper bound in Theorem 1.3 part (b) will give the upper bound in Theorem 1.2 part (b) for large np , but not when np is small.

3.1 Proof of lower bound

There is a simple reason why the modularity $q^*(G_{n,p})$ is bounded away from 0 whp when the average degree is bounded, namely that whp there is a linear number of isolated edges. First, here is a deterministic lemma.

► **Lemma 3.1.** *Let the graph G have $m \geq 2$ edges, and $i \geq \eta m$ isolated edges, where $0 < \eta \leq \frac{1}{2}$. Then $q_{\mathcal{C}}(G) \geq \eta$.*

Proof. Note first that if $i = m$ then $q_{\mathcal{C}}(G) = 1 - 1/m \geq \eta$. Thus we may assume that $i < m$, and so $i \leq m - 2$. Since there are in total $m - i$ edges in the components which are not isolated edges,

$$q_{\mathcal{C}}(G) \geq 1 - \frac{(m-i)^2}{m^2} - \frac{i}{m^2}.$$

Treating i as a continuous variable and differentiating, we see that the bound is an increasing function of i for $i \leq m - 1$; and so, setting $i = \eta m$,

$$q_{\mathcal{C}}(G) \geq 1 - (1 - \eta)^2 - \eta/m = \eta + \eta(1 - \eta - 1/m) \geq \eta,$$

as required. ◀

Assume that $1 \leq np \leq c_1$. Let X be the number of isolated edges in $G_{n,p}$. Then

$$\mathbb{E}[X] = \binom{n}{2} p(1-p)^{2n-4} = n \cdot \left(\frac{1}{2} + o(1)\right) np e^{-2np} \geq n \cdot \left(\frac{1}{2} + o(1)\right) c_1 e^{-2c_1},$$

since $f(x) = xe^{-2x}$ is decreasing for $x > \frac{1}{2}$. A simple calculation shows that the variance of X is $o((\mathbb{E}[X])^2)$; thus by Chebyshev's inequality, whp $X \geq n \cdot \frac{1}{3} c_1 e^{-2c_1}$. Similarly, whp $m = e(G_{n,p}) \leq \frac{2}{3} c_1 n$; and so whp $X/m \geq \frac{1}{2} e^{-2c_1}$. Finally, Lemma 3.1 shows that whp $q_{\mathcal{C}}(G_{n,p}) \geq \eta = \frac{1}{2} e^{-2c_1}$. This completes the proof of the lower bound.

3.2 Proof of upper bound

It is convenient to spell out the upper bound in Theorem 1.2(b) as the following lemma.

► **Lemma 3.2.** *Given constants $1 < c_0 < c_1$, there exists $\varepsilon = \varepsilon(c_0, c_1) > 0$ such that, if $c_0 \leq np \leq c_1$ for n sufficiently large, then whp $q^*(G_{n,p}) < 1 - \varepsilon$.*

For the proof of this lemma we use a result from [19] concerning edge expansion in the giant component. Define a (δ, η) -cut of $G = (V, E)$ to be a bipartition of V into V_1, V_2 such that both sets have at least $\delta|V|$ vertices and $e(V_1, V_2) < \eta|V|$. We need only the case $\delta = 1/3$.

Proof of Lemma 3.2. We employ double exposure. Let $G' \sim \mathcal{G}_{n, c_0/n}$. For each non-edge of G' resample with probability $p' = (p - c_0/n)/(1 - c_0/n)$ to obtain G , so $G \sim G_{n,p}$. Let \mathcal{A} be an optimal partition of G . Observe that whp $m = e(G) < c_1 n$, and then

$$1 - q^*(G) = \frac{1}{2m} \sum_{A \in \mathcal{A}} \left(e_G(A, \bar{A}) + \frac{\text{vol}_G(A)^2}{2m} \right) > \frac{1}{2c_1 n} \sum_{A \in \mathcal{A}} \left(e_{G'}(A, \bar{A}) + \frac{\text{vol}_{G'}(A)^2}{2c_1 n} \right).$$

Thus it suffices to show that whp, for each vertex partition \mathcal{A} ,

$$\sum_{A \in \mathcal{A}} \left(e_{G'}(A, \bar{A}) + \frac{\text{vol}_{G'}(A)^2}{2c_1 n} \right) \geq 2\varepsilon c_1 n. \quad (2)$$

We will now work solely with G' , so we shall drop the subscripts. Whp G' has a unique giant component H , such that H does not admit a $(1/3, \eta)$ -cut for a constant $\eta = \eta(c_0) > 0$ by [19] [Lemma 2], and such that $|V(H)| \sim (1 - t_0/c_0)n$ where $t_0 < 1$ satisfies $t_0 e^{-t_0} = c_0 e^{-c_0}$ [10]. Let F be the event that G' has a unique giant component H , such that H does not admit a $(1/3, \eta)$ -cut, and $|V(H)| \geq \frac{1}{2}(1 - t_0/c_0)n + 3$. Then the event F holds whp. Let W be a set of vertices such that $|W| \geq \frac{1}{2}(1 - t_0/c_0)n + 3$, and let F_W be the event that F holds and $V(H) = W$. To prove the lemma, it suffices to show that, conditioning on F_W holding, the inequality (2) holds with

$$\varepsilon = \min\{(1 - t_0/c_0)^2/36c_1^2, \eta(1 - t_0/c_0)/2c_1\}.$$

Let \mathcal{A} be any vertex partition which minimises the left side of (2), and let \mathcal{H} be the partition of the giant component H induced by \mathcal{A} , that is, \mathcal{H} consists of the parts $A \in \mathcal{A}$ with $A \cap W$ non-empty (since the induced subgraph on A is connected). Relabel \mathcal{H} as $\{W_1, \dots, W_h\}$ where $h \geq 1$ and $|W_1| \geq \dots \geq |W_h|$. We will restrict our attention to \mathcal{H} . There are two cases to consider.

Case 1. Suppose $|W_1| \geq |W|/3$. As the subgraph induced by W_1 is connected,

$$\text{vol}(W_1) \geq 2(|W_1| - 1) \geq (1 - t_0/c_0)n/3;$$

and so

$$\sum_{A \in \mathcal{A}} \frac{\text{vol}(A)^2}{2c_1 n} \geq \frac{\text{vol}(W_1)^2}{2c_1 n} \geq \frac{(1 - t_0/c_0)^2 n^2}{18c_1 n} \geq 2\varepsilon c_1 n,$$

which yields (2).

Case 2. Now suppose that $|W_i| < |W|/3$ for all $W_i \in \mathcal{H}$. We group the parts to make a bipartition $W = B_1 \cup B_2$ with B_1 and B_2 of similar size. We may for example start with B_1 and B_2 empty, consider the W_i in turn, and each time add W_i to a smaller of B_1 and B_2 . This clearly gives $||B_1| - |B_2|| < |W|/3$. Since there is no $(1/3, \eta)$ -cut of H in G' , we have $e(B_1, B_2) \geq \eta|W|$. But each edge between B_1 and B_2 lies between the parts of \mathcal{A} , and so

$$\sum_{A \in \mathcal{A}} e(A, \bar{A}) \geq 2e(B_1, B_2) \geq 2\eta|W| > \eta(1 - t_0/c_0)n \geq 2\epsilon c_1 n,$$

which again yields (2), and completes the proof. \blacktriangleleft

4 The $a(np)^{-1/2}$ lower bound on the modularity $q^*(G_{n,p})$

We consider a simple algorithm *Swap* which, given a graph G , runs in linear time (in time $O(n + m)$ if G has n vertices and m edges), and yields a balanced bipartition \mathcal{A} of the vertices. The theorem below shows that $q_{\mathcal{A}}(G_{n,p})$ yields a good lower bound for $q^*(G_{n,p})$.

► **Theorem 4.1.** *There are constants c_0 and $a > 0$ such that (a) if $p = p(n)$ satisfies $c_0 \leq np \leq n - c_0$ for n sufficiently large, then whp $q_{\mathcal{A}}(G_{n,p}) \geq \frac{1}{5} \sqrt{\frac{1-p}{np}}$; and (b) if $p = p(n)$ satisfies $1 \leq np \leq n - c_0$ for n sufficiently large, then whp $q_{\mathcal{A}}(G_{n,p}) \geq a \sqrt{\frac{1-p}{np}}$.*

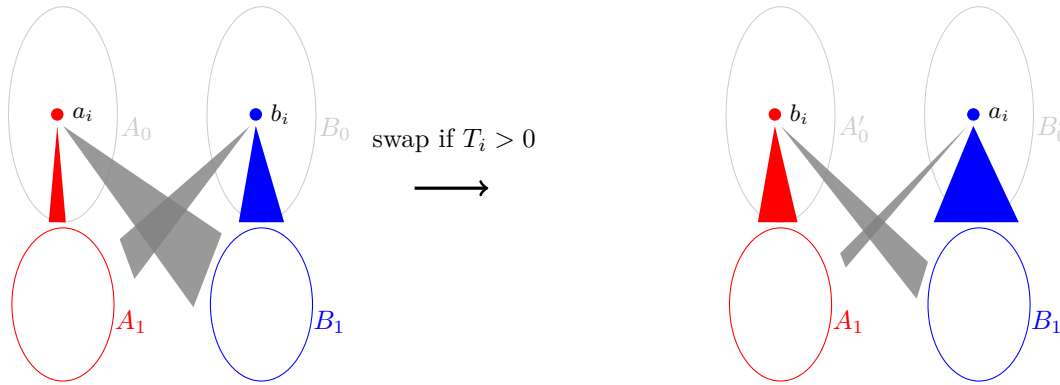
The idea of the proof of Theorem 4.1 is as follows. The algorithm *Swap* starts with a balanced bipartition of the vertex set into $A \cup B$, which has modularity very near 0 whp. By swapping some pairs (a_i, b_i) between A and B , whp we can increase the edge contribution significantly, without changing the distribution of the degree tax (and without introducing dependencies which would be hard to analyse). We give a sketch proof below but defer the full proof to the extended paper.

Proof of Theorem 4.1 (sketch of the main ideas). Let $n \geq 6$, and let $V = [n]$. We start with the initial bipartition \mathcal{A} of V into $A = \{j \in V : j \text{ is odd}\}$ and $B = \{j \in V : j \text{ is even}\}$. Let $k = k(n) = \lfloor n/6 \rfloor$. Let $V_0 = [4k]$, let $V_1 = \{4k + 1, \dots, 6k\}$ and let $V_2 = \{6k + 1, \dots, n\}$. Note that $0 \leq |V_2| \leq 5$: we shall essentially ignore any vertices in V_2 . Let $A_i = A \cap V_i$ and $B_i = B \cap V_i$ for $i = 0, 1, 2$. The six sets A_i, B_i are pairwise disjoint with union V . Currently V_0 is partitioned into $A_0 \cup B_0$: the algorithm *Swap* ‘improves’ this partition, keeping the other 4 sets fixed. For $i = 1, \dots, 2k$ let $a_i = 2i - 1$ and $b_i = 2i$, so $A_0 = \{a_1, \dots, a_{2k}\}$ and $B_0 = \{b_1, \dots, b_{2k}\}$. The way that we improve the partition $V_0 = A_0 \cup B_0$ is by swapping a_i and b_i for certain values i .

Consider the initial bipartition \mathcal{A} . Write G for $G_{n,p}$. It is not hard to show that whp $q_{\mathcal{A}}(G)$ is very near 0. For each $i \in [2k]$ let

$$T_i = e(a_i, B_1) - e(a_i, A_1) + e(b_i, A_1) - e(b_i, B_1),$$

and note that the random variables T_1, \dots, T_{2k} are iid. Observe that if $T_i > 0$ and we swap a_i and b_i between A_0 and B_0 (that is, replace A_0 by $(A_0 \setminus \{a_i\}) \cup \{b_i\}$ and similarly for B_0) then $e(A, B)$ decreases by T_i , so the edge contribution of the partition increases. The algorithm *Swap* makes all such swaps (looking only at possible edges between V_0 and V_1). For each $i \in [2k]$, let $(a'_i, b'_i) = (b_i, a_i)$ if we perform a swap, and let $(a'_i, b'_i) = (a_i, b_i)$ if not; and let $A'_0 = \{a'_1, \dots, a'_{2k}\}$ and $B'_0 = \{b'_1, \dots, b'_{2k}\}$. Let us call the resulting balanced bipartition $\mathcal{A}' = (A', B')$, where $A' = A'_0 \cup A_1 \cup A_2$ and $B' = B'_0 \cup B_1 \cup B_2$. We shall see that $q_{\mathcal{A}'}(G)$ is as required.



■ **Figure 1** An illustration of the constructed partition in the proof of Theorem 4.1.

Let $T^* = \sum_{i \in [2k]} |T_i|$. Observe that

$$e(A'_0, A_1) + e(B'_0, B_1) - (e(A'_0, B_1) + e(A_1, B'_0)) = T^*,$$

so

$$e(A'_0, B_1) + e(A_1, B'_0) = \frac{1}{2}e(V_0, V_1) - \frac{1}{2}T^*. \tag{3}$$

This is where \mathcal{A}' will gain over \mathcal{A} . The main technical part of the proof is to show that whp T^* is large and we leave this to the full version of the paper. We will also show that the degree tax for \mathcal{A}' has exactly the same distribution as for the initial bipartition \mathcal{A} , and it will follow that it is very close to $1/2$ whp. ◀

5 Upper bounds on modularity

In this section we prove the upper bound on $q^*(G_{n,p})$ in Theorem 1.3, which establishes both part (c) of Theorem 1.2, and the upper bound in part (b) of Theorem 1.2. In Section 5.1 we give bounds on the modularity of a graph G in terms of the eigenvalues of its normalised Laplacian $\mathcal{L}(G)$. In Section 5.2, these results are used, together with spectral bounds from [7] and [8], and a ‘robustness’ result on modularity, to complete the proof.

5.1 Spectral upper bound on modularity

The main task of this subsection is prove that the modularity of a graph is bounded above by the spectral gap of the normalised Laplacian. We begin with a definition. For an n -vertex graph G with adjacency matrix A_G and no isolated vertices define the *degrees matrix* D to be the diagonal matrix $\text{diag}(d_1, \dots, d_n)$ and the *normalised Laplacian* to be $\mathcal{L} = I - D^{-1/2}A_GD^{-1/2}$. Here $D^{-1/2}$ is $\text{diag}(d_1^{-1/2}, \dots, d_n^{-1/2})$. Denote the eigenvalues of \mathcal{L} by $0 = \lambda_0 \leq \dots \leq \lambda_{n-1} (\leq 2)$, see [6]. We call

$$\max_{i \neq 0} |1 - \lambda_i| = \max\{|1 - \lambda_1|, |\lambda_{n-1} - 1|\}$$

the *spectral gap* of G , and denote it be $\bar{\lambda}(G)$. (In terms of the eigenvalues $\tilde{\lambda}_0 \geq \dots \geq \tilde{\lambda}_{n-1}$ of $D^{-1/2}A_GD^{-1/2}$, we have $\tilde{\lambda}_i = 1 - \lambda_i$ and so $\bar{\lambda}(G) = \max_{i \neq 0} |\tilde{\lambda}_i| = \max\{|\tilde{\lambda}_1|, |\tilde{\lambda}_{n-1}|\}$.)

31:10 Modularity of Erdős-Rényi Random Graphs

► **Lemma 5.1.** *Let G be a graph with at least one edge and no isolated vertices. Then*

$$q_{\mathcal{A}}(G) \leq \bar{\lambda}(G) (1 - 1/k) \leq \bar{\lambda}(G)$$

for each k -part vertex partition \mathcal{A} , and so $q^*(G) \leq \bar{\lambda}(G)$.

The proof of Lemma 5.1 relies on a corollary of the Discrepancy Inequality, Theorem 5.4 of [6], which is an extension of the Expander-Mixing Lemma to non-regular graphs. Write $\bar{S} = V \setminus S$ where $V = V(G)$.

► **Lemma 5.2** (Corollary 5.5 of [6]). *Let G be a graph with at least one edge and no isolated vertices. Then for each $S \subseteq V$*

$$e(S, \bar{S}) \geq (1 - \bar{\lambda}(G)) \text{vol}(S) \text{vol}(\bar{S}) / \text{vol}(G).$$

Proof of Lemma 5.1. Let G have $m \geq 1$ edges. Let $\mathcal{A} = \{A_1, \dots, A_k\}$ be a vertex partition of G . Lemma 5.2 guarantees many edges between the parts of \mathcal{A} . The edge contribution satisfies

$$1 - q_{\mathcal{A}}^E(G) = \frac{1}{2m} \sum_i e(A_i, \bar{A}_i) \geq (1 - \bar{\lambda}) \frac{1}{4m^2} \sum_i \text{vol}(A_i) \text{vol}(\bar{A}_i);$$

and

$$\frac{1}{4m^2} \sum_i \text{vol}(A_i) \text{vol}(\bar{A}_i) = \frac{1}{4m^2} \sum_i \text{vol}(A_i) (2m - \text{vol}(A_i)) = 1 - q_{\mathcal{A}}^D(G).$$

Hence

$$1 - q_{\mathcal{A}}^E(G) \geq (1 - \bar{\lambda})(1 - q_{\mathcal{A}}^D(G)),$$

and so

$$q_{\mathcal{A}}(G) = q_{\mathcal{A}}^E(G) - q_{\mathcal{A}}^D(G) \leq \bar{\lambda}(1 - q_{\mathcal{A}}^D(G)) \leq \bar{\lambda}(1 - \frac{1}{k})$$

(since $q_{\mathcal{A}}^D(G) \geq 1/k$ by Lemma 2.1). This completes the proof. ◀

5.2 The $b(np)^{-1/2}$ upper bound on the modularity $q^*(G_{n,p})$.

We are now ready to prove the spectral upper bound for $q^*(G_{n,p})$. Let us restate the upper bound in Theorem 1.3 as a lemma.

► **Lemma 5.3.** *There is a constant b such that for $0 < p = p(n) \leq 1$*

$$q^*(G_{n,p}) \leq \frac{b}{\sqrt{np}} \quad \text{whp.}$$

Proof. Notice first that it suffices to show that there exist c_0 and b such that for $np \geq c_0$ whp $q^*(G_{n,p}) \leq b/\sqrt{np}$, and then replace b by $\max\{\sqrt{c_0}, b\}$.

For $p \gg \log^2 n/n$, the result follows directly from Lemma 5.1, and Theorem 3.6 of Chung, Vu and Lu [7] (see also (1.2) in [8]), which shows that

$$\bar{\lambda}(G_{n,p}) \leq 4(np)^{-1/2}(1 + o(1)) \quad \text{whp.}$$

For the remainder of the proof we assume that $c_0/n \leq p \leq 0.99$ for some large constant $c_0 \geq 1$. We will use the spectral bound in Lemma 5.1 on a subgraph H which is obtained from the random graph $G = G_{n,p}$ by deleting a small subset of the vertices (and the incident edges).

Following the construction in [8], let H be the induced subgraph of G obtained as follows.

- Initially set $H = G \setminus \{v \in V(G) : d_v < (n-1)p/2\}$.
- While there is a vertex $v \in V(H)$ with at least 100 neighbours in $V(G) \setminus V(H)$, remove v from H .

Let V' be the set of deleted vertices, and let E' be the set of deleted edges (the edges incident with vertices in V'). Then by Theorem 1.2 of Coja-Oghlan [8], assuming that c_0 is sufficiently large, there are positive constants c_1 and c_2 such that whp $|V'| \leq ne^{-np/c_2}$ and $\bar{\lambda}(H) \leq c_1(np)^{-1/2}$.

We want a bound on $|E'|$, not $|V'|$. By the proof of Corollary 2.3 in [8], whp in $G_{n,p}$ we have $\text{vol}(S) \leq 2np|S| + ne^{-np/1500}$ simultaneously for each set S of vertices. (The result is stated with $\text{vol}(S)$ replaced by $|N_G(S)|$, the number of neighbours of S outside S , but the proof actually shows the result for $\text{vol}(S)$.) Hence, noting also that $np \geq 1$ and setting $c_3 = \max\{c_2, 1500\}$, whp

$$|E'| \leq \text{vol}(V') \leq 2n^2pe^{-np/c_2} + ne^{-np/1500} \leq 3n^2pe^{-np/c_3} \leq e(G) \cdot 9e^{-np/c_3},$$

where the last inequality follows since whp $e(G) \geq n^2p/3$. By making c_0 larger if necessary, we can ensure that $9e^{-np/c_3} \leq (1/3)(np)^{-1/2}$, and so whp $|E'|/e(G) \leq (1/3)(np)^{-1/2}$. Now, by Lemma 5.1, whp

$$q^*(G \setminus E') = q^*(H) \leq \bar{\lambda}(H) \leq c_1(np)^{-1/2}.$$

One of the ‘robustness’ results in the full paper says that, if H is a graph and E' is a proper subset of the edges, then $|q^*(H) - q^*(H \setminus E')| \leq 3|E'|/e(H)$. Using this result, whp

$$q^*(G) \leq q^*(G \setminus E') + 3|E'|/e(G) \leq (c_1 + 1)(np)^{-1/2},$$

and the proof is complete. ◀

6 Concluding remarks

We have presented results on $q^*(G_{n,p})$, focussing on the three phases as the average degree moves past 1 and then grows to ∞ , and on the $\Theta((np)^{-1/2})$ result. The full paper [24] (as mentioned earlier) also contains corresponding results for $q^*(G_{n,m})$; and it contains some other results, including concentration for both $q^*(G_{n,p})$ and $q^*(G_{n,m})$.

There is further related work in progress: concerning the modularity of very dense graphs and random graphs, see [25]; concerning modularity and edge-sampling (it may be expensive to test if an edge is present), see [23]; and concerning extreme values of modularity (to set random results in context), see [22].

References

- 1 William Aiello, Fan Chung, and Linyuan Lu. A random graph model for power law graphs. *Experimental Mathematics*, 10(1):53–66, 2001. doi:10.1080/10586458.2001.10504428.
- 2 Aaron F. Alexander-Bloch, Nitin Gogtay, David Meunier, Rasmus Birn, Liv Clasen, Francois Lalonde, Rhoshel Lenroot, Jay Giedd, and Edward T. Bullmore. Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia. *Frontiers in Systems Neuroscience*, 4, 2010. doi:10.3389/fnsys.2010.00147.
- 3 James P. Bagrow. Communities and bottlenecks: Trees and treelike networks have high modularity. *Physical Review E*, 85(6):066118, 2012. doi:10.1103/PhysRevE.85.066118.
- 4 Ulrik Brandes, Daniel Delling, Marco Gaertler, Robert Görke, Martin Hoefer, Zoran Nikoloski, and Dorothea Wagner. On finding graph clusterings with maximum modularity. In *Graph-Theoretic Concepts in Computer Science*, pages 121–132. Springer, 2007. doi:10.1007/978-3-540-74839-7_12.

- 5 Ulrik Brandes, Daniel Delling, Marco Gaertler, Robert Gorke, Martin Hofer, Zoran Nikołoski, and Dorothea Wagner. On modularity clustering. *Knowledge and Data Engineering, IEEE Transactions on*, 20(2):172–188, 2008. doi:10.1109/TKDE.2007.190689.
- 6 Fan Chung. *Spectral graph theory*, volume 92. American Mathematical Soc. Providence, RI, 1997. doi:10.1090/cbms/092.
- 7 Fan Chung, Linyuan Lu, and Van Vu. The spectra of random graphs with given expected degrees. *Internet Mathematics*, 1(3):257–275, 2003. doi:10.1073/pnas.0937490100.
- 8 Amin Coja-Oghlan. On the Laplacian Eigenvalues of $G_{n,p}$. *Combinatorics, Probability and Computing*, 16:923–946, 2007. doi:10.1017/S0963548307008693.
- 9 Fabien De Montgolfier, Mauricio Soto, and Laurent Viennot. Asymptotic modularity of some graph classes. In *Algorithms and Computation*, pages 435–444. Springer, 2011. doi:10.1007/978-3-642-25591-5_45.
- 10 P. Erdős and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17–61, 1960.
- 11 Santo Fortunato. Community detection in graphs. *Physics Reports*, 486(3):75–174, 2010. doi:10.1016/j.physrep.2009.11.002.
- 12 Santo Fortunato and Marc Barthélemy. Resolution limit in community detection. *Proceedings of the National Academy of Sciences*, 104(1):36–41, 2007. doi:10.1073/pnas.0605965104.
- 13 Beate Franke and Patrick J. Wolfe. Network modularity in the presence of covariates. preprint *arXiv:1603.01214*, 2016.
- 14 Alan Frieze and Michał Karoński. *Introduction to random graphs*. Cambridge University Press, 2015. doi:10.1017/cbo9781316339831.
- 15 Roger Guimerà, Marta Sales-Pardo, and Luís A. Nunes Amaral. Modularity from fluctuations in random graphs and complex networks. *Physical Review E*, 70:025101, 2004. doi:10.1103/physreve.70.025101.
- 16 Ido Kanter and Haim Sompolinsky. Graph optimisation problems and the Potts glass. *Journal of Physics A*, 20(11):L673, 1987. doi:10.1088/0305-4470/20/11/001.
- 17 Dmitri Krioukov, Fragkiskos Papadopoulos, Maksim Kitsak, Amin Vahdat, and Marián Boguná. Hyperbolic geometry of complex networks. *Physical Review E*, 82(3):036106, 2010. doi:10.1103/physreve.82.036106.
- 18 Andrea Lancichinetti and Santo Fortunato. Limits of modularity maximization in community detection. *Physical Review E*, 84(6):066122, 2011. doi:10.1103/physreve.84.066122.
- 19 Malwina J. Luczak and Colin McDiarmid. Bisecting sparse random graphs. *Random Structures & Algorithms*, 18(1):31–38, 2001. doi:10.1002/1098-2418(200101)18:1<31::aid-rsa3>3.3.co;2-t.
- 20 Colin McDiarmid and Fiona Skerman. Modularity in random regular graphs and lattices. *Electronic Notes in Discrete Mathematics*, 43:431–437, 2013. doi:10.1016/j.endm.2013.07.063.
- 21 Colin McDiarmid and Fiona Skerman. Modularity of regular and treelike graphs. *Journal of Complex Networks*, 5, 2017. doi:10.1093/comnet/cnx046.
- 22 Colin McDiarmid and Fiona Skerman. Extreme values of modularity. In preparation, 2018.
- 23 Colin McDiarmid and Fiona Skerman. Modularity and edge-sampling. In preparation, 2018.
- 24 Colin McDiarmid and Fiona Skerman. Modularity of Erdős-Rényi random graphs. Manuscript, 2018.
- 25 Colin McDiarmid and Fiona Skerman. Modularity of very dense graphs. In preparation, 2018.

- 26 Mark E. J. Newman. *Networks: An Introduction*. Oxford University Press, 2010. doi:10.1093/acprof:oso/9780199206650.001.0001.
- 27 Mark E. J. Newman and Michelle Girvan. Finding and evaluating community structure in networks. *Physical Review E*, 69(2):026113, 2004. doi:10.1103/physreve.69.026113.
- 28 Mason A. Porter, Jukka-Pekka Onnela, and Peter J. Mucha. Communities in networks. *Notices of the AMS*, 56(9):1082–1097, 2009.
- 29 Liudmila Ostroumova Prokhorenkova, Paweł Prałat, and Andrei Raigorodskii. Modularity in several random graph models. *Electronic Notes in Discrete Mathematics*, 61:947–953, 2017. doi:10.1016/j.endm.2017.07.058.
- 30 Jörg Reichardt and Stefan Bornholdt. When are networks truly modular? *Physica D: Nonlinear Phenomena*, 224(1):20–26, 2006. doi:10.1016/j.physd.2006.09.009.
- 31 Fiona Skerman. *Modularity of Networks*. PhD thesis, University of Oxford, 2016.
- 32 Stojan Trajanovski, Huijuan Wang, and Piet Van Mieghem. Maximum modular graphs. *The European Physical Journal B-Condensed Matter and Complex Systems*, 85(7):1–14, 2012. doi:10.1140/epjb/e2012-20898-3.