



Modularity of minor-free graphs

Malgorzata Sulkowska, Michal Lasoń

► To cite this version:

Malgorzata Sulkowska, Michal Lasoń. Modularity of minor-free graphs. Journal of Graph Theory, 2022, 10.1002/jgt.22896 . hal-03462570

HAL Id: hal-03462570

<https://hal.science/hal-03462570>

Submitted on 3 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modularity of minor-free graphs

Michał Lason^{*} and Małgorzata Sulkowska^{*}

ABSTRACT. We prove that a class of graphs with an excluded minor and with the maximum degree sublinear in the number of edges is maximally modular, that is, modularity tends to 1 as the number of edges tends to infinity.

1. Introduction

1.1. Modularity. Modularity is a well-established parameter measuring the presence of community structure in the graph. It was introduced by Newman and Girvan in 2004 ([19]). Nowadays it is widely used as a quality function for community detection algorithms. The most popular heuristic clustering algorithms (check Louvain [5] or Leiden [20]) find the proper partition trying to maximize exactly this parameter. The definition of modularity is based on the comparison between the density of edges inside communities one observes in the graph and the density of edges one would expect if the edges of the graph were wired randomly preserving degree sequence. We make it precise just below.

Consider a simple undirected graph G with $|V(G)| = n$ and $|E(G)| = m$. Whenever the context is clear we write V and E for $V(G)$ and $E(G)$, respectively. For $v \in V$ by $\deg(v)$ denote the degree of a vertex v in G . For a subset of vertices $S \subseteq V$ define $E(S)$ to be the set of edges in G with both end-vertices within S and let $\deg(S) = \sum_{v \in S} \deg(v)$. The *modularity* of G is defined as follows.

DEFINITION 1 (Modularity, [19]). *Let G be a graph with at least one edge. For a partition \mathcal{A} of G into induced subgraphs define its modularity score on G as*

$$\text{mod}_{\mathcal{A}}(G) = \sum_{A \in \mathcal{A}} \left(\frac{|E(A)|}{|E(G)|} - \left(\frac{\deg(V(A))}{\deg(V(G))} \right)^2 \right).$$

Modularity of G is given by

$$\text{mod}(G) = \max_{\mathcal{A}} \text{mod}_{\mathcal{A}}(G).$$

Key words and phrases. Modularity, minor-free graph, sparse graph, edge separator, Cheeger's inequality.

Michał Lason was supported by Polish National Science Centre grant no. 2019/34/E/ST1/00087.

^{*}michalason@gmail.com; Institute of Mathematics of the Polish Academy of Sciences, ul. Śniadeckich 8, 00-656 Warszawa, Poland

^{*}malgorzata.sulkowska@pwr.edu.pl; Université Côte d'Azur, CNRS, Inria, I3S, France; Wrocław University of Science and Technology, Faculty of Fundamental Problems of Technology, Department of Fundamentals of Computer Science

Conventionally, a graph with no edges has modularity equal to 0. For a given partition \mathcal{A} the value $\sum_{A \in \mathcal{A}} \frac{|E(A)|}{|E(G)|}$ is called an *edge contribution* while $\sum_{A \in \mathcal{A}} \left(\frac{\deg(V(A))}{\deg(V(G))} \right)^2$ is a *degree tax*. A single summand of the modularity score is the difference between the fraction of edges within A and the expected fraction of edges within A if we considered a random multigraph on V with the degree sequence given by G .

It is easy to check that $0 \leq \text{mod}(G) < 1$, and also that adding or deleting isolated vertices from the graph does not impact its modularity.

In practice, the problem of community detection very often concerns complex networks, i.e., graphs modeling real-life systems. Since most complex networks are sparse, it is natural to investigate which classes of sparse graphs exhibit high modularity. Our paper addresses exactly this question - modularity of commonly considered subclasses of nowhere dense graphs, that is a class of graphs introduced by Nešetřil and Ossona de Mendez in [18] as capturing the notion of sparsity.

1.2. Related work. For a concise, up to date summary of modularity of various classes of graphs check the appendix of [16] by McDiarmid and Skerman from 2020.

Here we focus on modularity of commonly considered subclasses of nowhere dense graphs. First, recall the definition of *maximally modular* class of graphs.

DEFINITION 2 (Maximally modular class of graphs). *A class of graphs \mathcal{C} is maximally modular if for every $\varepsilon > 0$ there exists M_ε such that whenever G is a graph from \mathcal{C} with $m \geq M_\varepsilon$ edges, then $\text{mod}(G) > 1 - \varepsilon$.*

It is not hard to show (consult [7]) that a maximally modular class of graphs has the maximum degree sublinear in the number of edges. Hence, sublinear maximum degree is a necessary condition for a class of graphs to be maximally modular.

In 2018 McDiarmid and Skerman formulated the following sufficient condition for a class of graphs to be maximally modular. By $\Delta(G)$ and $\text{tw}(G)$ we denote respectively the maximum degree in G and the treewidth of G .

COROLLARY 3 ([15], Corollary 12). *For $m = 1, 2, \dots$ let G_m be a graph with m edges. If $\text{tw}(G_m) \cdot \Delta(G_m) = o(m)$ then $\text{mod}(G_m) \rightarrow 1$ as $m \rightarrow \infty$.*

The above result is tight in a sense that $o(m)$ can not be replaced by $O(m)$. To justify it McDiarmid and Skerman present two examples. First, let G be a star $K_{1,m}$ (with treewidth 1 and maximum degree m). Then, $\text{tw}(G) \cdot \Delta(G) = m$ and, by [7], $\text{mod}(G) = 0$. Second, let G be a random cubic graph on n vertices (thus with $m = 3n/2$ edges). Then, $\text{tw}(G) \cdot \Delta(G) = O(m)$ and with high probability $\text{mod}(G) \leq 0.79$ (see [11, 15]).

Thus, what follows from Corollary 3 is that bounded treewidth in addition to the maximum degree sublinear in the number of edges already guarantees that the class of graphs is maximally modular. On the other hand the example of a random cubic graph gives that classes of bounded degree graphs are not maximally modular. These conclusions already lead to the classification presented in Figure 1.

By Corollary 3 one can obtain also some partial results for planar graphs and for bounded genus graphs.

Indeed, the random planar graph G_n on n vertices has $\text{tw}(G_n) = O(\sqrt{n})$ (which follows from the separator theorem for graphs of bounded genus [10] and a recent

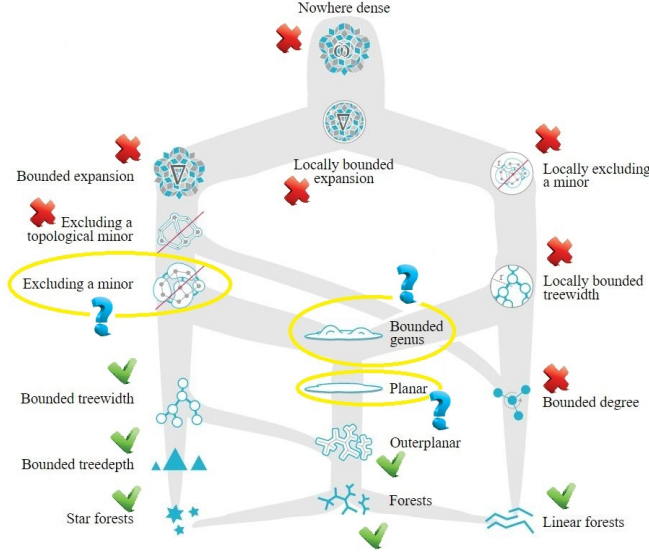


FIGURE 1. Maximally and non-maximally modular subclasses of nowhere dense graphs. All classes of graphs are considered to have the maximum degree sublinear in the number of edges. (Background picture by Felix Reidl, source: <https://tcs.rwth-aachen.de/~reidl/>.)

result by Dvorač and Norin [9] establishing the linear dependence between the treewidth and the separation number for graphs of bounded genus). Next, with high probability $|E(G_n)| = \Theta(n)$ and, by [14], with high probability $\Delta(G_n) = O(\log n)$. Thus, by Corollary 3 random planar graphs are maximally modular.

Similarly, again by [10, 9], for bounded genus graphs $\text{tw}(G_m) = O(\sqrt{m})$. Thus, a class of bounded genus graphs with $\Delta(G_m) = o(\sqrt{m})$ is maximally modular.

1.3. Our results. We prove the following.

THEOREM 4. *A class of graphs with an excluded minor and with the maximum degree sublinear in the number of edges is maximally modular.*

Since classes of graphs with bounded genus and the class of planar graphs are subclasses of graphs with an excluded minor, the above theorem resolves, in positive, all three questions marked in Figure 1 that remained unsolved. This way we achieve a complete classification of maximally modular classes among all commonly considered subclasses of nowhere dense graphs with sublinear maximum degree.

Our proof uses tools of spectral graph theory, in particular so-called Cheeger's Inequality and a recent important result by Biswal et al. ([4]) for graphs with an excluded minor.

2. Proof

We begin by presenting tools of spectral graph theory and results for minor-free graphs that will be used in this section.

Let $G = (V, E)$ be a simple undirected graph. Denote by A its *adjacency matrix* and by D its *diagonal degree matrix*. Recall that $L = D - A$ is the *Laplacian matrix* of G . When G has no isolated vertices, then $\mathcal{L} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}}$ is the *normalized Laplacian matrix* of G . Let $\lambda_i(L)$ and $\lambda_i(\mathcal{L})$ denote i^{th} smallest eigenvalue of L and \mathcal{L} , respectively. We concentrate on the second smallest eigenvalues, $\lambda_2(L)$ and $\lambda_2(\mathcal{L})$, as they carry information about the connectivity of G .

We will need to justify the existence of sufficiently small and reasonably balanced edge cuts in considered graphs. This will be done by Cheeger's inequality, which was first established for manifolds in 1970 by Cheeger [6], while the graph version is due to Alon and Milman [1, 2].

THEOREM 5 (Cheeger's Inequality, [2]). *When G has no isolated vertices, then*

$$\frac{\lambda_2(\mathcal{L})}{2} \leq \min_{\emptyset \neq S \subsetneq V} \frac{|E(S, V \setminus S)|}{\min\{\deg(S), \deg(V \setminus S)\}} \leq \sqrt{2\lambda_2(\mathcal{L})},$$

where $E(S, V \setminus S)$ denotes the set of edges in G between S and $V \setminus S$.

One of the first results for minor-free graphs shows that they have at most a linear number of edges in terms of the number of vertices.

THEOREM 6 ([13]). *A class of graphs with an excluded minor has at most a linear number of edges. That is, for every graph H there exists a constant c_H such that every graph on n vertices without a minor H has at most $c_H n$ edges.*

A recent important result by Biswal et al. for minor-free graphs gives an upper bound for the second smallest eigenvalue of the Laplacian matrix.

THEOREM 7 ([4], Theorem 5.3). *If G is K_h -minor-free and $|V| \geq c_1 h^2 \log h$, then $\lambda_2(L) \leq c_2 \frac{\Delta(G) h^6 \log h}{|V|}$ for some positive absolute constants c_1, c_2 .*

REMARK 8. *Notice that $\lambda_2(\mathcal{L}) \leq \lambda_2(L)$ for graphs without isolated vertices.*

PROOF. We have $\mathcal{L} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}} = (D^{-\frac{1}{2}} L D^{\frac{1}{2}}) D^{-1} =: L' D^{-1}$. Matrices L and L' are similar, hence have the same spectrum. Recall that $\lambda_1(\mathcal{L}) = \lambda_1(L) = 0$, so also $\lambda_1(L') = 0$. Now, we restrict linear maps \mathcal{L} and L' to the subspace, denoted by \ker^\perp , orthogonal to the kernel of \mathcal{L} and L' . Since D^{-1} is a shrinking linear map (module of every eigenvalue less or equal to 1) we get that the module of the smallest eigenvalue of $\mathcal{L}|_{\ker^\perp}$ is less or equal to the module of the smallest eigenvalue of $L'|_{\ker^\perp}$. That is, since these eigenvalues are real and nonnegative, $\lambda_2(\mathcal{L}) \leq \lambda_2(L') = \lambda_2(L)$. \square

PROPOSITION 9. *A class of graphs with an excluded minor, the maximum degree sublinear in the number of edges, and vertices weighted proportionally to their degree has a sublinear weighted edge separator.*

That is, for every graph H and $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that the following is true. Let G be a graph without minor H , with m edges, and maximum degree at most δm . A weight $w(v) = \frac{\deg(v)}{\deg(V)}$ is assigned to every vertex v of G . Then, there is a set of no more than εm edges in G whose deletion creates a graph in which the total weight of every connected component is smaller than ε .

PROOF. Fix a graph H and $\varepsilon > 0$, and let $h = |H|$. Now, choose $\delta > 0$ such that $\frac{\varepsilon}{2} \frac{1}{\delta} \frac{1}{c_H} \geq c_1 h^2 \log h$ and $(\lfloor \log_2 \frac{1}{\varepsilon} \rfloor + 2) 2 \sqrt{2c_2 c_H \frac{2}{\varepsilon} \delta h^6 \log h} < \varepsilon$.

Suppose $G(V, E)$ is a graph without minor H , with m edges, and maximum degree at most δm . In particular, $m \geq \frac{1}{\delta}$. We will show that the statement of the proposition holds for G , ε , and δ . Notice that without loss of generality we may assume that G does not have isolated vertices, as these do not impact weights.

Consider the following procedure which takes on input an induced subgraph $G'(V', E')$ of G with $2|E'| = \deg_{G'}(V') \geq \frac{\varepsilon}{2} \deg(V) = \varepsilon m$. Notice that by Theorem 6 we have $|V'| = |E'| \frac{|V'|}{|E'|} \geq \frac{\varepsilon}{2} m \cdot \frac{1}{c_H} \geq \frac{\varepsilon}{2} \frac{1}{\delta} \frac{1}{c_H} \geq c_1 h^2 \log h$. Clearly, G' is H -minor-free, hence also K_h -minor-free, thus by Theorem 7 since $|V'| \geq c_1 h^2 \log h$, we have that $\lambda_2(L') \leq c_2 \frac{\delta m \cdot h^6 \log h}{|V'|} = c_2 \frac{m}{|V'|} \delta h^6 \log h \leq c_2 \frac{|E'|}{|V'|} \frac{2}{\varepsilon} \delta h^6 \log h \leq c_2 c_H \frac{2}{\varepsilon} \delta h^6 \log h$. Thus, by Theorem 5 and Remark 8, we get that for some proper subset S of V' :

$$(1) \quad |E'(S, V' \setminus S)| \leq \min\{\deg_{G'}(S), \deg_{G'}(V' \setminus S)\} \sqrt{2c_2 c_H \frac{2}{\varepsilon} \delta h^6 \log h}.$$

The procedure deletes edges $E'(S, V' \setminus S)$ from G' , and returns as output all connected components of a resulting graph. These are also induced subgraphs of G .

Now, we begin a process starting with the set T consisting of the graph G and repeatedly apply the above procedure to elements of T (that is, induced subgraphs G' of G) satisfying $\deg_{G'}(V') \geq \frac{\varepsilon}{2} \deg(V)$.

Notice that since the degree of each element of the output is smaller than the degree of the input, the process has to end. Keep in mind that at any step of the process vertices of elements of T form a partition of the set V . Moreover, after first step of the process elements of T are connected induced subgraphs of G .

Suppose that at the end of the process the set T consists of connected induced subgraphs G_1, \dots, G_t . Denote by D the set of edges deleted during the process.

In order to count how many edges were deleted during the process in total, assign in a single procedure that started with G' edges $E'(S, V' \setminus S)$ to vertices in S proportionally to their G' -degree when $\deg_{G'}(S) \leq \deg_{G'}(V' \setminus S)$ and to $V' \setminus S$ otherwise. Now, by the inequality (1) in this single run of the procedure every vertex v in S got assigned at most $\deg_{G'}(v) \sqrt{2c_2 c_H \frac{2}{\varepsilon} \delta h^6 \log h}$ deleted edges when $\deg_{G'}(S) \leq \deg_{G'}(V' \setminus S)$ and 0 otherwise.

Notice that every single vertex v gets assigned nonzero deleted edges at most $\lfloor \log_2 \frac{\varepsilon}{2} \rfloor + 1 = \lfloor \log_2 \frac{1}{\varepsilon} \rfloor + 2$ times. Indeed, it happens when the degree of the induced subgraph to which v belongs gets at least halved. It starts from $\deg(V)$ and ends just after dropping below $\frac{\varepsilon}{2} \deg(V)$. Therefore, summing over all vertices, the total number of deleted edges $|D|$ is at most $(\lfloor \log_2 \frac{1}{\varepsilon} \rfloor + 2) 2m \sqrt{2c_2 c_H \frac{2}{\varepsilon} \delta h^6 \log h} < \varepsilon m$.

Now, connected components of the graph $G \setminus D$ are G_i 's. Every G_i has weight equal to $\frac{1}{\deg(V)} \deg_G(V(G_i)) \leq \frac{1}{\deg(V)} (\deg_{G_i}(V(G_i)) + |D|) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. \square

REMARK 10. Alon, Seymour, and Thomas [3] generalized the planar vertex separator theorem of Lipton and Tarjan [12] to minor-free graphs.

It is known (see [8] using a result of [17]) that a planar graph G has an edge separator of size $O(\sqrt{\Delta(G)|V(G)|})$. It would be interesting to generalize this theorem to minor-free graphs, as this would strengthen Proposition 9.

THEOREM 4. *Let \mathcal{C} be a class of graphs excluding a fixed minor H and with the maximum degree sublinear in the number of edges – that is, for every $\delta > 0$ there exists m_δ such that if $m \geq m_\delta$ and G is a graph from \mathcal{C} with m edges, then $\Delta(G) \leq \delta m$. Then, for every $\varepsilon > 0$ there exists M_ε such that if $m \geq M_\varepsilon$ and G is a graph from \mathcal{C} with m edges, then $\text{mod}(G) \geq 1 - \varepsilon$.*

PROOF. Fix a class of graphs \mathcal{C} as in the assumption and fix $\varepsilon > 0$. Now, choose $\delta := \delta(\frac{\varepsilon}{2})$ such that the assertion of Proposition 9 holds. We will show that if $m \geq M_\varepsilon := m_\delta$ and G is a graph from \mathcal{C} with m edges, then $\text{mod}(G) > 1 - \varepsilon$.

Indeed, then by Proposition 9 there is a set of no more than $\frac{\varepsilon}{2}m$ edges in G whose deletion creates a graph in which the total weight of every connected component is less than $\frac{\varepsilon}{2}$. Now, let \mathcal{A} be the set of those connected components and let D be the set of deleted edges.

Firstly, notice that

$$\sum_{A \in \mathcal{A}} \frac{|E(A)|}{|E(G)|} = \frac{|E(G)| - |D|}{|E(G)|} \geq 1 - \frac{\varepsilon}{2}.$$

Secondly, we have

$$\sum_{A \in \mathcal{A}} \left(\frac{\deg(V(A))}{\deg(V(G))} \right)^2 = \sum_{A \in \mathcal{A}} w(A)^2 < \sum_{A \in \mathcal{A}} \frac{\varepsilon}{2} w(A) = \frac{\varepsilon}{2}.$$

Concluding,

$$\text{mod}(G) \geq \sum_{A \in \mathcal{A}} \left(\frac{|E(A)|}{|E(G)|} - \left(\frac{\deg(V(A))}{\deg(V(G))} \right)^2 \right) > 1 - \varepsilon.$$

□

Acknowledgements

We would like to thank Fiona Skerman for suggesting the problem.

References

- [1] N. Alon. Eigenvalues and expanders. *J. Comb. Theory, Ser. B*, 6:83–96, 1986.
- [2] N. Alon and V. Milman. λ_1 , Isoperimetric inequalities for graphs, and superconcentrators. *J. Comb. Theory, Ser. B*, 38(1):73–88, 1985.
- [3] N. Alon, P. Seymour, and R. Thomas. A separator theorem for nonplanar graphs. *J. Am. Math. Soc.*, 3(4):801–808, 1990.
- [4] P. Biswal, J.R. Lee, and S. Rao. Eigenvalue bounds, spectral partitioning, and metrical deformations via flows. *J. ACM*, 57(3):13:1–13:23, 2010.
- [5] V.D. Blondel, J.L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *J. Stat. Mech. - Theory E*, 2008(10):P10008, 2008.
- [6] J. Cheeger. *A Lower Bound for the Smallest Eigenvalue of the Laplacian*. Princeton University Press, 1970.
- [7] F. de Montgolfier, M. Soto, and L. Viennot. Asymptotic modularity of some graph classes. In T. Asano, S.-I. Nakano, Y. Okamoto, and O. Watanabe, editors, *Algorithms and Computation - 22nd International Symposium, ISAAC 2011, Yokohama, Japan, December 5-8, 2011. Proceedings*, volume 7074 of *Lect. Notes Comput. Sc.*, pages 435–444. Springer, 2011.
- [8] K. Diks, H.N. Djidjev, O. Sykora, and I. Vrto. Edge separators of planar and outerplanar graphs with applications. *J. Algorithm.*, 14(2):258–279, 1993.
- [9] Z. Dvorák and S. Norin. Treewidth of graphs with balanced separations. *J. Comb. Theory, Ser. B*, 137:137–144, 2019.
- [10] J.R. Gilbert, J.P. Hutchinson, and R.E. Tarjan. A separator theorem for graphs of bounded genus. *J. Algorithm.*, 5(3):391–407, 1984.

- [11] L. Lichev and D. Mitsche. On the modularity of 3-regular random graphs and random graphs with given degree sequences, 2020. arXiv:2007.15574v1.
- [12] R.J. Lipton and R.E. Tarjan. A separator theorem for planar graphs. *SIAM J. Discrete Math.*, 36(2):177–189, 1979.
- [13] W. Mader. Homomorphieeigenschaften und mittlere kantendichte von graphen. *Math. Ann.*, 174(4):265–268, 1967.
- [14] C. McDiarmid and B. Reed. On the maximum degree of a random planar graph. *Comb. Probab. Comput.*, 17(4):591–601, 2008.
- [15] C. McDiarmid and F. Skerman. Modularity of regular and treelike graphs. *J. Complex Networks*, 6(4):596–619, 2018.
- [16] C. McDiarmid and F. Skerman. Modularity of Erdős-Rényi random graphs. *Random Struct. Algor.*, 57(1):211–243, 2020.
- [17] G.L. Miller. Finding small simple cycle separators for 2-connected planar graphs. *J. Comput. Syst. Sci.*, 32(3):265–279, 1986.
- [18] J. Nešetřil and P. Ossona de Mendez. On nowhere dense graphs. *Eur. J. Combin.*, 32(4):600–617, 2011.
- [19] M.E.J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69(2):026113, 2004.
- [20] V.A. Traag, L. Waltman, and N.J. van Eck. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep. UK*, 9(5233), 2019.