Bering IV, E.A., Conant, G. and Gaster, J.
Osaka J. Math.
55 (2018), 795–808

ON THE COMPLEXITY OF FINITE SUBGRAPHS OF THE CURVE GRAPH

Edgar A. BERING IV, Gabriel CONANT and Jonah GASTER

(Received July 27, 2017)

Abstract

We say a graph has property $\mathcal{P}_{g,p}$ when it is an induced subgraph of the curve graph of a surface of genus g with p punctures. Two well-known graph invariants, the chromatic and clique numbers, can provide obstructions to $\mathcal{P}_{g,p}$. We introduce a new invariant of a graph, the *nested complexity length*, which provides a novel obstruction to $\mathcal{P}_{g,p}$. For the curve graph this invariant captures the topological complexity of the surface in graph-theoretic terms; indeed we show that its value is $6g-6+2p$, i.e. twice the size of a maximal multicurve on the surface. As a consequence we show that large 'half-graphs' do not have $\mathcal{P}_{g,p}$, and we deduce quantitatively that almost all finite graphs which pass the chromatic and clique tests do not have $\mathcal{P}_{g,p}$. We also reinterpret our obstruction in terms of the first-order theory of the curve graph, and in terms of RAAG subgroups of the mapping class group (following Kim and Koberda). Finally, we show that large complete multipartite graphs cannot have $\mathcal{P}_{g,p}$. This allows us to compute the upper density of the curve graph, and to conclude that clique size, chromatic number, and nested complexity length are not sufficient to determine $\mathcal{P}_{g,p}$.

1. Statement of results

Let S indicate a hyperbolizable surface of genus g with p punctures (i.e. $2g+p>2$). The *curve graph* of S, denoted $\mathcal{C}(S)$, is the infinite graph whose vertices are isotopy classes of simple closed curves on S and whose edges are given by pairs of curves that can be realized disjointly. Let $\mathcal{P}_{g,p}$ indicate the property that a graph is an induced subgraph of the curve graph $\mathcal{C}(S)$. We are concerned with the following motivating question:

Question 1. Which finite graphs have $\mathcal{P}_{g,p}$? When is $\mathcal{P}_{g,p}$ obstructed?

The low complexity cases $\mathcal{P}_{0,3}$, $\mathcal{P}_{0,4}$, and $\mathcal{P}_{1,1}$ are trivial, so we assume further that $3g+p\geq 5$. See §2 for details and complete definitions.

Property $\mathcal{P}_{g,p}$ has been considered in different guises in the literature [8, 6, 7, 14, 11, 12, 13]. It is not hard to see that every finite graph has $\mathcal{P}_{g,0}$ for large enough g, though it is remarkable that there exist finite graphs which do not have $\mathcal{P}_{0,p}$ for any p [8, §2].[1] Question 1 above is especially salient when g and p are fixed, and we adopt this point of view in everything that follows.

There are few known obstructions to a graph G having property $\mathcal{P}_{g,p}$. The simplest is the presence of a clique of G that is too large, as the size of a maximal complete subgraph

2010 Mathematics Subject Classification. 57M15, 57M60, 05C75.

[1]The authors thank Sang-Hyun Kim for bringing this reference to our attention.

of $\mathcal{C}(S)$ is $3g-3+p$. A more subtle obstruction follows from a surprising fact proved by Bestvina, Bromberg, and Fujiwara: the graph $\mathcal{C}(S)$ has finite chromatic number [4, 11].

We introduce an invariant of a graph G which we call the 'nested complexity length' $\mathrm{NCL}(G)$ that controls the topological complexity of any surface whose curve graph contains G as an induced subgraph (see §4 for a precise definition). The following is our main result, providing a new obstruction to $\mathcal{P}_{g,p}$. In fact, our calculation applies equally well to the *clique graph* $\mathcal{C}^{cl}(S)$ of $\mathcal{C}(S)$, whose vertices are multicurves and with edges for disjointness.

Theorem 2. *We have* $\mathrm{NCL}(\mathcal{C}(S)) = \mathrm{NCL}(\mathcal{C}^{cl}(S)) = 6g-6+2p$.

The nested complexity length of a graph is obtained via a supremum over all *nested complexity sequences*, and while this definition is useful in light of the theorem above, we know of no non-exhaustive algorithm that computes the nested complexity length of a finite graph. Thus it is natural to ask:

Question 3. What is an algorithm to compute the nested complexity length of a finite graph? How can one find effective upper bounds?

As a starting point toward this question, Proposition 23 gives an upper bound for the nested complexity length of a graph G which is exponential in the maximal size of a complete bipartite subgraph of G.

We describe several corollaries of Theorem 2 below. The first concerns *half-graphs*, a family of graphs that has attracted study in combinatorics and model theory.

Definition 4. Given an integer $n \geq 1$ and a graph G, we say that G is a *half-graph of height n* if there is a partition $\{a_1, \ldots, a_n\} \cup \{b_1, \ldots, b_n\}$ of the vertices of G such that the edge $e(a_i, b_j)$ occurs if and only if $i \geq j$. The unique *bipartite* half-graph of height n is denoted H_n.

In Example 25, we observe that if G is a half-graph of height n then $\mathrm{NCL}(G) \geq n$. Because NCL is monotone on induced subgraphs, the following is an immediate consequence of Theorem 2.

Corollary 5. *If G is a half-graph of height n, with $n > 6g-6+2p$, then G does not have $\mathcal{P}_{g,p}$.*

Since H_n is 2-colorable and triangle-free, Corollary 5 implies:

Corollary 6. *Chromatic number and maximal clique size are not sufficient to determine if a finite graph has $\mathcal{P}_{g,p}$.*

In fact, a more dramatic illustration of Corollary 6 can be made quantitatively. By definition, $\mathcal{P}_{g,p}$ is a *hereditary graph property* (i.e. closed under isomorphism and induced subgraphs). Asymptotic enumeration of hereditary graph properties has been studied by many authors, resulting in a fairly precise description of possible ranges for growth rates [3, Theorem 1]. Combining Corollary 5 with a result of Alon, Balogh, Bollobás, and Morris [1], we obtain an upper bound on the asymptotic enumeration of $\mathcal{P}_{g,p}$. The argument for the following is in §4. Given $n > 0$, let $\mathcal{P}_{g,p}(n)$ denote the class of graphs with vertex set $[n] = \{1, \ldots, n\}$ satisfying $\mathcal{P}_{g,p}$.

Corollary 7. *There is an $\epsilon > 0$ such that, for large n, $|\mathcal{P}_{g,p}(n)| \leq 2^{n^{2-\epsilon}}$.*

The set of graphs on $[n]$ satisfying the clique and chromatic tests for $\mathcal{P}_{g,p}$ includes all $(3g+3-p)$-colorable graphs on $[n]$, and thus this set has size $2^{\Theta(n^2)}$. In particular, the upper bound above cannot be obtained from the clique or chromatic number obstructions to $\mathcal{P}_{g,p}$. This also strengthens the statement of Corollary 6: among the graphs on $[n]$ satisfying the clique and chromatic tests, the probability of possessing $\mathcal{P}_{g,p}$ tends to 0 as $n \to \infty$.

For *monotone graph properties* (i.e. closed under isomorphism and subgraphs), even more is known concerning asymptotic structure [2]. However, we can use Corollary 5 to show that in most cases, $\mathcal{P}_{g,p}$ is not monotone.

Corollary 8. *If $3g + p \geq 6$ then $\mathcal{P}_{g,p}$ is not a monotone graph property.*

Proof. If $3g + p \geq 6$ then S contains a pair of disjoint incompressible subsurfaces that support essential nonperipheral simple closed curves. It follows that the complete bipartite graph $K_{n,n}$ has property $\mathcal{P}_{g,p}$ for all $n \in \mathbb{N}$. However, the half-graph H_n is a subgraph of $K_{n,n}$. □

It is also worth observing that, for $3g + p \geq 6$, this result thwarts the possibility of using the Robertson-Seymour Graph Minor Theorem [20] to characterize $\mathcal{P}_{g,p}$ by a finite list of forbidden minors. Of course this would also require $\mathcal{P}_{g,p}$ to be closed under edge contraction, which is already impossible just from the clique number restriction. Indeed, edge contraction of the graphs $K_{n,n}$ produces arbitrarily large complete graphs.

Remark 9. Neither of the exceptional cases $\mathcal{P}_{0,5}$ and $\mathcal{P}_{1,2}$ are closed under edge contraction, as each contains a five-cycle but no four-cycles by Lemma 30. However, it remains unclear whether $\mathcal{P}_{0,5}$ and $\mathcal{P}_{1,2}$ are monotone graph properties; while Theorem 2 applies, complete bipartite graphs do not possess $\mathcal{P}_{g,p}$ in these cases.

Following Kim and Koberda, Question 1 is closely related to the problem of which *right-angled Artin groups* (RAAGs) embed in the mapping class group $\mathrm{Mod}(S)$ of S [14, 11, 12, 13]. If the graph G has $\mathcal{P}_{g,p}$, then $A(G)$ is a RAAG subgroup of $\mathrm{Mod}(S)$ [14, Theorem 1.1]. The converse is false in general [13, Theorem 3], but a related statement holds: if the RAAG $A(G)$ embeds as a subgroup of $\mathrm{Mod}(S)$, the graph G is an induced subgraph of the clique graph $\mathcal{C}^{cl}(S)$ [11, Lemma 3.3].

By exploiting a construction by Erdős of graphs with arbitrarily large girth and chromatic number, Kim and Koberda produce 'not very complicated' (precisely, cohomological dimension two) RAAGs that do not embed in $\mathrm{Mod}(S)$ [11, Theorem 1.2]. Rephrasing Theorem 2 in this context, the graphs H_n provide such examples which are 'even less complicated'.

Corollary 10. *If $A(G)$ is a RAAG subgroup of* $\mathrm{Mod}(S)$*, then*

$$\mathrm{NCL}(G) \leq 6g - 6 + 2p\,.$$

Moreover, for any g and p there exist bipartite graphs G so that $A(G)$ does not embed in $\mathrm{Mod}(S)$.

The nested complexity length of a graph G is closely related to the *centralizer dimension* of $A(G)$, i.e. the longest chain of nontrivially nested centralizers in the group (this algebraic invariant is discussed elsewhere in the literature [18]). In particular, it is straightforward from the definitions that the centralizer dimension of $A(G)$ is at least $\mathrm{NCL}(G)$. The possible centralizers of an element of a RAAG have been classified by Servatius [21], and the characterization there would seem to suggest that equality does not hold in general. Of course, this is impossible to check in the absence of a method to compute nested complexity length, and Question 3 arises naturally.

Our next corollary concerns the model theoretic behavior of $C(S)$. We focus on *stability*, one of the most important and well-developed notions in modern model theory. Given a first-order structure, stability of its theory implies an abstract notion of independence and dimension for definable sets in that structure (see Pillay [19] for details). Stability can also be treated as a combinatorial property obtained from half-graphs. Given an integer $k \geq 1$, we say that a graph G is *k-edge-stable* if it does not contain any half-graph of height $n \geq k$ as an induced subgraph. We can thus rephrase Corollary 5.

Corollary 11. *$C(S)$ is k-edge-stable for $k = 6g - 5 + 2p$.*

When considering the first-order theory of $C(S)$ in the language of graphs, this corollary implies that the edge formula (and thus any quantifier-free formula) is stable in the model theoretic sense [23, Theorem 8.2.3]. Whether arbitrary formulas are stable remains an intriguing open question, and would likely require some understanding of quantifier elimination for the theory of $C(S)$ in some suitable expansion of the graph language. This aspect of the nature of curve graphs remains unexplored, and stability is only one among a host of natural questions about their first-order theories to pursue.

On the other hand, edge-stability of $C(S)$ alone has strong consequences for the structure of large finite subgraphs of $C(S)$, via Szemerédi's regularity lemma [15, 22]. In particular, Malliaris and Shelah show that if G is a k-edge-stable graph, then the regularity lemma can be strengthened so that in Szemerédi partitions of large induced subgraphs of G, the bound on the number of pieces is significantly improved, there is no need for irregular pairs, and the density between each pair of pieces is within ϵ of 0 or 1 [16]. Thus a consequence of our work is that the class of graphs with $\mathcal{P}_{g,p}$ enjoys this stronger form of Szemerédi regularity.

Next we consider an explicit family of multipartite graphs.

DEFINITION 12. Given integers $r, t > 0$, let $K_r(t)$ denote the complete r-partite graph in which each piece of the partition has size t.

In Example 25, we show $\mathrm{NCL}(K_r(2)) = 2r$. Combined with the Erdős-Stone Theorem, this implies a general relationship (Proposition 28) between nested complexity length and the *upper density* $\delta(G)$ of an infinite graph G, i.e. the supremum over real numbers t such that G contains arbitrarily large finite subgraphs of edge density t. In the case of the curve graph $C(S)$, we show in Lemma 30 that $K_r(t)$ is obstructed from having $\mathcal{P}_{g,p}$ for large r. From this we obtain the following exact calculation of the upper density of the curve graph, which is proved in §6.

Theorem 13. *The upper density $\delta(C(S))$ is equal to* $1 - \dfrac{1}{g + \lfloor \frac{g+p}{2} \rfloor - 1}$.

REMARK 14. The exceptional cases $(g, p) = (0, 5)$ and $(1, 2)$ are again remarkable in that they imply $\delta(C(S)) = 0$ in Theorem 13. Thus any family of graphs $\{G_n\}$ with $\mathcal{P}_{g,p}$ and

$|V(G_n)| \to \infty$ in these exceptional cases must satisfy $|E(G_n)| = o(|V(G_n)|^2)$.

Question 15. Given $(g, p) = (0, 5)$ or $(1, 2)$, does there exist $\epsilon > 0$ such that, for any family of graphs $\{G_n\}$ with $\mathcal{P}_{g,p}$ and $|V(G_n)| \to \infty$ one has $|E(G_n)| = O(|V(G_n)|^{2-\epsilon})$?

Finally, we use the analysis of $\mathcal{P}_{g,p}$ for $K_r(t)$ to extend Corollary 6.

Corollary 16. *Chromatic number, maximal clique size, and nested complexity length are not sufficient to determine if a finite graph has $\mathcal{P}_{g,p}$.*

Proof. For $r = g + \lfloor \frac{g+p}{2} \rfloor$, the graph $K_r(2)$ does not have $\mathcal{P}_{g,p}$ by Lemma 30, but passes the clique, coloring, and NCL tests. Indeed the chromatic number of $C(S)$ is at least its clique number $3g - 3 + p$, which is greater than r (the clique and chromatic number of $K_r(2)$). Moreover, by Example 25 and Theorem 2, we have $\mathrm{NCL}(K_r(2)) = 2r \leq \mathrm{NCL}(C(S))$. □

2. Notation and conventions

We briefly list definitions and notation relevant in this paper, with the notable exception of nested complexity length (found in §4). For background and context see Farb and Margalit [10].

A *graph* G consists of a set of vertices $V(G)$ and an edge set $E(G)$ which is a subset of unordered distinct pairs from $V(G)$; we denote the edge between vertices v and w by $e(v, w) \in E(G)$. A subgraph $H \subset G$ is *induced* if $v, w \in H$ and $e(v, w) \in E(G)$ implies that $e(v, w) \in E(H)$. The *closed neighborhood* of a vertex $v \in V(G)$ is the set of vertices

$$N[v] = \{v\} \cup \{u \in V(G) : e(u, v) \in E(G)\}.$$

Given a graph G the right-angled Artin group corresponding to G, denoted $A(G)$, is defined by the following group presentation: the generators of $A(G)$ are given by $V(G)$ and there is a commutation relation $vw = wv$ for every edge $e(v, w) \in E(G)$.

Recall that we are concerned with the hyperbolizable surface S of genus g with p punctures, so we assume that $2g + p > 2$. The *mapping class group* of S, denoted $\mathrm{Mod}(S)$, is the group $\pi_0(\mathrm{Homeo}^+(S))$.

By a *curve* we mean the isotopy class of an essential nonperipheral embedded loop on S, and we refer to the union of curves which can be made simultaneously disjoint as a *multicurve*. The *curve graph* $C(S)$ is the graph consisting of a vertex for each curve on S, and so that a pair of curves span an edge when the curves can be realized disjointly. The *clique graph* $C^{cl}(S)$ of the curve graph is the graph obtained as follows: The vertices of $C^{cl}(S)$ correspond to cliques of $C(S)$ (i.e. multicurves), and two cliques are joined by an edge when they are simultaneously contained in a maximal clique (i.e can be realized disjointly). The curve graph is the subgraph of the clique graph induced by the one-cliques.

We strengthen the assumption on g and p above to $3g + p \geq 5$. When $(g, p) = (0, 3)$ the curve graph as defined above has no vertices. When $(g, p) = (0, 4)$ or $(1, 1)$, the curve graph has no edges, and the matter of deciding if a graph has $\mathcal{P}_{g,p}$ in these cases is trivial. The common alteration of the definition of these curve graphs yields the *Farey graph*. We make no comment on induced subgraphs of the Farey graph, though a comprehensive classification can be made.

Whenever we refer to a *subsurface* $\Sigma \subset S$ we make the standing assumption that Σ is a disjoint union of closed incompressible homotopically distinct subsurfaces with boundary of S, i.e. the inclusion $\Sigma \subset S$ induces an injection on the fundamental groups of components of Σ. We write $[\Sigma]$ for the isotopy class of Σ.

Given a subsurface $\Sigma \subset S$, we say that a curve γ is *supported on* Σ if it has a representative which is either contained in an annular component of Σ, or is a nonperipheral curve in a (necessarily non-annular) component. The curve γ is *disjoint from* Σ if it is has a representative disjoint from Σ. We say that γ is *transverse* to Σ, written $\gamma \pitchfork \Sigma$, if it is not supported on Σ and not disjoint from Σ. A multicurve γ is supported on (resp. disjoint from) Σ if each of its components is supported on (resp. disjoint from) Σ, and γ is transverse to Σ if at least one of its components is. Each of these above definitions applies directly for curves and isotopy classes of subsurfaces.

We note that the definitions above may be nonstandard, as they are made with our specific application in mind. For example, in our terminology the core of an annular component of Σ is both supported on and disjoint from the subsurface Σ.

Given a pair of subsurfaces Σ_1 and Σ_2, we write $[\Sigma_1] \leq [\Sigma_2]$ when every curve supported on Σ_1 is supported on Σ_2. If $[\Sigma_1] \leq [\Sigma_2]$ and $[\Sigma_1] \neq [\Sigma_2]$ (denoted $[\Sigma_1] < [\Sigma_2]$), we say that $[\Sigma_1]$ and $[\Sigma_2]$ are *nontrivially nested*.

Given a collection of curves C, we let $[\Sigma(C)]$ indicate the isotopy class of the minimal subsurface of S, with respect to the partial ordering just defined, that supports the curves in C. Concretely, a representative of $[\Sigma(C)]$ can be obtained by taking the union of regular neighborhoods of the curves in C (for suitably small regular neighborhoods that depend on the realizations of the curves in C) and filling in contractible complementary components of the result with disks.

A set of curves C supported on a subsurface Σ *fills* the subsurface if $[\Sigma(C)] = [\Sigma]$. Concretely, C fills Σ if the core of each annular component of Σ is in C, and if the complement in Σ of a realization of the remaining curves consists of peripheral annuli and disks. Given a subsurface Σ, we write $\xi(\Sigma)$ for the number of components of a maximal multicurve supported on Σ.

3. Topological complexity of subsurfaces

Since the ambient surface S is fixed throughout, it should not be surprising that any nontrivially nested chain of subsurfaces of S has bounded length. We make this explicit below, using $\xi(S)$ to keep track of 'how much' of the surface has been captured by subsurfaces from the chain. In fact, this is not quite enough to notice nontrivial nesting of subsurfaces, as annuli can introduce complications. We keep track of this carefully in Lemma 20 below. Throughout this section, Σ_1 and Σ_2 refer to a pair of subsurfaces of S.

Lemma 17. *If* $[\Sigma_1] \leq [\Sigma_2]$, *then* Σ_1 *is isotopic to a subsurface of* Σ_2.

The reader is cautioned that the converse is false: consider a one-holed torus subsurface T and let T' indicate the disjoint union of T with an annulus isotopic to ∂T. Though T' is isotopic to a subsurface of T, the curve ∂T is supported on T' but not on T. Thus $[T'] \nleq [T]$.

Proof. Choose a collection of curves C filling Σ_1. Since $[\Sigma_1] \leq [\Sigma_2]$, each curve in C is supported on Σ_2. It follows that there are small enough regular neighborhoods of representatives of the curves in C that are contained in Σ_2. Filling in contractible components of the complement produces a subsurface isotopic to Σ_1 inside Σ_2. □

Lemma 18. *If* $[\Sigma_1] \leq [\Sigma_2]$ *and* $[\Sigma_2] \leq [\Sigma_1]$, *then* $[\Sigma_1] = [\Sigma_2]$.

Proof. Fix representatives $\Sigma_1 \subset \Sigma_2$, using the previous lemma. Let C be a collection of curves that fill Σ_2. Since $[\Sigma_2] \leq [\Sigma_1]$ the curves in C can be realized on Σ_1 so that they have regular neighborhoods contained in Σ_1. Their complement in Σ_2 is a collection of disks and peripheral annuli, so their complement in Σ_1 must also be a collection of disks and peripheral annuli. Hence C fills Σ_1 and $[\Sigma_1] = [\Sigma(C)] = [\Sigma_2]$. □

Lemma 19. *Suppose that* $[\Sigma_1] < [\Sigma_2]$. *Then one of the following holds:*

(1) *we have* $\xi(\Sigma_1) < \xi(\Sigma_2)$, *or*

(2) *there exists a curve* α *which is the core of an annular component of* Σ_1 *but not disjoint from* Σ_2.

Proof. By Lemma 18 we have $[\Sigma_2] \not\leq [\Sigma_1]$. Thus there is a curve γ supported on Σ_2 that is not supported on Σ_1. If γ were the core of an annular component of Σ_2, then $\xi(\Sigma_1) < \xi(\Sigma_2)$ would be immediate. Likewise if γ were disjoint from Σ_1 then $\xi(\Sigma_1) < \xi(\Sigma_2)$ again.

We are left with the case that γ is transverse to Σ_1. Choose a boundary component α of Σ_1 intersected essentially by γ. Evidently, this curve must be supported on Σ_2. Since α is supported on Σ_2, either one has $\xi(\Sigma_1) < \xi(\Sigma_2)$ or α is supported as well on Σ_1. In this case, α is the core of an annular component of Σ_1 intersected essentially by γ. If α could be made disjoint from Σ_2 then its intersection with γ would be inessential, so we are done. □

Lemma 20. *Suppose that* $\emptyset \neq [\Sigma_1] < [\Sigma_2] < \ldots < [\Sigma_n]$ *is a nontrivially nested chain of subsurfaces of* S. *Then* $n \leq 6g - 6 + 2p$.

Proof. Choose k with $1 \leq k \leq n-1$, and let $c_k = \xi(\Sigma_k)$. We construct, inductively on k, a (possibly empty) multicurve ω. Evidently, we have $[\Sigma_1] < [\Sigma_2]$, so that by Lemma 19 either $c_1 < c_2$, or there is a curve α_1 isotopic to the core of an annular component of Σ_1 but not disjoint from Σ_2. In the second case, add α_1 to ω. Note that in the latter case there is a representative for α_1 which is contained in Σ_1.

We continue inductively. Since we have $[\Sigma_k] < [\Sigma_{k+1}]$, Lemma 19 guarantees that either $c_k < c_{k+1}$, or there is a curve α_k isotopic to the core on an annular component of Σ_k but not disjoint from Σ_{k+1}. Suppose that, for some $i < k$, the curve α_i is another component of ω supported on Σ_i. Since α_k has a representative disjoint from Σ_k, and $[\Sigma_i] < [\Sigma_k]$, α_k has a representative disjoint from Σ_i. Moreover, because α_i cannot be made disjoint from Σ_{i+1} but α_k can be (since $[\Sigma_{i+1}] \leq [\Sigma_k]$), the curves α_i and α_k are not isotopic. It follows that the curve α_k may be added to the multicurve ω so that ω remains a collection of disjoint curves, and so that its number of components increases by one.

At each step of the chain $[\Sigma_1] < [\Sigma_2] < \ldots < [\Sigma_n]$, either c_k strictly increases, or ω gains another component. Since $[\Sigma_1] \neq \emptyset$ and $[\Sigma_n] \leq [S]$, we have $1 \leq c_1$ and $c_n \leq \xi(S)$. The number of components of ω is also at most $\xi(S)$, so we conclude that

$$n \leq 2\xi(S) = 6g - 6 + 2p.$$

□

Finally, we will make use of a straightforward certificate that $[\Sigma_1] < [\Sigma_2]$.

Lemma 21. *Suppose that* $[\Sigma_1] \leq [\Sigma_2]$. *If α is a curve disjoint from Σ_1, β is a curve supported on Σ_2, and α and β intersect essentially, then* $[\Sigma_1] < [\Sigma_2]$.

Proof. Because α is disjoint from Σ_1, it has a representative disjoint from Σ_1. If β were supported on Σ_1, it would have a representative contained in Σ_1, contradicting the assumption that α and β intersect essentially. Thus β is a curve supported on Σ_2 but not supported on Σ_1.

□

4. The nested complexity length of a graph

The topological hypotheses of Lemma 21 suggest a natural combinatorial parallel, which we capture in the definition of nested complexity length.

Definition 22. Let G be a graph.

(1) Given $b_1, \ldots, b_n \in V(G)$, we say that $(b_1, \ldots, b_n)$ is a *nested complexity sequence* for G if for each $1 \leq k \leq n-1$ there is $a_k \in V(G)$ such that $b_1, \ldots, b_k \in N[a_k] \not\ni b_{k+1}$.

(2) The *nested complexity length of* G, denoted $\mathrm{NCL}(G)$, is given by

$$\mathrm{NCL}(G) := \sup\{n \mid (b_1, \ldots, b_n) \text{ is a nested complexity sequence for } G\}.$$

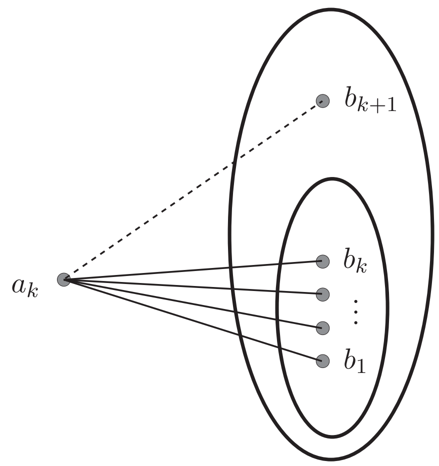

Fig. 1. A graph acquires some nested complexity. Dotted lines indicate edges that are necessarily absent.

Note that in the definition of a nested complexity sequence, a_k may be equal to b_i, for $1 \leq i \leq k$; see Figure 1 for a schematic in which this is not the case. To highlight this subtlety, and as a first step toward an answer to Question 3, we prove an upper bound for $\mathrm{NCL}(G)$ in terms of a maximal complete bipartite subgraph of G.

Proposition 23. *Let G be a graph. Fix $m, n > 0$ and suppose G does not contain a subgraph isomorphic to $K_{m,n}$. Then* $\mathrm{NCL}(G) \leq 2^{m+n+1} - 2$.

Proof. Given $k > 0$, let $s_k = 2^{k+1} - 2$. Note that $s_k = 2(s_{k-1} + 1)$. Set $N = 2^{m+n+1} - 1 = s_{m+n} + 1$. For a contradiction, suppose $(b_1, \ldots, b_N)$ is a nested complexity sequence for G, witnessed by $a_1, \ldots, a_{N-1}$.

We inductively produce values $j_1 < \ldots < j_{m+n} < N$ below, such that for all $1 \leq r \leq m+n$, $j_r \leq s_r$ and $a_{j_r} \neq b_i$ for all $i \leq j_r$. With these indices chosen, the set $\{b_{j_1}, \ldots, b_{j_m}\} \cup \{a_{j_{m+1}} \ldots, a_{j_{m+n}}\} \subseteq V(G)$ produces a (not necessarily induced) subgraph of G isomorphic to $K_{m,n}$, a contradiction.

If $a_1 \neq b_1$ then we set $j_1 = 1$. Otherwise, if $a_1 = b_1$ then b_1, b_2 are independent vertices in $N[a_2]$. So $a_2 \notin \{b_1, b_2\}$, and we set $j_1 = 2$.

Fix $1 \leq r < m + n$, and suppose we have defined j_r as above. If $a_{j_r+1} \neq b_i$ for all $i \leq j_r + 1$ then we let $j_{r+1} = j_r + 1$. Otherwise, let $k = j_r + 1$ and suppose $a_k = b_{i_k}$ for some $i_k \leq k$. We will find t such that $k < t \leq 2k$ and $a_t \neq b_i$ for any $i \leq t$. By induction, $2k \leq 2(s_r + 1) = s_{r+1} < N$, and so setting $j_{r+1} = t$ finishes the inductive step of the construction.

Suppose no such t exists. Then for all $k \leq t \leq 2k$, we have $a_t = b_{i_t}$ for some $i_t \leq t$. Fix $1 \leq s \leq k$. For any $0 \leq u < s$, we have $a_{k+u} = b_{j_{k+u}}$, and so $b_{k+u+1} \notin N[b_{j_{k+u}}]$. Therefore, for any $0 \leq u < s$, b_{k+u+1} and $b_{j_{k+u}}$ are independent vertices in $N[a_{k+s}]$, which means $a_{k+s} \notin \{b_{k+u+1}, b_{j_{k+u}}\}$. In other words, we have shown that for all $1 \leq s \leq k$,

$$j_{k+s} \notin \{k + 1, \ldots, k + s, j_k, j_{k+1}, \ldots, j_{k+s-1}\}.$$

It follows that $j_k, j_{k+1}, \ldots, j_{2k}$ are $k + 1$ distinct elements of $\{1, \ldots, k\}$. □

We make note of two useful properties of NCL that follow immediately from the definitions.

Proposition 24.

(1) *If H is an induced subgraph of G, then* $\mathrm{NCL}(H) \leq \mathrm{NCL}(G)$.
(2) *We have* $\mathrm{NCL}(G) \leq |V(G)|$.

We also give the following examples, which are heavily exploited in the results outlined in §1.

Example 25.

(1) Let G be a half-graph of height n. We may partition the vertices as $V(G) = \{a_1, \ldots, a_n\} \cup \{b_1, \ldots, b_n\}$, where $e(a_i, b_j) \in E(G)$ if and only if $i \geq j$. Then $(b_1, \ldots, b_n)$ is a nested complexity sequence for G, witnessed by $a_1, \ldots, a_{n-1}$. Therefore $\mathrm{NCL}(G) \geq n$.
(2) Let G be the multipartite graph $K_r(2)$. Let $V(G) = \{b_1, \ldots, b_{2r}\}$ where $e(b_{2k-1}, b_{2k}) \notin E(G)$ for $1 \leq k \leq r$. Set $a_1 = b_1$. For $2 \leq k \leq r$ let $a_{2k-2} = b_{2k}$ and $a_{2k-1} = b_{2k-1}$. Then $(b_1, \ldots, b_{2r})$ is a nested complexity sequence for G, witnessed by $a_1, \ldots, a_{2r-1}$. Combined with Proposition 24(2), we have $\mathrm{NCL}(G) = 2r$.

With the first example in hand, we immediately derive Corollary 5 from Theorem 2. We can also give the proof of Corollary 7.

Proof of Corollary 7. Given $k > 0$, let $\mathcal{U}(k)$ denote the class of graphs G for which there is a partition $V(G) = \{a_i : 1 \leq i \leq k\} \cup \{b_J : J \subseteq \{1, \ldots, k\}\}$ such that $e(a_i, b_J)$ holds if and only if $i \in J$. Any graph in $\mathcal{U}(k)$ contains an induced half-graph of height k (take $\{a_i : 1 \leq i \leq k\} \cup \{b_{J_i} : 1 \leq i \leq k\}$, where $J_i = \{i, \ldots, k\}$). By Corollary 5, every graph with $\mathcal{P}_{g,p}$ omits the class $\mathcal{U}(k)$ for $k > 6g - 6 + 2p$. The result now follows immediately from Theorem 2 of Alon, et al. [1]. □

5. Proof of Theorem 2

Recall that $[\Sigma(C)]$ refers to the minimal isotopy class, with respect to inclusion, of a subsurface of S that supports the curves in C.

Proof of Theorem 2. As $\mathcal{C}(S)$ is an induced subgraph of $\mathcal{C}^{cl}(S)$, Proposition 24(1) ensures that $\mathrm{NCL}(\mathcal{C}(S)) \leq \mathrm{NCL}(\mathcal{C}^{cl}(S))$. We proceed by showing that $6g-6+2p$ is simultaneously a lower bound for $\mathrm{NCL}(\mathcal{C}(S))$ and an upper bound for $\mathrm{NCL}(\mathcal{C}^{cl}(S))$.

For the lower bound, choose a maximal multicurve $\{\gamma_1, \ldots, \gamma_{3g-3+p}\}$. For each curve γ_i, choose a *transversal* η_i, i.e. a curve intersecting γ_i essentially and disjoint from γ_j for $j \neq i$. That such collections of curves exist is routine (e.g. a 'complete clean marking' in the language of Masur and Minsky [17, §2.5] is an even more restrictive example, see Figure 2).

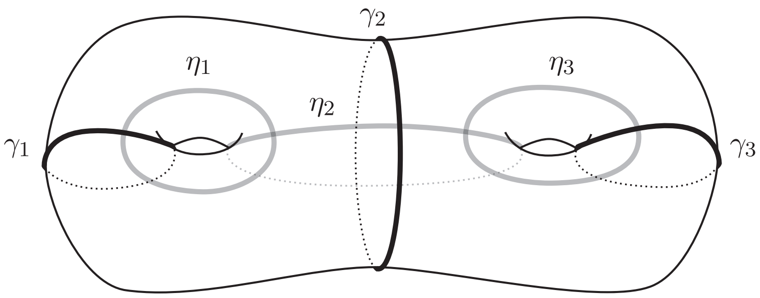

Fig. 2. A bold maximal multicurve $\{\gamma_1, \gamma_2, \gamma_3\}$ and a lighter set of transversals $\{\eta_1, \eta_2, \eta_3\}$.

For each $1 \leq i \leq 6g-6+2p$, let the curves α_i and β_i be given by

$$\beta_i = \begin{cases} \gamma_i & 1 \leq i \leq 3g-3+p \\ \eta_{i-(3g-3+p)} & 3g-3+p < i \leq 6g-6+2p\,, \end{cases}$$

$$\alpha_i = \begin{cases} \eta_{i+1} & 1 \leq i \leq 3g-4+p \\ \gamma_{i-(3g-4+p)} & 3g-4+p < i \leq 6g-7+2p\,. \end{cases}$$

It is straightforward to check that $(\beta_1, \ldots, \beta_{6g-6+2p})$ forms a nested complexity sequence for $\mathcal{C}(S)$, with witnessing curves $(\alpha_1, \ldots, \alpha_{6g-7+2p})$.

Towards the upper bound, suppose $(\beta_1, \ldots, \beta_n)$ is a nested complexity sequence for $\mathcal{C}^{cl}(S)$, so that there exists a vertex α_k in $\mathcal{C}^{cl}(S)$ with $\beta_1, \ldots, \beta_k \in N[\alpha_k] \not\ni \beta_{k+1}$ for each $1 \leq k \leq n-1$ as in Definition 22. For $1 \leq k \leq n$, let $B_k = \{\beta_1, \ldots, \beta_k\}$, and let $[\Sigma_k] = [\Sigma(B_k)]$.

Choose $k \leq n-1$. Because $B_k \subset B_{k+1}$, we have $[\Sigma_k] \leq [\Sigma_{k+1}]$; as $B_k \subset N[\alpha]$, α_k is disjoint from Σ_k; and since $\beta_{k+1} \in B_{k+1}$ we have that β_{k+1} is supported on Σ_{k+1}. Because $\beta_{k+1} \notin N[\alpha_k]$, α_k and β_{k+1} are independent vertices of $\mathcal{C}^{cl}(S)$, and so there are components of the multicurves α_k and β_{k+1} that intersect essentially. Lemma 21 applies, so that $[\Sigma_k] < [\Sigma_{k+1}]$. Thus $\emptyset \neq [\Sigma_1] < [\Sigma_2] < \ldots < [\Sigma_n]$ is a nontrivially nested chain of subsurfaces, and Lemma 20 implies that $n \leq 6g - 6 + 2p$. □

Remark 26. The upper bound can be strengthened for $\mathcal{C}(S)$ under the additional assumption that the antigraphs of the subgraphs induced on the vertices $\{b_1, \ldots, b_k\}$ are connected for each k (indeed, in this case the negative Euler characteristic of Σ_k is strictly less than that of Σ_{k+1}). In particular, if $n \geq 2g + p$ then the bipartite half-graph H_n does not have $\mathcal{P}_{g,p}$.

6. Obstructing $K_r(t)$ and the upper density of the curve graph

We turn to $K_r(t)$ and the upper density of curve graphs.

Definition 27. Let G be a graph.

(1) If $|G| = n > 1$ the *density of* G is

$$\delta(G) = \frac{|E(G)|}{\binom{n}{2}} .$$

(2) If G is infinite the *upper density of* G is

$$\delta(G) = \limsup_{H \subseteq G,\ |H| \to \infty} \delta(H) .$$

In other words, given an infinite graph G, $\delta(G)$ is the supremum over all $\alpha \in (0, 1]$ such that G contains arbitrarily large finite subgraphs of density at least α. Given a fixed $r > 0$, it is easy to verify $\lim_{t\to\infty} \delta(K_r(t)) = 1 - \frac{1}{r}$. It is a consequence of a remarkable theorem of Erdős and Stone that the graphs $K_r(t)$ witness the densities of all infinite graphs. We record this precisely in language most relevant for our application ([5, Ch. IV], [9]).

Erdős-Stone Theorem. *Fix $r > 0$. For any infinite graph G, if $\delta(G) > 1 - \frac{1}{r}$ then $K_{r+1}(t)$ is a subgraph of G for arbitrary t.*

Using this theorem, we obtain the following relationship between density and nested complexity length.

Proposition 28. *Let G be an infinite graph. If $\delta(G) < 1$ then*

$$\mathrm{NCL}(G) \geq \frac{2}{1 - \delta(G)}.$$

Proof. If this inequality fails then we have $\delta(G) > 1 - \frac{2}{\mathrm{NCL}(G)}$. By the Erdős-Stone Theorem, there is $r > \frac{1}{2}\mathrm{NCL}(G)$ such that $K_r(t)$ is a subgraph of G for arbitrarily large t. Let w be the size of the largest finite clique in G, which exists since $\delta(G) < 1$. If we consider a copy of $K_r(w + 1)$ in G, it follows that each piece of the partition contains a pair of independent vertices. Therefore $K_r(2)$ is an *induced* subgraph of G. By Proposition 24 and Example 25, $\mathrm{NCL}(G) \geq 2r$, which contradicts the choice of r. □

Note that if G is complete then $\mathrm{NCL}(G) = 1$. Therefore an infinite graph with upper density 1 need not have large nested complexity length.

Finally, we obstruct the multipartite graphs $K_r(t)$ from having $\mathcal{P}_{g,p}$ for large enough r and $t \geq 2$, and employ this fact in the proof of Theorem 13.

Lemma 29. *The maximum number of pairwise disjoint subsurfaces of S which are not annuli or pairs of pants is $g + \lfloor \frac{g+p}{2} \rfloor - 1$.*

Proof. It suffices to consider a pairwise disjoint sequence of subsurfaces in which each component is a four-holed sphere or a one-holed torus; any more complex subsurface can be cut further without decreasing the number of non-annular and non-pair of pants subsurfaces. Suppose there are T one-holed tori and F four-holed spheres. The dimension of the homology of S requires $0 \leq T \leq g$. Additivity of the Euler characteristic under disjoint union implies that $0 \leq T + 2F \leq 2g + p - 2$. Maximizing $T + F$ on this polygon is routine, and the solution is $T = g$ and $F = \lfloor \frac{g+p}{2} \rfloor - 1$. Moreover, it is straightforward to construct such a collection of subsurfaces of S. □

This is enough to obstruct $K_r(t)$ for large r.

Lemma 30. *For $t > 1$, $K_r(t)$ has $\mathcal{P}_{g,p}$ if and only if $r \leq g + \lfloor \frac{g+p}{2} \rfloor - 1$.*

Proof. Let $\ell = g + \lfloor \frac{g+p}{2} \rfloor - 1$. First, suppose $\Sigma_1, \ldots, \Sigma_\ell$ is a sequence of pairwise disjoint subsurfaces consisting of tori and four-holed spheres, as guaranteed by Lemma 29. The curves supported on the Σ_i induce $K_\ell(\infty)$ as a subgraph of $\mathcal{C}(S)$. Hence $K_r(t)$ is an induced subgraph of $\mathcal{C}(S)$ for all t and $r \leq \ell$. Conversely, suppose towards a contradiction that $K_{\ell+1}(2)$ is an induced subgraph of $\mathcal{C}(S)$, and let $V_1, \ldots, V_{\ell+1}$ be the partition of its vertices. For each $i \neq j$, both curves in V_i are disjoint from the curves in V_j, so the subsurfaces $\Sigma(V_i)$ and $\Sigma(V_j)$ are disjoint. Moreover, since the curves in V_i intersect, $\Sigma(V_i)$ is a connected surface that is not an annulus or a pair of pants. Thus $\Sigma(V_1), \ldots, \Sigma(V_{\ell+1})$ is a sequence of disjoint non-annular and non-pair of pants subsurfaces, contradicting Lemma 29. □

We can now prove Theorem 13.

Proof of Theorem 13. Let $\ell = g + \lfloor \frac{g+p}{2} \rfloor - 1$ again. By Lemma 30, $K_\ell(t)$ has $\mathcal{P}_{g,p}$ for all t, so that $\delta(\mathcal{C}(S)) \geq 1 - \frac{1}{\ell}$. If $\delta(\mathcal{C}(S)) > 1 - \frac{1}{\ell}$ then, as in the proof of Proposition 28, the Erdős-Stone Theorem and finite clique number of $\mathcal{C}(S)$ together imply that $K_{\ell+1}(2)$ is an *induced* subgraph of $\mathcal{C}(S)$, violating Lemma 30. □

ACKNOWLEDGEMENTS. The authors thank Tarik Aougab, Ian Biringer, Josh Greene, Sang-Hyun Kim, and Thomas Koberda for helpful comments and discussions.

References

[1] N. Alon, J. Balogh, B. Bollobás and R. Morris: *The structure of almost all graphs in a hereditary property*, J. Combin. Theory Ser. B **101** (2011), 85–110.

[2] J. Balogh, B. Bollobás and M. Simonovits: *The number of graphs without forbidden subgraphs*, J. Combin. Theory Ser. B **91** (2004), 1–24.

[3] J. Balogh, B. Bollobás and D. Weinreich: *The penultimate rate of growth for graph properties*, European J. Combin. **22** (2001), 277–289.

[4] M. Bestvina, K. Bromberg and K. Fujiwara: *Constructing group actions on quasi-trees and applications to mapping class groups*, Publ. Math. Inst. Hautes Études Sci. **122** (2015), 1–64.

[5] B. Bollobás: Extremal graph theory, London Mathematical Society Monographs **11**, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London-New York, 1978.

[6] J. Crisp and B. Wiest: *Embeddings of graph braid and surface groups in right-angled Artin groups and braid groups*, Algebr. Geom. Topol. **4** (2004), 439–472.

[7] J. Crisp and B. Wiest: *Quasi-isometrically embedded subgroups of braid and diffeomorphism groups*, Trans. Amer. Math. Soc. **359** (2007), 5485–5503.

[8] G. Ehrlich, S. Even and R.E. Tarjan: *Intersection graphs of curves in the plane*, J. Combinatorial Theory Ser. B **21** (1976), 8–20.

[9] P. Erdös and A.H. Stone: *On the structure of linear graphs*, Bull. Amer. Math. Soc. **52** (1946), 1087–1091.

[10] B. Farb and D. Margalit: A primer on mapping class groups, Princeton Mathematical Series **49**, Princeton University Press, Princeton, NJ, 2012.

[11] S.-H. Kim and T. Koberda: *An obstruction to embedding right-angled Artin groups in mapping class groups*, Int. Math. Res. Not. IMRN **2014**, 3912–3918.

[12] S.-H. Kim and T. Koberda: *Anti-trees and right-angled Artin subgroups of braid groups*, Geom. Topol. **19** (2015), 3289–3306.

[13] S.-H. Kim and T. Koberda: *Right-angled Artin groups and finite subgraphs of curve graphs*, Osaka J. Math. **53** (2016), 705–716.

[14] T. Koberda: *Right-angled Artin groups and a generalized isomorphism problem for finitely generated subgroups of mapping class groups*, Geom. Funct. Anal. **22** (2012), 1541–1590.

[15] J. Komlós and M. Simonovits: *Szemerédi's regularity lemma and its applications in graph theory*, In Combinatorics, Paul Erdös is eighty, Vol. 2 (Keszthely, 1993), Bolyai Soc. Math. Stud. **2**, János Bolyai Math. Soc., Budapest, 1996, 295–352.

[16] M. Malliaris and S. Shelah: *Regularity lemmas for stable graphs*, Trans. Amer. Math. Soc. **366** (2014), 1551–1585.

[17] H.A. Masur and Y.N. Minsky: *Geometry of the complex of curves. II, Hierarchical structure*, Geom. Funct. Anal. **10** (2000), 902–974.

[18] A. Myasnikov and P. Shumyatsky: *Discriminating groups and c-dimension*, J. Group Theory **7** (2004), 135–142.

[19] A. Pillay: An introduction to stability theory, Oxford Logic Guides **8**, The Clarendon Press Oxford University Press, New York, 1983.

[20] N. Robertson and P.D. Seymour: *Graph minors. XX. Wagner's conjecture*, J. Combin. Theory Ser. B **92** (2004), 325–357.

[21] H. Servatius: *Automorphisms of graph groups*, J. Algebra **126** (1989), 34–60.

[22] E. Szemerédi: *On sets of integers containing no k elements in arithmetic progression*, Collection of articles in memory of Juriĭ Vladimirovič Linnik, Acta Arith. **27** (1975), 199–245.

[23] K. Tent and M. Ziegler: A course in model theory, Lecture Notes in Logic **40**, Association for Symbolic Logic, La Jolla, CA, Cambridge University Press, Cambridge, 2012.

Edgar A. Bering IV
Department of Mathematics
Statistics, and Computer Science
University of Illinois at Chicago
Chicago, IL 60607
USA
e-mail: eberin2@uic.edu

Gabriel Conant
Department of Mathematics
University of Notre Dame
Notre Dame, IN 46556
USA
e-mail: gconant@nd.edu

Jonah Gaster
Department of Mathematics and Statistics
McGill University
Montreal, QC H3A 0B9
Canada
e-mail: jonah.gaster@mcgill.ca