

# Parameterized and Approximation Algorithms for the Load Coloring Problem

Florian Barbero, Gregory Gutin, Mark Jones, and Bin Sheng

Royal Holloway, University of London  
Egham, Surrey TW20 0EX, UK

---

## Abstract

Let  $c, k$  be two positive integers. Given a graph  $G = (V, E)$ , the  $c$ -LOAD COLORING problem asks whether there is a  $c$ -coloring  $\varphi : V \rightarrow [c]$  such that for every  $i \in [c]$ , there are at least  $k$  edges with both endvertices colored  $i$ . Gutin and Jones (IPL 2014) studied this problem with  $c = 2$ . They showed 2-LOAD COLORING to be fixed-parameter tractable (FPT) with parameter  $k$  by obtaining a kernel with at most  $7k$  vertices. In this paper, we extend the study to any fixed  $c$  by giving both a linear-vertex and a linear-edge kernel. In the particular case of  $c = 2$ , we obtain a kernel with less than  $4k$  vertices and less than  $8k$  edges. These results imply that for any fixed  $c \geq 2$ ,  $c$ -LOAD COLORING is FPT and the optimization version of  $c$ -LOAD COLORING (where  $k$  is to be maximized) has an approximation algorithm with a constant ratio.

**1998 ACM Subject Classification** F.2.2 Nonnumerical Algorithms and Problems

**Keywords and phrases** load Coloring, fixed-parameter tractability, kernelization

**Digital Object Identifier** 10.4230/LIPIcs.IPEC.2015.43

## 1 Introduction

Given a graph  $G = (V, E)$  and an integer  $k$ , the 2-LOAD COLORING Problem introduced in [1], asks whether there is a coloring  $\varphi : V \rightarrow \{1, 2\}$  such that for  $i = 1$  and  $2$ , there are at least  $k$  edges with both endvertices colored  $i$ . This problem is NP-complete [1], and Gutin and Jones studied its parameterization by  $k$  [9]. They proved that 2-LOAD COLORING is fixed-parameter tractable (FPT)<sup>1</sup> by obtaining a kernel with at most  $7k$  vertices. It is natural to extend 2-LOAD COLORING to any number  $c$  of colors as follows. Henceforth, for a positive integer  $p$ ,  $[p] = \{1, 2, \dots, p\}$ .

► **Definition 1** ( $c$ -LOAD COLORING). Let  $c$  be a positive integer. Given a positive integer  $k$  and a graph  $G = (V, E)$ , the  $c$ -LOAD COLORING problem asks whether there is a  $c$ -coloring  $\varphi : V \rightarrow [c]$  such that for every  $i \in [c]$ , there are at least  $k$  edges with both endvertices colored  $i$ . If such a coloring  $\varphi$  exists, we call  $\varphi$  a  $(c, k)$ -coloring of  $G$  and we write  $G \in (c, k)$ -LC.

The  $c$ -LOAD COLORING problem can be viewed as a subgraph packing problem: decide whether a graph  $G$  contains  $c$  disjoint  $k$ -edge subgraphs.

Observe first that  $G \in (1, k)$ -LC if and only if  $|E(G)| \geq k$ . In this paper, we consider  $c$ -LOAD COLORING parameterized by  $k$  for every fixed  $c \geq 2$ . Note that  $c$ -LOAD COLORING is NP-complete for every fixed  $c \geq 2$ . Indeed, we can reduce 2-LOAD COLORING to  $c$ -LOAD COLORING with  $c > 2$  by taking the disjoint union of  $G$  with  $c - 2$  stars  $K_{1,k}$ .

---

<sup>1</sup> For comprehensive introductions to parameterized algorithms and complexity, see recent monographs [4, 7]; [10, 11] are excellent recent survey papers on kernelization.



We prove that the problem admits a kernel with less than  $2ck$  vertices. Thus, for  $c = 2$  we improve the kernel result of [9]. To show our result, we introduce reduction rules, which are new even for  $c = 2$ . We prove that the reduction rules can run in polynomial time and that an irreducible graph with at least  $2ck$  vertices is in  $(c, k)$ -LC.

While there are many parameterized graph problems which admit kernels linear in the number of vertices, usually only problems on classes of sparse graphs admit kernels linear in the number of edges (since in such graphs the number of edges is linear in the number of vertices), see, e.g., [3, 7, 11]. To the best of our knowledge, only trivial  $O(k)$ -edge kernels for general graphs have been described in the literature, e.g., the kernel for MAX CUT parameterized by solution size (Prieto [12] improved the trivial result by obtaining a kernel with at most  $2k$  edges and at most  $k$  vertices). Thus, our next result is somewhat unusual:  $c$ -LOAD COLORING admits a kernel with  $O(k)$  edges for every fixed  $c \geq 2$ . Namely, the kernel has less than  $8k$  edges when  $c = 2$  and less than  $6.25c^2k$  edges when  $c > 2$ .

The optimization version of  $c$ -LOAD COLORING is as follows: for a graph  $G$ , find the maximum  $k$  such that  $G \in (c, k)$ -LC. We show that because of the above bounds on the number of edges in a kernel, this optimization problem, called the MAX  $c$ -LOAD COLORING problem, admits constant ratio approximation algorithms for any fixed  $c$ .

The paper is organized as follows. After providing additional terminology and notation on graphs in the remainder of this section, we show that the problem admits a kernel with less than  $2ck$  vertices in Section 2. Then, in Section 3, we prove an upper bound on the number of edges in a kernel for every  $c \geq 2$  and the corresponding approximation result for MAX  $c$ -LOAD COLORING. We improve our bound for  $c = 2$  in Section 4. The bound implies the approximation ratio of  $4 + \varepsilon$  for every  $\varepsilon > 0$ . We complete the paper with discussions in Section 5.

**Graphs.** For a graph  $G$ ,  $V(G)$  ( $E(G)$ , respectively) denotes the vertex set (edge set, respectively) of  $G$ ,  $\Delta(G)$  denotes the maximum degree of  $G$ ,  $n$  its number of vertices, and  $m$  its number of edges. A vertex  $u$  with degree 0 (1, respectively) is an *isolated vertex* (a *leaf-neighbor* of  $v$ , where  $uv \in E(G)$ , respectively). For a vertex  $x$  and a vertex set  $X$  in  $G$ ,  $N(x) = \{y : xy \in E(G)\}$  and  $N_X(x) = N(x) \cap X$ . For disjoint vertex sets  $X, Y$  of  $G$ , let  $G[X]$  be the subgraph of  $G$  induced by  $X$ ,  $E(X) = E(G[X])$  and  $E(X, Y) = \{xy \in E(G) : x \in X, y \in Y\}$ . For a coloring  $\varphi$ , we say that an edge  $uv$  is *colored  $i$*  if  $\varphi(u) = \varphi(v) = i$ .

## 2 Bounding Number of Vertices in Kernel

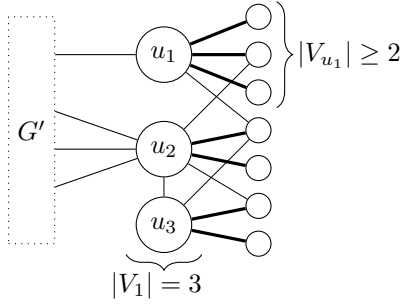
In this section, we show that  $c$ -LOAD COLORING admits a kernel with less than  $2ck$  vertices. The fact that  $(ck - 1)K_2$  is a No-instance suggests that this bound is likely to be optimal.

For any integer  $i \geq 1$  and  $\tau \in \{<, \leq, =, >, \geq\}$ ,  $K_{1,\tau i}$  denotes a *star*  $K_{1,j}$  such that  $j \tau i$  and  $j \geq 1$ . For instance,  $K_{1,\leq p}$  is a star with  $q$  edges,  $q \in [p]$ . Then, a  $K_{1,\tau i}$ -*graph* is a forest in which every component is a star  $K_{1,\tau i}$ , and a  $K_{1,\tau i}$ -*cover* of  $G$  is a spanning subgraph of  $G$  which is a  $K_{1,\tau i}$ -graph. We call any  $K_{1,\tau i}$ -graph a *star graph* and any  $K_{1,\tau i}$ -cover a *star cover*.

We first prove the bound for star graphs with small maximum degree.

► **Lemma 2.** *If  $G$  is a  $K_{1,<2k}$ -graph with  $n \geq 2ck$ , then  $G \in (c, k)$ -LC.*

**Proof.** Let  $G$  be a  $K_{1,<2k}$ -graph with  $n \geq 2ck$ . We prove the lemma by induction on  $c$ . The base case of  $c = 1$  holds since a  $K_{1,<2k}$ -graph has no isolated vertices. Indeed, this property implies  $G$  has at least  $\frac{V(G)}{2} \geq k$  edges.



■ **Figure 1** An overload from  $O_{3,2}$ .

Since all components of  $G$  are trees, for each one the number of vertices is one more than the number of edges. If there is a component  $C$ , with  $k \leq |E(C)| < 2k$ , we may color  $V(C)$  with the same color (then,  $G[V(C)] \in (1, k)$ -LC). Since we only used  $|V(C)| \leq 2k$  vertices,  $H = G - V(C)$  has at least  $2(c - 1)k$  vertices and so  $H \in (c - 1, k)$ -LC by the induction hypothesis. Thus,  $G \in (c, k)$ -LC.

We may assume that every component has less than  $k$  edges and let  $C_1, \dots, C_t$  be the components of  $G$ . Let  $b$  be the minimum nonnegative integer for which there exists  $I \subseteq [t]$  such that  $\sum_{i \in I} |E(C_i)| = k + b \geq k$ . Since there is no isolated vertex in a star graph,  $m \geq n/2 \geq ck$ , and thus such a set  $I$  exists. Observe that for any  $i \in I$ ,  $|E(C_i)| > b$ , as otherwise  $\sum_{j \in I \setminus \{i\}} |E(C_j)| = k + b - |E(C_i)| \geq k$ , a contradiction to the minimality of  $b$ . Since every component has less than  $k$  edges,  $b \leq k - 2$ . For a star  $(V, E)$ , the ratio  $\frac{|V|}{|E|}$  increases when  $|E|$  decreases. Thus, we have  $\sum_{j \in I} |V(C_j)| \leq \sum_{j \in I} |E(C_j)| \max_{h \in I} \left( \frac{|V(C_h)|}{|E(C_h)|} \right) \leq (k + b) \frac{b + 2}{b + 1}$ . But  $2k - (k + b) \frac{b + 2}{b + 1} = \frac{(k - 2 - b)b}{b + 1} \geq 0$ , and so  $\sum_{j \in I} |V(C_j)| \leq 2k$ . We may color the components  $C_i$ ,  $i \in I$ , by the same color. Again, we have that  $H = G - V(\bigcup_{i \in I} C_i)$  has at least  $2(c - 1)k$  vertices and so  $H \in (c - 1, k)$ -LC by the induction hypothesis. Thus,  $G \in (c, k)$ -LC. ◀

Since  $G \in (c, k)$ -LC whenever  $G$  has a subgraph  $H \in (c, k)$ -LC, we have that any graph with  $n \geq 2ck$  and a  $K_{1, < 2k}$ -cover is in  $(c, k)$ -LC. To decide the second property, we introduce a family  $(O_{i,k})_{i,k \in \mathbb{N}}$  of overloads.

► **Definition 3.** We call a pair  $(V_1, V_2)$  of disjoint vertex sets an *overload from  $O_{i,k}$*  if  $|V_1| = i$ ,  $N(v) \subseteq V_1$  for all  $v \in V_2$ , and for every  $u \in V_1$  there is a set  $V_u \subseteq N_{V_2}(u)$  such that  $|V_u| \geq k$  and for every pair  $u, v$  of distinct vertices of  $V_1$ ,  $V_u \cap V_v = \emptyset$  (see Fig. 1).

Note that if  $v$  is an isolated vertex, the pair  $(\emptyset, \{v\})$  is an overload from  $O_{0,k}$ . If a graph  $G$  has an overload  $(V_1, V_2)$  from  $O_{i,k}$ , then  $G[V_1 \cup V_2] \in (i, k)$ -LC: for each  $u \in V_1$ , color  $V_u \cup \{u\}$  with one color. However,  $G[V_1 \cup V_2] \notin (i + 1, k)$ -LC. Indeed, an edge can only be colored with one of  $|V_1| = i$  colors. So, in any coloring, an overload from  $O_{i,k}$  may give  $k$  edges for each of  $i$  colors but cannot bring any edge for all the other colors. From this observation, we deduce the following set of reduction rules. (Note that our reduction rules generalize the well-known Crown Reduction Rule. Similar, but different, reduction rules were used in [8].)

**Reduction rule  $R_{i,k}$ .** If an instance  $G$  for  $(c, k)$ -LC contains an overload  $(V_1, V_2)$  from  $O_{i,k}$ , delete all the vertices of  $V_1 \cup V_2$  from  $G$  and decrease  $c$  by  $i$ .

Since the existence of an overload from  $O_{i,k}$  for  $i \geq c$ , in a graph  $G$  implies  $G \in (c, k)$ -LC, we only consider  $R_{i,k}$  for  $i < c$ . We now show rules  $R_{i,k}$  are safe and can be applied in time polynomial in  $n$  (recall that  $c$  is fixed). We say that a graph is *irreducible for  $(c, k)$ -LC* if it is not possible to apply any rule  $R_{i,k}$ ,  $i < c$ , to the graph.

► **Lemma 4.** *Let  $G$  be a graph and  $G'$  be the graph obtained from  $G$  after applying reduction rule  $R_{i,k}$ . Then  $G \in (c, k)$ -LC if and only if  $G' \in (c - i, k)$ -LC.*

**Proof.** Let  $(V_1, V_2)$  be the overload from  $O_{i,k}$  used to map  $(G, c)$  to  $(G', c - i)$ . On the one hand, if  $G' \in (c - i, k)$ -LC, there exists a  $(c - i, k)$ -coloring of  $G'$  and together with a  $(i, k)$ -coloring of the overload, we obtain a  $(c, k)$ -coloring of  $G$ :  $G \in (c, k)$ -LC. On the other hand, observe that the vertices of  $V_2$  are isolated in  $G - V_1$ . Thus  $E(G - V_1) = E(G - V_1 - V_2) = E(G')$ . If  $G \in (c, k)$ -LC, in any  $(c, k)$ -coloring of  $G$ , there are at least  $c - |V_1| = c - i$  colors with no edge with endvertices in  $V_1$ . These colors must have their  $k$  edges in  $E(G - V_1) = E(G')$ . Thus  $G' \in (c - i, k)$ -LC. ◀

► **Lemma 5.** *One can decide whether Rule  $R_{i,k}$  is applicable to  $G$  in time  $O(n^{i+O(1)})$ .*

**Proof.** Generate all  $i$ -size subsets  $V_1$  of  $V(G)$ . For each  $V_1$ , construct the set  $V_2$  that includes every vertex outside  $V_1$  whose only neighbors are in  $V_1$ . If  $|V_2| \geq ik$ , construct the following bipartite graph  $B$ : the partite sets of  $B$  are  $V_1'$  and  $V_2$ , where  $V_1'$  contains  $k$  copies of every vertex  $v$  of  $V_1$  with the same neighbors as  $v$ . Observe that  $B$  has a matching covering  $V_1'$  if and only if  $R_{i,k}$  can be applied to  $G$  for the overload  $(V_1, V_2)$ . It is not hard to turn the above into an algorithm of runtime  $O(n^{i+O(1)})$ . ◀

In fact, the running time in Lemma 5 can be improved: in the journal version of this paper, we will show that  $O(n^{i+O(1)})$  can be replaced by  $O((cn)^2)$ .

Now, we want to show that a graph without any overloads has a star cover with small degree. If so, since an irreducible graph has no overload from  $O_{i,k}$ ,  $i \in [c - 1]$ , an irreducible graph for  $(c, k)$ -LC with at least  $2ck$  vertices would be in  $(c, k)$ -LC:

► **Lemma 6.** *Let  $G$  be a graph and  $k$  a positive integer. If  $G$  has no overload from  $O_{i,k}$  for any  $i \leq n$ , then  $G$  has a  $K_{1, \leq \max\{3, k\}}$ -cover.*

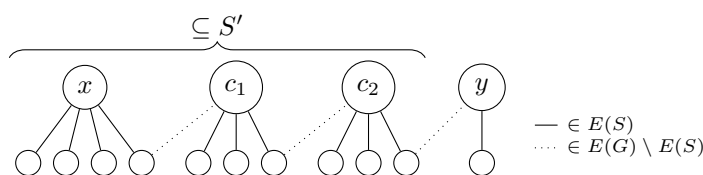
**Proof.** Let  $G$  be a graph with no overload from  $O_{i,k}$  for any  $i \leq n$ . We first show that  $G$  has a star cover. Since it is not possible to apply  $R_{0,k}$ ,  $G$  has no isolated vertex. By choosing a spanning tree of each component of  $G$ , we obtain a forest  $F$ . If a tree in  $F$  is not a star, it has an edge between two non-leaves. As long as  $F$  contains such an edge, delete it from  $F$ . Observe that  $F$  becomes a star cover of  $G$ . However, the number of leaves in each star of  $F$  is only bounded by  $\Delta(G)$ . We will show that among the possible star covers of  $G$ , there exists a  $K_{1, \leq \max\{3, k\}}$ -cover.

For each star cover  $F$ , we define the  $F$ -sequence  $(F_{\Delta(G)}, F_{\Delta(G)-1}, \dots, F_1)$ , where  $F_i$  is the number of stars with exactly  $i$  edges,  $i \in [\Delta(G)]$ . We say a star cover  $F$  is *smaller* than a star cover  $F'$  if and only if the  $F$ -sequence is smaller than the  $F'$ -sequence lexicographically, i.e. there exists some  $i \in [\Delta(G)]$  such that  $F_i < F'_i$  and for every  $j > i$ ,  $F_j = F'_j$ .

We select a star cover  $S$  of  $G$  which has the lexicographically minimum sequence, that is, for any star cover  $F$  of  $G$ , the  $S$ -sequence is smaller or equal to the  $F$ -sequence. Suppose that  $\Delta(S) > \max\{3, k\}$ . Let  $C_i$  ( $L_i$ , respectively) be the set of all the centers (leaves, respectively) of all stars of  $S$  isomorphic to  $K_{1,i}$ . We also define  $L_{\geq i} = \cup_{j \geq i} L_j$ . We will now prove two claims.

**Claim 1.** *There is no edge  $uv \in E(G) \setminus E(S)$  such that  $u \in L_{\geq 3}$  and  $v \in L_{\geq 1}$ .*

Indeed, suppose there exists one and let  $x, y$  be such that  $xu, yv \in E(S)$ . If  $v \in L_{\geq 2}$ , then by deleting edges  $xu, yv$  and adding edge  $uv$ , we do not create any isolated vertex but we decrease the size of the stars centered at  $x$  and  $y$ , and thus we get a smaller star cover than  $S$ , a contradiction. Otherwise,  $v$  is an endvertex of an independent edge, and by deleting



■ **Figure 2** An alternating path from  $x$  to  $y$  with  $\Delta(S) = 4$ .

edge  $xu$  and adding edge  $uv$ , we decrease the size of the star centered at  $x$ , and create a star  $K_{1,2}$  centered at  $v$ , which still induces a star cover smaller than  $S$ , a contradiction.

**Claim 2.** *Suppose  $S$  contains a star isomorphic to  $K_{1,i}$  and centered at vertex  $x$ , and a star isomorphic to  $K_{1,j}$  and centered at vertex  $y$ , such that  $i - j \geq 2$ . There is no path from  $x$  to  $y$  in which the odd edges are in  $E(S)$  and go from a center to a leaf, and the even edges are in  $E(G) \setminus E(S)$  and go from a leaf to a center. (see Fig. 2)*

Suppose there exists such a path. Then by deleting the odd edges of the path and adding the even ones, we do not create isolated vertices because  $x$  still has leaf-neighbors,  $y$  gets a neighbor, every transitional center keeps the same number of leaf-neighbors and the transitional leaves always go to a new center. This operation only decreases the size of star centered at  $x$  by 1 and increases the size of star centered at  $y$  by 1, giving us a lexicographically smaller star cover, a contradiction.

Let  $S'$  be the subgraph of  $S$  containing all stars  $K_{1,\Delta(S)}$  of  $S$ . While there is an edge  $uv \in E(G) \setminus E(S)$  such that  $u$  is a leaf of  $S'$  and  $v \in C_{\Delta(S)-1} \setminus S'$ , we add the star centered at  $v$  to  $S'$ . Let  $C'$  ( $L'$ , respectively) be the centers (leaves, respectively) in  $S'$ . Suppose there is an edge  $uv \in E(G) \setminus E(S)$  such that  $u \in L' \subseteq L_{\geq \Delta(S)-1} \subseteq L_{\geq 3}$  and  $v \in V(G) \setminus C'$ . By Claim 1,  $v \notin L_{\geq 1}$ . Since  $v \notin C_{\Delta(S)} \subseteq C'$  and since the above procedure has terminated,  $v \in C_j$  for some  $j$  such that  $\Delta(S) - j \geq 2$ . Now, by construction, there is an alternating path from a vertex in  $C_{\Delta(S)}$  to a vertex in  $C_j$  of the type described in Claim 2, which is impossible.

So, there is no edge  $uv \in E(G) \setminus E(S)$  such that  $u \in L'$  and  $v \notin C'$ . This means that for any  $u \in L'$ ,  $N(u) \subseteq C'$ . Furthermore, for each  $u \in C'$ , we can define  $V_u$  to be the leaves of the star centered at  $u$ , for which we have  $|V_u| \geq \Delta(S) - 1 \geq k$ . Thus,  $(C', L')$  is an overload from  $O_{|C'|,k}$ , which is impossible. ◀

Since we obtain the expected result, we can deduce our theorem:

► **Theorem 7.** *For  $k > 1$ , if  $G$  is irreducible for  $(c, k)$ -LC and has at least  $2ck$  vertices, then  $G \in (c, k)$ -LC. Furthermore, for any fixed  $c \geq 2$  and for any positive integer  $k$ ,  $c$ -LOAD COLORING admits a kernel with less than  $2ck$  vertices.*

**Proof.** Observe first that for every  $c \geq 2$ ,  $G \in (c, 1)$ -LC if and only if  $G$  has a matching with at least  $c$  edges. Since this property can be decided in polynomial time, we just need to consider the case when  $k > 1$ .

By Lemmas 4 and 5, there is a polynomial algorithm that reduces an instance  $(G, c)$  to an instance  $(G', c')$  such that  $c' \leq c$  and  $G'$  is irreducible for  $(c', k)$ -LC. Suppose the kernel  $G'$  has at least  $2c'k$  vertices, but  $G \notin (c', k)$ -LC. Then, observe that  $G'$  does not have an overload from  $O_{i,k}$ ,  $i \geq c'$ , and thus  $G'$  has a  $K_{1, \leq \max(3,k)}$ -cover by Lemma 6. Since  $k > 1$ , this star cover is a  $K_{1, < 2k}$ -cover and Lemma 2 implies that  $G' \in (c', k)$ -LC, a contradiction. So, if  $|V(G')| \geq 2c'k$ , then  $G' \in (c', k)$ -LC and  $G \in (c, k)$ -LC, hence we may conclude the kernel  $G'$  has less than  $2c'k \leq 2ck$  vertices. ◀

### 3 Bounding Number of Edges in Kernel

Let  $S(c)$  be the integer sequence defined by induction by  $S(1) = 1$ ,  $S(2c) = 4S(c)$  and  $S(2c+1) = 2S(c) + 2S(c+1)$ . This sequence is known as A073121 in the Online Encyclopedia of Integer Sequences [13] (see also [2]). We will use the following technical result.

► **Lemma 8.** *If  $c$  is even,  $S(c) \leq \frac{9c^2-4}{8}$ . For arbitrary  $c$ ,  $S(c) \leq \frac{9c^2-1}{8}$ .*

**Proof.** It is easy to check the base cases:  $S(1) = 1 = \frac{9(1)^2-1}{8}$ ,  $S(2) = 4 = \frac{9(2)^2-4}{8}$  and  $S(3) = 10 = \frac{9(3)^2-1}{8}$ . We now assume the claim holds for every  $c \leq 2c' - 1$  and we will prove it for  $c = 2c'$  and  $c = 2c' + 1$ .

For even value, we have:

$$S(2c) = 4S(c) \leq 4 \frac{9c^2 - 1}{8} = \frac{9(2c)^2 - 4}{8}.$$

For odd value, we have:

$$\begin{aligned} S(2c+1) &= 2(S(c) + S(c+1)) \\ &\leq 2 \frac{9c^2 + 9(c+1)^2 - 1 - 4}{8} = \frac{9(2c+1)^2 - 1}{8}. \end{aligned} \quad \blacktriangleleft$$

By using the kernel we proved in the previous section, we show that  $c$ -LOAD COLORING admits a kernel with less than  $(2S(c) + 4c^2 - 5c)k$  edges. Because of the upper bound on  $S(c)$  given by Lemma 8, the number of edges in a kernel may be bounded by  $6.25c^2k$ . We first prove a smaller bound for bipartite graphs.

► **Lemma 9.** *Let  $b(c, k, n) = S(c)k + (c-1)n$ . For every positive integer  $c$  and bipartite graph  $G$  with  $n$  vertices, if  $m \geq b(c, k, n)$  then  $G \in (c, k)$ -LC.*

**Proof.** We prove the lemma by induction on  $c$ . For the base case, observe that any graph with at least  $k = b(1, k, n)$  edges is in  $(1, k)$ -LC for every  $k$  and  $n$ . We now assume the claim holds for every  $c \leq 2c' - 1$  and we will prove it for  $c = 2c'$  and  $c = 2c' + 1$ .

Suppose that  $G = (A \cup B, E)$  is a bipartite graph with  $n$  vertices and at least  $b(c, k, n)$  edges, but  $G \notin (c, k)$ -LC. Let  $B_2$  be a maximal subset of  $B$  such that

$$|E(A, B_2)| < b(c - c', k, |A| + |B_2|) + b(c - c', k, |B_2|). \quad (1)$$

So, for any vertex  $u \in B \setminus B_2$ , the set  $B_2 \cup \{u\}$  doesn't satisfy (1). Such a set  $B_2$  exists since the empty set satisfies (1). Moreover, for any partition  $(A_1, A_2)$  of  $A$ , we know there exists  $i \in \{1, 2\}$  such that

$$|E(A_i, B_2 \cup \{u\})| \geq b(c - c', k, |A_i| + |B_2 \cup \{u\}|) \quad (2)$$

as otherwise, the linearity in  $n$  of  $b(c, k, n)$  implies a contradiction with the maximality of  $B_2$ :

$$\begin{aligned} |E(A, B_2 \cup \{u\})| &= |E(A_1, B_2 \cup \{u\})| + |E(A_2, B_2 \cup \{u\})| \\ &< b(c - c', k, |A_1| + |B_2 \cup \{u\}|) + b(c - c', k, |A_2| + |B_2 \cup \{u\}|) \\ &= b(c - c', k, |A| + |B_2 \cup \{u\}|) + b(c - c', k, |B_2 \cup \{u\}|). \end{aligned}$$

Let  $B_1 = B \setminus B_2$ ,  $A_1 = A$  and  $A_2 = \emptyset$ . We define the following inequalities.

$$|E(A_1, B_1)| < b(c', k, |A_1| + |B_1|) + |A_1| \quad (3)$$

$$|E(A_2, B_1)| < b(c', k, |A_2| + |B_1|) + |A_2|. \quad (4)$$

While (3) does not hold and (4) holds, we move an arbitrary vertex from  $A_1$  to  $A_2$ . Suppose eventually (3) and (4) are both false and let  $u$  be an arbitrary vertex in  $B_1$ . We deduce for both  $i = 1$  and  $i = 2$  that

$$|E(A_i, B_1 \setminus \{u\})| \geq b(c', k, |A_i| + |B_1|).$$

Thus, there exist disjoint vertex sets  $X$  and  $Y$  such that  $|E(X)| \geq b(c', k, |X|)$  and  $|E(Y)| \geq b(c - c', k, |Y|)$  (either  $X = A_1 \cup B_1 \setminus \{u\}$  and  $Y = A_2 \cup B_2 \cup \{u\}$ , or  $X = A_2 \cup B_1 \setminus \{u\}$  and  $Y = A_1 \cup B_2 \cup \{u\}$ ), depending on whether (2) holds for  $i = 1$  or  $i = 2$ . By taking a  $(c', k)$ -coloring of  $X$  and a  $(c - c', k)$ -coloring of  $Y$ , we have that  $G \in (c, k)$ -LC, a contradiction.

So, we may assume (3) eventually holds. If  $A_2 = \emptyset$ , then  $|E(A_2, B_1)| = 0$ . Otherwise, let  $v$  be the last vertex moved from  $A_1$  to  $A_2$ . Observe that

$$\begin{aligned} |E(A_2, B_1)| &\leq |E(A_2 \setminus \{v\}, B_1)| + |B_1| \\ &< b(c', k, |A_2 \setminus \{v\}| + |B_1|) + |A_2 \setminus \{v\}| + |B_1| \text{ (by (4))} \\ &< b(c', k, |A_2| + |B_1|) + |A_2| + |B_1|. \end{aligned} \tag{5}$$

In both cases, (5) holds and we can bound the number of edges in  $G$ :

$$\begin{aligned} |E(G)| &= |E(A, B_2)| + |E(A_1, B_1)| + |E(A_2, B_1)| \\ &< b(c - c', k, |A| + |B_2|) + b(c - c', k, |B_2|) \\ &\quad + b(c', k, |A_1| + |B_1|) + |A_1| \\ &\quad + b(c', k, |A_2| + |B_1|) + |A_2| + |B_1| \\ &\quad \text{(by inequalities (1),(3),(5))}. \end{aligned}$$

If  $c = 2c'$ , we have  $c - c' = c'$  and it is not hard to check that

$$|E(G)| < 4S(c')k + 2(c' - 1)n + n = b(c, k, n).$$

Otherwise,  $c = 2c' + 1$  and then  $c - c' = c' + 1$ . Thus,

$$\begin{aligned} |E(G)| &< 2S(c')k + 2S(c' + 1)k + 2(c' - 1)n \\ &\quad + |A| + 2|B_2| + |A_1| + |A_2| + |B_1| \\ &\leq S(2c' + 1)k + 2c'n = b(c, k, n). \end{aligned}$$

Thus, for  $c = 2c'$  and  $c = 2c' + 1$ , we have  $|E(G)| < b(c, k, n)$ , a contradiction. So, there is no bipartite graph with  $n$  vertices and at least  $b(c, k, n)$  edges such that  $G \notin (c, k)$ -LC.  $\blacktriangleleft$

We now generalize this lemma for any graph. We would like to find a partition  $(A, B)$  of  $V$  such that  $|E(A)| + |E(B)|$  is bounded, since  $|E(A, B)|$  is bounded.

**► Lemma 10.** *Let  $f(c, k, n) = (2S(c) - c)k + 2(c - 1)n$ . For every positive integer  $c$  and every graph  $G$  with  $n$  vertices, if  $m \geq f(c, k, n)$  then  $G \in (c, k)$ -LC.*

**Proof.** We prove the lemma by induction on  $c$ . For the base case, observe that any graph with at least  $k = f(1, k, n)$  edges is in  $(1, k)$ -LC for every  $k$  and  $n$ . We now assume the claim holds for every  $c \leq 2c' - 1$  and we will prove it for  $c = 2c'$  and  $c = 2c' + 1$ .

Consider a graph  $G$  with  $n$  vertices and at least  $f(c', k, n)$  edges, such that  $G \notin (c, k)$ -LC. We will first show that there exists a set  $A \subseteq V(G)$  such that  $f(c', k, |A|) \leq |E(A)| \leq f(c', k, |A|) + |A|$  (and thus  $G[A] \in (c', k)$ -LC). We may construct the set  $A$  as follows:

initially  $A = \emptyset$  and while  $|E(A)| < f(c', k, |A|)$ , add an arbitrary vertex of  $V(G) \setminus A$  to  $A$ . Let  $u$  be the last added vertex. Then

$$|E(A)| \leq |E(A \setminus \{u\})| + |A \setminus \{u\}| < f(c', k, |A \setminus \{u\}|) + |A \setminus \{u\}| < f(c', k, |A|) + |A|.$$

Let  $B = V(G) \setminus A$ . If  $G[B] \in (c - c', k)$ -LC, then  $G \in (c, k)$ -LC, a contradiction. So  $|E(B)| < f(c - c', k, |B|)$ . Furthermore,  $|E(A, B)| < b(c, k, n)$  by Lemma 9. Finally, we may bound  $|E(G)|$ . If  $c = 2c'$ , we have  $c - c' = c'$

$$\begin{aligned} |E(G)| &< f(c', k, |A|) + f(c', k, |B|) + b(2c', k, n) + |A| \\ &\leq (2S(2c') - 2c')k + (4c' - 2)n = f(c, k, n). \end{aligned}$$

Otherwise,  $c = 2c' + 1$  and  $c - c' = c' + 1$ . Thus,

$$\begin{aligned} |E(G)| &< f(c', k, |A|) + f(c' + 1, k, |B|) + b(2c' + 1, k, n) + |A| \\ &\leq (2S(2c' + 1) - (2c' + 1))k + 4c'n = f(c, k, n). \end{aligned}$$

Thus, in both cases  $|E(G)| < f(c, k, n)$ , as required.  $\blacktriangleleft$

► **Theorem 11.** *The  $c$ -LOAD COLORING Problem admits a kernel with less than  $f(c, k, 2ck) < 6.25c^2k$  edges.*

**Proof.** By Theorem 7, we can get a kernel with less than  $2ck$  vertices. Thus by Lemmas 10 and 8, we get a kernel such that  $|E(G)| < f(c, k, 2ck) < 6.25c^2k$ .  $\blacktriangleleft$

The size of this kernel may be optimal up to a constant factor. Indeed, the complete bipartite graph  $K_{c, ck-1}$  is an irreducible graph for  $(c, k)$ -LC with  $c^2k - c = O(c^2k)$  edges, but  $K_{c, ck-1} \notin (c, k)$ -LC. We can increase this lower bound by joining all  $c$  vertices on the smaller side of  $K_{c, ck-1}$ . The resulting graph is not in  $(c, k)$ -LC either, and it has  $c^2k + \frac{c(c-3)}{2}$  edges.

We now consider an approximation algorithm for the MAX  $c$ -LOAD COLORING problem: Given a graph  $G$  and integer  $c$ , we wish to determine the maximum  $k$ , denoted  $k_{opt}$ , for which  $G \in (c, k)$ -LC. Given an approximation algorithm, we define the approximation ratio  $r(c) = \frac{k_{opt}}{k}$ , where  $k$  is the output of the approximation algorithm.

Let  $K(c)k$  be an upper bound of the number of edges in a kernel for  $(c, k)$ -LC and let  $P(c) = \prod_{i=1}^c \frac{K(i)}{i}$ . By Theorem 11, we may have  $K(c) = 6.25c^2$ .

► **Theorem 12.** *There is a  $2^{c-1}P(c)$ -approximation algorithm for MAX  $c$ -LOAD COLORING.*

**Proof.** We prove the claim by induction on  $c$ . For  $c = 1$ , we have  $P(1) = 1$ . Assume the theorem is true for all  $c' < c$  and let  $G$  be an instance for  $c$ -LOAD COLORING with  $n$  vertices and  $m$  edges. We may assume that  $G$  has no isolated vertices. Clearly,  $k_{opt} \leq \frac{m}{c}$ . Consider  $k = \lfloor \frac{m}{K(c)} \rfloor$ .

If  $k = 0$ , then  $m < K(c)$  and we can find  $k_{opt}$  in  $O(1)$  time.

Now let  $k > 0$ . If  $n \leq 2ck$ , then by the proof of Theorem 11, since  $m \geq K(c)k$ ,  $G \in (c, k)$ -LC. So we return  $k$ , and  $\frac{k_{opt}}{k} \leq \frac{m}{ck} \leq \frac{K(c)(k+1)}{ck} \leq \frac{2K(c)}{c} \leq 2^{c-1}P(c)$ .

If  $n \geq 2ck$  and  $G$  is irreducible for  $(c, k)$ -LC, then by Theorem 7,  $G \in (c, k)$ -LC and we return  $k$  as above. If  $n \geq 2ck$  and  $G$  is not irreducible for  $(c, k)$ -LC, we can use Lemma 5 to reduce  $(G, c)$  to  $(G', c')$  with  $c' < c$ . By induction we may find  $k'$  such that  $k'_{opt} \leq 2^{c'-1}P(c')k'$ , where  $k'_{opt}$  is the optimal solution for MAX  $c'$ -LOAD COLORING on  $G'$ . Now consider three cases:

■  $k' \geq k$ . Then  $G' \in (c', k)$ -LC and so  $G \in (c, k)$ -LC. This case also leads to the above conclusion.



- $k'_{opt} \leq 2^{c'-1}P(c')k' < k$ . Because  $k'_{opt} + 1 \leq k$ , an overload from  $O_{c-c',k}$  is also an overload from  $O_{c-c',k'_{opt}+1}$ , therefore  $G'$  can be derived from  $G$  using a reduction rule for  $(c, k'_{opt} + 1)$ -LC. Since  $G' \notin (c', k'_{opt} + 1)$ -LC,  $G \notin (c, k'_{opt} + 1)$ -LC. Thus  $k_{opt} = k'_{opt}$ . The algorithm may output  $k'$  which satisfies  $k_{opt} = k'_{opt} \leq 2^{c'-1}P(c')k' \leq 2^{c-1}P(c)k$ .
- $k' < k \leq 2^{c'-1}P(c')k'$ . The algorithm gives  $k'$  as an approximation of  $k_{opt}$ . Then  $\frac{k_{opt}}{k'} \leq \frac{m}{ck'} \leq \frac{K(c)(k+1)}{ck'} \leq \frac{K(c)}{c} \frac{2k}{k'} \leq \frac{K(c)}{c} 2^{c'} P(c') \leq 2^{c-1}P(c)$ .

In every case, the approximation ratio is at most  $2^{c-1}P(c)$ . ◀

#### 4 Number of Edges in Kernel for $c = 2$

In this section, we look into the edge kernel problem for the special case when  $c = 2$ . By doing a refined analysis, we will give a kernel with less than  $8k$  edges for  $(2, k)$ -LC, which is a better bound than the general one. Henceforth, we assume that  $G$  is irreducible for  $(2, k)$ -LC, and just consider the case when  $|V(G)| < 4k$ , as we have proved that if  $|V(G)| \geq 4k$  then  $G \in (2, k)$ -LC.

► **Lemma 13.** *If  $G$  has at least  $3k - 2$  edges and every component in  $G$  has less than  $k$  edges then  $G \in (2, k)$ -LC.*

**Proof.** We consider colorings of the graph such that vertices in the same component are colored with the same color. Thus every edge in the graph is colored with 1 or 2. Denote the set of edges colored  $i$  with  $E_i, i = 1, 2$ . Among all possible colorings, choose a coloring of the graph such that  $|E_1| \geq |E_2|$  and  $||E_1| - |E_2||$  is minimum. Suppose  $|E_2| \leq k - 1$ , then  $|E_1| \geq 2k - 1, ||E_1| - |E_2|| > k$ . Changing the color of one component from 1 to 2, we get a new coloring of the graph. For the new coloring, denote the set of edges colored  $i$  with  $E'_i, i = 1, 2$ . Since each component has less than  $k$  edges,  $|E_1| > |E'_1| \geq k, |E'_2| \leq 2k - 2$ . So  $||E'_1| - |E'_2|| < ||E_1| - |E_2||$ , a contradiction. Therefore we have  $|E_1| \geq |E_2| \geq k$ , so  $G \in (2, k)$ -LC. ◀

If  $G$  has at least two components, each with at least  $k$  edges, it is obviously a Yes-instance. Therefore by Lemma 13, we may assume there is exactly one component  $C$  with at least  $k$  edges in the graph. Denote the total number of edges in  $G - V(C)$  with  $m'$ . Observe that if  $m' \geq k$ , trivially  $G \in (2, k)$ -LC. So assume that  $m' < k$ .

► **Lemma 14.** *If  $G$  is an irreducible graph for  $(2, k)$ -LC,  $m' < k$  and  $\Delta = \Delta(G) \geq 3k - 2m'$ , then  $G \in (2, k)$ -LC.*

**Proof.** Let  $u$  be one of the vertices with degree  $\Delta$  and  $N(u)$  its neighbors. Because the graph is reduced by Reduction Rule  $R_{1,k}$ ,  $u$  has at least  $2k - 2m'$  neighbors which are not leaves. Arbitrarily select  $k - m'$  vertices among them and for each one, select any neighbor but  $u$ . Color the selected vertices and  $G - V(C)$  by 1. By construction, there are at least  $k$  edges colored 1 and there are at most  $2k - 2m'$  colored vertices in  $N(u)$ . So there are at least  $k$  uncolored vertices in  $N(u)$ . We color them and  $u$  with 2. So  $G \in (2, k)$ -LC. ◀

► **Lemma 15.** *Let  $G$  be a graph with  $\Delta < 3k$  and  $|E(G)| \geq 8k$ , then  $G \in (2, k)$ -LC.*

**Proof.** By Lemma 13, we may assume there exists a connected component  $C$  with at least  $k$  edges. In  $C$ , choose a minimal set  $A \subseteq V(C)$  such that  $|A| \leq k + 1$  and  $|E(A)| = k + d \geq k$ . We may find such a set  $A$  in the following way. Select arbitrarily a vertex in  $C$  and put it into  $A$ , then keep adding to this set some neighbor of some vertex in  $A$  until  $|E(A)| = k + d \geq k$ . Since each time we select a neighbor of  $A$  we strictly increase  $|E(A)|$ ,  $|A| \leq k + 1$ . If there

is any vertex  $u \in A$  with  $|N_A(u)| \leq d$ , then  $A' = A \setminus \{u\}$  is a smaller vertex set such that  $|E(A')| \geq k$ . Thus, we may remove such vertices until  $|E(A)| = k + d$  and for each vertex  $u \in A$ ,  $|N_A(u)| > d$ . Denote  $B = V(G) \setminus A$ . We may assume  $|E(B)| < k$ , as otherwise  $G \in (2, k)$ -LC.

We now show that  $|A| + d \leq k + 3$ . Since every vertex  $u \in A$  has  $d_A(u) > d$ ,  $|E(A)| = \frac{1}{2} \sum_{u \in A} d_A(u) \geq \frac{d+1}{2} |A|$ . We have  $k + d = |E(A)| \geq \frac{d+1}{2} |A|$ , thus  $|A| \leq \frac{2(k+d)}{d+1}$ . Moreover as  $d \leq |A| - 1$ ,

$$d + |A| \leq 2|A| - 1 \leq \frac{4(k+d)}{d+1} - 1 < \frac{4k}{d+1} + 3.$$

If  $d \geq 3$ , we are done; otherwise  $d \leq 2$  and  $d + |A| \leq 2 + k + 1 = k + 3$ .

Let  $A_1, A_2, B_1, B_2$  be a partition of  $V(G)$  such that  $A = A_1 \cup A_2$ ,  $B = B_1 \cup B_2$ ,  $|A_2| = 1$  and  $|E(A, B_2)| < 2k$ . Such a partition is possible: let  $y = \operatorname{argmax}\{|N_B(u)| : u \in A\}$  and initially take  $A_1 = A \setminus \{y\}$ ,  $A_2 = \{y\}$ ,  $B_1 = B$ ,  $B_2 = \emptyset$ . Suppose  $|E(A_1, B_1)| \leq k + |A_1|$ . Then

$$\begin{aligned} |E(G)| &\leq |E(A)| + |E(B)| + |E(A_1, B_1)| + |E(A_2, B_1)| \\ &\leq (k + d) + (k - 1) + (k + |A| - 1) + \Delta \\ &\leq 7k + 1, \end{aligned}$$

a contradiction since  $|E(G)| > 8k$ . So,  $|E(A_1, B_1)| > k + |A_1|$ . We will consider two cases:  $\max\{|N_{B_1}(u)| : u \in A\}$  is greater than  $k$  or not.

If so, observe that  $|E(A_2, B_1)| = |E(\{y\}, B_1)| = \max\{|N_{B_1}(u)| : u \in A\} > k$ . Move all vertices of  $B_1 \setminus N(y)$  to  $B_2$ . We still have  $|E(\{y\}, B_1)| > k$  and  $|E(\{y\}, B_2)| = 0$ . Moreover  $B_1 \subseteq N(y)$ . If  $|E(A_1, B_2)| \geq k$ , then  $G$  is in  $(2, k)$ -LC, thus  $|E(A_1, B_2)| < k$ . While  $|E(\{y\}, B_1)| \geq k + 1$  and  $|E(A_1, B_1)| \geq k + |A_1|$ , move an arbitrary vertex from  $B_1$  to  $B_2$ . After each move,  $|E(\{y\}, B_1)| \geq k$  and  $|E(A_1, B_1)| \geq k$ , thus  $|E(A_2, B_2)| < k$  and  $|E(A_1, B_2)| < k$  as otherwise,  $G$  would be in  $(2, k)$ -LC.

Eventually, we have  $|E(A_1, B_1)| < k + |A_1|$  or  $|E(\{y\}, B_1)| = k$ . Suppose  $|E(A_1, B_1)| < k + |A_1|$ . Then

$$\begin{aligned} |E(G)| &\leq |E(A)| + |E(B)| + |E(A_1, B_1)| + |E(A_1, B_2)| + |E(\{y\}, B)| \\ &\leq (k + d) + (k - 1) + (k + |A_1| - 1) + (k - 1) + \Delta \\ &\leq 4k - 3 + (d + |A|) + \Delta \\ &< 8k, \end{aligned}$$

a contradiction. Thus,  $|E(A_1, B_1)| \geq k + |A_1|$  and  $|E(\{y\}, B_1)| = k$ . As  $B_1 \subseteq N(y)$ , we have  $|B_1| = k$ . We have found a new partition with the required properties and with  $\max\{|N_{B_1}(u)| : u \in A\} \leq |B_1| = k$ .

We now consider the case  $\max\{|N_{B_1}(u)| : u \in A\} \leq k$ . While there exists  $u \in B_1$  such that  $|E(A, B_2 \cup \{u\})| < 2k$ , move  $u$  from  $B_1$  to  $B_2$ . Then, (if and) while  $|E(A_1, B_1)| \geq k + |A_1|$  and  $|E(A_2, B_1)| < k + |A_2|$ , move an arbitrary vertex from  $A_1$  to  $A_2$ .

After all such moves, suppose that  $|E(A_1, B_1)| < k + |A_1|$ . If  $|A_2| = 1$ , we have  $|E(A_2, B_1)| \leq \max\{|N_{B_1}(u)| : u \in A\} \leq k$ , otherwise we moved some vertices from  $A_1$  to  $A_2$ . Let  $u$  be the last one. Since  $|E(A_2 \setminus \{u\}, B_1)| < k + |A_2 \setminus \{u\}|$ , we know  $|E(A_2, B_1)| \leq |E(A_2 \setminus \{u\}, B_1)| + \max\{|N_{B_1}(u)| : u \in A\} < k + |A_2| - 1 + k = 2k + |A_2| - 1$ . For both

cases,

$$\begin{aligned}
|E(G)| &= |E(A)| + |E(B)| + |E(A_1, B_1)| + |E(A_2, B_1)| + |E(A, B_2)| \\
&\leq (k + d) + (k - 1) + (k + |A_1| - 1) + (2k + |A_2| - 2) + (2k - 1) \\
&\leq 7k + d + |A| - 5 \\
&< 8k,
\end{aligned}$$

which is impossible.

So,  $|E(A_1, B_1)| \geq k + |A_1|$  which implies  $|E(A_2, B_1)| \geq k + |A_2|$ . For any vertex  $u \in B_1$ , we have  $|E(A_1, B_1 \setminus \{u\})| \geq k$  and  $|E(A_2, B_1 \setminus \{u\})| \geq k$  and we also obtain  $|E(A, B_2 \cup \{u\})| \geq 2k$ , i.e.  $E(A_1, B_2 \cup \{u\})$  or  $E(A_2, B_2 \cup \{u\})$  has at least  $k$  edges. Thus,  $G \in (2, k)$ -LC. ◀

The lemmas of this section and the fact that their proofs can be turned into polynomial algorithms, imply the following:

► **Theorem 16.** *If  $G$  is irreducible for  $(2, k)$ -LC and has at least  $8k$  edges, then  $G \in (2, k)$ -LC. Thus, 2-LOAD COLORING admits a kernel with less than  $8k$  edges.*

Since we have a better bound for the number of edges in a kernel when  $c = 2$ , we may get a better approximation when  $c = 2$ .

► **Theorem 17.** *For every  $\varepsilon > 0$ , there is a  $(4 + \varepsilon)$ -approximation algorithm for 2-MAX LOAD COLORING.*

**Proof.** Let  $G$  be an instance for 2-MAX LOAD COLORING with  $m = 8p + q$  edges, where  $0 \leq q < 8$ . Let  $k_{opt}$  be the optimal solution of 2-MAX LOAD COLORING on  $G$ , and observe that  $k_{opt} \leq \lfloor \frac{m}{2} \rfloor \leq 4p + 3$ . Let  $p_0 = \lceil \frac{3}{\varepsilon} \rceil$ . If  $p \leq p_0 - 1$  then we can find  $k_{opt}$  in  $O(1)$  time.

So assume that  $p \geq p_0$ . Note that  $\frac{k_{opt}}{p} \leq \frac{4p+3}{p} \leq 4 + \varepsilon$ . If  $G$  is irreducible for  $(2, p)$ -LC,  $G \in (2, p)$ -LC by Theorem 16, and so  $p$  gives the required approximation. We may assume that  $G$  is not irreducible for  $(2, p)$ -LC and reduce  $G$  to  $G'$ . If  $|E(G')| \geq p$ , then  $G' \in (1, p)$ -LC, and by Lemma 4,  $G \in (2, p)$ -LC. Again,  $p$  gives the required approximation.

Now assume that  $|E(G')| < p$  and let  $k'_{opt} = |E(G')|$  be the optimal solution of MAX 1-LOAD COLORING on  $G'$ . Then  $k'_{opt} + 1 \leq p$  and so an overload from  $O_{1,p}$  is also an overload from  $O_{1,k'_{opt}+1}$ . Thus,  $G'$  can be derived from  $G$  using a reduction rule for  $(2, k'_{opt} + 1)$ -LC. Since  $G' \notin (1, k'_{opt} + 1)$ -LC,  $G \notin (2, k'_{opt} + 1)$ -LC. Thus  $k_{opt} = k'_{opt} = |E(G')|$ . So let our algorithm output  $|E(G')|$  in this case. ◀

## 5 Discussions

To the best of our knowledge, we obtained the first nontrivial linear-edge kernel for a problem on general graphs. As we saw, such kernels can be used to obtain approximation algorithms. It would be interesting to obtain such kernels for other problems.

It is clear that our approximation algorithm is far from optimal; we will present an  $O(c)$ -approximation algorithm in the journal version of this paper.

Our linear-vertex kernel result implies an  $O^*(c^{2ck})$ -time algorithm for  $c$ -LOAD COLORING, which simply tests all the  $c$ -colorings of the kernel. However, it is possible that there is a subexponential FPT algorithm, since the problem NO  $c$ -LOAD COLORING (the complement of  $c$ -LOAD COLORING) has small, but not constant, forbidden minors and is minor-bidimensional (see [5, 6] for more information on forbidden minors and bidimensionality).

Let  $tw(G)$  denote the treewidth of  $G$ . The  $O^*(2^{tw(G)})$ -time algorithm for 2-LOAD COLORING from [9] can be generalized to an  $O^*(c^{tw(G)})$ -time algorithm for  $c$ -LOAD COLORING.

By [5, 6], if we require that the input  $G$  is  $H$ -minor-free for some fixed graph  $H$ , then we obtain an  $O^*(c^{\sqrt{ck}})$ -time algorithm. Unfortunately, there is no constant forbidden minor for NO  $c$ -LOAD COLORING as membership in  $(c, k)$ -LC requires at least  $ck$  edges.

We think that there exists a subexponential algorithm for  $c$ -LOAD COLORING. By Theorem 4.12 of [5], and since branchwidth is linked to the treewidth up to a constant factor, any graph  $G$  contains an  $(\Omega(\frac{tw(G)}{gen(G)}) \times \Omega(\frac{tw(G)}{gen(G)}))$ -grid as a minor, where  $gen(G)$  is the genus of  $G$ . Since the  $(r \times r)$ -grid is a forbidden minor for NO  $c$ -LOAD COLORING for  $r = \lceil \sqrt{(c+1)k} \rceil$ , we have  $tw(G) = O(\sqrt{ck} gen(G))$ . Thus, we obtain an  $O^*(c^{\sqrt{ck} gen(G)})$ -time algorithm to solve  $c$ -LOAD COLORING, which is subexponential for graphs of bounded genus.

**Acknowledgment.** Research of GG was partially supported by Royal Society Wolfson Research Merit Award. Research of BS was supported by China Scholarship Council.

---

### References

- 1 N. Ahuja, A. Baltz, B. Doerr, A. Prívativý, and A. Srivastav. On the minimum load coloring problem. *J. Discrete Algorithms*, 5(3):533–545, 2007.
- 2 J.-P. Allouche and J. Shallit. The ring of  $k$ -regular sequences, II. *Theor. Comput. Sci.*, 307(1):3–29, 2003.
- 3 H. L. Bodlaender, F. V. Fomin, D. Lokshtanov, E. Penninkx, S. Saurabh, and D. M. Thilikos. (meta) kernelization. In *Foundations of Computer Science, FOCS 2009*, pages 629–638. IEEE Computer Society, 2009.
- 4 M. Cygan, F. V. Fomin, L. Kowalik, D. Lokshtanov, D. Marx, M. Pilipczuk, M. Pilipczuk, and S. Saurabh. *Parameterized Algorithms*. Springer, 2015.
- 5 E. D. Demaine, F. V. Fomin, M. T. Hajiaghayi, and D. M. Thilikos. Subexponential parameterized algorithms on bounded-genus graphs and  $H$ -minor-free graphs. *J. ACM*, 52(6):866–893, 2005.
- 6 E. D. Demaine and M. T. Hajiaghayi. The bidimensionality theory and its algorithmic applications. *Comput. J.*, 51(3):292–302, 2008.
- 7 R. G. Downey and M. R. Fellows. *Fundamentals of Parameterized Complexity*. Springer, 2013.
- 8 F. V. Fomin, D. Lokshtanov, N. Misra, G. Philip, and S. Saurabh. Hitting forbidden minors: Approximation and kernelization. In *Symposium on Theoretical Aspects of Computer Science, STACS 2011*, volume 9 of *LIPICs*, pages 189–200. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2011.
- 9 G. Gutin and M. Jones. Parameterized algorithms for load coloring problem. *Inf. Process. Lett.*, 114(8):446–449, 2014.
- 10 S. Kratsch. Recent developments in kernelization: A survey. *Bulletin of the EATCS*, 113, 2014.
- 11 D. Lokshtanov, N. Misra, and S. Saurabh. Kernelization – preprocessing with a guarantee. In *The Multivariate Algorithmic Revolution and Beyond*, volume 7370 of *Lecture Notes in Computer Science*, pages 129–161. Springer, 2012.
- 12 E. Prieto. The method of extremal structure on the  $k$ -maximum cut problem. In *Theory of Computing 2005, Eleventh CATS 2005, Computing: The Australasian Theory Symposium, Newcastle, NSW, Australia, January/February 2005*, volume 41 of *CRPIT*, pages 119–126. Australian Computer Society, 2005.
- 13 J. Shallit, 2002. <http://oeis.org/A073121>.