Maximum colorful cliques in vertex-colored graphs

Giuseppe F. Italiano[2], Yannis Manoussakis[1], Nguyen Kim Thang[3*], and Hong Phong Pham[1]

[1] LRI, University Paris-Saclay, Orsay, France
[2] University of Rome Tor Vergata, Rome, Italy
[3] IBISC, University Paris-Saclay, Evry, France

Abstract. In this paper we study the problem of finding a maximum colorful clique in vertex-colored graphs. Specifically, given a graph with colored vertices, we wish to find a clique containing the maximum number of colors. Note that this problem is harder than the maximum clique problem, which can be obtained as a special case when each vertex has a different color. In this paper we aim to give a dichotomy overview on the complexity of the maximum colorful clique problem. We first show that the problem is NP-hard even for several cases where the maximum clique problem is easy, such as complement graphs of bipartite permutation graphs, complement graphs of bipartite convex graphs, and unit disk graphs, and also for properly vertex-colored graphs. Next, we provide a XP parameterized algorithm and polynomial-time algorithms for classes of complement graphs of bipartite chain graphs, complete multipartite graphs and complement graphs of cycle graphs, which are our main contributions.

1 Introduction

In this paper we deal with vertex-colored graphs, which are useful in various applications. For instance, the Web graph may be considered as a vertex-colored graph where the color of a vertex represents the content of the corresponding site (i.e., green for corporate sites, red for blogs, yellow for ecommerce sites, etc.) [4]. In a biological population, vertex-colored graphs can be used to represents the connections and interactions between species where different species have different colors. Other applications of vertex-colored graphs arise also in bioinformatics (Multiple Sequence Alignment Pipeline or for multiple Protein-Protein Interaction networks) [7], and in scheduling problems [18].

Given a vertex-colored graph, a *tropical subgraph* is a subgraph where each color of the initial graph appears at least once. Many graph properties, such as the domination number, the vertex cover number, independent sets, connected components, shortest paths, matchings, etc. can be studied in their tropical version. Tropical subgraphs find applications in many scenarios: in a (biological)

* Research supported by the ANR project OATA n° ANR-15-CE40-0015-01, Hadamard PGMO and DIM RFSI.

population, for instance, a tropical subgraph fully represents the (bio-)diversity of the population. There are many cases, however, where tropical subgraphs do not necessarily exist. Hence, one might be interested in the more general question of finding a *maximum colorful* subgraph, i.e., a subgraph with the maximum possible number of colors. For instance in biology this would represent a subgraph with the most diverse population. As a special case, a maximum colorful subgraph is tropical if it contains all colors. The notion of colorful subgraph is close to, but somewhat different from the *colorful* concept considered in [1, 13, 14], where neighbor vertices must have different colors. It is also related to the concepts of *color patterns* or *colorful* used in bio-informatics [8]. Note that in a *colorful* subgraph considered in our paper, two adjacent vertices may have the same color, i.e., the subgraph is not necessarily properly colored.

In this paper we are interested in finding maximum colorful cliques in vertex-colored graphs. Throughout, we let $G = (V, E)$ denote a simple undirected graph. Given a set of colors $\mathcal{C}$, $G^c = (V, E)$ denotes a vertex-colored graph whose vertices are (not necessarily properly) colored by one of the colors in $\mathcal{C}$. The number of colors of G^c is $|\mathcal{C}|$. Given a subset of vertices $U \subseteq V$, the set of colors of vertices in U is denoted by $\mathcal{C}(U)$. Moreover, we denote by $c(v)$ the color of vertex v and by $v(H, c)$ the number of vertices of H whose color is c. The set of neighbors of v is denoted by $N(v)$. More formally, in this paper we study the following:

Maximum Colorful Clique Problem (MCCP). Given a vertex-colored graph $G^c = (V, E)$, find a clique with the maximum number of colors of the original graph.

Related work. In the special case where each vertex has a distinct color, MCCP reduces to the maximum clique problem. The maximum clique problem has been widely studied in the literature and it is known to be NP-complete for general graphs, fixed-parameter intractable and hard to approximate. However, the maximum clique problem can be efficiently solved in polynomial time for several special classes of graphs, such as complement graphs of bipartite graphs, permutation graphs, comparability graphs, chordal graphs. All those are perfect graphs, and the maximum clique problem can be solved in polynomial time also in other non-perfect graphs such as circle graphs and unit disk graphs.

Another related problem, which has also been widely considered in the literature, is listing all maximal cliques in a graph. Clearly, if one can list all maximal cliques, one can also find a maximum clique, since a maximum clique must be maximal. In a similar vein, one can also find the maximum colorful clique, since any maximum colorful clique can be extended to a maximal clique. Therefore, MCCP is easy for all classes of graphs for which we can list in polynomial time all maximal cliques. Those graphs include chordal graphs, complete graphs, triangle-free graphs, interval graphs, graphs of bounded boxicity, and planar graphs.

Tropical subgraph and maximum colorful subgraph problems in vertex-colored graphs have been studied only recently. In particular, tropical subgraph problems in vertex-colored graphs such as tropical connected subgraphs, tropical dominating sets have been investigated in [9]. The maximum colorful matching problem [5], the maximum colorful path problem [6] and the maximum colorful cycles problem [12] have been studied, and several hardness results and polynomial-time algorithms were shown for different classes of graphs.

Our contributions. In this paper, we aim to give a dichotomy overview on the complexity of MCCP. First, we show that MCCP is NP-hard even for several cases where the maximum clique problem is known to be easy, such as complement graphs of bipartite permutation graphs, complement graphs of convex bipartite graphs, and unit disk graphs. Also, we show that MCCP is NP-hard for properly vertex-colored graphs. Next, we present polynomial-time algorithms for several classes of graphs. First, we prove that MCCP belongs to the class of XP parameterized algorithms. Second, we show that MCCP can be solved in polynomial time for complement graphs of bipartite chain graphs, which is a special case of complement graphs of bipartite permutation graphs (for which MCCP is NP-hard).

Our main contribution is polynomial-time algorithms for complete multipartite graphs and complement graphs of cycle graphs. A graph is called multipartite if its vertices can be partitioned into different independent sets. In a complete multipartite graph any two vertices in different independent sets are adjacent. To solve MCCP on complete multipartite graphs, we proceed as follows. We start with a maximum clique, by picking one vertex from each independent set. To compute a maximum colorful clique, we iterate through different maximum cliques by increasing at each step the number of colors, without decreasing the number of vertices. To do this efficiently, we define a special structure, called k-colorful augmentation, which might be of independent interest. The running time of our algorithm is $O(|\mathcal{C}|M(m+n,n))$, where $|\mathcal{C}|$ is the total number of colors and $M(m,n)$ is the time required for finding a maximum matching in a general graph with m edges and n vertices (Currently, $M(m,n) = O(\sqrt{n}m)$ [19])
.

A cycle graph is a graph that consists of a single cycle. Similar to the algorithm for complete multipartite graphs, we also investigate a special structure to obtain another better clique from the current cliques in a complement graph of a cycle graph and this yields a polynomial algorithm in this case.

Due to space limit, some results, proofs and details are omitted from this extended abstract and are deferred to the appendix.

2 Hardness results for MCCP

In this section, we present several NP-hardness results for MCCP. Specifically, we show that MCCP is NP-hard for the complement graphs of biconvex bipartite graphs and the complement graphs of permutation bipartite graphs. Note that the maximum clique problem can be solved in polynomial time for the complement graphs of bipartite graphs based on König's Theorem, and therefore the maximum clique problem can be also efficiently solved for the complement graphs of biconvex bipartite graphs and the complement graphs of permutation bipartite graphs. Next, we also show that MCCP is NP-hard for unit disk graphs, which are also easy cases for the maximum clique problem. We also prove that MCCP is NP-hard for properly vertex-colored graphs. The following lemma shows that MCCP is NP-hard for the complement graphs of bipartite permutation graphs. Recall that a graph is a bipartite permutation graph, if it is both bipartite and a permutation graph.

Lemma 1. *The maximum colorful clique problem is NP-hard for the complement graphs of permutation bipartite graphs.*

Proof. We reduce from the MAX-3SAT problem. Consider a boolean expression B in CNF with variables $X = \{x_1, \ldots, x_s\}$ and clauses $B = \{b_1, \ldots, b_t\}$. In addition, suppose that B constains exactly 3 literals per clause (actually, we may also consider clauses of arbitrary size). We show how to construct a vertex-colored graph G^c associated with any such formula B, such that, there exists a truth assignment to the variables of B satisfying t' clauses if and only if G^c contains a clique with t' distinct colors. Suppose that $\forall i, 1 \leq i \leq s$, the variable x_i appears in clauses $b_{i1}, b_{i2}, \ldots, b_{i\alpha_i}$ and $\overline{x_i}$ appears in clauses $b'_{i1}, b'_{i2}, \ldots, b'_{i\beta_i}$ in which $b_{ij} \in B$ and $b'_{ik} \in B$. Now a vertex-colored permutation bipartite graph G^c is constructed as follows. We give geometrical definition as the intersection graphs of line segments whose endpoints lie on two parallel lines L_1 and L_2.

We create first $\alpha_1 + \beta_1$ endpoints for two parallel lines as follows. Firstly, from left to right, we let α_1 endpoints of L_1 corresponding to pairs (x_1, b_{11}), (x_1, b_{12}), $\ldots$, $(x_1, b_{1\alpha_1})$, and next β_1 endpoints of L_1 will be $(\overline{x_1}, b'_{11})$, $(\overline{x_1}, b'_{12})$ $\ldots$, $(\overline{x_1}, b'_{1\beta_1})$. Conversely, on L_2, from left to right, we let first β_1 endpoints including $(\overline{x_1}, b'_{11})$, $(\overline{x_1}, b'_{12})$ $\ldots$, $(\overline{x_1}, b'_{1\beta_1})$, and next α_1 endpoints of L_2 will be (x_1, b_{11}), (x_1, b_{12}), $\ldots$, $(x_1, b_{1\alpha_1})$. This way, the segment of (x_1, b_{1i}) and the segment (x_1, b_{1j}) are in parallel and they do not intersect each other for $\forall i, j, 1 \leq i \neq j \leq \alpha_1$; similarly for each pair $(\overline{x_1}, b'_{1i})$ and $(\overline{x_1}, b'_{1j})$, for $\forall i, j, 1 \leq i \neq j \leq \beta_1$. However, it is easy to see that the pair of segment (x_1, b_{1i}) and $(\overline{x_1}, b'_{1j})$ intersect each other. Now we finish arrangement for clauses corresponding to the variable x_1 and $\overline{x_1}$. Next, we similarly create and arrange $\alpha_2 + \beta_2$ endpoints corresponding to the variable x_2 and $\overline{x_2}$ on L_1 and L_2. The segments of this section are separated with the segments of x_1 and $\overline{x_1}$ but they still guarantee the properties of intersection: the segment of (x_2, b_{2i}) and the segment (x_2, b_{2j}) are in parallel, similarly for $(\overline{x_2}, b'_{2i})$ and $(\overline{x_2}, b'_{2j})$ but the pair of segment (x_2, b_{2i}) and $(\overline{x_2}, b'_{2j})$ intersect each other.

By doing so, we created a permutation bipartite graph. In fact, it is a bipartite graph since edges are between (x_i, b_{ij}) and $(\overline{x_i}, b'_{ik})$. It is also a permutation graph since it is the intersection graph of line segments whose endpoints lie on two parallel lines. Now we color vertices as follows. We use color c_l for the vertex (x_i, b_{ij}) and $(\overline{x_i}, b'_{ik})$ if b_{ij} or b'_{ik} is the clause b_l of B. From this vertex-colored permutation bipartite graph, we focus on its complement graph. Clearly the obtained graph is the complement graph of a vertex-colored permutation bipartite graphs, denoted it by G^c. Now we claim that there exists a truth assignment to the variables of B satisfying t' clauses if and only if G^c contains a clique with t' distinct colors. Observe that in the complement graph G^c, there are no edges between (x_i, b_{ij}) and $(\overline{x_i}, b'_{ik})$ for $\forall i, 1 \le i \le s$ but there are all edges between (x_i, b_{ij}) and (x_i, b_{ik}), between (x_i, b_{ij}) and $(x_{i'}, b_{i'k})$, between (x_i, b_{ij}) and $(\overline{x_{i'}}, b_{i'k})$ in which $\forall i, i', 1 \le i \ne i' \le s$.

Now we extract a subgraph from a truth assignment to the variables of B satisfying t' clauses, as follows. For each $\forall i, 1 \le i \le s$, in the case that x_i is assigned true, then we choose all vertices (x_i, b_{ij}), $1 \le j \le \alpha_i$. Otherwise, we choose all vertices $(\overline{x_i}, b_{ij})$, $1 \le j \le \beta_i$. It is possible to see that this is a clique and this clique contains all vertices corresponding to satisfied clauses, so the number of colors of this cliques is equal to t'. Conversely, from a clique with t' colors we obtain an assignment as follows. Note that in this clique we can ot have both vertices (x_i, b_{ij}) and $(\overline{x_i}, b'_{ik})$. Therefore, it is possible to assign x_i as true if this clique contains vertices (x_i, b_{ij}), otherwise x_i is assigned as false. Clearly, this assignment is consistent. Since this clique has t' colors, it must contain vertices corresponding to t' different clauses. Thus, it is not difficult to see that the corresponding assignment leads to t' satisfied clauses. This complete our proof. □

Next, we show that MCCP is also NP-hard for the complement graphs of biconvex bipartite graphs. Recall that a bipartite graph (X, Y, E) is biconvex if there is an ordering of X and Y that fulfills the adjacency property, i.e., for every vertex $y \in Y$ (resp., $x \in X$), $N(y)$ (resp., $N(x)$) consists of vertices that are consecutive in the sorted ordering of X (resp., Y).

Lemma 2. *The maximum colorful clique problem is NP-hard for the complement graphs of biconvex bipartite graphs.*

Proof. We use the same notation as in Lemma 1, and reduce again from MAX-3SAT, as follows. We first create a vertex-colored biconvex bipartite graph (X, Y, E) from an instance of the MAX-3SAT problem such that, there exists a truth assignment to the variables of B satisfying t' clauses if and only if G^c contains a clique with t' distinct colors. We also assume that $\forall i, 1 \le i \le s$, the variable x_i appears in clauses $b_{i1}, b_{i2}, \ldots, b_{i\alpha_i}$ and $\overline{x_i}$ appears in clauses $b'_{i1}, b'_{i2}, \ldots, b'_{i\beta_i}$ in which $b_{ij} \in B$ and $b'_{ik} \in B$. Now each vertex of X represents a pair of (x_i, b_{ij}) with $1 \le j \le \alpha_i$ (x_i appears in the clause b_{ij}). Similarly, each vertex of Y represents a pair of $(\overline{x_i}, b'_{ik})$ with $1 \le k \le \beta_i$ ($\overline{x_i}$ appears in the clause b'_{ik}). Next, we sort an ordering over X from left to right as follows. The vertices of X are sorted as (x_1, b_{11}), (x_1, b_{12}), $\ldots$, $(x_1, b_{1\alpha_1})$, (x_2, b_{21}), (x_2, b_{22}), $\ldots$, $(x_2, b_{2\alpha_2})$, $\ldots$, (x_s, b_{s1}), (x_s, b_{s2}), $\ldots$, $(x_s, b_{s\alpha_s})$. Similarly, vertices over Y are sorted from left to right as follows: $(\overline{x_1}, b'_{11})$, $(\overline{x_1}, b'_{12})$, $\ldots$, $(\overline{x_1}, b'_{1\beta_1})$, $(\overline{x_2}, b'_{21})$, $(\overline{x_2}, b'_{22})$, $\ldots$, $(\overline{x_2}, b'_{2\beta_2})$, $\ldots$, $(\overline{x_s}, b'_{s1})$, $(\overline{x_s}, b'_{s2})$, $\ldots$, $(\overline{x_s}, b'_{s\beta_s})$. Now edges between X and Y are created by connecting each vertex (x_i, b_{ij}) to each vertex $(\overline{x_i}, b'_{ik})$ for $\forall j, 1 \le j \le \alpha_i$ and $\forall k, 1 \le k \le \beta_i$. Clearly the obtained graph is a biconvex bipartite graph since for each vertex (x_i, b_{ij}), all its neighbors, i.e., $(\overline{x_i}, b'_{i1})$, $(\overline{x_i}, b'_{i2})$, $\ldots$, $(\overline{x_i}, b'_{i\beta_i})$, are consecutive by the sorted ordering over Y and vice versa for each vertex $(\overline{x_i}, b'_{ik})$. Similar to the case of complement graphs of permutation bipartite graphs, we use the color c_l for the vertex (x_i, b_{ij}) and $(\overline{x_i}, b'_{ik})$ if b_{ij} or b'_{ik} is the clause b_l of B. Now we focus on the complement graph of this vertex-colored biconvex bipartite graph, denoted by G^c. It is not difficult to see that this graph has the same structure as the complement graphs of bipartite permutation graphs in Lemma 1. In other words, there are no edges between (x_i, b_{ij}) and $(\overline{x_i}, b'_{ik})$ for all $1 \leq i \leq s$ but there are all edges between (x_i, b_{ij}) and (x_i, b_{ik}), between (x_i, b_{ij}) and $(x_{i'}, b_{i'k})$, between (x_i, b_{ij}) and $(\overline{x_{i'}}, b_{i'k})$ in which $\forall i, i', 1 \leq i \neq i' \leq s$. So, exactly the same arguments used in Lemma 1 yield the proof. □

Now we prove that MCCP is NP-hard also for unit disk graphs and properly vertex-colored graphs, as shown in the following lemmas (proofs in the appendix).

Lemma 3. *The maximum colorful clique problem is NP-hard for unit disk graphs.*

Lemma 4. *The maximum colorful clique problem is NP-hard for properly vertex-colored graphs.*

3 Efficient algorithms for MCC

In this section we present several efficient algorithms for the maximum colorful clique problem. We start by proving that MCPP belongs to the class *XP* parameterized problems. Next, we show that MCCP can be efficiently solved for complement graphs of bipartite chain graphs. This is in contrast with the case of complement graphs of bipartite permutation graphs and complement graphs of biconvex bipartite graphs, for which we have shown in Section 2 that MCCP is NP-hard. Finally, we present our polynomial-time algorithms for MCCP in complete multipartite graphs and complement graphs of cycle graphs.

3.1 A XP parameterized algorithm for MCCP

Our algorithm is based on the following observation: each maximum colorful clique can be reduced to another maximum colorful clique in which each color appears at most once. Indeed, if a color c appears more than once in maximum colorful clique, we can maintain only one vertex of this color. By doing so, we can keep all colors of the original clique, and obtain a new maximum colorful clique where each color appears only once. From this observation, we can list all cases by trying each vertex from the set of vertices of a color for each subset of the original set of colors. Let $\mathcal{C} = \{c_1, c_2, \ldots, c_{|\mathcal{C}|}\}$ be the set of colors of the original graph G^c. For each color c_i, let n_i be the number of vertices of the graph with the color c_i and let denote $V(c_i)$ be the set of vertices of G^c of color c_i. Our *XP* parameterized algorithm for MCCP is as follows.

The following theorem shows the correctness of our *XP* parameterized algorithm for computing a maximum colorful cycle in a vertex-colored graph G^c.

Theorem 1. *Algorithm 1 computes a maximum colorful clique of G^c in time $O((\frac{n}{|\mathcal{C}|})^{|\mathcal{C}|})$ where $|\mathcal{C}|$ is the number of colors in G^c and n is the number of vertices of G^c.*

Algorithm 1 Maximum colorful clique in vertex-colored graphs.

$max \leftarrow 0$ /* the number of colors of a maximum colorful clique */
$H^* \leftarrow \emptyset$ /* the maximum colorful clique returned */
for $\{c_{i_1}, c_{i_2}, \ldots, c_{i_j}\} \subseteq \{c_1, c_2, \ldots, c_{|\mathcal{C}|}\}$ **do** `# Consider all subsets of colors of` $\mathcal{C}$
  for $\{v_1, v_2, \ldots, v_j\} \subseteq \{V(c_{i_1}) \times V(c_{i_2}) \times \ldots \times V(c_{i_j})\}$ **do**
    $H \leftarrow$ the induced graph of the vertices $\{v_1, v_2, \ldots, v_j\}$
    if H is a clique and $max < |\mathcal{C}(H)|$ **then**
      $H^* \leftarrow H$
    end if
  end for
end for
return H^* as a maximum colorful clique

Proof. The algorithm first consider all subsets of colors of the set $\mathcal{C}$. This takes $2^{|\mathcal{C}|}$ times in the outer loop. Inside each of these iterations (corresponding to the set $\{c_{i_1}, c_{i_2}, \ldots, c_{i_j}\}$), the algorithm considers all subsets of vertices of $\{V(c_{i_1}) \times V(c_{i_2}) \times \ldots \times V(c_{i_j})\}$ (each vertex from a set), to check the existence of clique. It is easy to see that this takes $O(|V(c_{i_1})| \times |V(c_{i_2})| \times \ldots \times |V(c_{i_j})|)$. By the Cauchy inequality, with $j \leq |\mathcal{C}|$ we have that $|V(c_{i_1})| \times |V(c_{i_2})| \times \ldots \times |V(c_{i_j})| \leq (\frac{\sum |V(c_i)|}{|\mathcal{C}|})^{|\mathcal{C}|}$. Therefore, the complexity of this algorithm is $O((\frac{n}{|\mathcal{C}|})^{|\mathcal{C}|})$. $\square$

3.2 An algorithm for MCCP for complement graphs of bipartite chain graphs

In this section, we show a polynomial algorithm for MCCP for complement graphs of bipartite chain graphs. Recall that a bipartite graph $G = (X, Y, E)$ is said to be a *chain bipartite* graph if its vertices of X can be linearly ordered such that $N(x_1) \supseteq N(x_2) \supseteq \ldots \supseteq N(x_{|X|})$. As a consequence, we also immediately obtain a similar linear ordering over Y such that $N(y_1) \supseteq N(y_2) \supseteq \ldots \supseteq N(y_{|Y|})$. It is known that these orderings over X and Y can be computed in $O(n)$ time. Here we will look for a maximum colorful clique in the complement graph of a vertex-colored bipartite chain graph $G^c = (X, Y, E)$. Let us denote this complement graphs by $\overline{G^c}$. First, we observe that in $\overline{G^c}$, we have that $N(x_1) \subseteq N(x_2) \subseteq \ldots \subseteq N(x_{|X|})$ and $N(y_1) \subseteq N(y_2) \subseteq \ldots \subseteq N(y_{|Y|})$. Let K be a maximum colorful clique of $\overline{G^c}$ in which the set of vertices of X and Y of K are denoted by X_K and Y_K, respectively. Then, we can convert K to another maximum colorful clique by exploiting the following lemma.

Lemma 5. *Let K be a maximum colorful clique of $\overline{G^c}$, let i and j be the minimum numbers such that $x_i \in X_K$ and $y_j \in Y_K$. Then there exists another maximum colorful clique K' where $V(K') = \{x_i, x_{i+1}, \ldots, x_{|X|}\} \cup \{y_j, y_{j+1}, \ldots, y_{|Y|}\}$.*

Proof. Let $x_{i'}$ be a vertex such that $x_{i'} \notin X_K$ and $i < i' < |X|$. Since $N(x_i) \subseteq N(x_{i'})$ and K is a clique, adding the vertex $x_{i'}$ yields a larger clique which contains the old clique. Similarly, it is possible to add vertices $y_{j'}$ to our clique such that $y_{j'} \notin Y_K$ and $j < j' < |Y|$. As a consequence, there exists another maximum colorful clique K' where $V(K') = \{x_i, x_{i+1}, \ldots, x_{|X|}\} \cup \{y_j, y_{j+1}, \ldots, y_{|Y|}\}$. □

The following algorithm, which computes a maximum colorful clique in complement graphs of vertex-colored bipartite chain graphs, is based directly on Lemma 5.

Algorithm 2 Maximum colorful clique in complement graphs of vertex-colored bipartite chain graphs.

$max \leftarrow 0$ /* the number of colors of a maximum colorful clique */
$H^* \leftarrow \emptyset$ /* the maximum colorful clique returned */
for $1 \leq i \leq |X|$ **do**
  for $1 \leq j \leq |Y|$ **do**
    $H \leftarrow$ the induced graph of the vertices $\{x_i, x_{i+1}, \ldots, x_{|X|}\} \cup \{y_j, y_{j+1}, \ldots, y_{|Y|}\}$
    if H is a clique and $max < |\mathcal{C}(H)|$ **then**
      $H^* \leftarrow H$
    end if
  end for
end for
return H^* as a maximum colorful clique

The proof of following theorem is immediate.

Theorem 2. *Algorithm 2 computes a maximum colorful clique of $\overline{G^c}$ in time $O(n^2)$.*

3.3 An algorithm for MCCP for complete multipartite graphs

In this section, we present our new algorithm for MCCP for vertex-colored complete multibipartite graphs. Recall that a k-partite graph is a graph whose vertices can be partitioned into k different independent sets. A k-partite graph is complete if there exists an edge from each vertex of an independent set to each vertex of another independent set. Observe that if we pick one vertex from each independent set we obtain a clique. Therefore, it is possible to reduce the original MCCP problem to the problem of choosing one vertex from each independent set such that the number of colors is maximized. Note that all independent sets have to be selected, otherwise one could add more vertices. From now on, we denote a set of such vertices as a maximum clique. Suppose that our complete multipartite graph has $\mathcal{N}$ independent sets and let us denote these independent sets by $I_1, I_2, \ldots, I_{\mathcal{N}}$. The main idea behind our algorithm is to create first a maximum clique K by randomly picking a vertex from each independent set. Next, we construct a maximum colorful clique from K by iteratively increasing the number of covered colors, without decreasing the number of vertices. To show how to accomplish this task, we need some further notation. Given a maximum clique K and a color c, let $v(K,c)$ be the total number of vertices of color c in K. We let c_0 be a color of the original graph which does not appear in the current maximum clique, i.e., $v(K,c_0)=0$. Now a set of pairs of vertices $\{(v'_1,v_1),\ (v'_2,v_2),\ \ldots,\ (v'_k,v_k)\}$ $(1\le k\le \mathcal{N})$ is called a *k-colorful augmentation w.r.t. a maximum clique* K if it satisfies the following properties:

1. For all $i=1,2,\ldots,k$, v_i is covered by K and v'_i is not covered by K.
2. The color $c(v'_1)=c_0$ is not presented in K, i.e., $v(K,c_0)=0$.
3. For all $i=1,2,\cdots,k-1$, $c(v_i)=c(v'_{i+1})=c_i$ and $v(K,c_i)=1$. (Note that $c_i\neq c_j$, for all $1\le i,j\le k-1$.)
4. The color $c(v_k)=c_k$ such that $v(K,c_k)\ge 2$ and $c_k\neq c_i$ for all $i=1,2,\cdots,k-1$.

Note that a k-colorful augmentation $\{(v'_1,v_1),\ (v'_2,v_2),\ \ldots,\ (v'_k,v_k)\}$ w.r.t. K provides a better solution for our problem. Indeed, if we replace the vertices $\{v_1,v_2,\ldots,v_k\}$ by the set of vertices $\{v'_1,v'_2,\ldots,v'_k\}$, we obtain a new maximum clique, which includes a new color $c(v'_1)=c_0$ and preserves the old colors $c(v_1),c(v_2),\ldots,c(v_k)$, i.e., the number of colors in the new clique increases by one. The following theorem is at the heart of our algorithm for finding a maximum colorful clique in a vertex-colored complete multipartite graph G^c.

Theorem 3. *Let K be a maximum clique in a vertex-colored complete multipartite graph G^c. Then, K admits a k-colorful augmentation w.r.t. K if and only if there exists another maximum clique K' such that $|\mathcal{C}(K')|>|\mathcal{C}(K)|$.*

Proof. First, assume that K admits a k-colorful augmentation. A k-colorful augmentation $\mathcal{P}=\{(v'_1,v_1),\ (v'_2,v_2),\ \ldots,\ (v'_k,v_k)\}$ w.r.t. K can be used to transform K into another clique K' such that $|\mathcal{C}(K')|>|\mathcal{C}(K)|$. Replacing vertices in K by vertices not in K' inside this k-colorful augmentation yields another maximum clique K', which increases the number of distinct colors used.

Conversely, suppose that there exists a maximum colorful clique K' such $|\mathcal{C}(K')|>|\mathcal{C}(K)|$. Since K and K' are maximum cliques, $|K|=|K'|=\mathcal{N}$, i.e., both K and K' contains one vertex from each independent set. Let S be the set of independent sets in which the vertex chosen by K is equal to the vertex chosen by K'. Clearly, the set of colors of vertices of K in S is equal to the set of colors of vertices of K' in S. Thus, from now we can ignore all those vertices. We first focus on vertices v'_1 of K' such that $c(v'_1)=c_0$ is not present in K, i.e., $v(K,c_0)=0$. Let us denote the set of those vertices by $V_1(K')$. From each vertex v'_1 of $V_1(K')$, we extend to a sequence of pairs of vertices of K and K', namely $\{(v'_1,v_1),\ (v'_2,v_2),\ \ldots,\ (v'_k,v_k)\}$ such that: (i) for all $i=1,2,\ldots,k$, $v_i\in K$ and $v'_i\in K'$, and (ii) for all $i=1,2,\cdots,k-1$, $c(v_i)=c(v'_{i+1})=c_i$ and $v(K,c_i)=1$. Clearly this extension is unique for each vertex v'_1. Note that the ending vertex in each of the extended sequences is the vertex v_k. Next, we will focus on those ending vertices.

- In the case that there exists an ending vertex v_k such that $v(K,c_k)\ge 2$ then this set of pair of vertices $\{(v'_1,v_1),\ (v'_2,v_2),\ \ldots,\ (v'_k,v_k)\}$ is *a k-colorful augmentation w.r.t. K.*

– Otherwise, for each v'_1 of $V_1(K')$ such that $v(K, c_0) = 0$ and $v(K', c_0) \geq 1$, we have that the ending vertex v_k satisfies $v(K, c_k) = 1$ and $v(K', c_k) = 0$. Note that $v(K', c_k) = 0$ since otherwise it is possible to extend this sequence. Clearly, except for the vertices of $V_1(K')$, then for each vertex v' of K' such that $v' \notin V_1(K')$ (denote this set of vertices by $V_2(K')$), we have that $c(v') \in \mathcal{C}(K)$, i.e., $v(K, c(v')) \geq 1$. Combining these properties of vertices in the sets $V_1(K')$, $V_2(K')$, we can deduce that $|\mathcal{C}(K')| \leq |\mathcal{C}(K)|$, a contradiction. This completes our proof. □

In order to complete our algorithm, we need a polynomial sub-routine for finding colorful augmentations. This is achieved with the following lemma.

Lemma 6. *Let K be a maximum clique in a vertex-colored complete multipartite graph G^c. The problem of finding a k-colorful augmentation with respect to K in G^c can be reduced in polynomial time to the problem of finding an M-augmenting path in another graph G'^c w.r.t. K.*

Proof. Since K is a maximum clique, each vertex of K is in an independent set I_i of G^c. Let us denote these vertices by $v_1, v_2, \ldots, v_{\mathcal{N}}$. W.l.o.g. assume that $v_i \in I_i$. To look for a k-colorful augmentation with respect to K in G^c, we construct a new graph $G' = (V', E')$ and a matching M as follows. Its vertex-set is defined as $V(G') = \{v \in V(G^c) : v(M, c(v)) \neq 1\} \cup \{v_\eta : v(M, \eta) = 1 \textit{ and } \eta \in \mathcal{C}\} \cup \{u_i | 1 \leq i \leq \mathcal{N}\} \cup \{z\}$, where u_i and z are new artificial vertices not in $V(G^c)$ and u_i is corresponding to the set I_i. Next the edge-set of G' is defined as $E(G') = \{(t(v), t(w)) : (v, w) \in E(G^c)\} \cup \{(u_i, v) : u_i \textit{ and } v \textit{ are covered by } I_i\} \cup \{(z, v) : v = t(v) \textit{ and } v \textit{ is covered by } K\}$, where $t(v) = v$ if $v(M, c(v)) \neq 1$, otherwise $t(v) = v_\eta$ if $v(M, \eta) = 1$ and $c(v) = \eta$.

Now we define our matching: $M = \{(u_1, v_1), (u_2, v_2), \ldots, (u_{\mathcal{N}}, v_{\mathcal{N}})\}$. Let M' be a subset of $E(G')$ such that $M' = \{(t(x), t(y)) : (x, y) \in M\}$. It is easy to see that M' is also a matching of G'.

Conversely, let $P = \{(v'_1, v_1), (v'_2, v_2), \ldots, (v'_k, v_k)\}$ be a *k-colorful augmentation w.r.t. K*, $1 \leq k \leq \mathcal{N}$, in which v_i and v'_i are in I_i, $1 \leq i \leq k$. From this k-colorful augmentation, we can define an M-augmenting path P' with respect to M in G' as $P = \{v'_1, u_1, v_{c_1}, u_2, v_{c_2}, u_3, \ldots, v_{c_{k-1}}, u_k, v_k, z\}$ in which $c_i = c(v_i)$. Recall that $v(K, c(v'_1)) = 0$ and $v(K, c(v_k)) \geq 2$. In conclusion, finding a k-colorful augmentation with respect to a maximum clique K in G^c is equivalent to finding an augmenting path P' with respect to M in G'. □

Our algorithm for MCCP in vertex-colored complete multipartite graphs derives immediately from Theorem 3 and Lemma 6.

The following theorem shows that Algorithm 3 runs in polynomial time.

Theorem 4. *Let G^c be a vertex-colored complete multipartite graph. Algorithm 3 computes a maximum colorful clique of G^c in time $O(|\mathcal{C}|M(m+n, n))$, where $O(M(m, n))$ is the time required to find a maximum matching in a general graph with n vertices and m edges.*

Algorithm 3 Maximum colorful clique in vertex-colored complete multipartite graphs.

1: $K \leftarrow$ any maximum clique of G^c /* pick each vertex from each independent set */
2: **while** a k-colorful augmentation w.r.t. K $\mathcal{P}$ is found **do**
3: $K \leftarrow K \oplus \mathcal{P}$ /* replace vertices of $\{v_i\}$ by vertices of $\{v'_i\}$ */
4: **end while**
5: **return** K as a maximum colorful clique

Proof. The initialization step of the algorithm is trivial. Next, after each iteration of the while loop in the algorithm, a new color is included into the new maximum clique. Thus, the maximum number of iterations of the while loop is $|\mathcal{C}|$ (the number of colors of G^c). Inside each iteration of the while loop, we can use Lemma 6 to look for a k-colorful augmentation, which requires $O(M(m,n))$ time. Note that the new graph G'^c has $O(n)$ vertices and $O(n+m)$ edges. In summary, the total running time of the algorithm is $O(|\mathcal{C}|M(m+n,n))$, as claimed. □

3.4 An algorithm for MCCP for complement graphs of cycle graphs

In the final section, we propose another algorithm for the case of complement graphs of cycle graphs. Related theorems and the algorithm can be found in the appendix.

References

[1] Akbari, S., Liaghat, V., Nikzad, A.: Colorful paths in vertex coloring of graphs. the electronic journal of combinatorics 18(1), P17 (2011)

[2] Akiyama, T., Nishizeki, T., Saito, Nobuji : NP-completeness of the Hamiltonian cycle problem for bipartite graphs. Journal of Information Processing 3 (2), 73–76 (1979)

[3] Bertossi, A.A.: Finding hamiltonian circuits in proper interval graphs. Information Processing Letters 17(2), 97–101 (1983)

[4] Bruckner, S., Hüffner, F., Komusiewicz, C., Niedermeier, R.: Evaluation of ILP-based approaches for partitioning into colorful components. In: International Symposium on Experimental Algorithms. pp. 176–187 (2013)

[5] Cohen, J., Manoussakis, Y., Pham, H., Tuza, Z.: Tropical matchings in vertex-colored graphs. In: Latin and American Algorithms, Graphs and Optimization Symposium (2017)

[6] Cohen, J., Italiano, G., Manoussakis, Y., Thang, N.K., Pham, H.P.: Tropical paths in vertex-colored graphs. In: Proc. 11th Conference on Combinatorial Optimization and Applications (2017, to appear)

[7] Corel, E., Pitschi, F., Morgenstern, B.: A min-cut algorithm for the consistency problem in multiple sequence alignment. Bioinformatics 26(8), 1015–1021 (2010)

[8] Fellows, M.R., Fertin, G., Hermelin, D., Vialette, S.: Upper and lower bounds for finding connected motifs in vertex-colored graphs. Journal of Computer and System Sciences 77(4), 799–811 (2011)
[9] Foucaud, F., Harutyunyan, A., Hell, P., Legay, S., Manoussakis, Y., Naserasr, R.: Tropical homomorphisms in vertex-coloured graphs. Discrete Applied Mathematics (to appear)
[10] Garey, M. R., Johnson, D. S., Stockmeyer, L: Some simplified NP-complete problems. In: 6th ACM Symposium on Theory of Computing (STOC '74). 47–63 (1974)
[11] Golumbic, M.C.: Algorithmic Graph Theory and Perfect Graphs. Annals of Discrete Mathematics 57,(2004)
[12] Italiano, G., Manoussakis, Y., Thang, N.K., Pham, H.P.: Maximum colorful cycles in vertex-colored graphs, submitted (2018)
[13] Li, H.: A generalization of the Gallai–Roy theorem. Graphs and Combinatorics 17(4), 681–685 (2001)
[14] Lin, C.: Simple proofs of results on paths representing all colors in proper vertex-colorings. Graphs and Combinatorics 23(2), 201–203 (2007)
[15] Louis Ibarra: A simple algorithm to find Hamiltonian cycles in proper interval graphs. Information Processing Letters 109(18), 1105–1108 (2009)
[16] Louis Ibarra: The clique-separator graph for chordal graphs. Discrete Applied Mathematics 157(8), 1737–1749 (2009)
[17] Mahadev, N.V.R., Peled, U.N.: Longest cycles in threshold graphs. Discrete Mathematics 135(1-3), 169–176 (1994)
[18] Marx, D.: Graph colouring problems and their applications in scheduling. Periodica Polytechnica Electrical Engineering 48(1-2), 11–16 (2004)
[19] Micali, S., Vazirani, V.V.: An $O(\sqrt{|V|}|E|)$ algorithm for finding maximum matching in general graphs. In: Proc. 21st Symposium on Foundations of Computer Science. pp. 17–27 (1980)
[20] Uehara, R., Valiente, G.: Linear structure of bipartite permutation graphs and the longest path problem. Information Processing Letters 103(2), 71–77 (2007)