SEMI-BALANCED GRAPH COLORINGS

JESPER JANSSON

Master's thesis
2002:E39

LUND UNIVERSITY

Centre for Mathematical Sciences
Mathematics

CENTRUM SCIENTIARUM MATHEMATICARUM

Abstract

We generalize the concept of a 2-coloring of a graph to what we call a *semi-balanced coloring* by relaxing a certain discrepancy condition on the shortest-paths hypergraph of the graph. Let G be an undirected, unweighted, connected graph with n vertices and m edges. We prove that the number of different semi-balanced colorings of G is: (1) at most $n+1$ if G is bipartite; (2) at most m if G is non-bipartite and triangle-free; and (3) at most $m+1$ if G is non-bipartite. Based on the above combinatorial investigation, we design an algorithm to enumerate all semi-balanced colorings of G in $O(nm^2)$ time.

Sammanfattning (in Swedish)

Vi generaliserar begreppet 2-färgning av en graf till något som vi kallar en *semi-balanserad färgning* genom att relaxera ett visst diskrepansvillkor på kortaste-vägarna-hypergrafen till grafen. Låt G vara en oriktad, oviktad, sammanhängande graf med n hörn och m kanter. Vi bevisar att antalet olika semi-balanserade färgningar av G är: (1) högst $n+1$ om G är bipartit; (2) högst m om G är icke-bipartit och triangelfri; och (3) högst $m+1$ om G är icke-bipartit. Vi presenterar även en algoritm baserad på ovanstående kombinatoriska resultat för att räkna upp alla semi-balanserade färgningar av G på $O(nm^2)$ tid.

Acknowledgments

The material presented in this Master of Science thesis in Mathematics also appears in the article:

- J. Jansson and T. Tokuyama. Semi-Balanced Colorings of Graphs – Generalized 2-Colorings Based on a Relaxed Discrepancy Condition. Submitted to *Graphs and Combinatorics.*

which was written in 2001–2002 together with Takeshi Tokuyama from Tohoku University, Sendai, Japan. The research was initiated when I visited the Laboratory for Design and Analysis of Information Systems at the Graduate School of Information Sciences (GSIS), Tohoku University, for three months in the autumn of 2001 and completed in 2002 at Lund University, Sweden.

I would like to thank Takeshi Tokuyama for taking good care of me in Sendai, for all valuable and interesting discussions, and for his kind permission to include our joint work in this thesis. A big arigato to all members of the laboratory for providing such a stimulating and friendly atmosphere; from the very first day of my stay, I felt lucky to be part of such a special environment. I would also like to thank Gert Almkvist for inspiration and helpful comments. Finally, I thank Sweden-Japan Foundation for financial support.

Contents

1 Introduction

Given a set V, a *coloring of* V is a mapping from V to $\{-1, 1\}$. For a graph $G = (V, E)$, a coloring π of the vertex set V is called a *2-coloring of* G if $\pi(x) \neq \pi(y)$ for every edge $\{x, y\}$ in E. We call a vertex which has been mapped to 1 (resp. -1) a *red* (resp. *blue*) vertex. A graph has a 2-coloring if and only if it is bipartite; in fact, by symmetry, a bipartite graph always has two different 2-colorings. A natural way to extend 2-colorings is by allowing k colors to be used, where k is any positive integer. Such a coloring is called a *k-coloring of G*. The number of possible k-colorings of a graph is given by its *chromatic polynomial*, and has been studied extensively (see, e.g., [11] or [13]). It is well known that for any fixed $k \geq 3$, a connected graph can have an exponential number (in $|V|$) of different k-colorings and that the problem of determining if a given graph has any k-colorings at all is NP-complete [13].

Another way to generalize 2-colorings is by relaxing the restriction on two adjacent vertices never being allowed to have the same color. For this purpose, we view the 2-coloring condition as a discrepancy condition on a special hypergraph induced by the graph called the *shortest-paths hypergraph*.

Discrepancy conditions

The red and blue vertices along any path in a 2-colored bipartite graph are always arranged in an alternating fashion. Thus, $-1 \leq \sum_{v \in P} \pi(v) \leq 1$ must hold for the set P of vertices on any path in the graph. This can be regarded as a discrepancy condition.

Discrepancy is a popular measure of uniformity and the quality of approximations, and has been used in combinatorics, geometry, and Monte-Carlo simulations [5, 9, 10]. It is defined as follows. Let $H = (V, \mathcal{F})$ be a hypergraph, where $\mathcal{F} \subseteq 2^V$. Given a coloring π of V, let $\pi(F) = \sum_{v \in F} \pi(v)$ for every $F \in \mathcal{F}$ and let $D_c(H, \pi) = \max_{F \in \mathcal{F}} |\pi(F)|$. The *combinatorial* (or *homogeneous*) *discrepancy* $D_c(H)$ is defined as $D_c(H) = \min_\pi D_c(H, \pi)$, where the minimum is taken over all possible colorings of V. In particular, if $D_c(H, \pi) \leq 1$ then π yields a coloring which is uniform in every hyperedge; this means that $-1 \leq \pi(F) \leq 1$ for every $F \in \mathcal{F}$. Such a π is called a *balanced coloring* of H. We call π a *semi-balanced coloring* of H if $-1 \leq \pi(F) \leq 2$ for every $F \in \mathcal{F}$. The *shortest-paths hypergraph induced by* G is the hypergraph $\mathcal{H}(G) = (V, \mathcal{P}_G)$, where $\mathcal{P}_G$ is the set of all shortest-path vertex sets in G. A 2-coloring of G is equivalent to a balanced coloring of $\mathcal{H}(G)$; hence, we generalize 2-colorings of G by considering semi-balanced colorings of $\mathcal{H}(G)$. A balanced (semi-balanced) coloring of $\mathcal{H}(G)$ is also called a balanced (semi-balanced) coloring of G.

Relation to a rounding problem

Our motivation for studying semi-balanced colorings comes from a conjecture in [2] called the *rounding conjecture*. Given a hypergraph $H = (V, \mathcal{F})$, where $\mathcal{F} \subseteq 2^V$, along with a real-valued function $\alpha : V \to [0,1]$, a *rounding of* α is any function from V to $\{0,1\}$. For every rounding β of α, define the linear discrepancy $D_\ell(H, \alpha, \beta) = \max_{F \in \mathcal{F}} |\alpha(F) - \beta(F)|$, where $\alpha(F) = \sum_{v \in F} \alpha(v)$ and $\beta(F) = \sum_{v \in F} \beta(v)$. Roundings with low linear discrepancy have several applications, including digital halftoning [1, 2, 3, 6, 12]. If, for a rounding β of α, it holds that $D_\ell(H, \alpha, \beta) < 1$ then β is called a *global rounding of* α *in* H. When $\mathcal{F} = \mathcal{P}_G$ for a graph G with real-valued vertex weights, a global rounding approximates the vertex weights by integral vertex weights so that the weight sum of each shortest path becomes either floor or ceiling of the original weight sum.

Now, the rounding conjecture states that if $G = (V, E)$ is a connected graph with n vertices and α is a function $V \to [0,1]$ then there are at most $n+1$ global roundings of α in the shortest-paths hypergraph $\mathcal{H}(G) = (V, \mathcal{P}_G)$, regardless of α. The rounding conjecture has been proved for some special types of graphs: if G is a path then $\mathcal{P}_G$ is a set of intervals; the corresponding rounding problem was studied by Sadakane *et al.* in [12]. This is a natural extension of the fact that a single real number (i.e., the case $n = 1$) has at most two roundings (floor and ceiling). The conjecture has also been proved for cycles, meshes, trees, and trees of cycles [2]. However, it seems difficult to prove in general, and it will be helpful to investigate other special cases. One such case is when the input α is restricted to $\alpha_{U+}(v) = 1/2 + \epsilon$ for every $v \in V$, where $0 < \epsilon < 1/n$; then, the number of global roundings in $\mathcal{H}(G)$ is precisely the number of semi-balanced colorings of G [1]. Thus, although our results so far on semi-balanced colorings provide weak evidence in support of the rounding conjecture, we hope that they will give some insight. Moreover, our algorithm in Section 6 might be useful when searching for a counterexample to the rounding conjecture.

Independent sets

If we only require that no blue vertices are adjacent to each other, the problem of 2-coloring the graph becomes equivalent to the problem of finding an independent set (also called a stable set) in the graph since any set of blue vertices then forms an independent set of G. However, the number of different independent sets of G can be very large, and we usually want one which satisfies some additional restrictions: the *maximum* independent set problem and the *minimum maximal* independent set problem are famous examples where the additional restriction is basically quantitative. These two optimization problems cannot be solved exactly, or even approximated within a factor of $|V|^{1-\epsilon}$ for any constant $\epsilon > 0$, in polynomial time unless some hypotheses concerning the computational hierarchy widely

[1] Given a rounding β of α_{U+}, define β' as $\beta'(v) = 2\beta(v) - 1$ for every $v \in V$. Then β is a global rounding in $\mathcal{H}(G)$ if and only if β' is a semi-balanced coloring of G.

believed to be true turn out to be false [4, 7, 8]. But by restricting the set of valid colorings to semi-balanced colorings, we obtain a class of independent sets with an imposed structural restriction which not only allows one member to be computed efficiently but in fact *all* members to be enumerated in polynomial time.

Although a semi-balanced coloring of G does not always exist, for any independent set W in G there is a supergraph G' of G obtained by adding suitable edges so that W becomes the set of blue vertices in one of the semi-balanced colorings of G'. Thus, the set of independent sets of G corresponds to the union of sets of semi-balanced colorings of supergraphs of G, yielding a covering structure of the set of independent sets. This further motivates us to study combinatorics and algorithms for semi-balanced graph colorings.

Our results

We show that if G is a connected graph with n vertices and m edges, then the number $\nu(G)$ of semi-balanced colorings of G is always polynomial in n. More precisely, we prove that $\nu(G)$ is: (1) at most $n+1$ if G is bipartite; (2) at most m if G is non-bipartite and triangle-free; and (3) at most $m+1$ if G is non-bipartite. Moreover, the semi-balanced colorings of G can be enumerated in $O(nm^2)$ time; thus, this version of the independent set problem is polynomial-time solvable.

2 Preliminaries

Let $H = (V, \mathcal{F})$ be a hypergraph, where $\mathcal{F} \subseteq 2^V$. A *coloring of* H is a mapping from V to $\{-1, 1\}$. For any coloring π of H and any $F \in \mathcal{F}$, let $\pi(F) = \sum_{v \in F} \pi(v)$.

Definition 2.1 *A coloring π of H is called a* balanced coloring of H *if for every $F \in \mathcal{F}$, it holds that:*

$$-1 \ \leq \ \pi(F) \ \leq \ 1.$$

π is called a semi-balanced coloring of H *if for every $F \in \mathcal{F}$, it holds that:*

$$-1 \ \leq \ \pi(F) \ \leq \ 2.$$

For the rest of this thesis, let $G = (V, E)$ be an undirected, unweighted, connected graph with n vertices and m edges.

Consider a path $\mathbf{p}$ in G connecting two vertices u and v. The set of all vertices on $\mathbf{p}$ (including u and v) is called the *vertex set of* $\mathbf{p}$ and is denoted by $F(\mathbf{p})$. If $\mathbf{p}$ is a shortest path between u and v, then $F(\mathbf{p})$ is a *shortest-path vertex set.* There may exist several different shortest paths between u and v, and hence each pair of vertices induces one or more shortest-path vertex sets. For any two vertices $u, v \in V$, $dist(u, v)$ denotes the length of a shortest path in G between u and v.

Given G, the *shortest-paths hypergraph induced by* G is the hypergraph $\mathcal{H}(G) = (V, \mathcal{P}_G)$, where $\mathcal{P}_G$ is the set of all shortest-path vertex sets in G. Our focus is on the semi-balanced colorings of $\mathcal{H}(G)$.

Definition 2.2 *A* coloring[2] of G *is a mapping* $\pi : V \to \{-1, 1\}$. *A vertex* v *in* V *is said to be colored* red *if* $\pi(v) = 1$, *or* blue *if* $\pi(v) = -1$.

A balanced (semi-balanced) coloring of the shortest-paths hypergraph $\mathcal{H}(G)$ is also called a balanced (semi-balanced) coloring of G.

Definition 2.3 *Let* π *be a coloring of* G *and* $\{u, v\} \in E$. *The edge* $\{u, v\}$ *is called* dangerous in π *if* $\pi(u) = \pi(v) = 1$, *i.e., if both of* u *and* v *are colored red.*

We say that an edge is "dangerous" rather than "dangerous in π" when there is no confusion about which coloring is being referred to.

Observation. *A balanced coloring can not contain any dangerous edges. Similarly, if* $\{u, v\} \in E$ *then a coloring in which both* u *and* v *are colored blue can never be a semi-balanced coloring of* G. *Furthermore, in any semi-balanced coloring, a shortest path between two vertices cannot include two dangerous edges.*

Definition 2.4 $\nu(G)$ *is the number of different semi-balanced colorings of* G.

It is easy to calculate $\nu(G)$ for certain types of graphs. For example, $\nu(G) = n+1$ if G is a tree since any semi-balanced coloring of a tree can have at most one dangerous edge and G has $n-1$ edges, and there are exactly two balanced colorings of G. Also, $\nu(G) = n+1$ if G is a complete graph because a semi-balanced coloring of a complete graph can have at most one blue vertex. If G is a cycle of length n, then $\nu(G) = 4$ if $n = 3$, $\nu(G) = n$ if n is odd and $n \geq 5$, $\nu(G) = n/2 + 2$ if $n \equiv 2$ (modulo 4), and $\nu(G) = 2$ if $n \equiv 0$ (modulo 4).

Not all graphs admit semi-balanced colorings. Figure 1 shows one such graph.

However, if we add an edge between the leftmost and the rightmost vertices in the graph in Figure 1, the coloring which makes the top and bottom vertices blue becomes a semi-balanced coloring. In general:

Proposition 2.5 *For any independent set* W *of* $G = (V, E)$, *there is a graph* $G' = (V, E')$ *such that* $E' \supseteq E$ *and* W *is the set of blue vertices in a suitable semi-balanced coloring of* G'.

[2] In traditional graph coloring terminology, a coloring of a graph often requires that adjacent vertices are assigned different colors; here, we refer to such a coloring as a balanced coloring in order to generalize the concept.

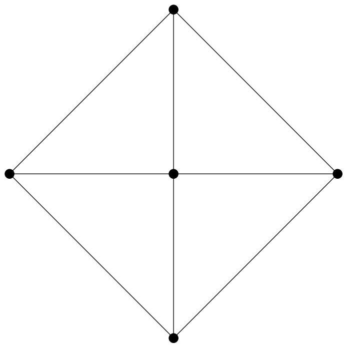

Figure 1: This graph has no semi-balanced coloring.

Proof: For every pair of vertices u and v in $V \setminus W$, if u and v are nonadjacent then add an edge between u and v. Call the resulting graph G'. Note that W is still an independent set in G'. Let π be the coloring of G' in which all vertices in W are colored blue and the rest red. Consider a shortest path $\mathbf{p}$ in G' between any two vertices u and v. If u and v belong to $V \setminus W$, then $\mathbf{p}$ consists of a single dangerous edge and $\pi(\mathbf{p}) = 2$. If one of u and v belongs to W and the other to $V \setminus W$, then $\mathbf{p}$ contains one blue vertex and one or two red vertices, i.e., $\pi(\mathbf{p}) = 0$ or 1. Similarly, if both of u and v belong to W, then $\pi(\mathbf{p}) = -1$ or 0 since no path contains two consecutive blue vertices. Hence, π is a semi-balanced coloring of G'. □

The rest of this thesis is devoted to proving the following enumerative combinatorial result and then designing a polynomial-time enumeration algorithm based upon it.

Theorem 2.6 *Let G be an undirected, unweighted, connected graph with n vertices and m edges. If G is bipartite, $\nu(G) \leq n+1$. If G is not bipartite, $\nu(G) \leq m+1$; moreover, if G is triangle-free, $\nu(G) \leq m$.*

Remark: Definition 2.1 is not generalized to allow $\pi(F)$ to belong to arbitrary intervals in order to guarantee that any graph has only a *polynomial* number (in its input size) of semi-balanced colorings. Generalizing to $[-2, 2]$ means that there can be an exponential number of valid colorings: take, e.g., the complete graph on n vertices; each vertex can be either blue or red, so there are 2^n such colorings. Similarly, generalizing to $[-1, 3]$ means that the star graph with n vertices has $2^{n-1} + 1$ valid colorings.

3 The bipartite case

We start with an upper bound on the number of semi-balanced colorings of a bipartite graph.

Proposition 3.1 *If G is a bipartite graph, then $\nu(G) \leq n+1$.*

Proof: Fix a spanning tree S of G. Any semi-balanced coloring of G is either a balanced coloring of S or a coloring of S with one or more dangerous edges. For each edge e in S, we claim that there is at most one semi-balanced coloring of G that makes e dangerous.

Suppose $e = \{u, v\} \in S$ is dangerous in a semi-balanced coloring of G. Since G is bipartite, it contains no odd cycles. Therefore, there is no vertex whose shortest distance in G to u equals its shortest distance in G to v. Thus, we can divide the vertices into two disjoint sets V_u and V_v so that V_u contains all vertices which are closer to u than v in G, and analogously for V_v. Let T_u and T_v be two shortest path trees (in G) of V_u and V_v rooted at u and v, respectively. We claim there is no dangerous edge in $T_u \cup T_v$. Assume that an edge $\{x, y\} \in T_u$ is dangerous, where x is the father of y in T_u. Let $\mathbf{p}$ be the path from u to y in T_u. Since $dist(u, y) < dist(v, y)$ and every edge in a path contributes 1 to its length, the path appending e to $\mathbf{p}$ is a shortest path between y and v. But this path has two dangerous edges, which is a contradiction. Thus, T_u and T_v must be colored in an alternating fashion (each node in T_u is colored red or blue depending on if its distance from u is even or odd, and similarly for T_v). This shows that there is a unique (if any) semi-balanced coloring of G in which e is dangerous.

Since S has $n-1$ edges, there are at most $n-1$ semi-balanced colorings of G which make at least one edge of S dangerous. S is a tree, so there are exactly two balanced colorings of S. Thus, we obtain the proposition. $\square$

4 The non-bipartite, triangle-free case

In this section, we assume that G is non-bipartite and triangle-free[3]. Although the triangle-free case is a special case, we investigate it in detail since it helps the reader understand our tools and strategy.

The following dominating relation is our key tool. It will be utilized later in an extended form for the case of general graphs.

Definition 4.1 (Dominating relation between edges) *For a pair of edges e, f in E, we say that e* dominates *f if we can write $e = \{u, r\}$ and $f = \{v, w\}$ so that $dist(r, v) = dist(r, w) = k$ and $dist(u, v) = dist(u, w) = k+1$, where k is an even integer. We denote by $e > f$ that e dominates f.*

[3] Triangle-free means that if two edges $\{u, v\}$ and $\{v, w\}$ belong to G then G cannot contain the edge $\{u, w\}$.

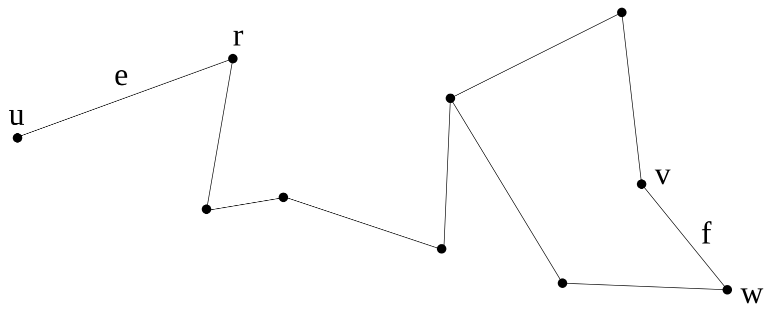

Figure 2: Edge e dominates edge f.

See Figure 2 for an example.

Lemma 4.2 *Let $e, f \in E$. If e is dangerous in a semi-balanced coloring π and $e > f$, then f is also dangerous in π.*

Proof: Let $e = \{u, r\}$ and $f = \{v, w\}$, where r is closer than u to f. Consider a shortest path $\mathbf{p}$ from r to v. By Definition 4.1, the path appending e to $\mathbf{p}$ is a shortest path from u to v. Hence, if there is a dangerous edge on $\mathbf{p}$, it contradicts the semi-balanced condition. Thus, the vertices along $\mathbf{p}$ are colored in an alternating fashion. Since $dist(r, v)$ is even, v has the same color as r, namely red. Similarly, w must be colored red, and hence f is dangerous. $\square$

Definition 4.3 *$D(G)$, the* dominance graph *of G, is a directed graph whose vertices are in one-to-one correspondence with the edges of G. For any two edges $e, f \in E$, there is a directed edge from e to f in $D(G)$ if and only if $e > f$.*

Given a coloring π of G, a vertex of $D(G)$ is called *dangerous in π* if the corresponding edge in G is dangerous in π.

Next, consider the decomposition of $D(G)$ into strongly connected components $C_1, C_2, \ldots, C_h$.

Corollary 4.4 *If a vertex in a strongly connected component C_i is dangerous in a semi-balanced coloring π, then all vertices belonging to C_i are dangerous in π. Furthermore, all elements in its transitive closure in $D(G)$ are also dangerous.*

To find an upper bound on $\nu(G)$, we need one more definition.

Definition 4.5 *For an edge e of E, a* regular coloring associated with e *is a semi-balanced coloring which makes all the vertices in the strongly connected component of $D(G)$ containing e dangerous and no other vertex dominating e dangerous.*

Lemma 4.6 *If G is not bipartite then any semi-balanced coloring of G is a regular coloring associated with some edge e in E.*

Proof: Let π be a semi-balanced coloring of G. π cannot be balanced since G is non-bipartite, so there is at least one dangerous edge. Let $E(\pi)$ be the set of dangerous edges. Consider the subgraph H induced by $E(\pi)$ in $D(G)$. By Corollary 4.4, H must be the union of some strongly connected components. Consider H as a directed acyclic graph on these strongly connected components, and pick a source component C. Then, for any element e in C, π is a regular coloring. □

Lemma 4.7 *For each edge $e \in E$, there is at most one regular coloring associated with e.*

Proof: Write $e = \{r, p\}$ and let C be the set of vertices in $D(G)$ which are reachable from e. We call the edges in G which are represented by vertices of C *predicted edges*. Construct a shortest path tree T of G rooted at r. In the construction of T, whenever two or more paths in G to the same vertex are of equal length, we apply a convention that a path containing a predicted edge is preferred; if two or more paths contain predicted edges, the one in which the predicted edge is nearer to r is preferred. For any vertex v, let $path(r, v)$ be the path in T from r to v.

Consider a regular coloring associated with e. By definition, e and all other predicted edges are dangerous. Below, we show that for every edge belonging to T, it is dangerous only if it is a predicted edge, implying that the set of vertices colored red is uniquely determined.

Suppose there exists a dangerous edge in T which is not predicted. Let $f = \{s, t\}$ be the nearest to r among such edges. Neither of s and t can be equal to r since otherwise f and e are adjacent dangerous edges, which is impossible in a semi-balanced coloring of a triangle-free graph. Without loss of generality, assume that s is the parent node of t in T, and let $k = dist(r, s)$. Note that k is even; otherwise, there would be another dangerous edge on $path(r, s)$, and the shortest path $path(r, t)$ would contain two dangerous edges.

Consider the path appending e to $path(r, t)$. This path has length $k + 2$, and contains two dangerous edges; hence, it cannot be a path with the shortest length. Thus, there is a path $\mathbf{p}$ in G between p and t whose length is less than $k + 2$. If it is less than or equal to k, the path obtained by appending e to $\mathbf{p}$ has length at most $k + 1$ from r to t. Moreover, it is preferred to the current path $path(r, t)$ in the lexicographic ordering; thus, we have a contradiction. Therefore, the path $\mathbf{p}$ has length $k + 1$. If $\mathbf{p}$ contains the edge f, then f dominates e. But because the coloring is a regular coloring associated with e, f can be dangerous only if f is in C and hence predicted.

Thus, we assume that $\mathbf{p}$ does not contain f. Since $\mathbf{p}$ has odd length, $\mathbf{p}$ has a dangerous edge $g = \{u, v\}$. We assume that u is nearer than v to p on the path. If $v = t$, we can again derive a contradiction because then g and $\{s, t\}$ are adjacent dangerous edges. Hence, $v \neq t$. The length ℓ of the path from p to u must be even since we cannot have a dangerous edge on $\mathbf{p}$ in the part from p to u, and both p and u are colored red. Consider the path $path(r, v)$ in T. The length of $path(r, v)$ must be $\ell + 1$ (if it is less than or equal to ℓ then the path connecting $path(r, v)$ to the part of $\mathbf{p}$ from v to t has length k or less, which contradicts that $path(r, t)$ is the shortest; if it is greater than or equal to $\ell + 2$ then the path appending e to the part of $\mathbf{p}$ from p to v is a shortest path with two dangerous edges). Thus, $path(r, v)$ has odd length, and hence contains a dangerous edge. If it is not predicted, it contradicts that f is the nearest edge among non-predicted dangerous edges on T. If it is predicted, the path connecting $path(r, v)$ and the part of $\mathbf{p}$ from v to t has length $k + 1$, and it is preferred to the current path $path(r, t)$, which is a contradiction.

Thus, we have proved that all dangerous edges on T are predicted, giving a unique way (if one exists) of assigning colors to the nodes of T. □

Proposition 4.8 *If G is a non-bipartite, triangle-free graph, then $\nu(G) \leq m$.*

Proof: G is non-bipartite, so any semi-balanced coloring of G must be a regular coloring associated with some edge in E by Lemma 4.6. Next, Lemma 4.7 implies that there are at most m semi-balanced colorings of G. □

5 The general non-bipartite case

In this section, we assume that G is non-bipartite.

A clique Q of G is called *maximal* if there is no other clique in G containing Q. A clique is called *submaximal* if it has at least two vertices and it is contained in a maximal clique which has one more vertex. For a clique Q in G, V_Q denotes the set of vertices of Q. The following lemma is immediate since two blue vertices can never be adjacent in a semi-balanced coloring.

Lemma 5.1 *Let Q be a maximal clique in a graph G. In any semi-balanced coloring of G, there is at most one vertex in V_Q colored blue.*

In a coloring of G, Q is called a *dangerous clique* if all vertices in V_Q are red. Lemma 5.1 implies that in any semi-balanced coloring, every maximal clique of at least three vertices is either dangerous or has a dangerous submaximal clique.

Lemma 5.2 *Let Q_1 and Q_2 be a pair of maximal cliques in G and let $W = V_{Q_1} \cap V_{Q_2}$. In any semi-balanced coloring of G, the following holds:*

(1) If $|W| \geq 2$, all vertices in W must be colored red.

(2) If $W = \{w\}$ and $|V_{Q_1}| \geq 3$ and $|V_{Q_2}| \geq 3$, the vertex w must be colored red if there is an edge between $V_{Q_1} - W$ and $V_{Q_2} - W$; otherwise, it must be colored blue.

(3) If $W = \{w\}$ and $|V_{Q_1}| \geq 3$ and $|V_{Q_2}| = 2$, the clique (indeed, the edge) Q_2 cannot be dangerous.

Proof: (1) Assume some vertex in W is colored blue. Then all other vertices in $V_{Q_1} \cup V_{Q_2}$ must be colored red. Since $|W| \geq 2$, there is a vertex w in W colored red. For any pair of vertices $x \in V_{Q_1} - W$ and $y \in V_{Q_2} - W$, the path $x \to w \to y$ cannot be a shortest path; hence the edge $\{x, y\}$ must be in the graph G. This means that G contains the complete bipartite graph between $V_{Q_1} - W$ and $V_{Q_2} - W$, and thus it has a complete graph on $V_{Q_1} \cup V_{Q_2}$, which is a contradiction.

(2) If there is a vertex $x \in V_{Q_1} - W$ and a vertex $y \in V_{Q_2} - W$ such that $\{x, y\}$ is an edge of G then the vertices $\{w, x, y\}$ constitute a triangle which is contained in a maximal clique Q_3 intersecting Q_1 at x and w. Therefore, w must be colored red by (1) above. If there are no edges between $V_{Q_1} - W$ and $V_{Q_2} - W$, then there are two red vertices $x \in V_{Q_1} - W$ and $y \in V_{Q_2} - W$ with $x \to w \to y$ being a shortest path from x to y. Hence, w must be colored blue.

(3) Suppose that Q_2 is an edge $\{w, y\}$. There is no edge between y and a vertex x in $V_{Q_1} - W$; otherwise, we have a triangle with node set $\{w, x, y\}$, contradicting the maximality of Q_2. Thus, for any vertex x in $V_{Q_1} - W$, the path $x \to w \to y$ is a shortest path. If Q_2 is dangerous, w and y are colored red so x must be colored blue. However, there can be at most one blue vertex in Q_1, yielding a contradiction since $V_{Q_1} - W$ contains at least two vertices. □

Because of Lemma 5.2, if two maximal cliques intersect at two or more vertices, we can fix the colors of all vertices in the intersection. Also, if two maximal cliques of size at least three intersect at one vertex, we can fix the color of that vertex.

We first remove those vertices and their incident edges from G. For any maximal clique Q, let $\tilde{Q}$ be the remaining part. Next, for each maximal clique Q of size two (i.e., edge) intersecting another clique of size greater than two, we set $\tilde{Q} = \emptyset$ and remove the corresponding edge but keep both endpoints of the edge if they have not been removed so far. Thus, we obtain a subgraph $\tilde{G}$ of G.

For a submaximal clique R in a maximal clique Q, $\tilde{R}$ denotes $R \cap \tilde{Q}$. We call $\tilde{Q}$ a *restricted clique* if Q is either maximal or submaximal.

Observe that if we give a coloring of $\tilde{Q}$ for each maximal clique Q having at least three vertices and determine the set of dangerous edges (i.e., red-colored cliques of size two), the coloring of G is uniquely determined.

To count the number of colorings, we use a function t defined below. For a maximal clique Q of size at least 3, there are at most $t(Q)$ ways to color the vertices in $\tilde{Q}$ so that $\tilde{Q}$ or some submaximal clique of $\tilde{Q}$ is dangerous.

Definition 5.3 *Let Q be a maximal clique with k vertices in $\tilde{Q}$. Define $t(Q)$ as:*

$$t(Q) = \begin{cases} k+1, & \text{if } Q \text{ has three or more vertices and } \tilde{Q} \neq \emptyset \\ 1, & \text{if } Q \text{ is an edge and } \tilde{Q} = Q \\ 0, & \text{if } Q \text{ is an edge and } \tilde{Q} \text{ has one vertex} \\ 0, & \text{if } \tilde{Q} = \emptyset \end{cases}$$

Lemma 5.4 $\sum_Q t(Q) \leq m+1$, *where the summation is taken over all maximal cliques of G.*

Proof: Assign $m+1$ tokens to G in such a way that every edge of G initially receives one token and one extra token is not assigned to any edge. To prove the lemma, we demonstrate how the tokens can be distributed among the maximal cliques of G so that every maximal clique Q is given $t(Q)$ tokens.

In the graph $\tilde{G}$, $\tilde{Q}$ for a maximal clique Q of size at least three cannot intersect any other restricted cliques and is therefore isolated. First, consider the case $Q \neq \tilde{Q}$. We can assume $|V_Q| \geq 3$ since otherwise $t(Q) = 0$. If $\tilde{Q}$ has k vertices, there are $m_1(Q) = k \cdot (|V_Q| - k)$ edges in G that are adjacent to vertices in $\tilde{Q}$. None of these edges are in another maximal clique since otherwise the endpoints of these edges should have been removed in $\tilde{G}$, too. $\tilde{Q}$ itself has $m_2(Q) = k(k-1)/2$ edges. Thus, we can take $m_1(Q) + m_2(Q)$ tokens from these edges and give to Q. It is easy to see that $m_1(Q) + m_2(Q) \geq t(Q)$.

Next, consider the cases where $Q = \tilde{Q}$. If Q has $k \geq 4$ vertices then $\tilde{Q}$ has $k(k-1)/2 > k+1 = t(Q)$ edges from which tokens can be taken. If Q has two vertices (i.e., Q is an edge e) then $t(Q) = 1$ and we assign to Q the token of the edge e itself. Finally, if Q is a triangle, it has three edges but $t(Q) = 4$. Suppose there are d such triangles; the sum of the corresponding $t(Q)$'s is $4d$. The triangles are isolated in $\tilde{G}$; however, G itself is connected. Hence, we can find $d-1$ edges which are adjacent to the triangles in G and removed in the construction of $\tilde{G}$. These edges' tokens have not previously been assigned to any maximal cliques, so we can use them together with the tokens of the $3d$ edges in the triangles. We then use the extra token, for a total of $(d-1) + 3d + 1 = 4d$ tokens. $\square$

Next, we define a dominating relation among maximal and submaximal cliques which generalizes Definition 4.1. Let $\mathcal{Q}$ be the set of all cliques which are maximal or submaximal.

Definition 5.5 (Dominating relation between cliques) *Let $Q_1, Q_2 \in \mathcal{Q}$. We say that Q_1* dominates *Q_2 if there exists an even integer k such that for every $v \in V_{Q_2}$, there is a vertex $r \in V_{Q_1}$ and a vertex $u \in V_{Q_1}$ for which $dist(r,v) = k$ and $dist(u,v) = k+1$. We write $Q_1 > Q_2$ if Q_1 dominates Q_2.*

Note that by Definition 5.5, a submaximal clique is dominated by a maximal clique containing it.

Lemma 5.6 *If Q_1 is dangerous in a semi-balanced coloring π and $Q_1 > Q_2$, then Q_2 is dangerous in π.*

Proof: Let k be an integer which satisfies the conditions in Definition 5.5. Let v be any vertex in V_{Q_2}, and let r and u be two vertices in V_{Q_1} such that $dist(r, v) = k$ and $dist(u, v) = k + 1$. Since $dist(u, v) = k + 1$, there is at least one shortest path from u to v which contains the edge $\{u, r\}$. See Figure 3. Both u and r are colored red in π, and since no shortest path from u to v can include two dangerous edges and the distance between r and v is even, v must be colored red. Thus, Q_2 is dangerous. □

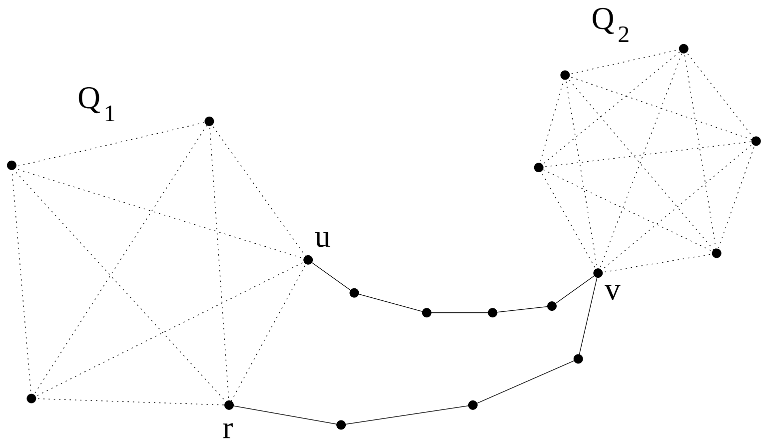

Figure 3: Illustrating the proof of Lemma 5.6.

Definition 5.7 *$D(G)$, the* dominance graph *of G, is a directed graph whose vertices are in one-to-one correspondence with $\mathcal{Q}$. For any two (maximal or submaximal) cliques $Q_1, Q_2 \in \mathcal{Q}$, there is a directed edge from Q_1 to Q_2 in $D(G)$ if and only if $Q_1 > Q_2$.*

Definition 5.8 *$D(\tilde{G})$, the* dominance graph *of $\tilde{G}$, is the directed graph obtained from $D(G)$ by identifying vertices associated with Q_1 and Q_2 if $\tilde{Q}_1 = \tilde{Q}_2$, and removing the vertex associated with Q if $\tilde{Q} = \emptyset$ or $Q - \tilde{Q}$ is known to contain a blue vertex in any semi-balanced coloring. Note that the last case may happen if the vertex is the intersection of two maximal cliques, each of size at least three.*

Corollary 5.9 *If a restricted clique $\tilde{Q}$ is dangerous in a semi-balanced coloring, the restricted cliques in its transitive closure in $D(\tilde{G})$ are also dangerous.*

Definition 5.10 *For a member $\tilde{Q}$ of a strongly connected component C of $D(\tilde{G})$, a* regular coloring associated with $\tilde{Q}$ *is a semi-balanced coloring which makes all restricted cliques in C dangerous and no other restricted clique dominating $\tilde{Q}$ dangerous.*

Lemma 5.11 *If G is not bipartite and if there is a clique Q of size more than two such that $\tilde{Q} \neq \emptyset$, then any semi-balanced coloring is a regular coloring associated with some restricted clique in G.*

Proof: Let π be a semi-balanced coloring of G. π cannot be balanced since G is non-bipartite, so there is at least one dangerous edge. Let $\mathcal{Q}(\pi)$ be the set of dangerous restricted cliques. Consider the subgraph H induced by $\mathcal{Q}(\pi)$ in $D(\tilde{G})$. By Corollary 5.9, H must be the union of some strongly connected components. Consider H as a directed acyclic graph on these strongly connected components, and pick a source component C. Then, for any element in C, π is a regular coloring. $\square$

Lemma 5.12 *For each restricted clique $\tilde{Q}$, there is at most one regular coloring associated with $\tilde{Q}$.*

Proof: Let C be the set of vertices in $D(\tilde{G})$ which are reachable from $\tilde{Q}$. The cliques in G which are represented by vertices of C are called *predicted cliques*, and edges in G belonging to predicted cliques are called *predicted edges*. Let Q be any maximal or submaximal clique whose restriction is $\tilde{Q}$, fix an ordering $\{v_1, v_2, \ldots, v_q\}$ of V_Q, and partition V into q groups $W_1, W_2, \ldots, W_q$ such that a vertex w in V is placed in W_i if v_i has the smallest index among all vertices in V_Q that are nearest to w in G. For every $i \in \{1, 2, \ldots, q\}$, construct a shortest path tree T_i on W_i rooted at v_i. In the construction of T_i, whenever two or more paths to the same vertex in W_i are of equal length, we apply a convention that a path containing a predicted edge is preferred; if two or more paths contain predicted edges, the one in which the predicted edge is nearer to v_i is preferred.

Consider a regular coloring associated with $\tilde{Q}$. By definition, $\tilde{Q}$ and all other predicted cliques are dangerous. Therefore, any edge of G contained in a predicted clique is dangerous. Below, we show that for every edge belonging to a tree T_i, it is dangerous only if it lies in a predicted clique, implying that the set of vertices colored red is uniquely determined.

Suppose there exists at least one dangerous, non-predicted edge in some T_i. Let $\{s, t\}$ be the dangerous, non-predicted edge in T_i which is closest to v_i, and assume without loss of generality that s is the parent node of t in T_i. Denote the distance between v_i and s by d, i.e., $dist(v_i, t) = dist(v_i, s) + 1 = d + 1$. and let $\mathbf{p}$ be the path in T_i from v_i to t. Note that d must be even since otherwise $\mathbf{p}$ has two or more dangerous edges. If $i \neq 1$ then $dist(v_1, s) > dist(v_i, s)$ and $dist(v_1, t) > dist(v_i, t)$ (otherwise, s and t would have been placed in W_1), so $dist(v_1, s) = d + 1$ and $dist(v_1, t) = d+2$. The path in G obtained by appending $\{v_1, v_i\}$ to $\mathbf{p}$ is therefore a shortest path with two dangerous edges, which is a contradiction. Thus, we have $i = 1$.

Next, for every $2 \leq j \leq q$, it holds that: (1) $dist(v_j, t) \leq d$ is impossible (since $t \in W_1$); (2) $dist(v_j, t) \geq d + 3$ is impossible (since there is a path of length $d + 2$ from v_j to t beginning with the edge $\{v_j, v_1\}$); (3) $dist(v_j, t) = d + 2$ is impossible (if $dist(v_j, t) = d + 2$ then the path $v_j \to v_1 \to s \to t$ of length $d + 2$ is a shortest path with two dangerous edges, which is a contradiction). Thus, $dist(v_j, t) = d+1$ for every $v_j \in V_Q$.

Let A be a dangerous maximal clique containing s and t, if one exists; otherwise, let A be a dangerous submaximal clique containing s and t. Consider any shortest path in G from a vertex $v_j \in V_Q$ to t. The path has odd length and consequently includes a dangerous edge $\{x, y\}$, where x is nearer than y to v_j. If $y = t$ then x also belongs to V_A (otherwise, for some $a \in V_A$ there would exist a shortest path $x \to y \to a$ consisting of two dangerous edges), and furthermore, $dist(v_j, x) = d$ and $dist(v_j, y) = d + 1$. Hence by Definition 5.5, if for every $v_j \in V_Q$ there exists a shortest path in G from v_j to t whose dangerous edge contains t, then A dominates $\tilde{Q}$. But then A must belong to C since the coloring is a regular coloring associated with $\tilde{Q}$ and A is dangerous; however, this means that the edge $\{s, t\}$ is contained in a predicted clique, which is a contradiction. Therefore, we may assume there exists a $k \neq 1$ such that there is no shortest path in G from v_k to t whose dangerous edge contains t.

Now, let $\mathbf{q}$ be a shortest path from v_k to t. $\mathbf{q}$ has a dangerous edge $\{x, y\}$, where x is nearer than y to v_k, satisfying $y \neq t$. Denote $dist(v_k, y)$ by ℓ. ℓ must be odd since v_k and y are colored red and the path from v_k to y through x contains exactly one dangerous edge. Let $\mathbf{r}$ be a shortest path from v_1 to y; its length is ℓ since if it was $\ell+1$ then the path consisting of the edge $\{v_1, v_k\}$ together with the shortest path from v_k to y through x would be a shortest path with two dangerous edges, and if $dist(v_j, y) = \ell - 1$ for any $v_j \in V_Q$ then $dist(v_j, t) = d$, which is a contradiction. This also shows that $y \in W_1$. Next, note that there is a dangerous edge on $\mathbf{r}$ since v_1 and y are colored red and ℓ is odd. In particular, the above holds for the shortest path $\mathbf{r}^*$ from v_1 to y consisting of edges in T_1; let e be the dangerous edge on $\mathbf{r}^*$.

The path in G of length $d+1$ from v_1 to t containing e is a shortest path. If e is not predicted, it contradicts the choice of $\{s, t\}$ because $\{s, t\}$ was the dangerous, non-predicted edge in T_1 nearest to v_1. If e is predicted, there is a contradiction because the path from v_1 to t in T_1 contains the non-predicted edge $\{s, t\}$ but a path containing e should have been preferred in the construction of T_1.

Thus, every dangerous edge in a shortest path tree belongs to some predicted clique, giving a unique way (if one exists) of assigning colors to V. $\square$

Now, we can derive our main result:

Proposition 5.13 *If G is a non-bipartite graph, then $\nu(G) \leq m+1$.*

Proof: If there is a clique Q of size more than two with $\tilde{Q} \neq \emptyset$, then every semi-balanced coloring of G must be a regular coloring associated with some restricted clique by Lemma 5.11. By Lemma 5.12, there is at most one regular coloring associated with each restricted clique. Thus, the number of semi-balanced colorings of G is at most $\sum_Q t(Q)$ (if it was greater, there would exist some maximal clique R with at least $t(R)+1$ regular colorings associated with restrictions of it or its submaximal cliques, contradicting that for each maximal clique Q, there are at most $t(Q)$ ways to color the vertices in $\tilde{Q}$ so that $\tilde{Q}$ or some submaximal clique of $\tilde{Q}$ is dangerous), and therefore at most $m+1$ by Lemma 5.4.

Otherwise, there is no clique Q of size more than two satisfying $\tilde{Q} \neq \emptyset$, in which case there may be a semi-balanced coloring of G which makes none of the restricted cliques dangerous. However, there is a most one such coloring since G is non-bipartite. In this case, $\sum_Q t(Q) \leq m$ since $\tilde{G}$ has no triangles. Thus, the number of semi-balanced colorings is bounded by $m+1$. □

6 An algorithm for enumerating semi-balanced colorings

6.1 Testing a coloring

To test if a given coloring satisfies the semi-balanced condition, we first observe the following:

Lemma 6.1 *Let $\mathbf{p}_1$ and $\mathbf{p}_2$ be two shortest paths in G from a vertex s to a vertex t. If π is a semi-balanced coloring of G then $\pi(\mathbf{p}_1)$ must be equal to $\pi(\mathbf{p}_2)$.*

Proof: Let $\mathbf{p}_1$ be the path $s \to a_1 \to a_2 \to \ldots \to a_l \to t$ and $\mathbf{p}_2$ the path $s \to b_1 \to b_2 \to \ldots \to b_l \to t$, and suppose that $\pi(\mathbf{p}_1) \neq \pi(\mathbf{p}_2)$. Assume without loss of generality that $\pi(\mathbf{p}_1) < \pi(\mathbf{p}_2)$. Since $\mathbf{p}_1$ and $\mathbf{p}_2$ have the same length, either both of $\pi(\mathbf{p}_1)$ and $\pi(\mathbf{p}_2)$ are even or both of $\pi(\mathbf{p}_1)$ and $\pi(\mathbf{p}_2)$ are odd. For any shortest path $\mathbf{q}$ in G, it holds that $\pi(\mathbf{q})$ belongs to $\{-1,0,1,2\}$. Thus, we have $\pi(\mathbf{p}_1)+2 = \pi(\mathbf{p}_2)$ and $\pi(\mathbf{p}_1) \in \{-1,0\}$. Set $k = \pi(\mathbf{p}_1)$. There are four cases to consider:

- If $\pi(s) = 1$ and $\pi(t) = 1$ then $\pi(a_1 \to a_2 \to \ldots \to a_l)$ equals $k-2$ which is always less than -1. Contradiction since $a_1 \to a_2 \to \ldots \to a_l$ is a shortest path in G.

- Similarly, if $\pi(s) = -1$ and $\pi(t) = -1$ then $\pi(b_1 \to b_2 \to \ldots \to b_l)$ equals $k+4$ which is always greater than 2. Contradiction.

- If $\pi(s) = 1$ and $\pi(t) = -1$ then $\pi(a_1 \to a_2 \to \dots \to a_l \to t)$ equals $k-1$ and $\pi(s \to b_1 \to b_2 \to \dots \to b_l)$ equals $k+3$. Both $k-1$ and $k+3$ cannot belong to $\{-1, 0, 1, 2\}$. Contradiction.
- The case with $\pi(s) = -1$ and $\pi(t) = 1$ is analogous to the previous case.

□

Next, we have:

Lemma 6.2 *Given a graph $G = (V, E)$ and a coloring π of G, we can decide whether π is a semi-balanced coloring of G or not in $O(nm)$ time.*

Proof: Since the number of shortest paths in G may be exponential in n, a naive algorithm that computes $\pi(\mathbf{p})$ for every hyperedge $\mathbf{p}$ in $\mathcal{P}_G$ is inefficient. Instead, we compute a shortest path tree T_u by breadth-first search in G for each vertex $u \in V$. In the construction of T_u, we associate with each vertex $w \in V$ two values $P[w]$ and $D[w]$. Initially, $P[u] = \pi(u)$ and $D[u] = 0$. Whenever an edge (v, w) is traversed in the breadth-first search, where v is the parent of w, if w has not been visited before then $u \to v \to w$ must be a shortest path and we set $P[w]$ to $P[v] + \pi(w)$ and $D[w]$ to $D[v] + 1$. If w has been visited before and $D[w]$ equals $D[v] + 1$ (i.e., if this is also a shortest path from u) then check if $P[w]$ equals $P[v] + \pi(w)$; if not then we can immediately answer "no" because if there are two shortest paths $\mathbf{p}_1$ and $\mathbf{p}_2$ from u to w and $\pi(\mathbf{p}_1) \neq \pi(\mathbf{p}_2)$ then π is not a semi-balanced coloring by Lemma 6.1. Thus, we do not need to store one value for every possible shortest path. Next, if in every tree T_u it holds that $P[w]$ equals -1, 0, 1, or 2 for every vertex w, we decide that π is semi-balanced and answer "yes", and otherwise answer "no". (In fact, we can stop the process and answer "no" directly if we discover that $P[w] \geq 3$ or $P[w] \leq -2$ for some vertex w in some tree T_u.)

For each vertex u, the breadth-first search from u takes $O(m)$ time and the values of $P[w]$ can be checked in $O(n)$ time, so the total time complexity is $O(nm)$. □

6.2 The enumeration algorithm

Here, we give an algorithm for enumerating all semi-balanced colorings of a given graph in $O(nm^2)$ time. The strategy is to generate a polynomial-sized set of candidate colorings that includes all the semi-balanced colorings and then test each candidate coloring for the semi-balanced condition in polynomial time using the method described in the proof of Lemma 6.2. The algorithm is outlined in Figure 4.

The algorithm first checks if G is bipartite by breadth-first search starting at an arbitrary vertex. If it is bipartite, the two balanced colorings of G are returned and then the proof of Proposition 3.1 provides a method for enumerating all remaining

Algorithm *Enumerate*
Input: An undirected, unweighted, connected graph $G = (V, E)$.
Output: All semi-balanced colorings of G.

if G is bipartite **then**
  Output the two balanced colorings of G.
  Fix a spanning tree S of G.
  For each edge e of S, compute the unique possible semi-balanced coloring of G in which e is dangerous as in the proof of Proposition 3.1 and then check if it is semi-balanced; if yes, output it.
else
  Compute $\tilde{G}$ and color the vertices in $G - \tilde{G}$ according to Lemma 5.2.
  Build $D(\tilde{G})$ from $\tilde{G}$ and locate all restricted cliques in $\tilde{G}$.
  for each restricted clique $\tilde{Q}$ **do**
    Let Q be a maximal or submaximal clique whose restriction is $\tilde{Q}$.
    Partition V into $W_1, ..., W_{|V_Q|}$; compute a shortest path tree T_i for each W_i.
    Color all vertices of $\tilde{G}$ which belong to predicted cliques red (use $D(\tilde{G})$), color the rest by alternating between blue and red while traversing T_i's.
    Check if the resulting coloring is semi-balanced; if yes, output it.
End *Enumerate*

Figure 4: An algorithm for enumerating all semi-balanced colorings of G.

semi-balanced colorings of G; as each candidate coloring is generated, test if it is semi-balanced using Lemma 6.2.

If G is not bipartite, the ideas from Section 5 are employed. The algorithm computes $\tilde{G}$ and assigns colors to all vertices in $G - \tilde{G}$ according to Lemma 5.2. (The colors of these vertices are uniquely determined for all semi-balanced colorings of G, and can therefore be fixed for the remainder of the algorithm.) To implement this step, we proceed as follows.

To start with, we locate intersections having at least two vertices between pairs of maximal cliques. Note that such an intersection must contain an edge.

Lemma 6.3 *An edge e is contained in just one maximal clique if and only if the union of all triangles containing e forms a clique.*

Proof: The if part is trivial. For proving the only-if part, consider a pair of vertices u_1 and u_2 forming triangles Δ_1 and Δ_2 with e, respectively. If $\{u_1, u_2\}$ is not an edge of G, any maximal clique containing Δ_1 must be different from any maximal clique containing Δ_2. □

For each edge e in G, we check whether it is contained in the intersection of two maximal cliques. From Lemma 6.3, it suffices to look at the set of triangles containing e and check whether the set of all vertices of triangles forms a clique or not. This can be done in $O(m)$ time for each edge e. Hence, it takes a total of $O(m^2)$ time.

Then, we check every vertex v not detected to be red in the previous stage if it is the intersection of two maximal cliques, each of size greater than two. The neighbors of v must be a disjoint union of cliques (otherwise, v would have been colored red previously) so this can be carried out in $O(m)$ time, i.e., in $O(nm)$ time in total.

Next, the algorithm builds $D(\tilde{G})$. We need to retrieve a maximal clique Q from a reduced clique $\tilde{Q}$ under the assumption that $\tilde{Q}$ or one of its submaximal cliques is dangerous. Below, we refer to the vertices in $G - \tilde{G}$ which are known to be red as *purple vertices*. Since we do not permit a path of length two consisting of two dangerous edges to become a shortest path, we obtain the following lemma.

Lemma 6.4 *Let Q be a maximal clique and let R be a submaximal clique of Q. If $\tilde{R}$ is dangerous, Q must be the clique consisting of all vertices in $\tilde{Q}$ together with the set of all purple vertices adjacent to R. If $\tilde{Q}$ is dangerous, Q must contain the clique Q' consisting of all vertices in $\tilde{Q}$ together with the set of all purple vertices adjacent to Q; furthermore, Q' is either Q or a submaximal clique of Q.*

Thus, we can retrieve Q or its submaximal clique corresponding to R from $\tilde{R}$ in polynomial time. To construct $D(\tilde{G})$, we have to check the dominating relation between such cliques. For this purpose, we compute the all-pairs shortest distances in G in $O(n^3)$ time; once we have computed the distances, it is possible to construct $D(\tilde{G})$.

Then, regular colorings are generated and tested. The proof of Lemma 5.12 gives a procedure for generating regular colorings. Each candidate coloring is tested using the method described in Lemma 6.2.

Constructing $\tilde{G}$ and $D(\tilde{G})$ takes $O(m^2) + O(nm) + O(n^3) = O(m^2 + n^3)$ time. At most $m+1$ candidate colorings are generated; for each one, it takes $O(nm)$ time to check if it is semi-balanced. Hence, the total time complexity of the algorithm is $O(nm^2)$.

Theorem 6.5 *Given an undirected, unweighted, connected graph with n vertices and m edges, all semi-balanced colorings of G can be enumerated in $O(nm^2)$ time.*

7 Concluding remarks

We have defined and studied the combinatorial concept of a semi-balanced coloring, obtained by generalizing 2-colorings. Motivated by our results, we state the following conjecture. If it is true, it means that $\nu(G)$ is maximized at the two extremes: when G is a tree and when G is a complete graph. (The only graph that the authors know of which satisfies $\nu(G) = m + 1$ is the triangle; incidentally, this graph also satisfies $\nu(G) = n+1$.) The conjecture is a special case of the rounding conjecture mentioned in Section 1.

Conjecture. *For any undirected, unweighted, connected graph G with n vertices, $\nu(G) \leq n+1$.*

As seen in Section 2, there are graphs for which $\nu(G) = 0$. We would like to know if there is some way to characterize all graphs with $\nu(G) > 0$.

Open problem. *Determine the class of graphs for which a semi-balanced coloring always exists.*

The algorithm presented in Section 6 may be a useful tool for resolving these two issues.

References

[1] T. Asano, N. Katoh, K. Obokata, and T. Tokuyama. Matrix rounding under the L_p-discrepancy measure and its application to digital halftoning. In *Proceedings of the 13th Annual ACM-SIAM Symposium on Discrete Algorithms* (SODA 2002), pages 896–904, 2002.

[2] T. Asano, N. Katoh, H. Tamaki, and T. Tokuyama. On the number of global roundings of a shortest-path hypergraph of a connected graph. In preparation.

[3] T. Asano, T. Matsui, and T. Tokuyama. Optimal roundings of sequences and matrices. *Nordic Journal of Computing*, 7:241–256, 2000.

[4] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, and M. Protasi. *Complexity and Approximation*. Springer-Verlag, 1999.

[5] J. Beck and V. T. Sös. Discrepancy theory. In R. Graham, M. Grötschel, and L. Lovász, editors, *Handbook of Combinatorics*, volume II, pages 1405–1446. Elsevier, 1995.

[6] B. Doerr. Lattice approximation and linear discrepancy of totally unimodular matrices. In *Proceedings of the 12th Annual ACM-SIAM Symposium on Discrete Algorithms* (SODA 2001), pages 119–125, 2001.

[7] M. M. Halldórsson. Approximating the minimum maximal independence number. *Information Processing Letters*, 46(4):169–172, 1993.

[8] J. Håstad. Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Mathematica*, 182:105–142, 1999.

[9] J. Matoušek. *Geometric Discrepancy: An illustrated guide*, volume 18 of *Algorithms and Combinatorics*. Springer-Verlag, 1999.

[10] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*, volume 63 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, 1992.

[11] R. Read and W. Tutte. Chromatic polynomials. In L. Beineke and R. Wilson, editors, *Selected Topics in Graph Theory*, volume 3, pages 15–42. Academic Press, 1988.

[12] K. Sadakane, N. Takki-Chebihi, and T. Tokuyama. Combinatorics and algorithms on low-discrepancy roundings of a real sequence. In *Proceedings of the 28th International Colloquium on Automata, Languages and Programming* (ICALP 2001), Lecture Notes in Computer Science 2076, pages 166–177. Springer-Verlag, 2001.

[13] B. Toft. Colouring, stable sets and perfect graphs. In R. Graham, M. Grötschel, and L. Lovász, editors, *Handbook of Combinatorics*, volume I, pages 233–288. Elsevier, 1995.

Master's Theses in Mathematical Sciences 2002:E39
ISSN 1404-6342

LUNFMA-3019-2002

Mathematics
Centre for Mathematical Sciences
Lund University
Box 118, SE-221 00 Lund, Sweden

http://www.maths.lth.se/