Colorations, Orthotopes, and a Huge Polynomial Tutte Invariant of Weighted Gain Graphs

Version of May 12, 2007

David Forge[1]
Laboratoire de recherche en informatique UMR 8623
Bât. 490, Université Paris-Sud
91405 Orsay Cedex, France
E-mail: forge@lri.fr

and

Thomas Zaslavsky[2]
Department of Mathematical Sciences
State University of New York at Binghamton
Binghamton, NY 13902-6000, U.S.A.
E-mail: zaslav@math.binghamton.edu

Abstract. A *gain graph* is a graph whose edges are labelled invertibly from a group. A *weighted gain graph* is a gain graph with vertex weights from a semigroup, where the gain group is lattice ordered and acts on the weight semigroup. For weighted gain graphs we establish basic properties and we present general dichromatic and tree-expansion polynomials that are Tutte invariants (they satisfy Tutte's deletion-contraction and multiplicative identities). Our dichromatic polynomial includes the classical graph one by Tutte, Zaslavsky's for gain graphs, Noble and Welsh's for graphs with positive integer weights, and that of rooted integral gain graphs by Forge and Zaslavsky. It is unusual in sometimes having uncountably many variables, in contrast to other known Tutte invariants that have at most countably many variables, and in not being itself a universal Tutte invariant of weighted gain graphs; that remains to be found.

An evaluation of our polynomial counts proper colorations of the gain graph when the vertex weights are lists of permissible colors from a color set with a gain-group action. When the gain group is $\mathbb{Z}^d$, the lists are order ideals in the integer lattice $\mathbb{Z}^d$, and there are specified upper bounds on the colors, then there is a formula for the number of bounded proper colorations that is a piecewise polynomial function, of degree $d|V|$, of the upper bounds. This example leads to graph-theoretical formulas for the number of integer lattice points in an orthotope but outside a finite number of affinographic hyperplanes, and for the number of $n \times d$ integral matrices that lie between two specified matrices but not in any of certain subspaces defined by simple row equations.

Mathematics Subject Classifications (2000): *Primary* 05C22; *Secondary* 05C15, 11P21.

Key words and phrases: Weighted gain graph, rooted gain graph, dichromatic polynomial, spanning-tree expansion, Tutte invariant, integral chromatic function, gain graph coloring, list

[1]Part of the research of this author was performed while visiting the State University of New York at Binghamton.

[2]Some of the work of this author was performed while visiting the Laboratoire de recherche en informatique, Université Paris-Sud, Orsay.

coloring, affinographic arrangement, subspace arrangement, lattice point enumeration, orthotope

Contents

1. Introduction

1.1. **Weights, gains, and dichromatics.** In 1999 Noble and Welsh introduced a dichromatic polynomial of graphs whose vertices are weighted by positive integers. The weights add during contraction, so that when an edge is contracted the new vertex weight is the sum of the endpoint weights of the edge. Having defined contraction they could formulate a deletion-contraction reduction formula for their polynomial, and they found the gratifying fact that the polynomial is a universal Tutte invariant. Curiously, it has infinitely many variables, one for each positive integer.

A few years later in [5] we introduced weighted integral gain graphs, which can be regarded as graphs whose edges are orientably labelled by integers, called the *gains* of the edges (orientability means that the gain negates if the edge is reoriented), and whose vertices are assigned integer weights which combine under contraction by taking the maximum (after an adjustment). These graphs were suggested by a problem of counting lattice points.

Reinterpreting the lattice points as proper colorations of a weighted integral gain graph led us to a chromatic function (not quite a polynomial) that counts proper colorations of the graph. Subsequently we discovered that our chromatic function is an evaluation of a dichromatic polynomial that has infinitely many variables (one for each integer) and satisfies Tutte invariance, though it is not universal.

When we learned of Noble and Welsh's work we noticed that both their and our graphs have weights taken from an abelian semigroup (the positive integers with addition, or all integers with maximization), and in our case there is an action of $\mathbb{Z}$, the group from which the gains are drawn, on the weight semigroup. In this paper we present a vast common generalization, in which the edges are labelled by a gain group that is lattice-ordered and acts upon an abelian weight semigroup; furthermore, there is a color filter that restricts the colors at each vertex, so that the number of proper colorations is a function of the choice of filter. Then there is a dichromatic polynomial with one variable for each potential weight, and in examples where we can define a proper coloration, the number of proper colorations, while not a polynomial itself, is an evaluation of the dichromatic polynomial.

With this machinery we extend our prior work on integral gain graphs.

First, taking the weight semigroup to be the class of order ideals in the integer lattice $\mathbb{Z}^d$ with the semigroup operation of set union and with the translation action of the additive gain group $\mathbb{Z}^d$ (Section 5.3) leads to a new definition of proper colorations of the gain graph; these colorations correspond to lattice points in $\mathbb{Z}^{d|V|}$ and can be counted in terms of the dichromatic polynomial.

Second, letting the weight semigroup be the class of sets of integers that are bounded below, we count integral proper list colorations of the gain graph (Section 5.2).

Finally, an *integral orthotope* is a rectangular parallelepiped whose edges are parallel to the coordinate axes and whose vertices have integer coordinates. An *affinographic hyperplane* is a hyperplane of the form $x_j = x_i + a$. Letting the weight semigroup be $\mathbb{Z}$ with maximization, and letting the color filter be an interval $(-\infty, M)$ in $\mathbb{Z}^n$, we obtain a formula (Theorem 1.1) for the number of integer points that lie in an integral orthotope but are not contained in any of a given arrangement (a finite set) of affinographic hyperplanes. This improves on [5], which dealt only with the case of a hypercube. We also get several generalizations of this result (see Theorem 1.2 and Section 5.4).

A brief outline is that, after introductory details, we discuss weighted gain graphs in Section 2, our new dichromatic polynomial in Section 3 and a Tutte-style tree expansion of it in Section 4, colorations in Section 5, weights without gains in Section 6, and in the last section some ideas about further development of the theory.

1.2. **Tutte invariance.** A function defined on objects O, such as graphs or matroids, having a ground set and operations of deletion and contraction of ground-set elements, is a *Tutte invariant* if it satisfies the three conditions:

- (A) Additivity: For every non-detachable element e, $f(O) = f(O \setminus e) + f(O/e)$.
- (M) Multiplicativity: The value of f on O is the product of its values on the components of O.
- (I) Invariance: If O and O' are isomorphic, then $f(O) = f(O')$.
- (U) Unitarity: f equals 1 on trivial objects.

The definitions of isomorphism, of components, of detachable element, and of triviality depend on the kind of object under consideration.

The most popular Tutte invariants have been those of graphs in which one takes a component to be a block and the detachable elements to be the loops and isthmi; edgeless graphs are trivial objects. These invariants are Tutte invariants of graphic matroids, and they extend to all matroids. Examples are the number of bases or maximal forests, the characteristic or chromatic polynomial (modulo a normalization), and the *Tutte polynomial*, a 2-variable polynomial $T_O(x, y)$ that is not only a Tutte invariant but is *universal*: every other Tutte invariant is obtained from the Tutte polynomial by fixing values for x and y. This kind of Tutte invariant of graphs was discovered by Tutte [10] and extended to matroids by Crapo [3]; universality on graphs is due to Tutte [10] and on matroids to Brylawski [1]. The polynomial arises from graphs by taking a component to be a block and the detachable elements to be the loops and isthmi. The trivial object is the empty matroid, corresponding to edgeless graphs.

A different kind of Tutte invariant of graphs appears if we take components to be connected components, the detachable elements to be the loops, and the empty graph to be trivial. (Tutte called these invariants *V-functions*.) For these Tutte invariants there is also a universal Tutte-invariant polynomial; this one has infinitely many variables, one for each nonnegative integer [10, 12].

In Noble and Welsh's weighted ordinary graphs the detachable elements are loops, components are connected components, and the empty graph is trivial. Once again, there is a universal Tutte-invariant polynomial with countably infinitely many variables, which they call $W_G(x_1, x_2, \ldots, y)$.

In the weighted integral gain graphs of [5] the detachable elements are also loops, components are connected components, and the empty graph is trivial. We found that the function counting proper colorations is a Tutte invariant, but it is far from universal.

1.3. **Something new.** Our new objects are *weighted gain graphs*. The edges are orientably labelled from a group and the vertices are weighted from a semigroup. There is a *total dichromatic polynomial* in which, in effect, the coloop variable x of the classical Tutte polynomial splits into a number of variables indexed by the semigroup. (Our polynomial is not strictly a Tutte polynomial but is a refinement and generalization of the two dichromatic polynomials of a gain graph [15, Section III.3], which in turn were based on Tutte's dichromatic polynomial of a graph [11].) The actual number of variables may be uncountably infinite, even in an interesting combinatorial problem (see the end of Section 5.3 in particular). We believe such a plethora of variables has never before been observed.

Our polynomial falls short of universality. That may be because we have not found a big enough polynomial or the right definitions of detachable elements and so forth, but Section 7.1, where we produce other Tutte invariants that cannot be evaluations of our polynomial (they appear to be something like quotients rather than evaluations of the dichromatic polynomial), leads us to believe that there is a more fundamental reason. There may, in fact, be no universal polynomial. This would be a new phenomenon for Tutte invariants, though it is known to occur with parametrized or colored Tutte invariants (cf. [16], for instance).

If as in Noble and Welsh's graphs all gains equal the group identity (so in effect there are no gains), we have weighted ordinary graphs. The number of independent variables in the total dichromatic polynomial, though, is the same. (See Section 6.)

Reexamined in light of the generality of semigroup weights, the problem of [5] of counting proper colorations of a weighted integral gain graph suggests the new problem of counting list colorations of a graph (either with or without edge gains) where the list is an upwardly infinite set of integers and the choice of colors is constrained by a variable upper bound m (Section 5.1). With a mild restriction on the lists, the number of proper colorations is an evaluation of the total dichromatic polynomial of a weighted integral gain graph whose weights are the vertex lists, in which there are uncountably many possible variables, one for each list. Each list variable is given a value that depends on m. With a slight further restriction on the lists the number of proper colorations becomes a polynomial function of m when m is sufficiently large. All this generalizes to a wide range of weighted graphs and gain graphs. An example is coloring a graph whose gains are in $\mathbb{Z}^d$ by colors from $\mathbb{Z}^d$ (this is the order-ideal example we referred to earlier), and there is a partial generalization to any lattice-ordered gain group with list coloring by colors in the group.

1.4. **Lattice points in orthotopes.** Our original reason for developing a theory of weighted gain graphs was an application to geometry. An *integral orthotope* is a rectangular parallelepiped with edges parallel to the coordinate axes. Our wish is to count the points of the integer lattice that lie in an orthotope but not in any of a certain arrangement of affinographic hyperplanes. We state two theorems of this kind here to give the flavor of our geometrical results. (Perhaps the statements seem rather technical; it might be helpful to read the technical definitions in Section 2 now.)

Suppose we have an arrangement $\mathcal{A}$ of affinographic hyperplanes. An *integral gain graph* is a graph whose gain group is $(\mathbb{Z}, +)$. We construct an integral gain graph Φ that has one vertex v_i for each coordinate and an edge ae_{ij} for each hyperplane $x_j = x_i + a$ in $\mathcal{A}$. This means there is an edge v_iv_j with gain a in the indicated direction; the gain of that edge in the other direction, from v_j to v_i, is $-a$. If we have an edge set B such that every circle in B has gain 0 (then B is called *balanced*), we define in each component B_k (including isolated vertices as components) a vertex t_k such that no path B_{wt_k} in B_k from any vertex w to t_k has negative gain. The gain of such a path is the sum of the constants a appearing in the equations of the corresponding hyperplanes $x_j = x_i + a$, each one oriented so that the path leads from v_i to v_j. For instance, if the path is $e_{14}e_{42}e_{28}$ with vertices $v_1, v_4, v_2, t = v_8$, and the corresponding hyperplanes are $x_4 = x_1 + a$, $x_4 = x_2 + b$, $x_8 = x_2 + c$, then the gain of the path is $a - b + c$. Finally, let g_k be the maximum gain of a path in B_k; this also equals the largest gain of paths that ends at t_k.

The first result generalizes the main theorem of [5], which applied only to hypercubes, where all $m_i = m$. We write $x^+ := \max(0, x)$, the positive part of the real number x.

Theorem 1.1. *Let $P := [0, m_1] \times \cdots \times [0, m_n]$ be an orthotope in $\mathbb{R}^n$ and let $\mathcal{A}$ be an arrangement of affinographic hyperplanes. The number of integer points in $P \setminus \bigcup \mathcal{A}$ equals*

$$\sum_{B \subseteq E:\ balanced} (-1)^{|B|} \prod_{B_k} \left(1 + \min_{v_i \in V(B_k)} [m_i + \varphi(B_{v_i t_k})] - g_k\right)^+,$$

where the product is over all components of B.

The proof is in Section 5.4. One can see that the count is a polynomial function of the arguments m_i when they are all sufficiently large.

For the second theorem, suppose that for each coordinate x_i we have a finite list L_i of possible integral values. We want to count lattice points in $L_1 \times \cdots \times L_n$ that are in none of the hyperplanes of the affinographic arrangement $\mathcal{A}$. Of course, this generalizes the preceding theorem, but the viewpoint is different and the formula is much more complex.

Theorem 1.2. *The number of these lattice points is given by the formula*

$$\sum_{B \subseteq E\colon \text{ balanced}} (-1)^{|B|} \prod_{B_k} \Big| \bigcap_{v_i \in V(B_k)} \big(L_i + \varphi(B_{v_i t_k})\big) \Big|,$$

where the product is over all components of B.

That is, we take the intersection of translates of the lists, governed by the gains of paths in the chosen balanced edge set B. The proof is in Section 5.4, which also has a common generalization of both results.

We even have a formula for the number of $n \times d$ integral matrices in an integral orthotope in $\mathbb{Z}^{nd}$ that do not lie in any of a class of d-codimensional subspaces whose equations compare rows of the matrix (Section 5.4).

2. Weighted gain graphs

2.1. **Graphs.** Edges of a graph $\Gamma = (V, E)$ are of four kinds. A *link* has two distinct endpoints; a *loop* has two coinciding endpoints. A *half edge* has one endpoint, and a *loose edge* has no endpoints. (Half and loose edges have a negligible role in this paper except in Section 2.4.) The set of loops and links is written E^*. Multiple edges are permitted. We write $n := |V|$ and $V = \{v_1, v_2, \ldots, v_n\}$. All our graphs have finite order and indeed are finite (except for root edges, when they appear). A *(connected) component* of Γ is a maximal connected subgraph that is not a loose edge; we do not count a loose edge as a component. The number of components is $c(\Gamma)$. For $S \subseteq E$, we denote by $c(S)$ the number of connected components of the spanning subgraph (V, S) (which we call the *components of* S) and by $\pi(S)$ the partition of V into the vertex sets of the various components. We write S_{vw} to denote any path in S from v to w (if one exists).

2.2. **Gain graphs.** A *gain graph* $\Phi = (\Gamma, \varphi)$ consists of a graph $\Gamma = (V, E)$, a group $\mathfrak{G}$ called the *gain group*, and an orientable function $\varphi : E^* \to \mathfrak{G}$, called the *gain mapping.* (Half and loose edges do not have gains.) The basic reference is [15, Part I]. "Orientability" means that, if e denotes an edge oriented in one direction and e^{-1} the same edge with the opposite orientation, then $\varphi(e^{-1}) = \varphi(e)^{-1}$. (It does not mean that Γ is directed; there is no fixed orientation of any edge.) We sometimes use the simplified notations e_{ij} for an edge with endpoints v_i and v_j, oriented from v_i to v_j, and ge_{ij} for such an edge with gain g; that is, $\varphi(ge_{ij}) = g$. (Thus ge_{ij} is the same edge as $g^{-1}e_{ji}$.) A *circle* is a connected 2-regular subgraph without half edges, or its edge set; for instance, a loop is a circle of length 1. We may write a circle C as a word $e_1 e_2 \cdots e_l$; this means the edges are numbered consecutively around C and oriented in a consistent direction. The gain of C is $\varphi(C) := \varphi(e_1)\varphi(e_2)\cdots\varphi(e_l)$; this is well defined up to conjugation and inversion, and in particular it is well defined whether the gain is the identity 1 or not. An edge set or subgraph is called *balanced* if every circle in it has gain 1 and it has no half edges. The notation $c(\Phi)$ means $c(\Gamma)$.

For $W \subseteq V$, the subgraph *induced* by W is notated $\Gamma{:}W$, or with gains $\Phi{:}W$. The edge set of $\Gamma{:}W$ consists of all edges that have at least one endpoint in W and no endpoint outside W; thus, half edges at vertices of W are included, but loose edges are not. If $S \subseteq E$, then $S{:}W$ means the subset of S induced by W.

Switching Φ by a *switching function* $\eta : V \to \mathfrak{G} : v_i \mapsto \eta(v_i)$ means replacing φ by

$$\varphi^\eta(e_{ij}) := \eta(v_i)^{-1}\varphi(e_{ij})\eta(v_j).$$

We write Φ^η for the switched gain graph (Γ, φ^η). It is clear that the switching action is an action of the group $\mathfrak{G}^V$ of switching functions on the set $\mathfrak{G}^E$ of gain functions on the underlying graph.

Consider a balanced edge set S. Let $S_{v_i v_j}$ denote a path in S from v_i to v_j, if one exists; the gain $\varphi(S_{v_i v_j})$ is independent of the particular path because S is balanced. There is a switching function η such that $\varphi^\eta\big|_S \equiv 1$ [15, Section I.5]; it is determined by any one value in each component of S through the formula

$$\eta(v_j) = \varphi(S_{v_j v_i})\eta(v_i), \tag{2.1}$$

We call η a *switching function for* S. Any two different switching functions for S, η and η', are connected by the relation

$$\eta' = \eta \cdot \alpha_W \tag{2.2}$$

for constants $\alpha_W \in \mathfrak{G}$, one for each $W \in \pi(S)$. (In fact, $\alpha_W = \eta(v_i)^{-1}\eta'(v_i)$ for any $v_i \in W$; this is easy to deduce from (2.1).) Thus, as long as the endpoints of an edge e_{ij} are in the same component of S, $\varphi^\eta(e_{ij}) = \eta(v_i)^{-1}\varphi(e_{ij})\varphi(S_{v_j v_i})\eta(v_i)$, which is uniquely determined up to conjugation. (If e_{ij} has endpoints in distinct components of S, then $\varphi^\eta(e_{ij}) = \eta(v_i)^{-1}\varphi(e_{ij})\eta(v_j)$ can be anything, since $\eta(v_i)$ and $\eta(v_j)$ are independently choosable elements of $\mathfrak{G}$.)

The operation of deleting an edge or a set of edges is obvious. The notation for Φ with $E \backslash S$ deleted, called the *restriction* of Φ to S, is $\Phi|S = (V, S, \varphi|_S)$. The number of components of S that are balanced is $b(\Phi|S)$ or briefly $b(S)$ (recall that this counts isolated vertices but not loose edges); $\pi_{\mathrm{b}}(S) = \pi_{\mathrm{b}}(\Phi|S)$ is the set $\{W \in \pi(S) : (S{:}W) \text{ is balanced}\}$; $V_0(S)$ is the set of vertices that belong to no balanced component of S; and $V_{\mathrm{b}}(S)$ denotes the set of vertices of balanced components, $V_{\mathrm{b}}(S) = V \setminus V_0(S)$.

Contraction is not so obvious. We take the definition from [15]. First, we describe how to contract a balanced edge set S. We first switch by η, any switching function for S; then we identify each block $W \in \pi(S)$ to a single vertex and delete S. The notation is Φ^η/S. This contraction depends on the choice of η, so Φ^η/S is well defined only up to switching. (Soon, however, we shall see how to single out a preferred switching function.)

For a general subset S we first delete the vertex set $V_0(S)$, then contract the remaining part of S, which is the union of all balanced components of S, and delete any remaining edges of S. Edges not in S that have one or more endpoints in $V_0(S)$ lose those endpoints but remain in the graph, thus becoming half or loose edges. So, the contraction has $V(\Phi/S) = \pi_{\mathrm{b}}(S)$ and $E(\Phi/S) = E \setminus S$.

A balanced edge set S is called *closed* if it contains every loose edge and any edge ge_{ij} whose endpoints are joined by an open path $P \subseteq S$ with the same gain (P has length 0 if $i = j$) is itself in S. This is equivalent to saying S equals its own closure; the closure of a balanced edge set S is given by

$$\mathrm{cl}(S) := S \cup \{e \notin S : e \text{ is contained in a balanced circle } C \subseteq S \cup \{e\}\} \cup \{\text{loose edges}\},$$

and is balanced [15, Proposition I.3.1]. The semilattice of all closed, balanced edge sets in Φ is written $\mathrm{Lat}^{\mathrm{b}}\,\Phi$.

2.3. **Ordered gain groups.** For the rest of this article we assume the gain group is lattice ordered. (It may be totally ordered; this case has some special features.)

We continue thinking of a balanced edge set S. The ordering singles out a particular switching function for S, the one for which the meet of its values on each block of $\pi(S)$ is the identity. We call this the *top switching function* and we write it η_S; it is what we use for switching throughout the rest of this article. Because S is balanced, the gain of every path S_{vw} is the same.

Lemma 2.1. *The top switching function η_S has the formula*

$$\eta_S(v) = \bigvee_w \varphi(S_{vw}),$$

where w ranges over vertices connected by S to v, and for its inverse

$$\eta_S(v)^{-1} = \bigwedge_w \varphi(S_{wv}).$$

Proof. We use the identity $(\alpha \wedge \beta)^{-1} = \alpha^{-1} \vee \beta^{-1}$.

We know two properties of η_S. As a switching function for S it satisfies (2.1). As a top switching function it satisfies $\bigwedge_{w \in W} \eta_S(w) = 1$. Equation (2.1) lets us rewrite this as

$$\bigwedge_{w \in W} \varphi(S_{wv})\eta_S(v) = 1.$$

Factoring out $\eta_S(v)$,

$$\eta_S(v) = \Big[\bigwedge_{w \in W} \varphi(S_{wv}) \Big]^{-1} = \bigvee \varphi(S_{wv}). \quad \square$$

In view of the importance of the meet of switching-function values, we define

$$\eta(X) := \bigwedge_{v \in X} \eta(v)$$

for $X \subseteq V$.

Now, to contract S we first switch by η_S; then we identify each block $W \in \pi(S)$ to a single vertex and delete S. The contraction Φ^{η_S}/S, which we call the *top contraction*, we usually write Φ/S for brevity. The contraction Φ/S is now a unique gain graph, because the gain-group ordering allows us to specify the switching function uniquely.

When the group is totally ordered, there is a *top vertex* in every component of S, a vertex t such that no path in S that begins at t has positive gain; this is any vertex for which $\eta_S(t) = 1$. Then the rule for defining η_S is that its minimum value on each block is the identity. A top vertex may also happen to exist when $\mathfrak{G}$ is not totally ordered. If t_i denotes a top vertex in the same component of S as v_i, then the top switching function has the formula

$$\eta_S(v_i) = \varphi(S_{v_i t_i}). \tag{2.3}$$

The gain function φ switched by η_S is given by the formula

$$\varphi^{\eta_S}(e_{ij}) = \varphi(S_{v_i t_i})^{-1}\varphi(e_{ij})\varphi(S_{v_j t_j}) = \varphi(S_{t_i v_i} e_{ij} S_{v_j t_j}). \tag{2.4}$$

2.4. **Weights.** Suppose we have an abelian semigroup $\mathfrak{W}$ (written additively) and a group $\mathfrak{G}$. We say $\mathfrak{G}$ *acts on* $\mathfrak{W}$ if each $g \in \mathfrak{G}$ has a right action on $\mathfrak{W}$ which is a semigroup automorphism satisfying the usual identities, i.e., $(hg)g' = h(gg')$ and $h1 = h$.

A *weighted gain graph* (Φ, h) is a gain graph Φ together with a *weight* function $h : V \to \mathfrak{W}$. We usually write $h_i := h(v_i)$. The way h transforms under switching is that

$$h_i^\eta = h^\eta(v_i) := h_i\eta(v_i)$$

(as if h_i were the gain of an edge oriented into v_i from an extra vertex at which η is the identity). Thus, the switching group $\mathfrak{G}^V$ has a right action on the set $\mathfrak{W}^V$ of weight functions. The contraction rule is that, first, we always contract with top switching; and if $W \in \pi_{\mathrm{b}}(S)$, then the weight function h_S in the contraction $(\Phi, h)/S$ is given by

$$h_S(W) := \sum_{v_i \in W} h_i^{\eta_S}.$$

If $R \subseteq S$ and $\pi_{\mathrm{b}}(R) = \pi_{\mathrm{b}}(S)$, then $h_R = h_S$.

If there is a top vertex t_i in the component $S{:}W$ that contains v_i, then $h_i^{\eta_S} = h_i\varphi(S_{v_i t_i})$ and $h_S(W) = \sum_{v_i \in W} h_i\varphi(S_{v_i t_i})$.

We have occasion to contract the induction $(\Phi, h){:}W := (\Phi{:}W, h|_W)$ of the entire weighted graph by the induction $S{:}W$ of an edge set, where $W \in \pi_{\mathrm{b}}(S)$; we ought to write this $((\Phi, h){:}W)/(S{:}W)$ but we simplify the notation to $(\Phi, h)/S{:}W$.

The next result states the fundamental properties of deletion and contraction of weighted gain graphs.

Proposition 2.2. *In a weighted gain graph* (Φ, h)*, let* $S \subseteq E$ *be the disjoint union of* Q *and* R*. Then*

$$((\Phi, h)/Q)/R = (\Phi, h)/S,$$
$$((\Phi, h)/Q) \setminus R = ((\Phi, h) \setminus R)/Q,$$
$$((\Phi, h) \setminus Q) \setminus R = (\Phi, h) \setminus S.$$

Proof. We may suppose Φ is connected. The two latter formulas are obvious.

The first one is not; indeed, in a purely technical sense it is false, since $V(\Phi/S) = \pi_{\mathrm{b}}(\Phi|S)$ while $V((\Phi/Q)/R) = \pi_{\mathrm{b}}(\Phi/Q|R)$; but it is correct if we identify $W \in \pi_{\mathrm{b}}(\Phi|S)$ with $W'' \in \pi_{\mathrm{b}}(\Phi/Q|R)$ in the natural way: W corresponds to $W'' = \{X \in \pi_{\mathrm{b}}(\Phi|Q) : X \subseteq W\}$ and conversely W'' corresponds to $W = \bigcup W'' = \{w \in V(\Phi) : w \in X \text{ for some } X \in W''\}$.

In proving the first formula, the first step is to show that we can assume S is balanced. It is a routine check to see that Φ/S and $\Phi/Q/R$ have the same half and loose edges. Since $V(\Phi/S) = \pi_{\mathrm{b}}(S)$, we have

$$(\Phi/S)^* = [(\Phi{:}U)/(S{:}U)]^* \quad\text{and}\quad (\Phi/Q/R)^* = \big[\big((\Phi{:}U)/(Q{:}U)\big)/(R{:}U)\big]^*$$

where $U := V_{\mathrm{b}}(S)$ and the superscript * denotes that loose and half edges are to be deleted. Since $S{:}U$ is balanced, both $Q{:}U$ and $R{:}U$ are also balanced. The weights of the contractions only appear on vertices of Φ/S so they depend only on vertices and edges in U; the same holds true for $\Phi/Q/R$. It follows that

$$((\Phi, h)/S)^* = [((\Phi, h){:}U)/(S{:}U)]^*$$

and

$$((\Phi, h)/Q/R)^* = \big[\big(((\Phi, h){:}U)/(Q{:}U)\big)/(R{:}U)\big]^*.$$

This proves that we may confine our attention to the balanced spanning subgraph $(U, S{:}U)$ in $\Phi{:}U$; thus, we may from now on assume S is balanced.

Let η'_S be the top switching function for $\Phi^{\eta_Q}|S$ and let η''_R be that for $\Phi/Q|R$. (Recall that Φ/Q means Φ^{η_Q}/Q.) The key to the proof is the factorization identity

$$\eta_S(v) = \eta_Q(v)\eta'_S(v), \tag{2.5}$$

which shows that the effect of η_S, which is to switch so $\varphi|_S$ becomes 1, can be divided into two stages: first switching by η_Q so that $\varphi|_Q$ becomes 1, and then switching by η'_S, which is constant on components of Q.

In proving (2.5), first we compare η_Q and η_S. Since they are two switching functions for Q, they are related by (2.2). Specifically, let $X \in \pi(Q)$ and $W \in \pi(S)$, with $X \subseteq W$; then $\eta_S(v) = \eta_Q(v)\alpha_X$ for $v \in X$. Taking the meet over X, $\eta_S(X) = \eta_Q(X)\alpha_X = 1\alpha_X$, so

$$\eta_S(v) = \eta_Q(v)\eta_S(X)$$

for $v \in X$. Next we show that $\eta_S(X) = \eta'_S(v)$. Define $\bar\eta := \eta_Q\eta'_S$. It is easy to verify that $\bar\eta$ is a switching function for S. Taking the meet over all $v \in W$, and taking note that $\eta'_S(v) = \eta'_S(X)$ for $v \in X$ because η'_S is constant on X, we find that

$$\begin{aligned}\bar\eta(W) &= \bigwedge_{v\in W} \eta_Q(v)\eta'_S(v) = \bigwedge_{X\in W''} \Big[\bigwedge_{v\in X} \eta_Q(v)\Big]\eta'_S(X) \\ &= \bigwedge_{X\in W''} 1\eta'_S(X) = \bigwedge_{v\in W} \eta'_S(v) = 1.\end{aligned}$$

Thus $\bar\eta$ is a top switching function for S and, as there is only one, it equals η_S. This proves (2.5).

From (2.5) it follows that $\varphi^{\eta_S} = (\varphi^{\eta_Q})^{\eta'_S}$ and $h^{\eta_S}(v) = (h^{\eta_Q})^{\eta'_S}(v)$, thus establishing that

$$(\Phi, h)^{\eta_S} = \big((\Phi, h)^{\eta_Q}\big)^{\eta'_S}. \tag{2.6}$$

Now we can analyze the process of contraction. We know from [15, Theorem I.4.7 and the proof of Theorem I.5.4] that $\Phi/Q/R \sim \Phi/S$ (where $\Phi_1 \sim \Phi_2$ means that each of them is a switching of the other). But the switching equivalence is really equality because the switching functions employed are related by Equation (2.5). Thus, $\Phi/S = \Phi/Q/R$.

The last step is to prove that weights contract properly. The key here is that contraction by Q commutes with two-stage switching, i.e.,

$$(h^{\eta_Q}/Q)^{\eta''_R} = (h^{\eta_Q})^{\eta'_S}/Q. \tag{2.7}$$

Observe that η'_S is constant on each $X \in \pi(Q)$ and its common value is $\eta''_R(X)$. Expanding both sides according to the definitions of switching and contraction, this is equivalent to

$$\Bigg(\sum_{w\in X} h^{\eta_Q}(w)\Bigg)\eta''_R(X) = \sum_{w\in X} h^{\eta_Q}(w)\eta'_S(w),$$

which is true because $\eta'_S(w) = \eta''_R(X)$. That concludes the proof of (2.7).

Equations (2.6) and (2.7) imply the double contraction formula through the sequence of transformations

$$\begin{aligned}
(\Phi,h)/Q/R &= \big((\Phi,h)^{\eta_Q}/Q\big)^{\eta''_R}/R && \text{by definition}\\
&= \big((\Phi,h)^{\eta_Q\eta'_S}/Q\big)/R && \text{by (2.7)}\\
&= \big((\Phi,h)^{\eta_S}/Q\big)/R && \text{by (2.6)}\\
&= (\Phi,h)^{\eta_S}/S
\end{aligned}$$

in the loose sense previously defined in terms of the correspondence $W \leftrightarrow W''$. □

Example 2.1 (*Weighted integral gain graphs; linearly ordered group weights*). Our original example [5] was that of *weighted integral gain graphs*, where the gain group is the additive group of integers and the weight semigroup is the integers with the operation of maximization. In other words, $\mathfrak{G} = (\mathbb{Z}, +)$ and $\mathfrak{W} = (\mathbb{Z}, \max)$.

A similar kind of example exists for every linearly ordered gain group, with $\mathfrak{W} = (\mathfrak{G}, \max)$ or $(\mathfrak{G}, \min)$.

Example 2.2 (*Semilattice weights*). To further generalize Example 2.1, let $\mathfrak{W}$ be a semilattice with a $\mathfrak{G}$-action; the semigroup operation is the semilattice operation. In an important example of this type there is a set $\mathfrak{C}$ on which there is a right action of the gain group; the weights are subsets of $\mathfrak{C}$, i.e., $\mathfrak{W} \subseteq \mathcal{P}(\mathfrak{C})$; and the semigroup operation is set intersection—so $\mathfrak{W}$ must be closed under intersection. (In Section 5 $\mathfrak{C}$ will be a color set and the weight $h_i \subseteq \mathfrak{C}$ will be treated as the list of colors possible for vertex v_i.)

Generalizing minimization, let $\mathfrak{C}$ be a partially ordered set and let the weights be order ideals in $\mathfrak{C}$. If $\mathfrak{C}$ is a meet semilattice, one may restrict the weights to be principal ideals. There are also the order duals of these examples.

When, on the other hand, $\mathfrak{C} = \mathfrak{G}$ with the right translation action, one may take $\mathfrak{W} = \mathcal{P}(\mathfrak{G})$, for instance, or the class of principal dual order ideals (since $\mathfrak{G}$ is a lattice), or the class of sets that have a lower bound (that is, all subsets of principal dual ideals). The dual of this last, with gain group $\mathbb{Z}$ was the prototype of weighted gain graphs, as we explain next.

2.5. **Rooted integral gain graphs.** The curious reader will be wondering how we came to semigroup weights. Originally, in [5], there was a *rooted integral gain graph*, which is an integral gain graph Ψ with a root vertex v_0 such that the gains of edges e_{0i} form an interval $(-\infty, h_i]$ in the gain group $\mathbb{Z}$, the infinite cyclic group. (This is the one exception to our assumption that gain graphs are finite.) In top switching of a balanced set S of nonroot edges of Ψ, one always takes $\eta_S(v_0) = 0$. This implies a rule for how the value h_i changes under switching. We used this rule in studying the chromatic function, which count proper colorations in an interval $(-\infty, m]$, of a rooted integral gain graph, as explained in Example 5.3.

An equivalent presentation omits the root and simply specifies an integral weight h_i on each vertex of $\Phi := \Psi \setminus v_0$. Then switching in Ψ, transferred to the rootless integral gain graph Φ, implies the rule $h^\eta = h + \eta$ for switching the weights h_i. That is the rule adopted and generalized in Section 2.4.

Similarly, contraction, defined in the standard gain-graphic way on Ψ and reinterpreted in terms of integral weights h_i on Φ, assigns to a set $W \in \pi_b(S)$ (as a vertex in Φ/S) a weight equal to the maximum weight of a vertex in W after switching; thus the weights belong to the set $\mathbb{Z}$ but with the semigroup operation of maximization instead of the group operation of addition. (Technically, contraction of a nonroot edge may create parallel root edges with the same gain, but for our purposes parallel root edges can be ignored.)

We noticed that one could get similar but more general conclusions about the chromatic function by allowing the root-edge gains to form fairly arbitrary sets H_i, instead of just intervals $(-\infty, h_i]$ as in the original rooted integral gain graphs, and especially if the gain set H_i is any subset of $\mathbb{Z}$ that is bounded above and has complement bounded below (see Section 5.1, where the weights are the complements of these gain sets). The effect of contraction on the gain sets is to take the union. Thus we had a new weight semigroup; instead of $\mathbb{Z}$, it consisted of all subsets that are bounded above and cobounded below. It became apparent that the semigroup can be treated independently of the gain group except for the action of the latter upon the former.

Thus, although the vertex weights can no longer be interpreted as gains of root edges, our thinking is based on the model of a rooted integral gain graph.

3. A Tutte-invariant polynomial

A function f defined on weighted gain graphs (with fixed gain group and weight semigroup) is a *Tutte invariant* if it satisfies the three conditions from the introduction:

(Ti) (Additivity) For every link e,

$$f(\Phi, h) = f(\Phi \setminus e, h) + f(\Phi/e, h/e),$$

where h/e denotes the contracted weight function.

(Tii) (Multiplicativity) The value of f on (Φ, h) is the product of its values on the components of (Φ, h).

(Tiii) (Invariance) If (Φ, h) and (Φ', h') are switching isomorphic, then $f(\Phi, h) = f(\Phi', h')$. (*Switching isomorphism* means an isomorphism of underlying graphs that preserves gains and weights up to switching.)

(Tiv) (Unitarity) $f(\varnothing) = 1$.

We present here an algebraic Tutte invariant. We need variables u_k for all $k \in \mathfrak{W}$; the collection of all u_k's is denoted by $\mathbf{u}$. The *total dichromatic polynomial* of a weighted gain graph is

$$Q_{(\Phi,h)}(\mathbf{u}, v, z) := \sum_{S \subseteq E} v^{|S|-n+b(S)} z^{c(S)-b(S)} \prod_{W \in \pi_{\mathrm{b}}(S)} u_{h_S(W)}, \tag{3.1}$$

where

$$h_S(W) := \sum_{w \in W} h^{\eta_S}(w).$$

This polynomial refines the balanced and ordinary dichromatic polynomials of a gain graph or biased graph [15, Section III.3], which are obtained by setting all $u_k = u$ and $z = 0$ or $z = 1$, respectively. (Thus the total polynomial with all $u_k = u$ fills a gap in the theory of [15, Part III] by unifying the balanced and ordinary polynomials.)

For a graph with no edges,

$$Q_{((V,\varnothing),h)} = \prod_{v_i \in V} u_{h_i}. \tag{3.2}$$

If e is a balanced loop or a loose edge,

$$Q_{(\Phi,h)} = (v+1)Q_{(\Phi\setminus e,h)}. \tag{3.3}$$

Theorem 3.1. *The total dichromatic polynomial $Q_{(\Phi,h)}(\mathbf{u}, v, z)$ is a Tutte invariant of weighted gain graphs.*

Proof. Invariance, unitarity, and multiplicativity are obvious. For additivity we follow the usual proof method, dividing up the terms of the defining sum into two parts: those sets S that do not contain the link e and those sets that do contain e. The sum of the former terms obviously equals $Q_{(\Phi,h)\setminus e}(\mathbf{u}, v, z)$ and the sum of the latter, we shall see, equals $Q_{(\Phi,h)/e}(\mathbf{u}, v, z)$.

A set $S \ni e$ contracts to a set $R = S \setminus e$ in Φ/e whose balanced components correspond to those of S. That is, if S_0 is a balanced component of S, then S_0 (if $e \notin S_0$ or S_0/e (if $e \in S_0$) is a balanced component of R, and vice versa. (This follows from [15, Lemma I.4.3].) So $b(\Phi/e|R) = b(\Phi|S)$. Since Φ/e has order $n-1$, the term of S in $Q_{(\Phi,h)}$ and that of R in $Q_{(\Phi,h)/e}$ are the same except for the factors $u_{h_S(W)}$ in the former and $u_{h_R(W'')}$ in the latter, where W'' is the block of $\pi_{\mathrm{b}}(\Phi/e|R)$ that corresponds to W. We want to show that these factors are equal, i.e., that $h_S(W) = h_R(W'')$. But the former is h/S and the latter is $(h/e)/R$, which we know by Proposition 2.2 to be equal.

It follows that Q satisfies additivity, so is a Tutte invariant. □

4. Tree expansion

We turn to an expression for the balanced dichromatic polynomial, $Q_{(\Phi,h)}(\mathbf{u}, v, 0)$, that depends on a linear ordering of the edge set. We fix one such ordering O and in terms of it we define a spanning-tree expansion similar to the Tutte polynomial of a matroid. The details are in Section 4.3, after some preliminary work with independent sets in semimatroids and gain graphs.

4.1. Activities in semimatroids.

A semimatroid is a generalization of a matroid that extends properties like rank and closure of the family of balanced edge sets in a gain graph. The theory was developed by Wachs and Walker in [13]. Just as with matroids, there are many equivalent ways to define a semimatroid. We define a semimatroid in terms of a matroid M_0 with ground set E_0 and a basepoint e_0. A subset of $E := E_0 \setminus e_0$ whose closure in M_0 does not contain e_0 is called *balanced*; the family of balanced sets is denoted by $\mathcal{P}_{\mathrm{b}}(M)$. The *semimatroid* M associated with (M_0, e_0) is the family of all balanced subsets of E with closure operator, rank function, closed or independent sets, circuits, and so forth the same as those of M_0 but restricted to balanced sets.

For instance, the independent sets of M are the ones of M_0 whose closures do not contain e_0. The closed sets of M are those of M_0 that do not contain e_0. A fundamental fact is that if S is balanced, the closure $\mathrm{cl}_0\, S$ (in M_0) is balanced. Consequently, the closure $\mathrm{cl}\, S$ in M equals $\mathrm{cl}_0\, S$. Also, any circuit in $\mathrm{cl}\, S$ is balanced. A maximal balanced independent set, that is, a maximal independent set of M, is called a *semibasis*.

If e_0 is a loop or coloop in M_0, then M is the matroid $M_0 \setminus e_0$. Otherwise, E is not balanced and M is not a matroid.

We develop some facts about independent sets, activities, and broken circuits in a semimatroid. Some of them are already known for matroids. We could not find an explicit source for exactly these results, but [4] and [6, Section 2] have theorems along similar lines. A reference for the fundamentals of activities in matroids is [1] or [2].

First, some basic definitions. Let F be independent in M_0. For a point $e \in (\mathrm{cl}_0 F) \setminus F$, there is a unique circuit contained in $F \cup e$; it is called the *fundamental circuit* of e with respect to F and denoted by $C_F(e)$. It is balanced if (but not only if) F is balanced. For a point $f \in F$, we call $\mathrm{cl}_0(F) \setminus \mathrm{cl}_0(F \setminus f)$ the *fundamental relative cocircuit* of f with respect to F, written $D_F(e)$. If F is balanced, the closures are in M so cl_0 can be replaced by cl.

We fix a linear ordering O of E. Consider an independent set F of the semimatroid M (that is, a balanced independent set of M_0). We say that a point e is *externally active* (in M) with respect to F if $e \notin F$ and e is the largest point in $C_F(e)$ (so only a point in $(\mathrm{cl}\, F) \setminus F$ can be externally active). A point e is *internally active* (in M) with respect to F if it is in F and it is the largest point in $D_F(e)$. A point that is not active is *internally inactive* if it belongs to F and *externally inactive* if it belongs to $(\mathrm{cl}\, F) \setminus F$. The sets of internally or externally active or inactive points with respect to F are denoted by $\mathrm{IA}(F)$, $\mathrm{EA}(F)$, $\mathrm{II}(F)$, $\mathrm{EI}(F)$. The number of externally active points is $\varepsilon(F)$. The number of internally active points is $\iota(F)$.

The definitions for M_0 are the same, except for the omission of the word "balanced", replacement of cl by cl_0, and the need to linearly order all of E_0. Thus we have $\mathrm{IA}_0(F)$, etc.; but when F is balanced, these are the same as $\mathrm{IA}(F)$, etc.

A *broken (balanced) circuit* is a (balanced) circuit with its largest element removed. Note that a set may be a broken circuit and balanced without being a broken balanced circuit; for an example let M_0 itself be a circuit and order E_0 so e_0 is largest; then E is balanced and a broken circuit, but there are no broken balanced circuits.

Lemma 4.1. *Let F be independent in the semimatroid M. Then $\mathrm{II}(F)$ is the union of all broken balanced circuits in F.*

Proof. If D is a broken balanced circuit in F, there is a point $e \in (\mathrm{cl}\, F) \setminus F$ which is maximal in its fundamental circuit $C_F(e) = D \cup \{e\}$. Any $f \in D$ is internally inactive because $e >_O f$ and $e, f \in (\mathrm{cl}\, F) \setminus F$. □

Lemma 4.2. *If F' is independent in the semimatroid M and $F \subseteq F'$, then $\mathrm{II}(F) \subseteq \mathrm{II}(F')$.*

Proof. Immediate from Lemma 4.1. □

Each point set S has a *minimal basis* $F(S)$, the basis that is lexicographically first in O; it is the one obtained by the greedy algorithm applied to S. It is balanced if and only if F is balanced. The next lemma says that the inverse of the mapping $S \mapsto F(S)$ partitions the power set of E_0 into intervals $[F, F \cup \mathrm{EA}_0(F)]$, one for each independent set F, and either all sets in the interval are balanced or all are unbalanced.

Lemma 4.3. *Let F be independent in M_0 and let $S \subseteq E_0$. For the minimal basis of S to be F, it is necessary and sufficient that $F \subseteq S \subseteq F \cup \mathrm{EA}_0(F)$. Further, $F \cup \mathrm{EA}_0(F)$ is balanced if and only if F is balanced.*

Proof. Assume F is the minimal basis of S and write $F = e_1 e_2 \cdots$ in increasing order in O. Every $e \in S \setminus F$ has a fundamental circuit $C_F(e)$. Suppose e is not externally active with respect to F, so that $C_F(e) = \cdots e e' \cdots$; let $e' = e_{k+1}$. The set $\{e_1, \ldots, e_k, e\}$ is independent because the only circuit it could contain is $C_F(e)$, but $e' \in C_F(e) \setminus \{e_1, \ldots, e_k, e\}$. Consider the greedy algorithm for finding F. After choosing $e_1, \ldots, e_k$, the next point chosen cannot be e', because e (or some other point different from e') would be preferred as it has not been chosen, it precedes e' in the ordering, and $\{e_1, \dots, e_k, e\}$ is independent. Thus, $e_{k+1} \neq e'$. This is a contradiction. Therefore, e must be externally active.

Assume $F \subseteq S \subseteq F \cup \mathrm{EA}_0(F)$. Thus, F is a basis for S; we want to show it is minimal. Let $e \in S \setminus F$ and write $C_F(e) = e_1 \cdots e_k e_{k+1}$ in the ordering O; then $e = e_{k+1}$. In the greedy algorithm for constructing the minimal basis $F(S)$, each point $e_1, \dots, e_k, e_{k+1}$ is considered in order for inclusion. Let $F_i(S)$ be the set of points that have already been chosen when e_i is considered for inclusion. If e_i is not then chosen for $F(S)$, it is because $e_i \in \operatorname{cl} F_i(S)$. If e_i is chosen, then $e_i \in F_{i+1}(S)$. Thus, all of $e_1, \dots, e_k \in \operatorname{cl} F_{k+1}(S)$. It follows that $e \in \operatorname{cl} F_{k+1}(S)$, so $e \notin F(S)$. This shows that no point of $\mathrm{EA}_0(F)$ can belong to the minimal basis $F(S)$; hence, $F(S) \subseteq F$ and by comparing ranks we see that $F(S) = F$.

The last part of the lemma follows because $\mathrm{EA}_0(F) \subseteq \operatorname{cl}_0 F$, which is balanced if F is balanced. □

Suppose we already have a balanced independent set F that we want to extend to a semibasis. We can do that by applying the *reverse greedy algorithm*. That means we take $E \setminus F$ and scan down it from the largest point (in the ordering O) to the smallest, adding a point to the independent set whenever the resulting set remains independent and balanced. The set obtained in this way we call the *maximal semibasis extension*, $T(F)$. It is clear that $T(F)$ is a semibasis.

Lemma 4.4. *For an independent set F in M, $T(F)$ has the following properties:*

(i) $F \supseteq \mathrm{II}(T(F))$ *and* $T(F) \setminus F \subseteq \mathrm{IA}(T(F))$.
(ii) $\mathrm{II}(F) = \mathrm{II}(T(F))$.
(iii) $\mathrm{EA}(F) \subseteq \operatorname{cl}(\mathrm{II}(T(F))$.
(iv) $\mathrm{EA}(F) = \mathrm{EA}(\mathrm{II}(T(F))) = \mathrm{EA}(T(F))$*; thus,* $\varepsilon(F) = \varepsilon(T(F))$.

Proof. In (i) the two statements are obviously equivalent; we prove the latter. Suppose we have a balanced independent set F' and a point $e \notin \operatorname{cl} F'$; call e *F'-tolerable* if $F' \cup \{e\}$ is balanced. Write $T(F) \setminus F = e_k \cdots e_1$ in increasing order, so that each e_i is the largest $F \cup \{e_1, \dots, e_{i-1}\}$-tolerable point. Since

$$e_i \notin \operatorname{cl}(T(F) \setminus e_i) \supseteq \operatorname{cl}(F \cup \{e_1, \dots, e_{i-1}\}),$$

e_i is larger than any other $F \cup \{e_1, \dots, e_{i-1}\}$-tolerable point not in $\operatorname{cl}(T(F) \setminus e_i)$. That is, it is externally active.

In Part (ii), $\mathrm{II}(F) \subseteq \mathrm{II}(T(F))$ by Lemma 4.2. To prove the reverse containment, apply the same lemma to Part (i) to conclude that $\mathrm{II}(F) \supseteq \mathrm{II}(\mathrm{II}(T(F)))$; the latter equals $\mathrm{II}(T(F))$ by Lemma 4.1.

In (iii), e is maximal in $C_F(e)$ for $e \in \mathrm{EA}(F)$. By Lemma 4.1 and Part (ii), $C_F(e) \setminus \{e\} \subseteq \mathrm{II}(F) = \mathrm{II}(T(F))$.

For (iv), suppose we have two balanced independent sets, $F_1 \subseteq F_2$. Obviously $\mathrm{EA}(F_1) \subseteq \mathrm{EA}(F_2)$, because $(\operatorname{cl} F_1) \setminus F_1 \subseteq (\operatorname{cl} F_2) \setminus F_2$. If $\mathrm{EA}(F_2) \subseteq \operatorname{cl} F_1$, then $\mathrm{EA}(F_2) \subseteq \mathrm{EA}(F_1)$; thus the two EAs are equal. Now apply this fact to $F_1 = \mathrm{II}(T(F_2))$ and $F_2 = F$ or $T(F)$, recalling (iii). □

4.2. **Activities in gain graphs.** When we come to gain graphs, the semimatroid we need is that associated with the balanced edge sets of Φ. Here, M_0 is the *complete lift matroid*

$L_0(\Phi)$, which is the matroid on $E_0 := E(\Phi) \cup \{e_0\}$ with rank function

$$\mathrm{rk}(S) = \begin{cases} n - c(S) & \text{if } S \text{ is balanced,} \\ n - c(S) + 1 & \text{if } S \text{ is unbalanced} \end{cases}$$

[15, Section II.4]. In a way, the complete lift matroid generalizes the usual graphic matroid $G(\Gamma)$, since when Φ is balanced, e_0 is a coloop and $G(\Gamma) = L_0(\Phi) \backslash e_0$. We call the semimatroid associated with $L_0(\Phi)$ and e_0 the *semimatroid of graph balance* of Φ.

(For those concerned with loose and half edges: In this section we treat a loose edge e as a balanced loop and a half edge as an unbalanced loop since in the matroid the two types behave exactly the same.)

Here is how the previous discussion of semimatroids applies to gain graphs. A balanced circle is the same thing as a semimatroid circuit, i.e., it is a matroid circuit (in $L_0(\Phi)$) that is a balanced edge set. A forest F is the same as a balanced independent set; its closure defined in graphical terms is

$$\mathrm{cl}(F) := F \cup \{e \notin F : F \cup \{e\} \text{ contains a balanced circle}\}.$$

The reason is that $C_F(e)$, if it exists, must be balanced so it is a balanced circle; it is called the *fundamental circle* of e with respect to F.

A *broken balanced circle* is a balanced circle with its largest edge removed.

Applying the general semimatroid definitions to balanced independent sets in $L_0(\Phi)$, an edge e is externally active with respect to a forest F when $F \cup e$ contains a balanced circle C (which has to be the fundamental circle) and e is the largest edge in C. For an edge $e \in F$, in $F \setminus e$ one component of F is divided into two; the fundamental relative cocircuit of e with respect to F is the set $D_F(e)$ of edges $f \in E$ that join these two into one and such that $F \cup f$ is balanced. So, e is internally active with respect to F when it is in F and it is the largest edge in $D_F(e)$. (Our definitions of activity differ from the usual ones for graphs because the latter ignore balance.) To clarify these ideas we give two descriptive lemmas; the first is a translation of matroid theory but the second is particular to gain graphs.

Lemma 4.5. *Let F be a forest in Φ. Then $\mathrm{II}(F)$ is the union of all broken balanced circles in F.*

Proof. The circuits that make broken balanced circles in F are contained in $\mathrm{cl}\, F$, which is balanced. The matroid circuits in a balanced set are the balanced circles. Thus, in F a broken balanced circle is the same as a broken circuit. Apply Lemma 4.1. □

Lemma 4.6. *Suppose Φ is a gain graph with no balanced digons. Let F be a forest in Φ. Then $\mathrm{EA}(F)$ is the set of all edges $e \notin F$ such that $e \in (\mathrm{cl}\, F) \setminus F$ and $C_F(e) \setminus e$ is a broken balanced circle.*

Proof. If an edge $e \notin F$ is externally active, it is in $(\mathrm{cl}\, F) \setminus F$ and it is maximal in $C_F(e)$. The latter implies that $C_F(e) \setminus e$ is a broken balanced circle.

To prove the converse, assume $C_F(e)$ exists and $D := C_F(e) \setminus e$ is a broken balanced circle. Either e is maximal in $C_F(e)$, so e is externally active, or $D \neq \varnothing$ and there is an edge $e' \notin F$, other than e, such that $C_F(e') \setminus e' = D$. Then e and e' are parallel links with the same endpoints. Because they form a digon in $\mathrm{cl}\, F$, which is balanced, they form a balanced digon, contrary to the assumption. Consequently, e' cannot exist. □

4.3. **The forest expansion.** The *forest expansion* of (Φ, h) is

$$F_{(\Phi,h),O}(\mathbf{u}, y) := \sum_{F} y^{\varepsilon(F)} \prod_{W\in\pi(F)} u_{h_F(W)}, \tag{4.1}$$

summed over all forests F of Φ.

Theorem 4.7. *The forest expansion is independent of O. Indeed,*

$$F_{(\Phi,h),O}(\mathbf{u}, y) := Q_{(\Phi,h)}(\mathbf{u}, y-1, 0).$$

Proof. Let us expand. In each sum, S is restricted to balanced edge sets that satisfy the stated conditions.

$$\begin{aligned} Q_{(\Phi,h)}(\mathbf{u}, v, 0) &= \sum_{S} v^{|S|-\mathrm{rk}(S)} \prod_{W\in\pi(S)} u_{h_S(W)} \\ &= \sum_{F \text{ forest}} \sum_{\substack{S\supseteq F \\ F(S)=F}} v^{|S\setminus F|} \prod_{W\in\pi(S)} u_{h_S(W)} \\ &= \sum_{F \text{ forest}} \sum_{F\subseteq S\subseteq F\cup \mathrm{EA}(F)} v^{|S\setminus F|} \prod_{W\in\pi(F)} u_{h_F(W)} \end{aligned}$$

by Lemma 4.3, because $\pi(S) = \pi(F)$, and because $\mathrm{EA}(F) \subseteq \mathrm{cl}(F)$ so that $h_S(W) = h_{\mathrm{cl}(F)}(W) = h_F(W)$,

$$= \sum_{F \text{ forest}} (v+1)^{|\,\mathrm{EA}(F)|} \prod_{W\in\pi(F)} u_{h_F(W)}. \qquad \square$$

We hoped for a spanning-tree expansion analogous to Tutte's for graphs, but we could not find one. The problem is that semibases do not span the matroid. When the semimatroid of graph balance is a matroid, as when Φ is balanced, a semibasis is a basis; then for a basis T and an independent set F, $T(F) = T$ if and only if $\mathrm{II}(T) \subseteq F \subseteq T$. This property lets us replace the sum over forests by a sum over spanning trees. We did not find an analogous property of semibases.

5. Coloring

A proper coloration is a way of assigning to each vertex an element of a color set, subject to exclusion rules governed by the edges. The subject of [5] was the problem of counting integral lattice points not contained in specified integral affinographic hyperplanes (see Section 5.4). We solved it by reinterpreting lattice points as proper colorations of a $\mathbb{Z}$-weighted integral gain graph. In this section we develop a theory of proper colorations of all weighted gain graphs. We begin with list coloring, where the weight of a vertex is a finite list of of possible colors (Section 5.1). We then go on to infinite lists with an additional constraint regarded as a variable, e.g., upper bounds that make the effective list finite (Section 5.2); this generalizes ordinary graph k-coloring, in which the list is $\{1, 2, 3, \ldots\}$ restricted by the variable upper bound k. Finally, we apply the general definition to multidimensional integral gain groups and weights, which have the geometrical meaning of counting integer lattice points that lie in a given rectangular parallelepiped but not in any of a family of integral affinographic subspaces (all of which will be explained).

A general notion of proper coloring of gain graphs was developed in [15, Section III.5] (and there called zero-free coloring). There is a *color set* $\mathfrak{C}$, which is any set upon which the gain group $\mathfrak{G}$ has a right action such that the only element of $\mathfrak{G}$ that has any fixed points is the identity. A *coloration* of Φ is any function $x : V \to \mathfrak{C}$. The set of *improper* edges of x is

$$I(x) := \{e_{ij} : x_j = x_i\varphi(e_{ij})\}.$$

The coloration is *proper* if $I(x) = \varnothing$. A basic fact from [15, Section III.5] is that an improper edge set is balanced and closed. For completeness we give the easy proof here.

Lemma 5.1. *The improper edge set $I(x)$ of a coloration is balanced and closed.*

Proof. First we prove balance. Suppose a circle $e_{01}e_{12}\cdots e_{l-1,l} \subseteq I(x)$, where $v_0 = v_l$. Impropriety of the edges implies that $x_l = x_0\varphi(e_{01}e_{12}\cdots e_{l-1,l})$. Thus $x_0 = x_l$ is a fixed point of $\varphi(e_{01}e_{12}\cdots e_{l-1,l})$. By our overall hypothesis that the action is proper, the circle has gain 1. Thus, $I(x)$ is balanced.

Suppose now that e is an edge from v to w in the closure of $I(x)$. Since $I(x)$ is balanced, there is a path $e_{12}\cdots e_{l-1,l}$ in $I(x)$ connecting the endpoints of e (that is, $v_1 = v$ and $v_l = w$) whose gain $\varphi(e_{12}\cdots e_{l-1,l}) = \varphi(e)$. Since $x_l = x_{l-1}\varphi(e_{l-1,l}) = \cdots = x_1\varphi(e_{12})\cdots\varphi(e_{l-1,l}) = x_1\varphi(e_{12}\cdots e_{l-1,l}) = x_1\varphi(e)$, e is improper. Hence, $e \in I(x)$; that is, $I(x)$ is closed. □

In contrast to [15], in this paper we have an infinite color set. We use weights in various ways to limit the possible colorations to a finite set. We have especially in mind two kinds of example. In the first, the group and the color set are both $\mathbb{Z}$ and the color lists are arbitrary subsets of $\mathbb{Z}$ that are bounded below and whose complements are bounded above; but there is a variable upper bound m on the possible colors; thus the number of proper colorations is a function of m. We call this *open-ended list coloring*. In the second, the group and color set are both $\mathbb{Z}^d$, the color lists are dual order ideals in $\mathbb{Z}^d$, and there is a variable upper bound $\mathbf{m}_i$ on the colors that can be used at vertex v_i. (This problem has a nice geometrical interpretation.) We wish to cover both of these examples, as well as similar ones, in a way that exposes to view the essential features; therefore we generalize considerably.

5.1. **List coloring.** A simple kind of list coloring is the basis of all our methods of coloring a weighted gain graph. The idea is to let $\mathfrak{W}$ be any class of subsets of the color set $\mathfrak{C}$ that is closed under intersection and the $\mathfrak{G}$-action; these subsets can be used as vertex color lists.

In list coloring a contracted weight has the formula

$$h_B(W) = \bigcap_{v_i \in W} h_i\eta_B(v_i).$$

(Recall that if $v_i \in W$ and it happens that W has a top vertex t_i, then $\eta_B(v_i) = \varphi(B_{v_i t_i})$.)

We need to switch colorations. If x is a coloration of (Φ, h) and η is a switching function, we define x^η by

$$x^\eta(v_i) := x_i\eta(v_i),$$

the result of the gain-group action on x_i.

Proposition 5.2. *If in (Φ, h) not all vertex lists h_i are finite, then the number of proper colorations is either zero or infinite. If all lists are finite, then the number of proper colorations equals*

$$\sum_{B \in \operatorname{Lat^b}\Phi} \mu(\varnothing, B) \prod_{W \in \pi(B)} |h_B(W)| = \sum_{B \subseteq E:\ \textit{balanced}} (-1)^{|B|} \prod_{W \in \pi(B)} |h_B(W)|,$$

where μ *is the Möbius function of* $\operatorname{Lat^b}\Phi$.

For the Möbius function of a poset see, i.a., [8, 9]; note that $\mu(\varnothing, B) = 0$ if the empty set is not closed, that is, if Φ has a balanced loop or a loose edge. The two sums are equal because $\mu(\varnothing, B) = \sum\{(-1)^{|B'|} : \operatorname{cl}(B') = B\}$ if B is balanced and closed, and $\pi(B') = \pi(B)$.

Proof. First let us suppose h_n is infinite. If there is any proper coloration x, then $x_1, \ldots, x_{n-1}$ prevent x_n from taking on only finitely many possible values, because Φ is finite. There is an infinite number of permitted possible choices of $x_n \in h_n$.

Now we assume all lists are finite. To prove the first part of the formula we use Möbius inversion over $\operatorname{Lat^b}\Phi$ as in [8, p. 362] or [14, Theorem 2.4]. (The second part has a similar proof by inversion over the class of balanced edge sets.) Throughout the proof B denotes an element of $\operatorname{Lat^b}\Phi$. Consider all colorations of (Φ, h), proper or not; let $f(B)$ be the number of colorations x such that $I(x) = B$ and let $g(B)$ be the number of colorations such that $I(x) \supseteq B$. By Lemma 5.1 each coloration is counted in one $f(B)$, so

$$g(A) = \sum_{B \supseteq A} f(B),$$

from which by Möbius inversion

$$f(A) = \sum_{B \supseteq A} \mu(\varnothing, B) g(B).$$

Setting $A = \varnothing$, the total number of proper colorations equals

$$\sum_{B} \mu(\varnothing, B) g(B).$$

We show by a bijection that $g(B)$ is the number of all colorations of $(\Phi, h)/B$, which clearly equals $\prod_{W \in \pi(B)} |h_B(W)|$. Let η_B be the top switching function for B. It is easy to see that switching a coloration x of (Φ, h) gives a coloration of $(\Phi, h)^{\eta_B}$ that is constant on components of B, and conversely. Therefore, if $W \in \pi(B)$, y_W, defined as the common value of $x_i^{\eta_B}$ for every $v_i \in W$, belongs to $h_i^{\eta_B}$ for every $v_i \in W$. When we contract (Φ, h) by B, $y_W \in \bigcap_{v_i \in W} h_i^{\eta_B} = h_B(W)$, so we get a well-defined coloration y of $(\Phi, h)/B$.

Conversely, for any $B \in \operatorname{Lat^b}\Phi$, a coloration y of $(\Phi, h)/B$ pulls back to a coloration of $(\Phi, h)^{\eta_B}$ by $x_i = y_W$ where $v_i \in W \in \pi(B)$. Then switching back to (Φ, h) we have a coloration $x^{\eta_B^{-1}}$ of (Φ, h) whose improper edge set contains B. Since it is clear that these correspondences are inverse to each other, the bijection is proved. □

The last part of the proof can be strengthened to yield a formula for proper colorations of contractions.

Proposition 5.3. *If all vertex lists* h_i *are finite and* B *is a balanced edge set, then the number of colorations of* (Φ, h) *whose balanced edge set equals* B *equals the number of proper colorations of* $(\Phi, h)/B$.

The proof is a simple modification of the evaluation of $g(B)$ in the previous proof, and is also a simple generalization of the evaluation of $f(B)$ in the proof of [5, Theorem 3.3]. (In [5] we accidentally wrote $f(B)$ when we meant $g(B)$; but that led us to write a proof of Proposition 5.3 in the special situation of [5]. We thank Seth Chaiken for pointing out the error in [5].)

Let $\mathcal{P}_{\text{fin}}(\mathfrak{C})$ be the class of finite subsets of $\mathfrak{C}$, and let us call a weighted gain graph with weights in $\mathcal{P}_{\text{fin}}(\mathfrak{C})$ *finitely list weighted*. Then the total dichromatic polynomial has a variable u_h for each finite subset $h \subseteq \mathfrak{C}$.

Theorem 5.4. *If (Φ, h) is finitely list weighted, then the number of proper colorations equals $(-1)^n Q_{(\Phi,h)}(\mathbf{u}, -1, 0)$ evaluated at $u_h = -|h|$.*

Proof. The proof is by comparing the second formula of Proposition 5.2 to the definition of $Q_{(\Phi,h)}$. □

Call a *signed Tutte invariant* any function that satisfies (Tii–iv) and the modified form of (Ti),

(Ti′) (Subtractivity) For every link e,

$$f(\Phi, h) = f(\Phi \setminus e, h) - f(\Phi/e, h/e).$$

It is clear that f is a signed Tutte invariant if and only if $(-1)^{|V|} f$ is a Tutte invariant.

Corollary 5.5. *Given a gain group $\mathfrak{G}$ and a color set $\mathfrak{C}$, the number of proper colorations is a signed Tutte invariant of finitely list weighted gain graphs with gains in $\mathfrak{G}$.*

Proof. The family of finitely list weighted $\mathfrak{G}$-gain graphs is closed under deletion and contraction because finiteness of lists is preserved by those operations. Apply Theorem 5.4. □

Example 5.1 (*Finite lists*). If $\mathfrak{C}$ is partially ordered, take $\mathfrak{W}$ to consist of all finite order ideals, or all finite intervals. The special case $\mathfrak{C} = \mathbb{Z}^d$ is our main example.

5.2. **Filtered lists.** In examples the color lists for the vertices are not always finite. A general picture is that there is a list h_i for each vertex, which is some subset of $\mathfrak{C}$, and there is also a set $M_i \subseteq \mathfrak{C}$ that acts as a filter of colors: a color must lie not only in its vertex list but also in M_i. Thus we have a function $\chi_{(\Phi,h)}(M)$ defined for $M := (M_1, \ldots, M_n) \in \mathcal{P}(\mathfrak{C})^n$ whose value is the number of proper colorations of (Φ, h) using only colors in M_i at vertex v_i. This is the *list chromatic function* of (Φ, h). An example, of course, is the quantity of Proposition 5.2, which equals $\chi_{(\Phi,h)}(\mathfrak{C}^n)$ (finite lists with no filtering). That proposition has the following extension. We define switching of a color filter M and its contraction $M_B(W)$ just as for weights, so that

$$M_B(W) := \bigcap_{v_i \in W} M_i \eta_B(v_i).$$

Proposition 5.6. *If every intersection $M_i \cap h_i$ is finite, then*

$$\chi_{(\Phi,h)}(M) = \sum_{B \in \mathrm{Lat}^{\mathrm{b}} \Phi} \mu(\varnothing, B) \prod_{W \in \pi(B)} |h_B(W) \cap M_B(W)|.$$

Proof. In Proposition 5.2 replace h_i by $h_i \cap M_i$. □

Since color filters contract like weights, we can form a *doubly weighted* gain graph by taking new weights (h_i, M_i), provided we define a $\mathfrak{G}$-invariant semigroup $\mathfrak{M}_0$ from which the double weights are drawn. To that end, let

$$\mathfrak{M}_0 := \{(h', M') : h' \in \mathfrak{W},\ M' \subseteq \mathfrak{C},\ \text{and } M' \cap h' \text{ is finite}\}.$$

The semigroup operation is componentwise intersection, i.e., $(h', M') \cap (h'', M'') := (h' \cap h'', M' \cap M'')$. Given this weight semigroup, there is a doubly weighted total dichromatic polynomial, which we write $Q_{(\Phi,h),M}(\mathbf{u},v,z)$, with $M = (M_1,\dots,M_n)$, to emphasize the different roles of h and M. The variables are now $u_{h',M'}$ for each pair $(h',M') \in \mathfrak{M}_0$, and of course v and z, and the formula for the doubly weighted polynomial is

$$Q_{(\Phi,h),M}(\mathbf{u},v,z) = \sum_{S\subseteq E} v^{|S|-n+b(S)} z^{c(S)-b(S)} \prod_{W\in\pi_{\mathrm{b}}(S)} u_{h_S(W),M_S(W)},$$

where, as usual, $h_S := h^{\eta_S}$ and $M_S := M^{\eta_S}$. It is easy to see that every $(h_S(W), M_S(W)) \in \mathfrak{M}_0$ if every $(h_i, M_i) \in \mathfrak{M}_0$.

Theorem 5.7. *If (Φ,h) and $M_1,\dots,M_n \subseteq \mathfrak{C}$ are such that the filtered list $M_i \cap h_i$ is finite for each vertex v_i, then the list chromatic function $\chi_{(\Phi,h)}(M_1,\dots,M_n)$ is obtained from $(-1)^n Q_{(\Phi,h),M}(\mathbf{u},-1,0)$ by setting $u_{h',M'} = -|M'\cap h'|$ for each $h' \in \mathfrak{W}$ and $M' \subseteq \mathfrak{C}$.*

Proof. Like that of Theorem 5.4, but from Proposition 5.6. ☐

Fix the gain group $\mathfrak{G}$, color set $\mathfrak{C}$, and weight subsemigroup $\mathfrak{W} \subseteq (\mathcal{P}(\mathfrak{C}),\cap)$. Consider any weighted gain graph (Φ,h) with gains in $\mathfrak{G}$ and weights (functioning as vertex color lists) in $\mathfrak{W}$. Allow any collection of color filters $(M_1,\dots,M_n)$ for which all $h_i \cap M_i$ are finite; we call $((\Phi,h),M)$ a *finitely filtered, list weighted gain graph.* This gives us a list chromatic function $\chi_{(\Phi,h)}(M_1,\dots,M_n)$ that is always a well defined nonnegative integer. Thinking of $((\Phi,h),M)$ as a (doubly) weighted gain graph, we have deletions and contractions and we can ask about Tutte invariance.

Corollary 5.8. *The list chromatic function $\chi_{(\Phi,h)}(M_1,\dots,M_n)$ is a signed Tutte invariant of finitely filtered, list weighted gain graphs.*

Proof. As we noted, the class of list weighted gain graphs with suitable arguments is closed under deletion and contraction. Now, apply Theorem 5.7. ☐

Example 5.2 (*Locally finite join semilattice*). Suppose $\mathfrak{C}$ is a locally finite join semilattice. We may take $\mathfrak{W}$ to be the set of principal dual order ideals $\langle z\rangle^*$ and let the M_i range over principal ideals. The join operation makes $\mathfrak{W}$ an intersection semigroup, as the intersection $\langle z\rangle^* \cap \langle z'\rangle^*$ is the principal dual ideal $\langle z \vee z'\rangle^*$; and $\mathfrak{W}$ is clearly $\mathfrak{G}$-invariant. The intersection $M_i \cap h_i$ will always be finite so the preceding proposition and theorem apply.

We may weaken the assumptions. Let $\mathfrak{W}$ be the class of all subsets of $\mathfrak{C}$ that have a lower bound; that is, subsets of principal dual ideals. And let M_i be any subset with an upper bound; that is, a subset of a principal ideal. Then $M_i \cap h_i$ is necessarily finite so the preceding results hold good. That is, we admit as filters all upper-bounded subsets of $\mathfrak{C}$.

An example of this kind is that in which $\mathfrak{C} = \mathbb{Z}^d$; it is the topic of the next subsection.

5.3. **Open-ended list coloring in an integer lattice.** To introduce the main applications we turn once again to the original example from [5] but with a slight change in viewpoint.

Example 5.3 (*Open-ended interval coloring*). The gain group is $\mathbb{Z}$ and the weights can be treated as upper intervals, (h_i,∞) at vertex v_i. A *proper m-coloration* of (Φ,h) is a function $x : V \to \mathbb{Z}$ such that each $x_i \in (h_i,\infty)$ and all $x_i \leq m$. One can think of this as a list coloration in which the list for each vertex is an interval that grows with m. The *integral chromatic function* $\chi^{\mathbb{Z}}_{(\Phi,h)}(m)$ of [5] counts proper m-colorations. This function is obtained from $Q(\mathbf{u},-1,0)$ by substituting $u_k = -\max(m-k,0)$. Hence it is a signed Tutte invariant, as we showed in [5] directly from its counting definition. We showed in [5] that it is eventually a monic polynomial of degree $n = |V|$, and that it is a sum of simple terms that appear successively as m increases.

We generalize this example in several ways: to higher-dimensional coloring, to upper bounds that depend on the vertex, and to arbitrary vertex lists.

For higher-dimensional coloring the color set $\mathfrak{C}$ is the d-dimensional integer lattice $\mathbb{Z}^d$ with the componentwise partial ordering and the gain group $\mathfrak{G}$ is the additive group $\mathbb{Z}^d$ acting on $\mathfrak{C}$ by translation. A coloration is any $\mathbf{x} : V \to \mathbb{Z}^d$. The weight semigroup $\mathfrak{W}$ is either of the classes

$$\mathfrak{W}_1 := \{h' \subset \mathbb{Z} : h' \text{ is bounded below}\},$$

that is, h' is contained in a cone $\langle \mathbf{a} \rangle^* = \bigtimes_{k=1}^{d} [a_k, \infty)$ for some $\mathbf{a} \in \mathbb{Z}^d$, and

$$\mathfrak{W}_2 := \{h' \in \mathfrak{W}_1 : \text{ for some } \mathbf{a} \in \mathbb{Z}^d,\ h' \subseteq \langle \mathbf{a} \rangle^* \text{ and } \langle \mathbf{a} \rangle^* \setminus h' \text{ is bounded above}\},$$

that is, h' is all of $\langle \mathbf{a} \rangle^*$ except for a finite subset. Both classes $\mathfrak{W}_1$ and $\mathfrak{W}_2$ are closed under intersection of pairs and under translation. For $h' \in \mathfrak{W}_1$ we define $\mathbf{h}'$ as the meet of the members of h'; when $h' \in \mathfrak{W}_2$ and $d > 1$ this is the only possible $\mathbf{a}$, but in dimension $d = 1$ it is the largest possible $\mathbf{a}$ and also the smallest element of h'.

The color filters M_i are principal order ideals $\langle \mathbf{m}_i \rangle = \bigtimes_{k=1}^{d} (-\infty, m_{ik}]$ for $\mathbf{m}_i \in \mathbb{Z}^d$ (we write $\mathbf{m}_i := (m_{i1}, \ldots, m_{id})$), so the list chromatic function has domain $\mathbb{Z}^d$ and is defined by

$$\begin{aligned} \chi_{(\Phi,h)}(\mathbf{m}) := \text{ the number of } \mathbf{x} : V \to \mathbb{Z}^d \text{ such that each } \mathbf{x}_i \in h_i,\ \mathbf{x} \le \mathbf{m}, \\ \text{and } \mathbf{x}_j \ne \mathbf{x}_i + \varphi(e_{ij}) \text{ for each edge } e_{ij}, \end{aligned}$$

where $\mathbf{m} := (\mathbf{m}_1, \ldots, \mathbf{m}_n)$. This number is a function of one variable m_{ik} for each $i = 1, \ldots, n$ and $k = 1, \ldots, d$ and is finite for each $\mathbf{m}$. Our general theory shows that $\chi_{(\Phi,h)}(\mathbf{m})$ is an evaluation of the total dichromatic polynomial; and now we can generalize Example 5.3. In the present case a switching function is $\eta : V \to \mathbb{Z}^d$ and the contraction formula for weights takes the form

$$h_B(W) = \bigcap_{v_i \in W} (h_i + \eta_B(v_i))$$

if B is a balanced edge set, where $h_i + \mathbf{a}$ denotes the translation of h_i by $\mathbf{a}$. In the total dichromatic polynomial there is one variable $u_{h',\mathbf{m}'}$ for each $h' \in \mathfrak{W}_1$ and $\mathbf{m}' \in (\mathbb{Z}^d)^n$. (Variables with empty $h' \cap \langle \mathbf{m}' \rangle$ can be omitted.)

Theorem 5.9. *With all $h_i \in \mathfrak{W}_1$, the list chromatic function $\chi_{(\Phi,h)}(\mathbf{m})$ is obtained from $(-1)^n Q_{(\Phi,h),M}(\mathbf{u}, -1, 0)$ by setting $u_{h',\mathbf{m}'} = -\big|h' \cap \langle \mathbf{m}' \rangle\big|$ for each $h' \in \mathfrak{W}_1$ and $\mathbf{m}' \in \mathbb{Z}^d$.*

Proof. A corollary of Theorem 5.7, since the intersections $h_i \cap M_i$ are finite. □

Corollary 5.10. *The list chromatic function is a signed Tutte invariant of gain graphs with gain group $\mathbb{Z}^d$ and weight lists belonging to $\mathfrak{W}_1$.*

Proof. A special case of Corollary 5.8. □

For $h_i \in \mathfrak{W}_1$ define $\mathbf{h}_i^- := \mathbf{h}_i - (1, \ldots, 1)$ and $H_i := \langle \mathbf{h}_i \rangle^* \setminus h_i$, and let $\hat{\mathbf{h}}_i := \bigvee H_i$, except that $\hat{\mathbf{h}}_i = \mathbf{h}_i^-$ if $H_i = \varnothing$. Clearly, $\hat{\mathbf{h}}_i$ is defined in $\mathbb{Z}^d$ if and only if H_i is finite, that is, $h_i \in \mathfrak{W}_2$. For vertices v_i and v_j, let

$$\alpha_{ji} := \bigvee_{P_{ji}} \varphi(P_{ji}),$$

where P_{ji} ranges over all paths in Φ from v_j to v_i, and let $\alpha_j := \bigvee_i \alpha_{ji}$, the least upper bound of the gains of all paths that begin at v_j.

A function $p(y_1, \ldots, y_r)$ is a *piecewise polynomial* if it is defined on a domain in $\mathbb{R}^r$ that is a union of a finite number of closed, full-dimensional sets D_σ, each containing infinitely many integer points, on each of which p is a polynomial $p_\sigma(y_1, \ldots, y_r)$. By saying p has leading term $y_1 \cdots y_r$, or has degree at most 1 in each variable, we mean that each p_σ has that property. This definition is chosen to suit the following result. The numbers h_{ik} are the components of $\mathbf{h}_i$.

Theorem 5.11. *Assume Φ has no balanced loops or loose edges. Suppose all $h_i \in \mathfrak{W}_2$. Define*

$$q_k(B,W) := \min_{v_i \in W} \big(m_{ik} + \eta_B(v_i)\big) - \max_{v_i \in W} \big(h_{ik} + \eta_B(v_i)\big) + 1$$

for $B \in \operatorname{Lat^b} \Phi$ and $W \in \pi(B)$, and

$$p(\mathbf{m}) := \sum_{B \in \operatorname{Lat^b} \Phi} \mu(\varnothing, B) \prod_{W \in \pi(B)} \left(\prod_{k=1}^{d} q_k(B,W) - \Big| \bigcup_{v_i \in W} \big(H_i + \eta_B(v_i)\big) \Big| \right),$$

which is a piecewise polynomial function of the nd variables m_{ik} having degree at most 1 in each variable and leading term $\prod_{i-1}^{n} \prod_{k=1}^{d} m_{ik}$. The list chromatic function $\chi_{(\Phi,h)}(\mathbf{m})$ equals $p(\mathbf{m})$ for all $\mathbf{m} \geq \mathbf{m}_0$, where

$$\mathbf{m}_{0i} := \bigvee_{j=1}^{n} \big(\hat{\mathbf{h}}_j + \alpha_{ji}\big).$$

In the theorem, $r = nd$, the variables are $x_{11}, x_{12}, \ldots, x_{1d}, x_{21}, \ldots, x_{nd}$, and the domains D_σ are the sets on which each of the d sets (one for each fixed $k \leq d$) of shifted variables $m_k(B,W) := \min_{v_i \in W} \big(m_{ik} + \eta_B(v_i)\big)$ (one variable for each B and each $W \in \pi(B)$) assumes a particular weakly increasing order, since the orderings of these variables determine exactly which polynomial $p(\mathbf{m})$ is.

Proof. We apply the formula of Proposition 5.6 in the form

$$\sum_{B \in \operatorname{Lat^b} \Phi} \mu(\varnothing, B) \prod_{W \in \pi(B)} |h_B(W) \cap M_B(W)|,$$

which shows that the theorem is true when the range of $\mathbf{m}$ is such that $|h_B(W) \cap M_B(W)|$ is a polynomial of degree 1 in each m_{ik} such that $v_i \in W$. Rewrite the expression:

$$\begin{aligned} h_B(W) \cap M_B(W) &= \bigcap_{v_i \in W} (h_i + \eta_B) \cap \bigcap_{v_i \in W} \langle \mathbf{m}_i + \eta_B(v_i) \rangle \\ &= \bigcap_{v_i \in W} \big([\langle \mathbf{h}_i \rangle^* \setminus H_i] + \eta_B(v_i)\big) \cap \bigcap_{v_i \in W} \langle \mathbf{m}_i + \eta_B(v_i) \rangle \\ &= \Big(\Big[\bigvee_{v_i \in W} \mathbf{h}_i, \bigwedge_{v_i \in W} \mathbf{m}_i \Big] + \eta_B(v_i)\Big) \setminus \bigcup_{v_i \in W} \big(H_i + \eta(v_i)\big) \end{aligned} \tag{5.1}$$

since $h_i = \langle \mathbf{h}_i \rangle^* \setminus H_i$ and $M_i = \langle \mathbf{m}_i \rangle$, where $[\mathbf{x}, \mathbf{y}]$ denotes an interval in the lattice $\mathbb{Z}^d$.

Let us see what natural conditions are sufficient for piecewise polynomiality (with the specified leading term) when $\mathbf{m}$ is large. It should be true for each factor in each term.

Consider $W = \{v_i\}$, which is a component when $B = \varnothing$. We can compute the factor corresponding to this W; it is

$$\begin{aligned}\big|\big([\mathbf{h}_i, \mathbf{m}_i] + \eta_B(v_i)\big) \setminus \big(H_i + \eta_B(v_i)\big)\big| &= \big|[\mathbf{h}_i, \mathbf{m}_i] \setminus \big(H_i \cap \langle \mathbf{m}_i \rangle\big)\big| \\ &= \prod_{k=1}^{d} (m_{ik} - h_{ik} + 1)^+ - \big|H_i \cap \langle \mathbf{m}_i \rangle\big|.\end{aligned} \tag{5.2}$$

This is not a piecewise polynomial function unless, firstly, $\mathbf{m}_i \geq \mathbf{h}_i^-$, and secondly, $|H_i \cap \langle \mathbf{m}_i \rangle|$ is a constant. The way to ensure the latter is for H_i to be contained in $\langle \mathbf{m}_i \rangle$, or equivalently $\mathbf{m}_i \geq \hat{\mathbf{h}}_i$. We assume this from now on.

Treating in the same manner each factor in Equation (5.1), rewritten as

$$|h_B(W) \cap M_B(W)| = \Big[\bigvee_{v_i \in W} \big(\mathbf{h}_i + \eta_B(v_i)\big), \bigwedge_{v_i \in W} \big(\mathbf{m}_i + \eta_B(v_i)\big) \Big] \setminus \bigcup_{v_i \in W} \big(H_i + \eta_B(v_i)\big),$$

we see that piecewise polynomiality is ensured if

$$\bigwedge_{v_i \in W} \big(\mathbf{m}_i + \eta_B(v_i)\big) \geq \bigvee_{v_i \in W} \bigvee \big(H_i + \eta_B(v_i)\big) = \bigvee_{v_i \in W} \big(\hat{\mathbf{h}}_i + \eta_B(v_i)\big).$$

This is equivalent to having $\mathbf{m}_i + \eta_B(v_i) \geq \hat{\mathbf{h}}_j + \eta_B(v_j)$ for every $v_i, v_j \in W$, or, rewriting again, $\mathbf{m}_i \geq \hat{\mathbf{h}}_j + \eta_B(v_j) - \eta_B(v_i)$. Now recall from Equation (2.1) that $\eta_B(v_j) - \eta_B(v_i) = \varphi(B_{ji})$. Here B_{ji} is any path in B from v_j to v_i. As B is any balanced, closed set, we can take B to be the closure of any path P_{ji} from v_j to v_i in Φ. Thus, to ensure piecewise polynomiality we require that $\mathbf{m}_i \geq \hat{\mathbf{h}}_j + \varphi(P_{ji})$ for every path P_{ji}; that is, $\mathbf{m}_i \geq \hat{\mathbf{h}}_j + \alpha_{ji}$.

We have found a sufficient condition on $\mathbf{m}$ for $\chi_{(\Phi,h)}(\mathbf{m})$ to be a piecewise polynomial function. The term of highest degree is that for which $\pi(B)$ has the most components; that is n components when $B = \varnothing$. The corresponding term is the product of the factors associated with singleton sets W; these factors are all monic piecewise polynomials of total degree d, as we see in Equation (5.2). □

The proof suggests that the theorem's lower bound on $\mathbf{m}$ is essential; for any other choice of $\mathbf{m}$, $p(\mathbf{m})$ will not agree with $\chi_{(\Phi,h)}(\mathbf{m})$. One reason is that, if $\mathbf{m} \not\geq \mathbf{h}$, then $p(\mathbf{m})$ does not extract the positive part of $q_k(B, W)$. A more subtle one is that the constant term $\big|\bigcup_{v_i \in W} \big(H_i + \eta_B(v_i)\big)\big|$ in the factor of W assumes that $\bigcup_{v_i \in W} \big(H_i + \eta_B(v_i)\big)$ is contained in $\big\langle \bigwedge_{v_i \in W} \big(\mathbf{m}_i + \eta_B(v_i)\big)\big\rangle$. However, we have not tried to prove necessity of the lower bound, and there might be exceptions. We have also not tried to decide whether the domain on which $\chi_{(\Phi,h)}(\mathbf{m})$ is a piecewise polynomial with the right leading term (though not necessarily agreeing with p) is larger than $\langle \mathbf{m}_0 \rangle^*$.

It is clear, though, why h_i has to be bounded below and its cone complement H_i must be bounded above. If some h_i has no lower bound, then $\chi_{(\Phi,h)}(m)$ will be infinite. Even when each h_i is bounded below, if some H_i has no upper bound then $\chi_{(\Phi,h)}(m)$ will not become a polynomial when m is large.

In the one-dimensional case, where the gain group and color set are $\mathbb{Z}$,

$$\mathfrak{W}_2 = \{h' \subset \mathbb{Z} : h' \text{ is bounded below and } \mathbb{Z} \setminus h' \text{ is bounded above}\}$$

and $\mathbf{m} = (m_1, \ldots, m_n) \in \mathbb{Z}^n$. In this case $\mathbf{h}_i = \min h_i$ and $\hat{\mathbf{h}}_i = \max(\mathbb{Z} \setminus h_i)$; and α_{ji} is the largest gain of a path in Φ from v_j to v_i.

Corollary 5.12. *For an integral gain graph without balanced loops or loose edges, and with all weights* $h_i \in \mathfrak{W}_2$*,* $\chi_{(\Phi,h)}(m_1,\ldots,m_n)$ *is a monic polynomial in the n variables* m_i *for large enough* m_i*'s, linear in each variable and with leading term* $m_1 \cdots m_n$*. Polynomiality holds when all*

$$m_i \geq m_{0i} := \max_{j=1,\ldots,n} \big[\max(\mathbb{Z} \setminus h_j) + \alpha_{ji} \big].$$

The theorem simplifies when all $\mathbf{m}_i$ equal a common value $\mathbf{m}' \in \mathbb{Z}^d$.

Corollary 5.13. *For a* $\mathbb{Z}^d$*-gain graph* Φ *with no balanced loops or loose edges, suppose all weights* $h_i \in \mathfrak{W}_2$*. Define*

$$\bar{q}_k(B,W) := m'_k - \max_{v_i,v_j \in W} \big(h_{ik} + \eta_B(v_i)_k - \eta_B(v_j)_k \big) + 1$$

for $B \in \operatorname{Lat}^{\mathrm{b}} \Phi$*,* $W \in \pi(B)$*, and* $\mathbf{m}' \in \mathbb{Z}^d$*. For large enough* $\mathbf{m}' \in \mathbb{R}^d$*, the list chromatic function* $\chi_{(\Phi,h)}(\mathbf{m}',\ldots,\mathbf{m}')$ *equals*

$$\bar{p}(\mathbf{m}') := \sum_{B \in \operatorname{Lat}^{\mathrm{b}} \Phi} \mu(\varnothing, B) \prod_{W \in \pi(B)} \left(\prod_{k=1}^{d} \bar{q}_k(B,W) - \Big| \bigcup_{v_i \in W} \big(H_i + \eta_B(v_i) \big) \Big| \right),$$

a polynomial function of the d variables m'_k *having degree at most n in each variable and leading term* $\prod_{k=1}^d (m'_k)^n$*. The equation* $\chi_{(\Phi,h)}(\mathbf{m}',\ldots,\mathbf{m}') = p(\mathbf{m}')$ *holds true for all*

$$\mathbf{m}' \geq \bigvee_{j=1}^{n} \big(\hat{\mathbf{h}}_j + \alpha_j \big).$$

Now we begin to justify the claim that a total dichromatic polynomial connected with an interesting combinatorial problem has an uncountable number of variables. (We assume the reader finds the list chromatic function with gains in $\mathbb{Z}$ or $\mathbb{Z}^d$ interesting, or this argument fails! The geometrization in the next subsection may add to the interest.) The number of variables $u_{h',\mathbf{m}'}$ when the weight semigroup is $\mathfrak{W}_1$ (and $\mathbf{m}' \in \mathbb{Z}^d$) is $|\mathfrak{M}_0| = |\mathfrak{W}_1| \cdot |\mathbb{Z}^d|$. As $\mathfrak{W}_1$ contains every subset of the natural numbers, its cardinality is that of the continuum.

On the other hand, $|\mathfrak{W}_2| = \aleph_0$, for $\mathfrak{W}_2$ is a countable union of countable sets. We see this by describing $h \in \mathfrak{W}_2$ as an ordered pair $(\mathbf{h}', h' \setminus \langle \mathbf{a} \rangle^*)$. There is a countable number of pairs $(\mathbf{a}, X)$ of this type, for each $\mathbf{a}$, and the number of integer vectors $\mathbf{a}$ is countable.

5.4. **Arrangements of affinographic hyperplanes and affinographic matrix subspaces.** First we prove the geometrical theorems stated in the introduction. Then we generalize them to matrix subspaces. We restate Theorem 1.1 in a more sophisticated but equivalent form. As usual, Φ is the gain graph corresponding to $\mathcal{A}$; t_k is a top vertex of the component B_k of B, whose vertex set is W_k; and μ is the Möbius function of the semilattice $\operatorname{Lat}^{\mathrm{b}} \Phi$.

Theorem 5.14. *With P and* $\mathcal{A}$ *as in Theorem 1.1, the number of integer points in* $P \setminus \bigcup \mathcal{A}$ *equals*

$$\chi_{(\Phi,0)}(m_1,\ldots,m_n) = \sum_{B \in \operatorname{Lat}^{\mathrm{b}} \Phi} \mu(\varnothing, B) \prod_{W_k \in \pi(B)} \big(1 + \min_{v_i \in W} [m_i + \varphi(B_{v_i t_k})] - g_k \big)^{+}.$$

Proof. The lattice points to be counted are simply proper colorations in disguise. In Proposition 5.6 take the list for v_i to be $h_i = [0, \infty)$ and the filter to be $M_i = (-\infty, m_i]$, and then sort through the definitions. E.g., from Equation (2.3), if v_i is in the component B_k of B, then $\eta_B(v_i) = \varphi(B_{v_i t_k})$. Also, if $W = V(B_k)$, then $h_B(W) = [g_k, \infty)$ where g_k is the largest gain of a path in B_k, and $M_B(W) = (-\infty, \min_{v_i \in W} m_i + \varphi(B_{v_i t_k})]$. Hence, the factor in Proposition 5.6 equals

$$\left|[g_k, \min_{v_i \in W} m_i + \varphi(B_{v_i t_k})]\right| = \left(1 + \min_{v_i \in W}[m_i + \varphi(B_{v_i t_k})] - g_k\right)^+.$$

Thus we have Theorem 1.1. Theorem 5.14 follows by the formula for μ given at Proposition 5.2. □

Proof of Theorem 1.2. This follows directly from Proposition 5.2 and the formula for $\eta_B(v_i)$. □

We state one more theorem, a combination of the previous two. Here we have a list L_i of nonnegative integral permitted values for each coordinate, which may be infinite, and we also have an upper bound m_i, which we treat as a variable. Again we let $P := [0, m_1] \times \cdots \times [0, m_n]$ where the m_i are integers.

Theorem 5.15. *The number of points in $P \cap (L_1 \times \cdots \times L_n)$ but not in any of the hyperplanes of the arrangement $\mathcal{A}$ equals*

$$\sum_{B \in \operatorname{Lat}^{\mathrm{b}} \Phi} \mu(\varnothing, B) \prod_{B_k} \Bigg| \bigcap_{v_i \in V(B_k)} \big((L_i \cap [0, m_i]) + \varphi(B_{v_i t_k})\big) \Bigg|,$$

where the product is over all components of B.

Proof. The proof is similar to that of Theorem 1.2. □

When L_i has finite complement in the nonnegative integers, then for sufficiently large variables m_i this count is a polynomial in the variables, just as in Theorem 1.1.

Theorem 5.11 allows us to count integer matrices that are contained in an orthotope but not in any of a finite set of subspaces that are determined by affinographic equations.

Write $\mathbb{Z}^{n \times d}$ for the lattice of $n \times d$ integer matrices and $\mathbb{R}^{n \times d}$ for the real vector space that contains them; if X is a matrix, we write $\mathbf{x}_i = (x_{i1}, \ldots, x_{id})$ for the ith row vector, an element of $\mathbb{R}^d$. An integral orthotope $[H, M]$, where H and M are integer matrices with $H \leq M$, is the convex polytope given by the constraints $H \leq X \leq M$ in $\mathbb{R}^{n \times d}$.

We call a subspace determined by an equation of the form $\mathbf{x}_j = \mathbf{x}_i + \mathbf{a}$ *row-affinographic*, and *integral* if $\mathbf{a}$ is an integral vector in $\mathbb{R}^d$. (The name "affinographic" comes from the fact that such a subspace is an affine translate of a graphic subspace, i.e., a subspace defined by lists of equal coordinates, in this case by the equation $\mathbf{x}_j = \mathbf{x}_i$.) A finite set $\mathcal{S}$ of such subspaces is an *integral row-affinographic subspace arrangement.*

We want to know the number N of integral matrices in an integral orthotope $[H, M]$ but not in any of the subspaces of $\mathcal{S}$. This number is given by Theorem 5.11. Rather than translate the theorem into purely geometrical language, which seems unnatural, we explain how to set up a weighted gain graph (Φ, h) to which it applies, thereby getting the formula

$$N = \chi_{(\Phi, h)}(\mathbf{m}_1, \ldots, \mathbf{m}_n).$$

There is one vertex for each row of the matrices; thus, $V = \{v_1, \ldots, v_n\}$. There is one edge for each subspace; that with equation $\mathbf{x}_j = \mathbf{x}_i + \mathbf{a}$ becomes an edge from v_i to v_j with gain

$\mathbf{a}$ (in that direction; the gain from v_j to v_i is $-\mathbf{a}$). The weight of v_i is the cone $\langle \mathbf{h}_i \rangle^*$. An integral matrix X in the orthotope becomes a coloration, the color of v_i being the ith row vector $\mathbf{x}_i$. It is now clear that an integral matrix that we wish to count is precisely the same as a proper coloration of (Φ, h) that satisfies the upper bound $(\mathbf{m}_1, \ldots, \mathbf{m}_n)$.

6. Graphs without gains: Noble and Welsh generalized

We think of a graph without gains (and with no loose or half edges) as having all gains 1 (or 0 if the gain group is additive). It is instructive to see what our results say here. We write Γ for $\Phi = (\Gamma, 1)$ to emphasize that, the gains being fixed, the only significant datum is the graph. Since the graph is balanced, $b(S) = c(S)$ and $\pi_{\mathrm{b}}(S)$ is a partition of V for every edge set S.

A $\mathfrak{W}$-*weighted graph* is a pair (Γ, h) where $h : V \to \mathfrak{W}$. There is no need for switching; thus contraction is ordinary graph contraction together with contraction of h to

$$h(W) = \sum_{v_i \in W} h_i$$

for $W \in \pi(S)$, where summation means the semigroup operation and the subscript S is superfluous because there is no switching. The total dichromatic polynomial becomes

$$Q_{(\Gamma,h)}(\mathbf{u}, v, z) = \sum_{S \subseteq E} v^{|S|-n+c(S)} \prod_{W \in \pi(S)} u_{h(W)} \tag{6.1}$$

with tree expansion

$$= \sum_T (v+1)^{\varepsilon(T)} \sum_{\substack{F \subseteq T \\ F \supseteq \bar{\Pi}(T)}} \prod_{W \in \pi(F)} u_{h(W)}. \tag{6.2}$$

Observe that z drops out; thus we write $Q_{(\Gamma,h)}(\mathbf{u}, v)$ for this polynomial.

These graphs with weights but no gains subsume the weighted graphs (Γ, ω) of Noble and Welsh [7], which have positive integral vertex weights. Indeed, their work largely inspired our generalization to semigroup weights. At first we had the total dichromatic polynomial only for weighted integral gain graphs with integral weights, but we compared their definitions to ours and noticed remarkable analogies. Noble and Welsh's weights add: $\omega(W) = \sum_{w \in W} \omega(w)$, while the weights on weighted integral gain graphs maximize. Our polynomial $Q_{(\Gamma,h)}$ (for weighted integral gain graphs) and the polynomial $W_{(\Gamma,\omega)}$ of [7] have virtually the same variables (if one makes simple substitutions) and satisfy the same Tutte relations (Ti–Tiii), initial conditions (3.2), and loop reduction identity (3.3). We had to suspect a common generalization. This paper is the result.

The theorem without gains is stronger than our broader results.

Theorem 6.1. *Given an abelian semigroup $\mathfrak{W}$, the polynomial-valued function $(\Gamma, h) \mapsto Q_{(\Gamma,h)}(\mathbf{u}, v)$ of $\mathfrak{W}$-weighted graphs is universal with the properties* (Ti), (Tiii), (Tiv), (3.2), *and* (3.3); (Tii) *holds; and there is a tree expansion as in* (6.2).

Proof. The proof is like that of Noble and Welsh. □

The treatment of coloring in Section 5 applies to ordinary graphs, without gains, simply by taking $\varphi \equiv 1$, the identity. The only differences are that every edge set is balanced, so $\mathrm{Lat}^{\mathrm{b}}\,\Phi$ is the class of all closed edge sets (sets B such that any edge whose endpoints are connected by B is itself in B), and that $h_B(W)$ becomes simply $h(W) = \bigcap_{v_i \in W} h_i$. Also, the weight semigroup $\mathfrak{W}$ need only be closed under intersection, as there is no group action constraining it.

Taking gain group $\mathbb{Z}^d$ as in Section 5.3, so that the graph can be treated as having all zero gains, we have the following corollary of Theorem 5.11; the notation is that of the theorem.

Corollary 6.2. *Let (Γ, h) be a weighted graph with weights $h_i \in \mathfrak{W}_2$. Assume Γ has no balanced loops or loose edges. For large enough $\mathbf{m}$, $\chi_{(\Gamma,h)}(\mathbf{m})$ is a monic polynomial function of the nd variables m_{ik}, having degree 1 in each variable and highest-degree term $\prod_{i=1}^n \prod_{k=1}^d m_{ik}$. Polynomiality holds when all $\mathbf{m}_i \geq \bigvee_{j=1}^n \hat{\mathbf{h}}_j$.*

When the weights are principal dual ideals $\langle \mathbf{h}_i \rangle$, the lower bound on $\mathbf{m}_i$ is $\bigvee_j \mathbf{h}_j^-$.

7. *Caveat lector*

7.1. **Modular gains.** We know we have not found the universal Tutte invariant of weighted gain graphs. (In defining a Tutte invariant here, we take the detachable edges to be those that are not links. This is equivalent to allowing contraction only of balanced edge sets, and it is not precisely the same definition as used elsewhere, e.g., in [15, Part III]. We take the components to be the connected components.)

Choosing an example from [5] (which the reader may skip; we are about to generalize it), the modular chromatic function $\chi_\Phi^{\mathrm{mod}}(m)$ of a rooted integral gain graph is a Tutte invariant (once it has been multiplied by $(-1)^n$) but it is not obtainable as an evaluation of the total dichromatic polynomial. One can prove this by noting that $\chi_\Phi^{\mathrm{mod}}(m) = 0$ when Φ has a loop with gain divisible by m, but the dichromatic polynomial of (Φ, h) cannot distinguish loops with different nonzero gains.

One gets a generalization by starting with a weighted gain graph whose gain group $\mathfrak{G}$ has a nontrivial normal subgroup $\mathfrak{N}$, which could even be $\mathfrak{G}$ itself. Ignore the weights and take the gains modulo $\mathfrak{N}$. This gives a gain graph $\Phi_{\mathrm{mod}\,\mathfrak{N}}$ with gain group $\mathfrak{G}/\mathfrak{N}$ whose total dichromatic polynomial is

$$Q_{\Phi_{\mathrm{mod}\,\mathfrak{N}}}(u, v, z) := \sum_{S \subseteq E} u^{b_\mathfrak{N}(S)} v^{|S|-n+c(S)} z^{c(S)-b_\mathfrak{N}(S)},$$

where we write $b_\mathfrak{N}$ to emphasize that we count balanced components in $\Phi_{\mathrm{mod}\,\mathfrak{N}}$. This polynomial is a Tutte invariant. When $z = 0$ it also has a tree expansion. Let

$$T_{\Phi_{\mathrm{mod}\,\mathfrak{N}},O}(x, y) := \sum_T x^{\iota(T)} y^{\varepsilon(T)},$$

summed over all maximal forests T, where O is, as usual, a linear ordering of the edge set and ι and ε are the internal and external activities in $\Phi_{\mathrm{mod}\,\mathfrak{N}}$.

For instance, when the gain group is $\mathbb{Z}^d$, we can take the gains modulo a positive integer vector $\mathbf{m} = (m_1, \ldots, m_d) \in \mathbb{Z}_{>0}^d$. This gives a gain graph $\Phi_{\mathrm{mod}\,\mathbf{m}}$ with gain group that is an integral torus $\mathbb{Z}_{m_1} \times \cdots \times \mathbb{Z}_{m_d}$.

Theorem 7.1. *The total dichromatic polynomial of $\Phi_{\mathrm{mod}\,\mathfrak{N}}$ is a Tutte invariant of gain graphs. The balanced dichromatic polynomial $Q_{\Phi_{\mathrm{mod}\,\mathfrak{N}}}(u, v, 0)$ has the spanning-tree expansion*

$$Q_{\Phi_{\mathrm{mod}\,\mathfrak{N}}}(u, v, 0) = u^{c(\Phi)} T_{\Phi_{\mathrm{mod}\,\mathfrak{N}},O}(u+1, v+1),$$

valid for every linear ordering O.

Proof. The first part is obvious, since taking gains modulo $\mathfrak{N}$ commutes with deletion and contraction of links.

The proof of the second part begins as in that of Theorem 4.7; we obtain

$$Q_{\Phi_{\operatorname{mod}\mathfrak{N}}}(u,v,0) = \sum_{T} \sum_{\substack{F\subseteq T\\ T(F)=T}} u^{c(F)}(v+1)^{\varepsilon(F)}.$$

Now we observe that $c(F) = n - |F|$. Furthermore, when $F \subseteq T$, $T(F) = T$ if and only if $F \supseteq \mathrm{II}(T)$. If F does contain $\mathrm{II}(T)$, then $\mathrm{EA}(F) = \mathrm{EA}(T)$, because $\mathrm{II}(F)$ is the union of all broken circles contained in F, the same holds for T, and the fact that $T = T(F)$ implies that F and T have the same internally active elements. Thus,

$$\begin{aligned}
Q_{\Phi_{\operatorname{mod}\mathfrak{N}}}(u,v,0) &= \sum_{T}(v+1)^{\varepsilon(T)} \sum_{\substack{F\subseteq T\\ F\supseteq \mathrm{II}(T)}} u^{n-|F|} \\
&= \sum_{T}(v+1)^{\varepsilon(T)} \sum_{\substack{T\setminus F\subseteq T\\ T\setminus F\subseteq \mathrm{IA}(T)}} u^{c(T)+|T\setminus F|} \\
&= u^{c(\Gamma)} \sum_{T}(v+1)^{\varepsilon(T)}(u+1)^{\iota(T)},
\end{aligned}$$

because $c(T) = c(\Gamma)$. This proves the tree expansion formula. □

The idea of producing a new Tutte invariant by taking the total dichromatic polynomial with gains that are quotients modulo a normal subgroup of the gain group can also be applied to an unweighted gain graph. It appears to complicate the problem of finding the universal Tutte invariant.

7.2. **Dichromatic overgeneralization.** Suppose we tried to generalize $Q_{(\Phi,h)}(\mathbf{u},v,z)$ to a more powerful Tutte invariant by taking indeterminates $u_{(\Phi,h)/S:W}$, dependent on the isomorphism type of $(\Phi,h)/S{:}W$. As we saw in the last step of the proof, in order to have a Tutte invariant we would have to say that the indeterminate was not changed by adding or subtracting any loop or half edge e. That is, u could depend only on the vertex and weight; but being an isomorphism invariant, it could only depend on the weight. It follows that our polynomial Tutte invariants cannot be generalized in this direction.

What we do see as a necessary generalization, if there is any hope of finding the universal Tutte invariant, is to take account of loops (and half edges). In our total dichromatic polynomial loops and half edges play no role. We have explored a more general dichromatic polynomial with variables corresponding to loops, but there appear to be inescapable relations among these variables that suggest the universal invariant cannot be a true polynomial but instead lies in a quotient of a polynomial ring.

References

[1] Thomas H. Brylawski, A decomposition for combinatorial geometries. *Trans. Amer. Math. Soc.* **171** (1972), 235–282. MR 46 #8869. Zbl. 224.05007, 248.05019.

[2] Thomas Brylawski and James Oxley, The Tutte polynomial and its applications. In Neil White, ed., *Matroid Applications*, Ch. 6, pp. 123–225. Encyc. Math. Appl., Vol. 40. Cambridge Univ. Press, Cambridge, 1992. MR 93k:05060. Zbl. 769.05026.

[3] Henry H. Crapo, The Tutte polynomial. *Aequationes Math.* **3** (1969), 211–229. MR 41 #6705. Zbl. 197, 502b (e: 197.50202).

[4] Gwihen Etienne and Michel Las Vergnas, External and internal elements of a matroid basis. *Discrete Math.* **179** (1998), 111–119. MR 98m:05038. Zbl. 887.05012.

[5] David Forge and Thomas Zaslavsky, Lattice point counts for the Shi arrangement and other affinographic hyperplane arrangements. *J. Combin. Theory Ser. A* **114** (1) (2007), 97–109.

[6] W. Kook, V. Reiner, and D. Stanton, Combinatorial Laplacians of matroid complexes. *J. Amer. Math. Soc.* **13** (1) (2000), 129–148. MR 2001e:05028. Zbl. 940.05021.

[7] S.D. Noble and D.J.A. Welsh, A weighted graph polynomial from chromatic invariants of knots. Symposium à la Mémoire de François Jaeger (Grenoble, 1998). *Ann. Inst. Fourier (Grenoble)* **49** (3) (1999), 1057–1087. MR 2000h:05066. Zbl. 917.05025.

[8] Gian-Carlo Rota, On the foundations of combinatorial theory: I. Theory of Möbius functions. *Z. Wahrsch. verw. Gebiete* **2** (1964), 340–368. MR 30 #4688. Zbl. 121, 24f (e: 121.02406).

[9] Richard P. Stanley, *Enumerative Combinatorics*, Vol. 1. Wadsworth & Brooks/Cole, Monterey, Calif., 1986. MR 87j:05003. Zbl. 608.05001.

[10] W.T. Tutte, A ring in graph theory. *Proc. Cambridge Philos. Soc.* **43** (1947), 26–40. MR 8, 284k. Zbl. 31, 418c (e: 031.41803).

[11] W.T. Tutte, A contribution to the theory of chromatic polynomials, *Canadian J. Math.* **6** (1954), 80–91. MR 15, 814c. Zbl. 55, 171a (e 055.17101).

[12] W.T. Tutte, *Graph Theory*. Encyc. Math. Appl., Vol. 21. Addison–Wesley, Reading, Mass., 1984. MR 87c:05001. Zbl. 554.05001. Reprinted by Cambridge Univ. Press, Cambridge, 2001. MR 2001j:05002. Zbl. 964.05001.

[13] Michelle L. Wachs and James W. Walker, On geometric semilattices. *Order* **2** (1986), 367–385. MR 87f:06004. Zbl. 589.06005.

[14] Thomas Zaslavsky, Signed graph coloring. *Discrete Math.* **39** (1982), 215–228. MR 84h:05050a. Zbl. 487.05027.

[15] Thomas Zaslavsky, Biased graphs. I. Bias, balance, and gains. II. The three matroids. III. Chromatic and dichromatic invariants. *J. Combin. Theory Ser. B* **47** (1989), 32–52; **51** (1991), 46–72; **64** (1995), 17–88. MR 90k:05138; 91m:05056; 96g:05139. Zbl. 714.05057; 763.05096; 857.05088.

[16] Thomas Zaslavsky, Strong Tutte functions of matroids and graphs. *Trans. Amer. Math. Soc.* **334** (1992), 317–347. MR 93a:05047. Zbl. 781.05012.