Annals of Combinatorics Manuscript-Nr.
(will be inserted by hand later)

The Polytope of Win Vectors

J. E. Bartels[1] *, J. Mount[2], and D. J. A. Welsh[1] **

[1] Mathematical Institute
University of Oxford
27-29 St Giles
Oxford, OX1 3LB
United Kingdom
{bartels,dwelsh}@maths.ox.ac.uk

[2] Arris Pharmaceutical
385 Oyster Pt., Blvd., Suite 3
South San Francisco, CA 94080
USA
jmount@arris.com

Received: 23.12.96 / Accepted: 14.02.97

Abstract. Imagine a graph as representing a fixture list with vertices corresponding to teams and the number of edges joining u and v representing the number of games in which u and v have to play each other. Each game ends in a win, loss or tie and we say a vector $\boldsymbol{w} = (w_1, \ldots, w_n)$ is a *win vector* if it represents the possible outcomes of the games, with w_i denoting the total number of games won by team i. We study combinatorial and geometric properties of the set of win vectors and in particular we consider the problem of counting them. We construct a fully polynomial randomised approximation scheme for their number in dense graphs. To do this we prove that the convex hull of the set of win vectors of G forms an integral polymatroid and then use volume approximation techniques.

Subject Classifications. 05C20,05C90

Keywords. score vector – polytope – polymatroid – competition – random generation – approximate counting – fully polynomial randomised approximation scheme

1. Introduction

Let a graph G be given and imagine G as representing a fixture list with vertices corresponding to teams and the number of edges joining u and v representing the number of games in which u and v have to play each other. Each game

* Supported by the "DAAD Doktorandenstipendium des zweiten Hochschulsonderprogrammes HSPII/AUFE"
** Partially supported by RAND-REC EC US030

ends in a win, loss or tie and we say a vector $\boldsymbol{w} = (w_1, \ldots, w_n)$ is a *win vector* if it represents the possible outcomes of the games, with w_i denoting the total number of games won by team i.

Win vectors are a generalisation of score vectors which arise naturally in the study of competitions. Tournaments have long been a subject of study in combinatorics, see for example [15, 16], and have diverse applications in statistics, economics and biology [14, 8, 6].

Win vectors also have applications in biology. For example, they arise when studying populations where individuals compete for mating partners, food or territory. In this context the underlying graph is hardly ever complete but is usually defined by boundary conditions of the habitat in question such as its geography or coalitions between its individuals. In addition to that, it is usually possible that competitions between individuals do not take place at all or end in a tie. When studying such a system biologists often face the problem of having to decide whether observed outcomes of competitions behave randomly or have properties which suggest that certain individuals or groups are better adapted to their environment [11, 3, 13]. This leads naturally to the problem of having to generate a random win vector and counting the number of possible win vectors for a given underlying graph.

In this paper we address some of these problems. Viewed from a statistical perspective the algorithmic problems in this paper have a similar flavour to those addressed by Dyer et al. in [4]. Kannan et al. [9] have recently studied the related problem of randomly generating a realising orientation for a fixed score vector (for definitions see below) of the complete graph.

The outline of the paper is as follows. In Section 2 we give the basic definitions and some examples. Then, in Section 3 we give our results on the combinatorial and geometrical properties of the set of win vectors. In Section 4 we describe a method to quickly generate win vectors almost uniformly at random for graphs which fulfill certain denseness conditions. In Section 5 we briefly address complexity theoretic questions before we describe how to use this generator in an approximation scheme which will give good estimates of the number of win vectors of those graphs.

2. Definitions and Examples

Our terminology is standard. $G = (V, E)$ is a undirected, labelled, finite graph. For a vertex $v \in V$ we denote by $d(v)$ the number of edges incident to it. We call a graph G on n vertices α-*dense* if for all vertices v we have that $d(v) \geq \lfloor \alpha n \rfloor$ for some fixed $\alpha \in (0, 1)$. An orientation ω of G is an assignment of directions to all edges of G. A suborientation is an orientation of a subset of the edgeset E. For an edge $e = (u, v) \in E$ we denote by $(u \to v)$ (respectively $(v \to u)$) its directed versions. For any orientation ω of G we denote by $d^+_\omega(v)$ (respectively $d^-_\omega(v)$) the number of outgoing (ingoing) edges incident at vertex v. For a given orientation ω of G the *score vector* of ω is $\boldsymbol{s}(\omega) = (d^+_\omega(v_1), \ldots, d^+_\omega(v_n)) \in \mathbb{N}^V$. We let S_G be the set of all score vectors of G. A *win vector* $\boldsymbol{w} \in \mathbb{N}^V$ of G is a vector $\boldsymbol{w}$ such that $\boldsymbol{w}$ is a score vector for at least one orientation of some subgraph $G' = (V, E')$ with $E' \subseteq E$. Furthermore we let W_G be the set of all win vectors of G. For a vector $\boldsymbol{v} = (v_1, \ldots, v_n)$ and $A \subseteq V = [n]$ we define

$\boldsymbol{v}(A) = \sum_{i \in A} v_i$. Let $G \backslash e$ (respectively G/e) denote the graph obtained from G by deleting (contracting) $e \in E$.
We start off with some examples.

Example 1. $K_{1,n-1}$, the star on n vertices.
Here we can derive an easy formula for $|W_{K_{1,n-1}}|$. Going from $K_{1,n-2}$ to $K_{1,n-1}$ we have three choices of what to do with the new edge e when extending a win vector from $K_{1,n-2}$ to $K_{1,n-1}$. Clearly, if we delete e we obtain for each win vector of $K_{1,n-2}$ a win vector of $K_{1,n-1}$. If we orient e out of the new vertex then again every win vector of $K_{1,n-2}$ yields a new win vector for $K_{1,n-1}$ that is where the new component is equal to one. If we orient e into the new vertex then we only get a new win vector for $K_{1,n-1}$ from a score vector of $K_{1,n-2}$, since if not all edges are used then the induced graph is isomorphic to a star $K_{1,j}, j < n-2$. But there are 2^{n-2} score vectors for $K_{1,n-2}$. Thus we get the recursion

$$|W_{K_{1,n-1}}| = 2|W_{K_{1,n-2}}| + 2^{n-2}$$

from which we obtain

$$|W_{K_{1,n-1}}| = (n+1)2^{n-2}.$$

□

Example 2. P_n, the path on n vertices.
It is not difficult to derive a recursion formula for $|W_{P_n}|$. We get

$$|W_{P_n}| = |W_{P_{n-1}}| + \sum_{i=1}^{n} |W_{P_{n-i}}|$$

with recursion base $|W_{P_1}| = 1$ and $|W_{P_0}| = 1$. By standard generating function techniques we get

$$|W_{P_n}| = \left\lfloor \frac{1}{\sqrt{5}} \left(\frac{3+\sqrt{5}}{2} \right)^n \right\rfloor.$$

□

Example 3. K_n, the complete graph on n vertices.
We order the win vectors according to the sum of their components $r = \sum_{i \in V} x_i$. For each r, $0 \leq r \leq n(n-1)$ the nonzero entries of a win vector $\boldsymbol{w}$ of component sum r form a partition of the integer r. If we denote by $P(k,i)$ the number of ways of partitioning the integer k into exactly i summands then, after accounting for the choice of zero entries and permutations of nonzero entries we get

$$|W_{K_n}| = \sum_{k=0}^{n(n-1)} \sum_{i=1}^{k} P(k,i) i! \binom{n}{i}.$$

Moreover, and this is a fact we shall need later, the numbers $P(k,i)$ can be computed in time polynomial in n by determining the respective coefficients in the generating function

$$\sum_{k,i \geq 0} P(k,i) x^k y^i = \frac{1}{(1-yx)(1-yx^2)(1-yx^3)\cdots},$$

see [23]. □

However all these examples are very special and we believe that in general computing $|W_G|$ is a hard problem.

3. Structural results

We define $\mathcal{W}_G$ to be the convex hull of all win vectors of G, that is $\mathcal{W}_G := \operatorname{conv}(W_G)$. This convex polytope $\mathcal{W}_G \subset \mathbb{R}_+^V$ is an *integral polymatroid* as we show below. We first recall the definition of a polymatroid from [21].

Definition 1. *A polymatroid $\mathcal{P}$ is a pair (S, r) where S, the* ground set *is a non-empty finite set and r, the* ground set rank function *is a function: $2^S \to \mathbb{R}_+$ satisfying*

$$\begin{aligned}
&r(\emptyset) = 0\\
&A \subseteq B \subseteq S \Rightarrow r(A) \leq r(B)\\
&A, B \subseteq S \Rightarrow r(A \cup B) + r(A \cap B) \leq r(A) + r(B);
\end{aligned}$$

and the vectors $\boldsymbol{x} \in \mathbb{R}_+^S$ such that $\boldsymbol{x}(A) \leq r(A)$ for all $A \subseteq S$ are the independent vectors *of $\mathcal{P}$.*

Note that every subvector of an independent vector is independent. For the sake of brevity, we identify the polymatroid $\mathcal{P} = (S, r)$ with its set of independent vectors if the ground set is obvious from the context.
For $\mathcal{W}_G$ we now have the following

Proposition 1. *$\mathcal{W}_G$ is the set of independent vectors of a polymatroid with ground set V. The rank function is given by $r : 2^V \to \mathbb{N}$ where $r(U)$ is the number of edges having at least one endpoint in U.*

Proof. The function $r(\cdot)$ can clearly be written as

$$r(U) = \operatorname{int}(U) + \operatorname{ext}(U),$$

where $\operatorname{int}(U)$ is the number of edges in the subgraph induced by U and $\operatorname{ext}(U)$ is the number of edges joining U to $V \setminus U$. By $\operatorname{ext}_A(U)$ for $A \subseteq V$ we denote the number of edges joining U to $A \setminus U$.
The first two condition on r are obvious from the definition. For the submodularity condition let $A, B \subseteq S$ be given. We then have

$$\begin{aligned}
r(A \cup B) &= r(A) + r(B) - \operatorname{ext}_{B\setminus A}(A \setminus B) - \operatorname{ext}_{A \setminus B}(A \cap B) -\\
&\quad \operatorname{ext}_{B\setminus A}(A \cap B) - \operatorname{ext}_{V\setminus(A\cup B)}(A \cap B) - \operatorname{int}(A \cap B)\\
r(A \cap B) &= \operatorname{int}(A \cap B) + \operatorname{ext}_{A\setminus B}(A \cap B) +\\
&\quad \operatorname{ext}_{B\setminus A}(A \cap B) + \operatorname{ext}_{V\setminus(A\cup B)}(A \cap B).
\end{aligned}$$

Substituting this into the definition above yields submodularity of r. We note that equality holds for sets A and B when there are no edges going from $A \setminus B$ to $B \setminus A$. It remains to be shown that $\mathcal{W}_G$ is indeed the set of independent vectors for the polymatroid defined by r. For this let $\boldsymbol{v}$ first be a win vector with realising orientation ω. Now observe that, for $A \subseteq V$, $v(A)$ is equal to the number of oriented edges of ω which have their tail in A. Since this is less than or equal to the number of edges incident to A we have $v(A) \leq r(A)$ and we

conclude that every win vector is r-independent. Since the independent vectors of a polymatroid form a convex polytope (see Theorem 3 in [21]) we conclude that indeed all vectors of $\mathcal{W}_G$ are r-independent.
For the reverse inclusion now let $\boldsymbol{v}$ be an r-independent vector. We first assume that $\boldsymbol{v}$ is a vertex of the polytope of r-independent vectors. Below, in the proof of Proposition 3, we give a characterisation of the vertices of the independence polytope defined by the rank function r as win vectors with an acyclic realising orientation. In particular it follows that all vertices of the polymatroid are win vectors. Taking the convex hull now gives the identity of $\mathcal{W}_G$ with this independence polytope. Thus $\mathcal{W}_G$ is indeed the integral polymatroid defined by r. □

The set of independent vectors of a polymatroid forms a convex polytope in $\mathbb{R}_+^V$ which can be described as the intersection of the following halfspaces

$$\begin{array}{rcl} \sum_{i\in U} x_i & \leq & r(U) \text{ for all } U \subseteq V \\ x_i & \geq & 0. \end{array} \tag{1}$$

For a polymatroid (S,r) with independent vectors P the *basepolytope* is the convex polytope $P \cap \{\sum_{i\in S} x_i = r(S)\}$. This polytope determines the polymatroid uniquely. Clearly $r(V) = |E|$ in the case of $\mathcal{W}_G$ and hence we have that the maximal, integer elements of $\mathcal{W}_G$ are exactly the score vectors of G.

Proposition 2. *The polytope $\mathcal{W}_G$ is uniquely determined by and uniquely determines the* score polytope $\mathcal{S}_G := \; conv(S_G)$.

We continue with some results on the combinatorial structure of $\mathcal{S}_G$ and $\mathcal{W}_G$. First we have

Proposition 3. *The vertices of $\mathcal{W}_G$ are exactly the set of win vectors which have a unique realising orientation ω. This orientation ω is an acyclic orientation of a subgraph G' of G which is obtained by deleting the edges with both endpoints in U for some $U \subseteq V$.*

Proof. Recall from [5] that the vertices of a polymatroid $\mathcal{P} = (S,r)$ can be characterised as follows:
Let $B = \{s_1, \ldots, s_{|B|}\}$ be an ordered subset of S and let $B_i = \{s_1, \ldots, s_i\}$ for $0 \leq i \leq |B|$ with $B_0 = \emptyset$. Then B *generates* the vertex $\boldsymbol{v}_B$ given by

$$\begin{array}{rcll} v_{s_1} & = & r(s_1) & \\ v_{s_i} & = & r(B_i) - r(B_{i-1}) & \text{for } 1 < i \leq |B| \\ v_s & = & 0 & \text{for } s \notin B. \end{array}$$

In our particular case, we see that this win vector $\boldsymbol{v}_B$ can be achieved by a greedy orientation procedure which orients the incident edges of the vertices in the order prescribed by B. It is not difficult to see that this orientation must be acyclic. The claim now follows from observing that by dropping the edges which are not incident to vertices in B, that is $U = V \setminus B$, we obtain a unique (sub)orientation ω of G which realises $\boldsymbol{v}_B$. □

From this we obtain the vertices of $\mathcal{S}_G$.

Corollary 1. *The vertices of $\mathcal{S}_G$ are the score vectors realised by the acyclic orientations of G.*

Proof. Just observe that vertices of $\mathcal{S}_G$ have modulus $|E|$. The corollary then follows. □

For the edges of $\mathcal{W}_G$ we have

Proposition 4. *Two vertices of $\mathcal{W}_G$ are adjacent iff either they differ in exactly one component or they have the same modulus and their realising orientations differ in the orientation of exactly one edge.*

Proof. This can be proved using the general characterisation of adjacency on polymatroids given by Topkis [20]. □

Specialising this to maximal elements only, we have for $\mathcal{S}_G$

Corollary 2. *Two vertices of $\mathcal{S}_G$ are adjacent iff their realising acyclic orientations differ in exactly one directed edge.*

Our final result concerning the combinatorial structure of $\mathcal{S}_G$ and $\mathcal{W}_G$ gives a full description of their facets.

Proposition 5. *An inequality $\sum_{i \in U} x_i \leq r(U)$ or $x_j \geq 0$ of Eq. (1) determines a facet of $\mathcal{W}_G$ iff deleting the edges with exactly one endpoint in U (respectively those incident to j) from E increases the number of connected components of G by one.*

Proof. Since $\mathcal{W}_G$ is the intersection of the halfspaces determined by the inequalities of Eq. (1) it follows that each inequality determines a (possibly empty) face of $\mathcal{W}_G$. We note that if an inequality $\sum_{i \in U} x_i \leq r(U)$ holds then this fixes the orientation of the edges of E which have exactly one endpoint in U. Thus all score vectors for which this inequality holds will only have realisations which agree on those edges. We are now interested in the dimension of the affine space spanned by these score vectors. It is not difficult to see that the score vectors of a graph with k components lie in a $|V| - k$ dimensional affine subspace of $\mathbb{R}_+^V$. Thus an inequality of Eq. (1) determines a $|V| - k - 1$ dimensional affine subspace iff deleting the edges with exactly one endpoint in U from E increases the number of connected components of G by one. □

Specialising to $\mathcal{S}_G$ we obtain the following combinatorial description of its facets.

Corollary 3. *Let C be a cut of G whose removal disconnects a connected component of G into exactly two parts V_1 and V_2. Let $F^{\rightarrow}$ (respectively $F^{\leftarrow}$) be the set of score vectors realised by orientations which have the edges of C directed from V_1 to V_2 (respectively from V_2 to V_1). Then $conv(F^{\rightarrow})$ and $conv(F^{\leftarrow})$ are two facets of $\mathcal{S}_G$. Furthermore every facet of $\mathcal{S}_G$ can be described by such a cut C.*

Similar but more complicated results hold for the k-faces of $\mathcal{S}_G$.
The results above give a relatively detailed picture of the combinatorial structure of $\mathcal{S}_G$ and $\mathcal{W}_G$. Our main interest in $\mathcal{W}_G$ and $\mathcal{S}_G$ is because of their integer lattice properties. We are going to explore this next. We state first

Proposition 6. *The set all integer points in $\mathcal{W}_G$ is exactly the set W_G of win vectors.*

Proof. It is obvious from the definition that $W_G \subset \mathcal{W}_G$. For the reverse inlusion let $\boldsymbol{w}$ be an integer point in $\mathcal{W}_G$. We need to construct a (sub)orientation of G which realises $\boldsymbol{w}$. From $G = (V, E)$ we construct a capacitated, directed graph $F_{\boldsymbol{w}} = (S \cup V \cup \{s, t\}, E')$. The vertex set consists of V, two distinct new vertices s and t, and the set S which consists of vertices v_e which are in one-to-one correspondence with the edges $e \in E$ of G. The edgeset E' is constructed from E as follows. For every $e = (u, v) \in E$ we have two directed edges $(v_e \to u)$ and $(v_e \to v)$ of capacity 1 in E'. We call these edges *type 1 edges.* To this we add for each $v_e \in S$ the edge $(s \to v_e)$ of capacity 1 and for each $v \in V$ the edge $(v \to t)$ of capacity w_v. These are *type 2 edges.* An integer s-t-flow in $F_{\boldsymbol{w}}$ corresponds now in a natural way to an orientation of G and a flow of capacity $\boldsymbol{w}(E)$ gives a realising orientation of $\boldsymbol{w}$. To complete the proof, we need to show that there is no edgecut $C \subseteq E'$ in $F_{\boldsymbol{w}}$ whose capacity is less than $\boldsymbol{w}(E)$. By the Max-Flow Min-Cut Theorem we are then guaranteed to find a flow of capacity $\boldsymbol{w}(E)$ and thus have a realising orientation for $\boldsymbol{w}$.

Assume C is a minimal minimum cut in $F_{\boldsymbol{w}}$. Without loss of generality we can assume that C does not contain any type 1 edge. For if C contained $(v_e \to u)$ which disconnects the path $(s \to v_e)(v_e \to u)(u \to t)$ then we could replace it by the edge $(s \to v_e)$ which has the same capacity and disconnects this path as well. Thus C consists only of type 2 edges. If C consists solely of edges incident to either s or t then we have $|E| \geq \boldsymbol{w}(E)$ for C's capacity in the first case and $\boldsymbol{w}(E)$ in the latter. Hence the claim would follow for these two cases. So assume finally that C does not contain all edges incident to t and let $(u_1 \to t), \ldots, (u_i \to t)$ be the edges of $F_{\boldsymbol{w}}$ incident to t and not in C. Let $U = \bigcup_i \{u_i\}$. All the paths from s to t through these edges have to have at least one edge in C. Since C contains only type 2 edges it follows that for each $u_i \in U$ and all edges e incident to u_i in G the edge $(s \to v_e)$ (which has capacity one) of $F_{\boldsymbol{w}}$ must be in C. But there are $r(U) \geq \boldsymbol{w}(U)$ of those edges. Accounting for the remaining edges of C which are incident to t and have capacity w_v we conclude that the capacity of C is larger than or equal to $\boldsymbol{w}(E)$ in this case, too.

Finally we remark that constructing an orientation for $\boldsymbol{w}$ of G by finding a maximum s-t-flow in $F_{\boldsymbol{w}}$ can be done by standard techniques in polynomial time. □

We now turn to the basepolytope $\mathcal{S}_G$. Somewhat surprisingly there is a bijection between the integer points of $\mathcal{S}_G$ and the forests of G. This seems to have been noticed first by Stanley in [19] where he attributes the original idea to Zaslavsky.

Proposition 7. *The number of score vectors equals the number of forests of G.*

Kleitman and Winston [10] give a proof of this result using a depth-first-search algorithm to construct a bijection which is dependent on an *a priori* fixed order. A conceptually simpler proof can be given based on Tutte-Grothendieck invariants.

Turning to the volume of $\mathcal{S}_G$, it is easy to check that if G has k components then $\mathcal{S}_G$ is a $|V| - k$ dimensional convex polytope in $\mathbb{R}_+^V$. The following is a consequence of $\mathcal{S}_G$ being a unimodular zonotope, see [18] and a general formula for the volume of zonotopes, see [17].

Proposition 8. *If G is connected then the relative volume $vol_{|V|-1}(\mathcal{S}_G)$ of $\mathcal{S}_G$ equals the number of spanning trees of G.*

We do not know of any similar results for $|W_G|$ or have any interpretation of $\mathcal{W}_G$'s volume. We can however give some inequalities. To derive them we use the concept of polymatroid duality which is in many respects analogous to matroid duality.

We say $\boldsymbol{c} \in \mathbb{R}_+^V$ *bounds* $P \subset R_+^V$ iff $\boldsymbol{c} \geq \boldsymbol{x}$ for all $\boldsymbol{x} \in P$. For a polymatroid (S, r) with independence polytope P and any vector $\boldsymbol{c} \in \mathbb{R}_+^S$ which bounds P we define

$$r^{\boldsymbol{c}}(A) := \boldsymbol{c}(A) + r(S \setminus A) - r(S) \qquad \text{for } A \subseteq S.$$

It is routine to check that $r^{\boldsymbol{c}}$ satisfies the conditions of Definition 1. We call $(S, r^{\boldsymbol{c}})$ the $\boldsymbol{c}$*-dual* of (S, r) and denote its independence polytope by $P^{\boldsymbol{c}}$. The *vector rank* of $\boldsymbol{x} \in R_+^S$ in (S, r) is defined by

$$\|\boldsymbol{x}\| := \min_{A \subseteq S} \{\boldsymbol{x}(A) + r(S \setminus A)\}$$

and we say that $\boldsymbol{x}$ is a *spanning vector* of (S, r) iff $\|\boldsymbol{x}\| = r(S)$. We will use the following result of [21].

Lemma 1. *$\boldsymbol{x} \in P^{\boldsymbol{c}}$ iff $\boldsymbol{c} - \boldsymbol{x}$ is spanning for P.*

The next result shows that $\mathcal{W}_G$ is self $\boldsymbol{c}$-dual for a suitably chosen $\boldsymbol{c}$.

Proposition 9. *Let $\boldsymbol{d}$ be the degree vector of G. Then we have*

$$(\mathcal{W}_G)^{\boldsymbol{d}} = \mathcal{W}_G.$$

Proof. According to the definition of $r^{\boldsymbol{d}}$ we have

$$\begin{aligned} r^{\boldsymbol{d}}(A) &= \boldsymbol{d}(A) + r(V \setminus A) - r(V) \\ &= \boldsymbol{d}(A) + |E| - \operatorname{int}(A) - |E| \\ &= 2\operatorname{int}(A) + \operatorname{ext}_{V \setminus A}(A) - \operatorname{int}(A) \\ &= r(A). \end{aligned}$$

□

From this we can get the following inequalities.

Proposition 10. *For a connected graph G we have*

$$\begin{aligned} 2|W_G| - |S_G| &\leq \prod_{v \in V} (d(v) + 1), \text{ and} \\ vol_{|V|}(\mathcal{W}_G) &\leq \frac{1}{2} \prod_{v \in V} d(v). \end{aligned}$$

Proof. From Lemma 1 we know that there is a bijection between the independent vectors of $(\mathcal{W}_G)^{\boldsymbol{d}}$ and the spanning vectors of $\mathcal{W}_G$ which are subvectors of $\boldsymbol{d}$. Proposition 9 shows that this is in effect a bijection between the points in $\mathcal{W}_G$ and the spanning vectors of $\mathcal{W}_G$ which are subvectors of $\boldsymbol{d}$. Observe that if $\boldsymbol{x} \in \mathcal{W}_G$ is spanning then we have $\boldsymbol{x} \in \mathcal{S}_G$. Furthermore if $\boldsymbol{x}$ is integer then so is $\boldsymbol{d} - \boldsymbol{x}$. Thus we have a bijection between W_G and $(\{\boldsymbol{x} \in \mathbb{N}_0^V : \boldsymbol{x} \leq \boldsymbol{d} \text{ and } \boldsymbol{x} \text{ is spanning}\} \setminus W_G) \cup S_G$. Twice the number of win vectors of G minus $|S_G|$ is therefore less

than or equal to the number of integer subvectors of $\boldsymbol{d}$.
The inequality for the volume follows from this bijection with the observation that image and preimage overlap only in a set of measure zero. Their union is contained in the box of $\mathbb{R}_+^V$ having $\boldsymbol{d}$ as "upper right hand" corner. □

Both inequalities can be tight.

4. Random Generation of Win Vectors

Suppose we know or can estimate the probability p of a draw between each pair of players and also assume that if it is not a draw each player has an equal probability of winning. Then the probability of obtaining any particular win vector $\boldsymbol{w}$ of $G = (V, E)$ is given by

$$\mathsf{Prob}\ [\boldsymbol{w}] = p^{|E|-\boldsymbol{w}(E)}(1-p)^{\boldsymbol{w}(E)}|\Omega_{\boldsymbol{w}}|$$

where $\Omega_{\boldsymbol{w}}$ is the set of suborientations which realise $\boldsymbol{w}$.
For some graphs this probability measure will be uniform over the set of win vectors of the same weight. For general graphs however it will be highly non uniform over W_G. Thus the "naive" approach for generating a win vector uniformly by randomly deleting edges and orienting the remaining ones randomly is doomed to fail.
What is mildly curious is that with respect to this measure there is the following interpretation of the expected number of win vectors of G.

Proposition 11. *If G represents a fixture list in which each pair (i, j) joined by an edge has independently probability p of producing a tie, then the expected number of win vectors is given in terms of the Tutte polynomial $T(G; x, y)$ by*

$$\mathsf{E}_p[|W_G|] = (1-p)^{|V|-1} T\left(G; \frac{2-p}{1-p}, 1\right).$$

Proof. Given $G = (V, E)$ then we have

$$\mathsf{E}_p[|W_G|] = \sum_{A \subseteq E} p^{|E \setminus A|}(1-p)^{|A|}|S_{G|A}|$$

since every edge not corresponding to a tie must be given an orientation. We now use the following identity from [22]. For any θ and any matroid $M = (E, r)$

$$\sum_{A \subseteq E} \theta^{|A|}(x-1)^{r(E)-r(A)} T(M|A; x, y) = (\theta+1)^{|E|-r(E)}\theta^{r(E)} T(M; X, Y)$$

where $X - 1 = (x-1)(\theta+1)/\theta$ and $Y - 1 = (y-1)\theta/(\theta+1)$. Using the fact that $|S_{G|A}| = T(M(G|A); 2, 1)$ and defining $x = 2$, $\theta = (1-p)/p$, $y = 1$ then gives

$$\mathsf{E}_p[|W_G|] = (\theta+1)^{-r(E)}\theta^{r(E)} T(G; 2 + 1/\theta, 1).$$

When G is connected $r(E) = |V| - 1$ and the result follows. □

Note, when $p = 0$ Proposition 11 gives our earlier Proposition 7, since we are only counting score vectors. However when $p = 1/2$ we get a new, somewhat strange interpretation of $T(G; 3, 1)$. Other consequences of the above follow from known results of [7] and [1], namely,

- Computing $\mathsf{E}_p[|W_G|]$ exactly is #**P**-hard.
- There is a fully polynomial randomised approximation scheme which approximates $\mathsf{E}_p[|W_G|]$ to within a given ratio in polytime if G is dense.

Our approach to the random generation problem is different from the "naive" one. First of all we are interested in obtaining a *uniform* or *almost uniform* distribution. To do this we exploit the geometrical structure of $\mathcal{W}_G$ and make use of known methods for the almost uniform generation of points inside convex polytopes. Throughout the rest of this section we assume that the graph G is α-dense for some fixed $\alpha > 0$.
If $U(\cdot)$ is the uniform distribution over Ω then we call a probability distribution π on a probability space with ground set Ω, ϵ-*uniform* iff

$$\frac{1}{2} \sum_{x \in \Omega} |\pi(x) - U(x)| \leq \epsilon$$

respectively, if Ω is not discrete,

$$\frac{1}{2} \int_{x \in \Omega} |\pi(x) - U(x)| \leq \epsilon.$$

Consider now the following algorithm.
Algorithm `GEN` :

1. Set $b := \lceil \alpha n/2 \rceil$
2. Generate a point $\boldsymbol{p}$ almost uniformly at random in the blown up polytope $(1 + 1/b)\mathcal{W}_G$.
3. Map $\boldsymbol{p} \to \lfloor \boldsymbol{p} \rfloor$, the integer "bottom left hand corner" of the unit cube containing $\boldsymbol{p}$.
4. If $\lfloor \boldsymbol{p} \rfloor$ belongs to $\mathcal{W}_G$ then accept it as our lattice point and stop. If it lies outside $\mathcal{W}_G$ then repeat steps 2-4 for at most $-\log_2(1 - e^{-2/\alpha})$ times. Stop with output "Fail" if still unsuccessful after that many repetitions.

We now claim

Theorem 1. *For G α-dense,* `GEN` *generates win vectors ϵ-uniformly at random within time polynomial in $n^{1/\alpha}$ and ϵ^{-1}. The probability of* `GEN` *stopping with output "Fail" is smaller than 1/2.*

We prove Theorem 1 by a series of lemmata. First we need to check that the blown up polytope $(1 + 1/b)\mathcal{W}_G$ contains all unit cubes which have nonempty intersection with $\mathcal{W}_G$. Only then can we guarantee that all integer points of $\mathcal{W}_G$ are images of a unit cube in step 3. By $\mathcal{C}_l$ we denote the cube of side-length l in $\mathbb{R}_+^V$, in other words

$$\mathcal{C}_l := \{\boldsymbol{x} : 0 \leq \boldsymbol{x} \leq (l, l, \ldots, l)\} \subset \mathbb{R}_+^V.$$

Lemma 2. *$\mathcal{C}_b \subseteq \mathcal{W}_G$, with $b = \alpha n/2$ as defined in step 1 of* `GEN` .

Proof. Suppose $\mathcal{W}_G$ does not contain the point $(b, \ldots, b)$. Then there exists some subset $U \subseteq V$ for which the constraint $\sum_{i \in U} x_i \leq r(U)$ is violated. In other words

$$r(U) < \sum_{i \in U} x_i = |U| \frac{\alpha n}{2}.$$

But

$$r(U) = \frac{1}{2} \sum_{i \in U} u(i) + \sum_{i \in U} (d(i) - u(i))$$

where $u(i)$ is the degree of vertex i within the subgraph of G induced by U. Hence

$$r(U) = \sum_{i \in U} d(i) - \frac{1}{2} \sum_{i \in U} u(i) \geq \frac{1}{2} \sum_{i \in U} d(i)$$

and since $u(i) \leq d(i)$ we get

$$r(U) \geq \frac{1}{2} |U| \alpha n$$

which is a contradiction. Thus $\mathcal{W}_G$ contains $(b, \ldots, b)$ and since $\mathcal{W}_G$ is closed under taking subvectors all of $\mathcal{C}_b$. □

A *lattice unit cube* is a set $\mathcal{C} \subset \mathbb{R}_+^V$ for which there exists an integer lattice point $\boldsymbol{x} \in \mathbb{N}^V$ such that $\mathcal{C} = \boldsymbol{x} + \mathcal{C}_1$. Here "+" denotes the usual Minkowski sum.
Now we can show

Lemma 3. *The blown up polytope $(1 + 1/b)\mathcal{W}_G$ contains all lattice unit cubes which have nonempty intersection with $\mathcal{W}_G$.*

Proof. Suppose $\boldsymbol{x}$ is a point such that $\lfloor \boldsymbol{x} \rfloor \in \mathcal{W}_G$. Then $\boldsymbol{x} \in \mathcal{W}_G + \mathcal{C}_1$. By Lemma 2 we have that $\mathcal{C}_b \in \mathcal{W}_G$. Hence

$$\mathcal{W}_G + (1/b)\mathcal{C}_b \subseteq (1 + 1/b)\mathcal{W}_G$$

and since $\mathcal{C}_1 \subseteq (1/b)\mathcal{C}_b$

$$\mathcal{W}_G + \mathcal{C}_1 \subseteq \mathcal{W}_G + (1/b)\mathcal{C}_b \subseteq (1 + 1/b)\mathcal{W}_G.$$

Hence $\boldsymbol{x} \in (1 + 1/b)\mathcal{W}_G$ as required. □

This shows that under the map of step 3 all integer points of $\mathcal{W}_G$ receive the same measure: exactly one unit volume segment of $(1 + 1/b)\mathcal{W}_G$. Therefore, if we can generate a point in $(1 + 1/b)\mathcal{W}_G$ almost uniformly at random then we can do the same for the integer points in $\mathcal{W}_G$. But this generation in the blown up polytope can be done in time $O^*(n^5)$ by techniques of [12].
By the proof of Proposition 6 we know that checking membership in $\mathcal{W}_G$ in step 4 reduces to an instance of the maximum flow problem and can thus be solved in polynomial time.
It remains to be shown that the rejection probability of the generator is not too high.

Lemma 4. *The rejection probability at step 4 of* `GEN` *is at most $(1 - e^{-2/\alpha})$ and the overall failure probability of* `GEN` *is less than $1/2$.*

Proof. The probability of getting a lattice point inside $\mathcal{W}_G$ at step 3 is at least

$$\mathrm{vol}(\mathcal{W}_G)/\ \mathrm{vol}((1+1/b)\mathcal{W}_G) = 1/(1+1/b)^n.$$

Now $(1+1/b)^{bn/b} \leq e^{n/b}$. Hence we generate a lattice point inside $\mathcal{W}_G$ with probability of at least $e^{-n/b}$. In order for this to be frequent enough for the generator to have polynomial running time we need $b \in O(n)$. Here we have $b = \alpha n/2$ and thus we get a success probability of at least $e^{-2/\alpha}$. By making at most $-\log_2(1-e^{-2/\alpha})$ repetitions we can boost the success probability to at least $1/2$ as required. □

5. Approximating $|W_G|$

From the proof of Proposition 6 we conclude that the membership problems for $\mathcal{W}_G$ and $\mathcal{S}_G$ are polytime solvable. Since these sets are trivially nonempty there is no obvious complexity theoretic reason why determining $|W_G|$ or $|S_G|$ should be hard. For the problem of exactly determining $|S_G|$ it follows from [7] that this problem is $\#P$-complete. Furthermore Annan [2] proved that even in the dense case there is no polytime algorithm which counts forests, and hence score vectors, unless $\mathbf{RP} = \mathbf{NP}$. These results and the close connection between $\mathcal{S}_G$ and $\mathcal{W}_G$ lead us to believe that counting win vectors is a computationally difficult problem even for dense instances. We pose this as an open problem.

Problem 1. Is computing the number of win vectors of a graph $\#\mathbf{P}$-complete?

We now describe how to use the generator of the last section to approximate the number of win vectors of G. To do this we reduce the problem of approximately counting win vectors to that of randomly generating them and then use a well known ratio reduction technique. We write

$$|W_G| = \frac{|W_G|}{|W_{G+e_1}|} \times \frac{|W_{G+e_1}|}{|W_{G+e_1+e_2}|} \times \cdots \times \frac{|W_{G+e_1+\cdots+e_{l-1}}|}{|W_{K_n}|} \times |W_{K_n}|,$$

where $\{e_1, \ldots, e_l\} = E(K_n) \setminus E(G)$. We note that all graphs in this reduction chain are at least α-dense.
From Example 3 we have a formula for $|W_{K_n}|$ which is computable in time polynomial in n. Knowing $|W_{K_n}|$ and having good approximations for the ratios will then yield a good approximation of $|W_G|$ with high probability.
To approximate the ratios

$$r_i := |W_{G+e_1+\cdots+e_i}|/|W_{G+e_1+\cdots+e_{i+1}}|$$

we interpret them as the expectation of the indicator function

$$I_i(\boldsymbol{x}) := \begin{cases} 1 & \boldsymbol{x} \in W_{G+e_1+\cdots+e_i} \cap W_{G+e_1+\cdots+e_{i+1}} \\ 0 & \text{otherwise} \end{cases}$$

under the uniform distribution. Then we have $r_i = \mathsf{E}\,[I_i] = \mathsf{Prob}\,[\{I_i = 1\}]$.
For each $1 \leq i \leq l$ we will use `GEN` to draw a sample of $t = \lceil 198 l \epsilon^{-2} \rceil$ independent win vectors $\{\boldsymbol{x_1}, \ldots, \boldsymbol{x_t}\}$ from an $(\epsilon/12l)$-uniform distribution π_i over $W_{G+e_1+\cdots+e_{i+1}}$. The non-uniformity will introduce a bias in the estimation and

we therefore define $\tilde{r}_i := \mathsf{E}_\pi[I_i]$. We use the sample to compute an estimate for $\tilde{r}_i$ by $X_i := (I_i(\boldsymbol{x_1}) + \cdots + I_i(\boldsymbol{x_t}))/t$.
Our final estimate for $|W_G|$ will then be $\tilde{W} := |W_{K_n}| \prod_{i=1}^{l} X_i$.
We prove the correctness of this approach by a series of lemmata. First we address the problem of bias of $\tilde{r}_i$ which arises since we cannot sample uniformly from W_G.

Lemma 5. *If π_i is $(\epsilon/12l)$-uniform over $W_{G+e_1+\cdots+e_{i+1}}$ then we have $(1-\epsilon/4l)r_i \leq \tilde{r}_i \leq (1+\epsilon/4l)r_i$.*

Proof. Since I_i is 0/1-valued $(\epsilon/12l)$-uniformity yields $|r_i - \tilde{r}_i| \leq \epsilon/12l$. As we show below we have that $r_i \geq 1/3$. Thus $\epsilon/12l \geq |r_i - \tilde{r}_i| = r_i|1 - \tilde{r}_i/r_i| \geq (1/3)|1 - \tilde{r}_i/r_i|$ which is equivalent to $(1-\epsilon/4l)r_i \leq \tilde{r}_i$. The other inequality follows as easily. □

We next prove the bound used in the last proof.

Lemma 6. *For $e \notin E$ we have $|W_G| < |W_{G+e}| \leq 3|W_G|$.*

Proof. The first inequality is obvious. The second inequality follows from the observation that for every win vector $\boldsymbol{x}$ of G we can add e in two different orientations to a realising orientation of $\boldsymbol{x}$ thus obtaining at most two different win vectors of $G+e$. Accounting for $\boldsymbol{x}$ as well this then gives at most three win vectors of $G+e$ for each win vector of G. This inequality can be tight, for example if e is an edge between two isolated vertices of G. □

Next we show that $\tilde{W}$ has got the right expectation.

Lemma 7. $(1-\epsilon/3)|W_G| \leq \mathsf{E}\,[\tilde{W}] \leq (1+\epsilon/3)|W_G|$.

Proof. We have $\mathsf{E}\,[\tilde{W}] = \prod_{i=1}^{l} \mathsf{E}\,[X_i] = \prod_{i=1}^{l} \tilde{r}_i$. The claim now follows from the bounds in Lemma 5 together with the observation that for $0 \leq \epsilon \leq 1$ we have that $e^{\epsilon/4} \leq 1 + \epsilon/3$ and $(1-\epsilon/4)^l \geq (1-\epsilon/3)$. □

And finally we show that with high probability $\tilde{W}$ is close to its expectation.

Lemma 8. *With probability larger than 3/4 we have $(1-\epsilon)|W_G| \leq \tilde{W} \leq (1+\epsilon)|W_G|$.*

Proof. By Chebychev's inequality we have

$$\mathsf{Prob}\,[|\tilde{W} - \mathsf{E}\,[\tilde{W}]| > (\epsilon/3)\mathsf{E}\,[\tilde{W}]] \leq \frac{9}{\epsilon^2}\frac{\mathsf{Var}\,[\tilde{W}]}{(\mathsf{E}\,[\tilde{W}])^2} \leq \frac{1}{4}$$

provided we can show $\mathsf{Var}\,[\tilde{W}]/(\mathsf{E}\,[\tilde{W}])^2 \leq \epsilon^2/36$. Write

$$\begin{aligned}
\frac{\mathsf{Var}\,[\tilde{W}]}{(\mathsf{E}\,[\tilde{W}])^2} &= \frac{\mathsf{E}\,[\tilde{W}^2] - (\mathsf{E}\,[\tilde{W}])^2}{(\mathsf{E}\,[\tilde{W}])^2} \\
&= \prod_{i=1}^{l}\left(\frac{(\mathsf{E}\,[X_i])^2 + \mathsf{E}\,[X_i^2] - (\mathsf{E}\,[X_i])^2}{(\mathsf{E}\,[X_i])^2}\right) - 1 \\
&= \prod_{i=1}^{l}\left(1 + \frac{\mathsf{Var}\,[X_i]}{(\mathsf{E}\,[X_i])^2}\right) - 1. \qquad (2)
\end{aligned}$$

Now we have $\mathsf{E}\,[X_i] = \mathsf{E}\,[I_i] = \tilde{r}_i$ and $\mathsf{Var}\,[X_i] = \mathsf{Var}\,[I_i]/t \leq \mathsf{E}\,[I_i]/t = \tilde{r}_i/t$ since $\mathsf{Var}\,[I_i]$ takes values in $(0,1)$. Using Lemma 5 we get $2\tilde{r}_i \geq r_i$ and with Lemma 6 this yields $\mathsf{Var}\,[X_i]/(\mathsf{E}\,[X_i])^2 \leq 6/t$. Therefore we have the required bound

$$\begin{aligned} \frac{\mathsf{Var}\,[\tilde{W}]}{(\mathsf{E}\,[\tilde{W}])^2} &\leq \left(1+\frac{6}{t}\right)^l - 1 \\ &\leq e^{6l/t} - 1 = e^{\epsilon^2/36} - 1 \\ &\leq \epsilon^2/36. \end{aligned}$$

Thus with probability at least 3/4 we have that $\tilde{W}$ is $(1 \pm \epsilon/3)$-close to $\mathsf{E}\,[\tilde{W}]$. Combining this with Lemma 7 we arrive at the claim of the lemma. □

Summarising we finally have

Theorem 2. *For an α-dense graph G we can approximate $|W_G|$ to within ratio $(1 \pm \epsilon)$ with probability at least 3/4 by using at most $\lceil 198 l^2 \epsilon^{-2} \rceil = O(n^4 \epsilon^{-2})$ samples from* `GEN` .

Thus, having established polynomial running time for `GEN` above, the overall running time for this approximation scheme will be polynomial, too.

6. Conclusion

As shown above, the collection of win vectors can be regarded as a polymatroid analogue of the score vectors. They are in bijective correspondence with the integral independent vectors of a very natural polymatroid defined on graphs, in the same way as score vectors are in bijective correspondence with the independent sets of the cycle matroid of a graph. It seems plausible that our techniques would also give an approximation scheme for the number of score vectors by putting disjoint balls of uniform size around the lattice points in the basepolytope. It is somewhat curious that though using a completely different approach in our approximation scheme, we have again needed the same density condition as Annan [2] and Alon et al. [1] needed in theirs for the number of forests (= independent sets of the cycle matroid) and other evaluations of the Tutte polynomial of a graph. Finding new approaches which would allow these density conditions to be discarded is an important and challenging area for further research.

References

1. Alon, N., Frieze, A., and Welsh, D. Polynomial time randomized approximation schemes for Tutte-Grothendieck invariants: the dense case. *Random Structures and Algorithms 6*, 4 (1995), 459–478.
2. Annan, J. D. A randomised approximation algorithm for counting the number of forests in dense graphs. *Combinatorics, Probability and Computing 3* (1994), 273–283.
3. Crawley, M., and May, R. Population dynamics and plant community structure: Competition between annuals and perennials. *Journal of Theoretical Biology 125* (1987), 475–489.
4. Dyer, M., Kannan, R., and Mount, J. Sampling contingency tables. *Random Structures and Algorithms* (to appear).

5. EDMONDS, J. Submodular functions, matroids, and certain polyhedra. In *Combinatorial Structures and their Applications (Proc. Calgary Internat. Conf., Calgary, Alta., 1969)*. Gordon and Breach, New York, 1970, pp. 69–87.
6. GIBBONS, J. D., OLKIN, I., AND SOEBEL, M. Baseball competitions: Are enough games played? *American Statistician 32* (1978), 89–95.
7. JAEGER, F., VERTIGAN, D., AND WELSH, D. On the computational complexity of the Jones and Tutte polynomials. *Mathematical Proceedings of the Cambridge Philosophical Society 108* (1990), 35–53.
8. JECH, T. The ranking of incomplete tournaments: A mathematicians guide to popular sports. *American Mathematical Monthly 90* (1983), 246–266.
9. KANNAN, R., TETALI, P., AND VEMPALA, S. Simple Markov chain algorithms for generating bipartite graphs and tournaments. preprint, 1996.
10. KLEITMAN, D. J., AND WINSTON, K. J. Forests and score vectors. *Combinatorica 1* (1981), 49–54.
11. KULLMANN, H. personal communication, 1996.
12. LOVASZ, L., AND SIMONOVITS, M. Random walks in a convex body and an improved volume algorithm. *Random Structures and Algorithms 4*, 4 (1993), 359–412.
13. MCGILCHRIST, C. A., AND TRENBATH, B. R. A revised analysis of plant competition experiments. *Biometrics* (1971), 659–671.
14. MEAD, R. Competition experiments. *Biometrics* (1979), 41–54.
15. MOON, J. *Topics on Tournaments*. Holt, Rinehart and Winston, New York, 1968.
16. REID, K., AND BEINEKE, L. Tournaments. In *Selected Topics in Graph Theory*, L. Beineke and R. Wilson, Eds. Academic Press, New York, 1978, pp. 169–204.
17. SHEPHARD, G. Combinatorial properties of associated zonotopes. *Canadian Journal of Mathematics 26*, 2 (1974), 302–321.
18. STANLEY, R. *Enumerative Combinatorics*. Cambridge University Press, New York, 1986.
19. STANLEY, R. P. Decompositions of rational convex polytopes. *Annals of Discrete Mathematics 6*, 6 (1980), 333–342.
20. TOPKIS, D. M. Adjacency on polymatroids. *Mathematical Programming 30* (1984), 229–237.
21. WELSH, D. J. A. *Matroid Theory*. No. 8 in London Mathematical Society Monographs. (New series). Academic Press, London, 1976, ch. 18: Convex Polytopes Associated with Matroids, pp. 333–356.
22. WELSH, D. J. A. Counting colourings and flows in random graphs. In *Combinatorics - Paul Erdös is Eighty*, vol. 2 of *Bolyai Society Mathematical Studies*. Bolyai Society, Budapest, 1996, pp. 491–505. Conference in Keszthely (Hungary), 1993.
23. WILF, H. S. *Generatingfunctionology*. Academic Press, London, 1990.

This article was processed by the author using the LaTeX style file *cljour1* from Springer-Verlag.