WEIGHTED COUNTING OF INTEGER POINTS IN A SUBSPACE

ALEXANDER BARVINOK AND GUUS REGTS

July 2018

Abstract. Given complex numbers $w_1, \ldots, w_n$, we define the weight $w(X)$ of a set $X$ of non-negative integer $n$-vectors as the sum of $w_1^{x_1} \cdots w_n^{x_n}$ over all vectors $(x_1, \ldots, x_n)$ in $X$. We present an algorithm, which for a set $X$ of 0-1 vectors defined by a system of homogeneous linear equations with at most $r$ variables per equation and at most $c$ equations per variable, computes $w(X)$ within relative error $\epsilon > 0$ in $(rc)^O(\ln n - \ln \epsilon)$ time provided $|w_j| \leq \beta (r \sqrt{c})^{-1}$ for an absolute constant $\beta > 0$ and all $j = 1, \ldots, n$. A similar algorithm is constructed for computing the weight of a set of non-negative integer vectors satisfying linear constraints and the weight of a linear code over $\mathbb{F}_p$. Applications include counting weighted perfect matchings in hypergraphs, counting weighted graph homomorphisms, computing weight enumerators of linear codes with sparse code generating matrices, computing the partition function of the ferromagnetic Potts model at low temperatures and computing the hard-core model at large fugacity on biregular bipartite graphs.

1. Weighted counting of integer points

(1.1) Weight of a set of integer points. Let $\mathbb{Z}_+$ denote the set of non-negative integers. Let us fix complex numbers $w_1, \ldots, w_n$, referred to as weights in what follows. We define the weight $w(x)$ of a vector $x \in \mathbb{Z}_+^n$ by

$$w(x) = w_1^{\xi_1} \cdots w_n^{\xi_n} \quad \text{where} \quad x = (\xi_1, \ldots, \xi_n).$$

Here we agree that $0^0 = 1$, so that $w(x)$ is a continuous function of $w_1, \ldots, w_n$ for a fixed $x$.

We define the weight of a finite set $X \subset \mathbb{Z}_+^n$ by

(1.1.1) $$w(X) = \sum_{x \in X} w(x) = \sum_{x \in X, x=(\xi_1, \ldots, \xi_n)} w_1^{\xi_1} \cdots w_n^{\xi_n}.$$
Given $X \subset \mathbb{Z}^n_+$, the value of $w(X)$ as a function of $w_1, \ldots, w_n$ is also known as the partition function or generating function of $X$.

Our first main result is as follows.

**Theorem.** Let $A = (a_{ij})$ be an $m \times n$ integer matrix, let $\nu_1, \ldots, \nu_n$ be positive integers and let us define $X \subset \mathbb{Z}^n_+$ by

$$X = \{x \in \mathbb{Z}^n_+, x = (\xi_1, \ldots, \xi_n) : \sum_{j=1}^n a_{ij} \xi_j = 0 \text{ for } i = 1, \ldots, m$$

$$\text{and } 0 \leq \xi_j \leq \nu_j \text{ for } j = 1, \ldots, n\}.$$  

Suppose that the number of non-zero entries in every row of $A$ does not exceed $r$ for some $r \geq 2$ and that the number of non-zero entries in every column of $A$ does not exceed $c$ for some $c \geq 1$. There is an absolute constant $\alpha > 0$ such that if $w_1, \ldots, w_n \in \mathbb{C}$ are weights satisfying

$$|w_j| \leq \frac{\alpha}{r \sqrt{c}} \text{ for } j = 1, \ldots, n,$$

then

$$w(X) \neq 0.$$  

One can choose $\alpha = 0.46$.

Geometrically, the set $X$ in Theorem 1.2 is the set of integer points in a polytope (that is the intersection of an axis-aligned parallelepiped and a subspace). The function $w(X)$, under different names, was investigated in great many papers, see, for example, [BV97] and [Ba08]. We are interested in efficient algorithms to compute $w(X)$ approximately. Theorem 1.2 implies that such an efficient algorithm exists for a non-trivial range of weights $w_1, \ldots, w_n$ provided the matrix $A$ is sufficiently sparse (that is, $r$ and $c$ are sufficiently small), even when the dimension $n$ of the ambient space is allowed to be large (for the dimension $n$ fixed in advance, the value of $w(X)$ can be computed in polynomial time exactly, cf. [Ba08]). This connection of the sparsity condition for $A$ (frequent in applications and easily verified) and the computational complexity of $w(X)$ appears to be new.

**Computing $w(X)$**. Theorem 1.2 implies that $w(X)$ can be efficiently approximated as long as the weights $w_j$ satisfy a slightly stronger inequality,

$$|w_j| \leq \frac{\beta}{r \sqrt{c}} \text{ for } j = 1, \ldots, n$$

for any $\beta < \alpha$, fixed in advance, so one can choose $\beta = 0.45$. We describe the connection below, see also Section 1.2 of [Ba16].

Without loss of generality we assume that the matrix $A$ has no zero rows and no zero columns (although this assumption is not needed in this section, it will be
relevant later in Section 5). Indeed, zero rows of $A$ can be ignored and if, say, the $n$-th column of $A$ is zero, we have

$$w(X) = \left( \sum_{i=0}^{\nu_n} w_n^i \right) w(\hat{X}),$$

where $\hat{X} \subset \mathbb{Z}_{+}^{n-1}$ is the set defined by the system $\hat{A}x = 0$, where $\hat{A}$ is the $m \times (n-1)$ matrix obtained from $A$ by deleting the $n$-th column.

For a $\zeta \in \mathbb{C}$, let $\zeta w_1, \ldots, \zeta w_n$ be the scaling of the weights and let $w(X; \zeta)$ be the corresponding weight of $X$ so that $w(X; 1) = w(X)$ while $w(X; 0) = 1$ (note that $0 \in X$). Theorem 1.2 implies that as long as the weights $w_j$ satisfy (1.3.1), we have

$$w(X; \zeta) \neq 0 \quad \text{provided} \quad |\zeta| \leq \frac{\alpha}{\beta} =: \gamma.$$  

Note that $\gamma > 1$.

Let us choose a continuous branch of $f(\zeta) = \ln w(X; \zeta)$ for $|\zeta| \leq \gamma$ and let

$$(1.3.3) \quad T_s(\zeta) = f(0) + \sum_{k=1}^{s} \frac{f^{(k)}(0)}{k!} \zeta^k$$

be the Taylor polynomial of $f$ of some degree $s$ computed at $\zeta = 0$. Since (1.3.2) holds and $w(X; \zeta)$ is a polynomial of degree at most

$$N = \nu_1 + \ldots + \nu_n$$

in $\zeta$, we have

$$|f(1) - T_s(1)| \leq \frac{N}{(s + 1) \gamma^s (\gamma - 1)},$$

see Lemma 2.2.1 of [Ba16]. Using that $\gamma > 1$, we conclude that to approximate $f(1) = \ln w(X)$ within an additive error $\epsilon > 0$ by $T_s(1)$, it suffices to choose $s = O(\ln N - \ln \epsilon)$, where the implied constant in the “$O$” notation depends only on $\gamma$. We say then that $e^{f(1)}$ approximates $w(X)$ within relative error $\epsilon$.

We have $f(0) = 0$ and computing $f^{(k)}(0)$ for $k = 1, \ldots, s$ reduces to computing

$$(1.3.4) \quad \frac{d^k}{d\zeta^k} w(X; \zeta) \bigg|_{\zeta=0} \quad \text{for} \quad k = 1, \ldots, s$$

in $O(s^2)$ time. Indeed, it is not hard to see that the values $f^{(k)}(0)$ are the solutions of a non-degenerate triangular system of linear equations with right hand side given by (1.3.4), see Section 2.2.2 of [Ba16]. Furthermore,

$$\frac{d^k}{d\zeta^k} w(X; \zeta) \bigg|_{\zeta=0} = k! \sum_{x \in X, \atop \sum_{i=1}^{n} \xi_i = k} w_{\xi_1} \cdots w_{\xi_n},$$

for $x = (\xi_1, \ldots, \xi_n)$.
so computing (1.3.4) reduces to the inspection of all points $x \in X$, $x = (\xi_1, \ldots, \xi_n)$, satisfying $\xi_1 + \ldots + \xi_n \leq s$, which can be done through the exhaustive search in $mn^{O(s)}$ time. Given that $s = O(\ln N - \ln \epsilon)$, this produces an algorithm approximating $w(X)$ within a relative error $\epsilon > 0$ in quasi-polynomial $n^{O(\ln N - \ln \epsilon)}$ time, where the implied constant in the “$O$” notation depends only on $\gamma$ in (1.3.2). In Section 5 we show that we can compute $f^{(k)}(0)$ in (1.3.3) faster, in $(rc)^{O(\ln N - \ln \epsilon)}$ time. In particular, if $r$ and $c$ are fixed in advance and $\nu_j = 1$ for all $j$ (that is, we count 0-1 vectors), we obtain a polynomial time approximation algorithm.

Next, we consider enumerating 0-1 vectors in arbitrary polytopes, not necessarily containing the origin.

(1.4) Non-homogeneous linear equations in 0-1 vectors. We interpret a vector $x = (\xi_1, \ldots, \xi_n)$ as a column $n$-vector. Let $A$ be an $m \times n$ integer matrix as above, let $b$ be an integer $m$-vector and let $X = \{x \in \{0, 1\}^n : Ax = b\}$ be the set of 0-1 vectors satisfying a system of linear equations with matrix $A$. In general, it is an NP-hard problem to decide whether $X$ is empty, so there is no hope to compute $w(X)$ efficiently.

Suppose, however, that we are presented with a point $y \in X$, $y = (\eta_1, \ldots, \eta_n)$. Every point $x \in X$ can be uniquely written as $x = y + z$, $z = (\zeta_1, \ldots, \zeta_n)$, where $Az = 0$ and $\zeta_j \in \{-1, 0\}$ if $\eta_j = 1$ and $\zeta_j \in \{0, 1\}$ if $\eta_j = 0$. Let $a_1, \ldots, a_n$ be the columns of $A$ and let $\tilde{A}$ be the matrix obtained from $A$ by replacing $a_j$ with $-a_j$ whenever $\eta_j = 1$. Let

$$Z = \{z \in \{0, 1\}^n : \tilde{A}z = 0\}.$$ 

Hence every point $x \in X$, $x = (\xi_1, \ldots, \xi_n)$, can be uniquely written as $\xi_j = \eta_j + \sigma_j \zeta_j$, where for $z = (\zeta_1, \ldots, \zeta_n)$ we have $z \in Z$ and

$$\sigma_j = \begin{cases} 1 & \text{if } \eta_j = 0, \\ -1 & \text{if } \eta_j = 1. \end{cases}$$

Then, for the weight of $Z$, we have

$$w(Z) = \sum_{z \in Z} \prod_{j=1}^n w_j^{\zeta_j} = \sum_{x \in X} \prod_{j=1}^n w_j^{\sigma_j(\xi_j - \eta_j)}$$

(1.4.1)

$$= \sum_{x \in X} \prod_{j: \xi_j \neq \eta_j} w_j.$$ 

For $x \in \{0, 1\}^n$, $x = (\xi_1, \ldots, \xi_n)$, let

$$\text{dist}(x, y) = |\{j : \xi_j \neq \eta_j\}|$$

\[4\]
be the Hamming distance between $x$ and $y$.

In particular, if we choose

$$w_1 = \ldots = w_n = \omega$$

for some $\omega$, we get

$$w(Z) = \sum_{x \in X} \omega^{\text{dist}(x,y)}. \tag{1.4.2}$$

Assuming that every row of $A$ contains not more than $r \geq 2$ non-zero entries and every column of $A$ contains not more than $c \geq 1$ non-zero entries, we conclude that the sum (1.4.2) can be computed within relative error $\epsilon > 0$ in $(rc)^{O(\ln n - \ln \epsilon)}$ time provided

$$|\omega| \leq \frac{\beta}{r\sqrt{c}},$$

where $\beta > 0$ is an absolute constant (one can choose $\beta = 0.45$). If $r$ and $c$ are fixed in advance, we have a polynomial time approximation algorithm of $(n/\epsilon)^{O(1)}$ complexity.

In the next section we consider combinatorial applications of our result. We first consider a variation of Theorem 1.2 that applies to codes.

**1.5 Weight of a code.** Let $\kappa > 1$ be an integer. We consider $n$-vectors $x = (\xi_1, \ldots, \xi_n)$ with coordinates $\xi_j$ taking values in the set $\{0, \ldots, \kappa - 1\}$, which we interpret as the set $\mathbb{Z}/\kappa\mathbb{Z}$ of remainders modulo $\kappa$. Given $n$ complex numbers $w_1, \ldots, w_n$, we define the weight $w(x)$ of a vector $x \in (\mathbb{Z}/\kappa\mathbb{Z})^n$ by

$$w(x) = \prod_{j: \xi_j \neq 0} w_j \quad \text{for} \quad x = (\xi_1, \ldots, \xi_n)$$

and the weight $w(X)$ of a set $X \subset (\mathbb{Z}/\kappa\mathbb{Z})^n$ by

$$w(X) = \sum_{x \in X} w(x)$$

(we agree that the weight of the zero vector is 1).

We obtain the following result.

**1.6 Theorem.** Let $A = (a_{ij})$ be an $m \times n$ integer matrix and let us define a set $X \subset (\mathbb{Z}/\kappa\mathbb{Z})^n$ by

$$X = \left\{ x \in (\mathbb{Z}/\kappa\mathbb{Z})^n, \ x = (\xi_1, \ldots, \xi_n) : \sum_{j=1}^n a_{ij}\xi_j \equiv 0 \mod \kappa \quad \text{for} \quad i = 1, \ldots, m \right\}.$$
Suppose that the number of non-zero entries in every row of A does not exceed \( r \) for some \( r \geq 2 \) and that the number of non-zero entries in every column of A does not exceed \( c \) for some \( c \geq 1 \). There is an absolute constant \( \alpha > 0 \) such that if \( w_1, \ldots, w_n \in \mathbb{C} \) are weights satisfying

\[
|w_j| \leq \frac{\alpha}{(\kappa - 1)r\sqrt{c}} \quad \text{for} \quad j = 1, \ldots, n,
\]

then

\[
w(X) \neq 0.
\]

One can choose \( \alpha = 0.46 \).

As in Section 1.3, we obtain an algorithm of \( (rc)^{O(\ln \kappa n - \ln \epsilon)} \) complexity to approximate \( w(X) \) within relative error \( \epsilon > 0 \) provided

\[
|w_j| \leq \frac{\beta}{(\kappa - 1)r\sqrt{c}} \quad \text{for} \quad j = 1, \ldots, n,
\]

where \( \beta < \alpha \) is fixed in advance (we can choose \( \beta = 0.45 \)). For \( r \) and \( c \) fixed in advance, the algorithm has polynomial \( (\kappa n/\epsilon)^{O(1)} \) complexity. We consider an application to computing weight enumerators of codes in Section 2.4.

Organization. We deduce Theorem 1.2 and Theorem 1.6 from a general result asserting that

\[
\int_{\mathbb{T}^m} e^{p(z)} \, d\mu \neq 0,
\]

for some Laurent polynomials \( p : \mathbb{T}^m \to \mathbb{C} \) on the torus \( \mathbb{T}^m \) endowed with a product probability measure \( \mu \) (see Theorem 3.2 and Corollary 3.3 below). After that, the proofs of Theorems 1.2 and 1.6 are completed in a more or less straightforward way in Section 4.

In Section 5, we provide details of an approximation algorithm for \( w(X) \). We do not discuss an analogous algorithm for codes in Theorem 1.6 as it is very similar. We first consider some concrete combinatorial applications of these results in Section 2 below.

2. Combinatorial applications

(2.1) Perfect matchings in hypergraphs. A hypergraph \( H = (V, E) \) is a finite set \( V \) of vertices together with a collection \( E \) of non-empty subsets \( V \), called edges of the hypergraph. The degree of a vertex \( v \) is the number of edges \( e \in E \) that contain \( v \). A perfect matching in \( H \) is a set of pairwise disjoint edges \( e_1, \ldots, e_n \), such that \( e_1 \cup \ldots \cup e_n = V \). Let us introduce a 0-1 variable \( x_e \) for each \( e \in H \). We encode a collection of edges of \( H \) by a 0-1 vector, where

\[
x_e = \begin{cases} 
1 & \text{if } e \text{ is in the collection} \\
0 & \text{otherwise}.
\end{cases}
\]
Then $e_1, \ldots, e_n$ is a perfect matching if and only if

$$\sum_{e: v \in e} x_e = 1 \quad \text{for all} \quad v \in V. \quad (2.1.1)$$

In the system (2.1.1) the number of variables per equation is the maximum degree $d$ of a vertex of $H$ and the number of equations per variable is the maximum cardinality $k$ of an edge. As is known, it is an NP-complete problem to find if a given hypergraph contains a perfect matching provided $k \geq 3$, see, for example, Problem SP1 in [A+99]. However, as follows from Section 1.4, given one perfect matching $M_0$, we can efficiently approximate a certain statistics over all perfect matchings $M$ of $H$, namely the sum

$$\sum_{M \in \mathcal{M}(H)} \omega_{\text{dist}(M_0, M)}, \quad (2.1.2)$$

where $\mathcal{M}(H)$ is the set of all perfect matchings, $\text{dist}(M_0, M)$ is the Hamming distance between matchings, that is, the number of edges where the matchings differ and

$$|\omega| \leq \frac{\beta}{d\sqrt{k}}.$$

The complexity of the algorithm approximating (2.1.2) within relative error $\epsilon > 0$ is $(dk)^{O(\ln |E| - \ln \epsilon)}$. If $d$ and $k$ are fixed in advance, the algorithm achieves polynomial $(|E|/\epsilon)^{O(1)}$ complexity. This can be contrasted with the fact that knowing one solution of a problem generally does not help to find another or to count all solutions, cf. [Va79] and [VV86].

As is shown in [Ba18] that if the hypergraph is uniform and $k$-partite, that is, we have $V = V_1 \cup \ldots \cup V_k$ with pairwise disjoint $V_1, \ldots, V_k$ such that $|V_1| = \ldots = |V_k| = n$ and every edge $e \in E$ contains exactly one vertex from each $V_i$, then one can efficiently approximate (2.1.2) under the weaker condition

$$|\omega| \leq \frac{\beta}{\sqrt{d-1}}$$

for any $\beta < 1$, fixed in advance.

(2.2) Counting graph homomorphisms. Let $G_1 = (V_1, E_1)$ be an undirected graph without loops or multiple edges and let $G_2 = (V_2, E_2)$ be an undirected graph without multiple edges, but possibly with loops. We assume that $V_2 = \{1, \ldots, n\}$ and assume that $G_1$ and $G_2$ are both connected. A map $\phi: V_1 \rightarrow V_2$ is called a homomorphism if $\phi(u)$ and $\phi(v)$ span an edge of $G_2$ whenever $u$ and $v$ span an edge of $V_1$. If $V_2$ is the complete graph without loops then every homomorphism $\phi: G_1 \rightarrow G_2$ is naturally interpreted as a coloring of the vertices of $G_1$ with a set of $n$ colors such that no two vertices spanning an edge of $G_1$ are colored with the
same color (such colorings are called *proper*). As is known, for any fixed \( n \geq 3 \), it is an NP-complete problem to decide wether a given graph admits a proper \( n \)-coloring, see for example, Problem GT5 in [A+99]. Our goal is to encode all homomorphisms \( \phi : G_1 \to G_2 \) that map a fixed vertex \( a \in V_1 \) to a fixed vertex, say \( n \), of \( G_2 \) as the set of 0-1 solutions to a system of linear equations.

We say that vertices \( u, v \in V_1 \) are *neighbors* if \( \{u, v\} \in E_1 \). We orient the edges of \( G_1 \) arbitrarily, so that an edge of \( G_1 \) is an ordered pair of neighbors \( (u, v) \). Let us introduce 0-1 variables \( x_{ij} \) indexed by (now directed) edges \( (u, v) \in E_1 \) and ordered pairs \( 1 \leq i, j \leq n \) such that \( \{i, j\} \in E_2 \) (we may have \( i = j \)). The idea is to use the variables \( x_{ij} \) to encode a map \( \phi : V_1 \to V_2 \), so that

\[
(2.2.1) \quad x_{ij} = \begin{cases} 
1 & \text{if } \phi(u) = i \text{ and } \phi(v) = j, \\
0 & \text{otherwise.}
\end{cases}
\]

For every pair of neighbors \( \{u, v\} \in E_1 \) and every vertex \( i \in V_2 \) we define the sum

\[
S_{i}^{u,v} = \sum_{j: \{i,j\} \in E_2} x_{ij}^{uv} \quad \text{if } (u, v) \in E_1 \quad \text{and}
\]

\[
(2.2.2) \quad S_{i}^{u,v} = \sum_{j: \{i,j\} \in E_2} x_{ji}^{vu} \quad \text{if } (v, u) \in E_1
\]

and for every \( u \in V_1 \) and every \( i \in V_2 \), we introduce the equations:

\[
(2.2.3) \text{Fix } u \in V_1 \text{ and } i \in V_2. \text{ For all neighbors } v \text{ of } u, \text{ the sums } S_{i}^{u,v} \text{ are equal.}
\]

The idea, of course, is that the sums (2.2.2) are all equal to 1 if \( \phi(u) = i \) and equal to 0 if \( \phi(u) \neq i \). Next, we encode the condition \( \phi(a) = n \) by the following system of equations:

\[
(2.2.4) \text{For all neighbors } v \text{ of } a, \quad S_{n}^{a,v} = 1 \quad \text{and} \quad S_{j}^{a,v} = 0 \quad \text{for } j \neq n.
\]

Now we claim that for every 0-1 solution \( \{x_{ij}^{uv}\} \) of the system (2.2.3)–(2.2.4), for any vertex \( u \in V_1 \), there is a unique vertex \( i_u \in V_2 \) such that the following equations hold:

\[
(2.2.5) \text{For all neighbors } v \text{ of } u \text{ we have} \quad S_{i_u}^{u,v} = 1 \quad \text{and} \quad S_{j}^{u,v} = 0 \quad \text{for } j \neq i_u.
\]

Then for the map \( \phi : V_1 \to V_2 \) defined by \( \phi(u) = i_u \) the conditions (2.2.1) are satisfied.
Clearly, if a choice \( u \mapsto i_u \) exists, it is unique. Because of (2.2.4), the equations (2.2.5) hold for \( u = a \) and \( i_u = n \). Since \( G_1 \) is connected, it suffices to show whenever (2.2.5) holds for some vertex \( u \) then for every neighbor \( w \) of \( u \) we can define \( i_w \in V_2 \) so that (2.2.5) holds with \( u \) replaced by \( w \) throughout. Indeed, let \( w \) be a neighbor of \( u \) such that \((u, w) \in E_1\). It follows by (2.2.5) that there exists \( i_w \) such that

\[ x^{uw}_{iu} = 1 \quad \text{and} \quad x^{uw}_{jk} = 0 \quad \text{whenever} \quad j \neq i_u \quad \text{or} \quad k \neq i_w. \]

From (2.2.3) it follows that for any neighbor \( v \) of \( w \), we have

\[ S^{w,v}_{i_w} = S^{w,u}_{i_w} = 1 \quad \text{and} \quad S^{w,v}_{j} = S^{w,u}_{j} = 0 \quad \text{for} \quad j \neq i_w, \]

as required. The case of neighbors \( w \) of \( u \) such that \((w, u) \in E_1\) is handled similarly.

As we are interested to keep the system (2.2.3)–(2.2.4) as sparse as possible, we arrange the equations (2.2.3) as follows: for a given \( u \in V_1 \), we list the neighbors \( v \) of \( u \) in some order \( v_1, \ldots, v_m \) and then equate \( S^{u,v_k}_{i} = S^{u,v_{k+1}}_{i} \) for \( k = 1, \ldots, m - 1 \). When the chosen vertex \( a \) is a neighbor, we list it first. This way the system (2.2.3)–(2.2.4) has not more than \( 2d_2 \) variables per equation, where \( d_2 \) is the largest degree of a vertex of \( G_2 \), and not more than 4 equations per variable.

Suppose that we are given a homomorphism \( \phi : G_1 \to G_2 \) satisfying the constraint \( \phi(a) = n \) for a fixed vertex \( a \) of \( G_1 \) and a fixed vertex \( n \) of \( G_2 \). As in Section 1.4, for an \( \omega \in \mathbb{C} \) we consider the sum

\[
(2.2.6) \quad \sum_{\psi : \psi(a) = n} \omega^{2 \text{dist}(\phi, \psi)},
\]

where \( \psi \) ranges over all graph homomorphisms satisfying \( \psi(a) = n \) and \( \text{dist}(\phi, \psi) \) is the number of edges where \( \phi \) and \( \psi \) disagree. As follows from Section 1.4, we can approximate (2.2.6) within relative error \( \epsilon > 0 \) in \( d_2^{O(\ln |E_1| + \ln |E_2| - \ln \epsilon)} \) time provided

\[
(2.2.7) \quad |\omega| \leq \frac{\gamma}{d_2}
\]

for some absolute constant \( \gamma > 0 \) (we can choose \( \gamma = 0.1 \)). If the largest degree \( d_2 \) of a vertex of \( G_2 \) is fixed in advance, we obtain a polynomial time approximation algorithm.

Suppose that \( G_2 \) is the complete graph with \( n \) vertices and no loops, so that a homomorphism \( G_1 \to G_2 \) is interpreted as a proper \( n \)-coloring of \( G_1 \) and \( d_2 = n - 1 \). If \( n > d_1 \), where \( d_1 \) is the largest degree of a vertex of \( G_1 \), it is trivial to come
up with a homomorphism (proper $n$-coloring) $\phi : G_1 \rightarrow G_2$ having a prescribed value on a prescribed vertex. In this case, the sum (2.2.6) is taken over all proper $n$-colorings $\psi$ of $G_2$ and each coloring is counted with weight exponentially small in the number of edges of $G_1$ whose coloring differ under $\phi$ and $\psi$. If we could choose $\omega = 1$ in (2.2.6), we would have counted all proper $n$-colorings of $G_1$ with $n > d_1$ colors, a notoriously difficult problem, see [Vi00] and [CM18, D+18] for a randomized polynomial time approximation algorithm for counting $n$-colorings assuming that $n > (11/6)d_1$.

(2.3) Ramifications and limitations. Given a pair of graphs $G_1$ and $G_2$ as in Section 2.2, let us modify $G_2$ to a graph $\hat{G}_2$ by adding an extra vertex $n + 1$ with a loop and connected to all other vertices of $\hat{G}_2$. Then there is always a homomorphism $\phi : G_1 \rightarrow \hat{G}_2$ which sends every vertex of $G_1$ to the newly added vertex $n + 1$. In this case the sum (2.2.6) with $G_2$ replaced by $\hat{G}_2$ and $n$ replaced by $n + 1$ is interpreted as the sum over all homomorphisms of the induced subgraphs of $G_1$ to $G_2$. By connecting vertex $n + 1$ only to some vertices of $G_2$, we can prescribe the behavior of $\psi$ on the “interface” consisting of the edges of $G_1$ with $\psi$ defined on one endpoint and not defined on the other.

We remark that even though the dependence on the maximum degree of $G_1$ appears absent in the conditions in (2.2.7), in special cases we do see this dependence. Given an undirected graph $G = (V, E)$, a set $S \subset V$ of vertices is called independent if no two vertices of $S$ span an edge of $G$ (we agree that $S = \emptyset$ is always independent). The independence polynomial of $G$ is a univariate polynomial defined by

\[ p_G(\lambda) = \sum_{S \subset V} \lambda^{|S|}, \]

(2.3.1)

see for example, Chapter 6 of [Ba16]. Let us choose $G_2$ to be the graph with set $V_2 = \{1, 2\}$ of vertices, with a loop at 2 and an edge connecting 1 and 2, let $G_1$ be a $d$-regular graph and let $\phi$ be the map sending every vertex of $G_1$ to 2. One can observe that (2.2.6) is equal to $p_{G_1 - a}(\lambda)$ for $\lambda = \omega^{2d}$ (where $a$ is some fixed vertex of $G_1$). Theorem 1.2 can only guarantee that if $|\lambda|$ is of order $2^{-d}$, then this polynomial does not vanish at $\lambda$. Shearer [Sh85] however proved that it suffices to have $|\lambda|$ of order $1/d$ for the independence polynomial not to vanish at $\lambda$. This shows that even though our results are quite general, they are not optimal in certain special cases.

(2.4) Computing weight enumerators of linear codes. If $\kappa$ is a prime, the set $\mathbb{Z}/\kappa\mathbb{Z}$ is identified with the finite field $\mathbb{F}_\kappa$ with $\kappa$ elements and $(\mathbb{Z}/\kappa\mathbb{Z})^n$ is the $n$-dimensional vector space over $\mathbb{F}_\kappa$. A set $X \subset \mathbb{F}_\kappa^n$ is called a code. The univariate polynomial

\[ p_X(z) = 1 + \sum_{k=1}^{n} p_k(X)z^k, \]

is
where \( p_k(X) \) is the number of vectors in \( X \) with exactly \( k \) non-zero coordinates, is called the \textit{weight enumerator} of \( X \), see for example, Chapter 3 of [Li99].

Suppose that \( X \subset \mathbb{F}_k^n \) is defined by a system of linear equations
\[
(2.4.1) \quad X = \{ x \in \mathbb{F}_k^n : \ A x = 0 \},
\]
where \( A = (a_{ij}) \) is an \( m \times n \) matrix with entries \( a_{ij} \in \mathbb{F}_k \). Hence \( X \subset \mathbb{F}_k^n \) is a subspace, called a \textit{linear code}. Generally, it is hard to compute \( p_X(z) \) as it is hard to determine the smallest \( k \geq 1 \) with \( p_k(X) \neq 0 \), see [B+78] and [BN90].

Suppose now that the number of non-zero entries in every row of \( A \) does not exceed \( r \geq 2 \) and the number of non-zero entries in every column of \( A \) does not exceed \( c \geq 1 \). Let us define weights
\[
w_1 = \ldots = w_n = z
\]
for some \( z \in \mathbb{C} \). Then
\[
w(X) = p_X(z)
\]
and Theorem 1.6 implies that \( p_X(z) \neq 0 \) provided \( |z| \leq \alpha/(\kappa - 1)r\sqrt{c} \) and that \( p_X(z) \) can be approximated within relative error \( \epsilon > 0 \) in \( (rc)^{O(\ln n - \ln \epsilon)} \) time, provided \( |z| \leq \beta/(\kappa - 1)r\sqrt{c} \), where \( \beta < \alpha \) is fixed in advance. Again, if \( r \) and \( c \) are fixed in advance, we obtain an algorithm of polynomial \( m(n/\epsilon)^{O(1)} \) complexity. Linear codes \( X \) (typically binary, that is for \( \kappa = 2 \)) for which the number of non-zero entries in each row of the matrix \( A \) in (2.4.1) is small are called \textit{low-density parity-check} codes. They have many desirable properties and are of considerable interest, cf. Section 11 of [MM09]. If, in addition, the binary matrix \( A \) has the same number \( r \) of 1s in every row and the same number \( c \) of 1s in every column, the code \( X \) is called \textit{regular}.

Let \( C = X^\perp \), \( C \subset \mathbb{F}_k^n \), be the subspace (linear code) spanned by the rows of \( A \) (we say that \( A \) is the \textit{generator matrix} of \( C \)). The MacWilliams identity for the weight enumerators of \( p_X \) and \( p_C \) (see Theorem 3.5.3 of [Li99]) states that
\[
p_X(z) = \frac{1}{\kappa^{\dim C}}(1 + (\kappa - 1)z)^np_C\left(\frac{1 - z}{1 + (\kappa - 1)z}\right).
\]
It follows that
\[
p_C\left(\frac{1 - z}{1 + (\kappa - 1)z}\right) \neq 0 \quad \text{provided} \quad |z| \leq \frac{\alpha}{(\kappa - 1)r\sqrt{c}}
\]
and that the value of
\[
p_C\left(\frac{1 - z}{1 + (\kappa - 1)z}\right)
\]
can be efficiently approximated provided
\[
|z| \leq \frac{\beta}{(\kappa - 1)r\sqrt{c}}.
\]
In other words, the weight enumerator $p_C(z)$ of a linear code $C$ with a sparse code generator matrix is non-zero and can be efficiently approximated provided $|z| = 1 + O(1/r\sqrt{c})$, where $r$ is an upper bound on the number of non-zero entries in every row, $c$ is an upper bound on the number of non-zero entries in every column of the matrix and the implied constant in the “$O$” notation is absolute (in particular, it does not depend on $\kappa$).

One notable example of such a code with a sparse generating matrix is the binary cut code consisting of the indicators of cuts in a given graph $G = (V, E)$ with set $V$ of vertices and set $E$ of edges, see Section 1.9 of [Di05] and [BN90], that is, indicators of subsets $E_S \subset E$ consisting of the edges with one endpoint in $S \subset V$ and the other in $V \setminus S$. The rows of the code generating matrix are parameterized by vertices $v \in V$ of the graph, the columns are parameterized by the edges $e$ of the graph and the $(v, e)$ entry of the matrix is 1 if $v$ is an endpoint of $e$ and 0 otherwise (hence each row is the indicator of the cut associated with the corresponding vertex). We observe that the code generating matrix of a cut code contains at most $d(G)$ non-zero entries in every row, where $d(G)$ is the largest degree of a vertex of $G$, and exactly two non-zero entries in every column. The obtained algorithm for computing the weight of a cut code achieves roughly the same approximation as the algorithms of [PR17a] and of Chapter 7 of [Ba16], where we approach computing weights of cuts via the graph homomorphism partition function.

**2.5) Ferromagnetic Potts model at low temperatures.** Let $G = (V, E)$ be a connected undirected graph, without loops or multiple edges. Given a real $\beta > 0$ and an integer $\kappa > 1$, we consider the sum

$$
(2.5.1) \quad P_{G, \kappa}(\beta) = \sum_{\phi: V \rightarrow \{0, \ldots, \kappa-1\}} \exp \left\{ \beta \sum_{\{u, v\} \in E} \delta_{\phi(u)\phi(v)} \right\},
$$

where

$$
\delta_{ij} = \begin{cases} 
1 & \text{if } i = j, \\
0 & \text{if } i \neq j.
\end{cases}
$$

The expression (2.5.1) is known as the partition function of the ferromagnetic (since $\beta > 0$) Potts model with $\kappa$ colors, see, for example, [FV18]. Here the numbers $0, 1, \ldots, \kappa - 1$ are interpreted as colors: we color the vertices of $G$ with $\kappa$ colors in all possible ways, and each edge of $G$ with identically colored endpoints contributes to the inner sum. The number $\beta$ plays the role of the inverse temperature. Using cluster expansions, it was shown in [H+18] that for some induced subgraphs $G$ of the lattice $\mathbb{Z}^d$ the sum (2.5.1) can be approximated in polynomial time provided $\beta > \beta_0(d, \kappa)$ for some constant $\beta_0$ (that is, at sufficiently low temperatures). Here we deduce this result for a wide family of graphs and an explicit bound on $\beta_0$ from our Theorem 1.6.
First, we rewrite (2.5.1) in the form
\[ P_{G,\kappa}(\beta) = \exp^{\beta |E|} \sum_{\phi: V \rightarrow \{0, \ldots, \kappa - 1\}} \prod_{\{u,v\} \in E} w(\phi(u), \phi(v)) \]
(2.5.2)
where \( w(i,j) = w_\beta(i,j) = \exp^{\beta \delta_{ij} - 1} \).

Since \( \beta > 0 \), we have \(|w(i,j)| \leq 1\) and \( w(i,j) = 1 \) if and only if \( i = j \).

Next, we write the sum in (2.5.2) in the form \( w(X) \), where \( X \) is the set in Theorem 1.6. For that, we interpret colors \( 0, 1, \ldots, \kappa - 1 \) as remainders modulo \( \kappa \). We direct the edges of \( G \) in an arbitrary way and with every, now directed, edge \( (u,v) \) we associate a variable \( x_{uv} \) taking values in \( \mathbb{Z}/\kappa\mathbb{Z} \). The intended meaning of the variables \( x_{uv} \) is that
\[ x_{uv} \equiv \phi(v) - \phi(u) \mod \kappa \]
for all \( (u,v) \in E \), so that \( x_{uv} \equiv 0 \) if and only if the endpoints of the edge \( \{u,v\} \) are colored with the same color. Given a set \( \{x_{uv} : (u,v) \in E\} \), a solution \( \phi : V \rightarrow \mathbb{Z}/\kappa\mathbb{Z} \) to the system (2.5.3) exists, if and only if \( \{x_{uv}\} \) satisfy a system of linear equations, constructed as follows: we pick a cycle \( C \) in \( G \), orient it arbitrarily, and write
\[ \sum_{\{u,v\} \in C : (u,v) \text{ is co-oriented with } C} x_{uv} - \sum_{\{u,v\} \in C : (u,v) \text{ is counter-oriented with } C} x_{uv} \equiv 0 \mod \kappa. \]
(2.5.4)
Moreover, since \( G \) is connected, as long as the equations (2.5.4) are satisfied, the system (2.5.3) has exactly \( \kappa \) solutions, that differ by a shift by an element of \( \mathbb{Z}/\kappa\mathbb{Z} \).

Let \( X \subset (\mathbb{Z}/\kappa\mathbb{Z})^E \) be the set of solutions of the system (2.5.4). We introduce a weight \( w_{uv} = \exp^{-\beta} \) for each coordinate \( x_{uv} \) with \( (u,v) \in E \) and write (2.5.2) as
\[ P_{G,\kappa}(\beta) = \kappa \exp^{\beta |E|} w(X), \]
where \( X \) is the set of solutions to the system (2.5.4).

The equations (2.5.4) are not independent: it suffices to write (2.5.4) for a set of cycles \( C \) that generate the homology group \( H_1(G;\mathbb{Z}) \). In view of Theorem 1.6, we would like to choose the generators \( C \) of \( H_1(G;\mathbb{Z}) \) so that the number of edges in each cycle \( C \) does not exceed some \( r \geq 2 \) and the number of cycles \( C \) containing a given edge does not exceed some \( c \geq 1 \), for the smallest possible values of \( r \) and \( c \). Then we can approximate the partition function \( P_{G,\kappa}(\beta) \) of (2.5.1)–(2.5.2) provided
\[ \beta \geq 0.8 + \ln \left( (\kappa - 1)r\sqrt{c} \right) > -\ln 0.45 + \ln \left( (\kappa - 1)r\sqrt{c} \right), \]
and for fixed \( r \) and \( c \), we get a polynomial time approximation algorithm.
For example, suppose that \( G \) is an induced subgraph of the integer lattice \( \mathbb{Z}^d \) (with \( d \geq 2 \)) constructed as follows. Given a point \( (a_1, \ldots, a_d) \in \mathbb{Z}^d \), we call the set
\[
\{(x_1, \ldots, x_d) : a_k \leq x_k \leq a_k + 1 : k = 1, \ldots, d\}
\] an elementary cube. We take finitely many elementary cubes whose union \( U \) is a simply connected subset of \( \mathbb{R}^d \) and let \( G \) be the induced subgraph with vertices in \( U \). Then there is a system of generators of \( H_1(G; \mathbb{Z}) \) consisting of cycles with \( r = 4 \) edges each and such that every edge of \( C \) belongs to at most \( c = 2(d - 1) \) cycles (we choose the cycles on the boundary of 2-dimensional faces of the elementary cubes comprising \( U \)). Hence for such a graph \( G \), we obtain a polynomial time approximation algorithm for \( P_{G,\kappa}(\beta) \) provided \( \beta \geq 2.6 + \ln \left( (\kappa - 1)\sqrt{d-1} \right) > \ln \frac{4\sqrt{2}}{0.45} + \ln \left( (\kappa - 1)\sqrt{d-1} \right) \).

### (2.6) The hard-core model at large fugacity.

For a graph \( G = (V,E) \) and \( \lambda \in \mathbb{C} \) recall the definition of the independence polynomial \( p_G(\lambda) \), cf. (2.3.1). This is sometimes also called the partition function of the hard-core model. The parameter \( \lambda \) is known as the fugacity.

The problem of (approximately) computing the number of independent sets in a bipartite graph is considered to be computationally hard. It is the basis of the class of \#\text{BIS} hard problems, and it is known that to approximate the \( p_G(\lambda) \) on bipartite graphs of maximum degree \( d \) is a \#\text{BIS} hard problem, provided \( \lambda > \frac{(d-1)^d-1}{(d-2)^d} \) [C+16].

In [J+18] it was however shown that for \( d \geq 3 \) there exists \( \lambda^* = \lambda^*(d) > 0 \) such that for all \( \lambda > \lambda^* \) and all \( d \)-regular, bipartite, expander graphs \( G \), the value of \( p_G(\lambda) \) can be approximated in polynomial time. Here we will use Theorem 1.2 to show that for each fixed \( d_1, d_2 \in \mathbb{N} \) such that \( d_2 - d_1 \geq 2 \), there exists \( \lambda_0 = \lambda_0(d_1,d_2) > 0 \) such that for all \( \lambda > \lambda_0 \) and any biregular, bipartite graph with degrees \( d_1, d_2 \), we can approximate \( p_G(\lambda) \) in polynomial time.

To this end, let us fix a biregular bipartite graph \( G = (V,E) \) with degrees \( d_1 \) and \( d_2 \geq d_1 + 2 \). We write \( V = L \cup R \) for the bipartition and we assume that each vertex in \( L \) has degree \( d_1 \) and each vertex in \( R \) has degree \( d_2 \). For an independent set \( I \) we write \( I_L := I \cap L \) and \( I_R := I \cap R \).

We wish to encode \( p_G(\lambda) \) as the weight, \( w(X) \), of a suitably defined set \( X \). We direct all edges from \( L \) to \( R \), thus making \( G \) a directed graph. We associate to each vertex \( v \in V \) a 0-1 variable \( x_v \) and to each edge \((u,v) \in E\) a 0-1 variable \( x_{uv} \). Let \( X \) be the solution set to the following system of equations:
\[
-x_u + x_v + x_{uv} = 0 \quad \text{for each directed edge } (u,v) \in E.
\]

Any \( x \in X \) uniquely corresponds to an independent set \( I \) of \( G \). Indeed, let \( I \) be the set of vertices \( u \in L \) for which \( x_u = 0 \) and vertices \( v \in R \) for which \( x_v = 1 \).
Then for \( u \in I_L \) none of its neighbor will be contained in \( I \) since for each edge \((u, v)\), the value of \( x_v \) is forced to be zero. Similarly for any \( v \in I_R \), none of its neighbor will be contained in \( I \) since for each edge \((u, v)\), the value of \( x_u \) is forced to be 1. Hence the set \( I \) is independent. Conversely, if \( I \) is an independent set, setting

\[
\begin{align*}
x_u &= \begin{cases} 0 & \text{if } u \in I_L, \\ 1 & \text{if } u \in L \setminus I_L, \end{cases} \\
x_v &= \begin{cases} 0 & \text{if } v \in I_R, \\ 1 & \text{if } v \in R \setminus I_R \end{cases} \\
x_{uv} &= \begin{cases} 0 & \text{if } u \in I_L \text{ or } v \in I_R, \\ 1 & \text{if } u \in L \setminus I_L \text{ and } v \in R \setminus I_R, \end{cases}
\end{align*}
\]

gives a solution to (2.6.1).

Next, we introduce weights \( w_u \) for the coordinates \( x_u \) with \( u \in L \), weights \( w_v \) for the coordinates \( x_v \) with \( v \in R \) and weights \( w_{uv} \) for the coordinates \( x_{uv} \) with \((u, v) \in E\) as follows:

\[
\begin{align*}
w_u &= \omega^{d_2-d_1-1} \quad \text{for } u \in L, \\
w_v &= \omega \quad \text{for } v \in R \quad \text{and} \\
w_{uv} &= \omega \quad \text{for } (u, v) \in E.
\end{align*}
\]

For a solution \( x \in X \) corresponding to an independent set \( I \), we then have

\[
\begin{align*}
w(x) &= \left( \prod_{v \in L \setminus I_L} \omega^{(d_2-d_1-1)} \right) \left( \prod_{\{u,v\} \in E \setminus I_{uv}} \omega \right) \\
&= \omega^{|L \setminus I_L|((d_2-d_1-1)+d_1|L|-d_1|L|-d_2|I_R|+|I_R|} \\
&= \omega^{(d_2-1)|L|} \cdot \omega^{(1-d_2)|I|}.
\end{align*}
\]

In other words, for the weight of \( X \), we have

\[
w(X) = \omega^{(d_2-1)|L|} p_G \left( \frac{1}{\omega^{d_2-1}} \right).
\]

Now, since in (2.6.1), the number of variables per equation is 3 and the number of equations per variable is at most \( d_2 \), it follows from Theorem 1.2 that if

\[
|\lambda| \geq \left( 6.7 \sqrt{d_2} \right)^{d_2-1} > \left( \frac{3 \sqrt{d_2}}{0.45} \right)^{d_2-1},
\]

then \( p_G(\lambda) \neq 0 \) and moreover that we can efficiently approximate \( p_G \) (in polynomial time if \( d_2 \) is fixed in advance).
We moreover note that with a similar argument, for a $d$-regular bipartite graph $G = (L \cup R, E)$, we can efficiently approximate the sum

$$\sum_{I \subset L \cup R} \lambda^{|I \cap L|}$$

for large $\lambda$. This is somewhat similar in spirit to a result of van den Berg and Steiff [BS94], who showed that for the integer lattice $\mathbb{Z}^d$, assigning $\lambda_1 > 0$ to vertices with even coordinate sum and $\lambda_2 > 0$ to vertices with odd coordinate sum, for all but a countable set of pairs $(\lambda_1, \lambda_2)$ the associated Gibbs measure is unique.

3. Integrating over the torus

We begin our preparations to prove Theorems 1.2 and 1.6.

(3.1) Laurent polynomials on the torus. Let

$$S^1 = \{ z \in \mathbb{C} : |z| = 1 \}$$

be the unit circle in the complex plane and let

$$T^m = S^1 \times \ldots \times S^1$$

be the direct product of $m$ copies of $S^1$ (torus), endowed with the product measure $\mu = \mu_1 \times \ldots \times \mu_m$, where $\mu_i$ is a Borel probability measure on the $i$-th copy of $S^1$. We consider Laurent polynomials $p : T^m \rightarrow \mathbb{C}$,

(3.1.1) $$p(z_1, \ldots, z_m) = \sum_{a \in A} \gamma_a z^a$$

as random variables on $T^m$. Here $A \subset \mathbb{Z}^m$ is a finite set of integer vectors, $\gamma_a \in \mathbb{C}$ for all $a \in A$ and

$$z^a = z_1^{\alpha_1} \cdots z_m^{\alpha_m} \quad \text{provided} \quad a = (\alpha_1, \ldots, \alpha_m),$$

where $z_i^0 = 1$. We are interested in conditions on the coefficients $\gamma_a$ which ensure that $\mathbb{E} e^p \neq 0$.

For $a \in A$ we define the support of $a$ by

$$\text{supp } a = \{ i : \alpha_i \neq 0 \} \quad \text{where} \quad a = (\alpha_1, \ldots, \alpha_m).$$

Consequently, $|\text{supp } a|$ is the number of non-zero coordinates of $a \in \mathbb{Z}^m$. In this section, we prove the following main result.
(3.2) Theorem. Let $p : \mathbb{T}^m \to \mathbb{C}$ be a Laurent polynomial as in (3.1.1). Suppose that for some $0 \leq \theta_1, \ldots, \theta_m < 2\pi/3$, we have

$$
2 \sum_{a \in A : \gamma_a \in \text{supp } a} |\gamma_a| \prod_{j \in \text{supp } a} \frac{1}{\cos(\theta_j/2)} \leq \theta_i \quad \text{for } i = 1, \ldots, m.
$$

Then

$$
E e^p \neq 0.
$$

By choosing $\theta_i$ in a particular way, we obtain the following corollary.

(3.3) Corollary. There exists an absolute constant $\kappa > 0$ such that if $p : \mathbb{T}^m \to \mathbb{C}$ is a Laurent polynomial as in (3.1.1) and

$$
|\text{supp } a| \leq c \quad \text{for all } a \in A
$$

and some $c \geq 1$ and

$$
\sum_{a \in A : \gamma_a \in \text{supp } a} |\gamma_a| \leq \frac{\kappa}{\sqrt{c}} \quad \text{for } i = 1, \ldots, m.
$$

Then

$$
E e^p \neq 0.
$$

One can choose $\kappa = 0.56$.

The proof is somewhat similar to that of [Ba17] for $E e^p$ where $p : \{-1,1\}^m \to \mathbb{C}$ is a polynomial on the Boolean cube.

We start with a simple lemma (a discrete version of this lemma was suggested by Bukh [Bu15]).

(3.4) Lemma. Let $f : \Omega \to \mathbb{C}$ be a random variable and let $0 \leq \theta < 2\pi/3$ be a real number such that $f(\omega) \neq 0$ for all $\omega \in \Omega$ and the angle between any two complex numbers $f(\omega_1) \neq 0$ and $f(\omega_2) \neq 0$ considered as vectors in $\mathbb{R}^2 = \mathbb{C}$ does not exceed $\theta$. Suppose further that $E |f| < +\infty$. Then

$$
|E f| \geq \left( \cos \frac{\theta}{2} \right) E |f|.
$$

Proof. First, we claim that 0 does not lie in the convex hull of vectors $f(\omega) \in \mathbb{C} = \mathbb{R}^2$. Otherwise we conclude by the Carathéodory Theorem that 0 is a convex combination of some 3 vectors $f(\omega_1)$, $f(\omega_2)$ and $f(\omega_3)$ and the angle between some two of them is at least $2\pi/3$, which is a contradiction. Hence the vectors $f(\omega)$ lie
in some convex cone (angle) $K \subset \mathbb{C}$ measuring at most $\theta$ and with vertex at 0. Let $L : \mathbb{R}^2 \to \mathbb{R}^2$ be the orthogonal projection onto the bisector of $K$. Then

$$|E f| \geq |L(E f)| = |E L(f)| = E |L(f)| \geq E \left( |f| \cos \frac{\theta}{2} \right) = \left( \cos \frac{\theta}{2} \right) E |f|.$$ 

Here the first (reading from left to right) inequality follows since the length of the orthogonal projection of a vector does not exceed the length of the vector, the next identity follows since $L$ is a linear operator, the next inequality follows since for all $z \in K$ the vectors $L(z)$ are non-negative multiples of each other, the next inequality follows since

$$|L(z)| \geq \left( \cos \frac{\theta}{2} \right) |z| \quad \text{for all} \quad z \in K$$

and the final identity follows since the expectation is a linear operator. \hfill \Box

(3.5) **Proof of Theorem 3.2.** For a function $f : \mathbb{T}^m \to \mathbb{C}$ and a subset $I \subset \{1, \ldots, m\}$, we denote by $E_I f$ the conditional expectation of $f$ obtained by integrating $f$ over the variables $z_i$ with $i \in I$. Hence if $f$ is a function of $z_1, \ldots, z_m$ and $I \subset \{1, \ldots, m\}$ then $h_I = E_I f$ is a function of $z_i$ for $i \notin I$. In particular, $h_I = f$ if $I = \emptyset$ and $h_I = E f$ if $I = \{1, \ldots, m\}$. If $I$ consists of a single element $i$, we write $E_i f$ instead of $E_{\{i\}} f$. We denote

$$I^c = \{1, \ldots, m\} \setminus I$$

the complement of $I$. We will consider functions $f = e^p$ where $p : \mathbb{T}^m \to \mathbb{C}$ is a Laurent polynomial.

For $0 \leq \theta_1, \ldots, \theta_m < 2\pi/3$, we denote by $P_m(\theta_1, \ldots, \theta_m)$ the set of $m$-variate Laurent polynomials $p$ for which the inequalities (3.2.1) hold. Note that the condition $p \in P_m(\theta_1, \ldots, \theta_m)$ is a finite system of linear inequalities for $|\gamma_a|, a \in A$.

Let us choose $p \in P_m(\theta_1, \ldots, \theta_m)$, let us fix some values $z_i \in \mathbb{S}^1$ for $I \subset \{1, \ldots, m\}$ and consider $p$ as a function of $z_i$ for $i \notin I$. It is not hard to see that $p \in P_{m-|I|}(\theta_i : i \notin I)$.

We prove by induction on $m$ the following statements.

**Statement 1**. For any $p \in P_m(\theta_1, \ldots, \theta_m)$, we have $E e^p \neq 0$. Moreover, suppose that $p, q \in P_m(\theta_1, \ldots, \theta_m)$ are two Laurent polynomials that differ in at most one monomial, so that the polynomial $p$ is obtained from $q$ by multiplying the coefficient $\gamma_b$ of some $2^b$ by some $\zeta \in \mathbb{S}^1$. Then the angle between $E e^p \neq 0$ and $E e^q \neq 0$ does not exceed

$$2|\gamma_b| \prod_{i \in \text{supp } b} \frac{1}{\cos(\theta_i/2)}.$$

**Statement 2**. Let $p \in P_m(\theta_1, \ldots, \theta_m)$ be a Laurent polynomial. Let $I = \{1, \ldots, m\} \setminus \{i\}$ for some $1 \leq i \leq m$ and let $h_I(z_i) = E_I e^p$. Then for any $z'_i, z''_i \in \mathbb{S}^1$, etc.
we have \( h_I(z'_i) \neq 0 \), \( h_I(z''_i) \neq 0 \) and the angle between the two complex numbers does not exceed \( \theta_i \).

We start by proving Statement 2₁. Then

\[
p(z) = \sum_{a \in A} \gamma_a z^a \text{ for some finite } A \subset \mathbb{Z}
\]

is a univariate Laurent polynomial. For any \( z \in S^1 \), we have

\[
\left| \arg e^{p(z)} \right| \leq |\Im p(z)| \leq |p(z)| \leq \sum_{a \in A} |\gamma_a| \leq \frac{1}{2} \theta_1
\]

and the result is immediate.

Next, we prove that Statements 2ₙ for \( s \leq m \) imply Statement 1ₘ.

Let us choose \( p \in P_m(\theta_1, \ldots, \theta_m) \). For a set \( I \subset \{1, \ldots, m\} \), let

\[
h_I(z_i : i \notin I) = E_{I} e^{p}.
\]

Assuming that \( I \neq \{1, \ldots, m\} \), let us pick an \( i \notin I \). Then

\[
h_{I \cup \{i\}} = E_i h_I.
\]

Let us fix variables \( z_j \in S^1 \) with \( j \notin I \cup \{i\} \) arbitrarily and consider \( p \) as a Laurent polynomial from \( P_r(\theta_k : k \in I \cup \{i\}) \) with \( r = |I| + 1 \). Thus \( h_I \) is a function of a single variable \( z_i \in S^1 \) and by Statement 2ᵣ for any two \( z'_i, z''_i \in S^1 \), the angle between \( h_I(z'_i) \neq 0 \) and \( h_I(z''_i) \neq 0 \) does not exceed \( \theta_i \). It follows from Lemma 3.4 that \( h_{I \cup \{i\}}(z_j : j \notin I \cup \{i\}) \neq 0 \) and, moreover,

\[
|h_{I \cup \{i\}}| = |E_i h_I| \geq \left( \prod_{j \in J} \cos \frac{\theta_j}{2} \right) E_i |h_I| > 0.
\]

Iterating, we obtain

\[
(3.5.1) \quad |h_{I \cup J}| = |E_J h_I| \geq \left( \prod_{j \in J} \cos \frac{\theta_j}{2} \right) E_J |h_I| > 0 \quad \text{provided } J \cap I = \emptyset.
\]

In particular, choosing \( J = \overline{I} \), we obtain that \( E e^{p} \neq 0 \).

Suppose now that \( p, q \in P_m(\theta_1, \ldots, \theta_m) \) where \( p \) is obtained from \( q \) by replacing a single monomial \( \gamma_b z^b \) by \( \gamma_b \zeta z^b \) for some \( \zeta \in S^1 \). Let us fix all the remaining coefficients of \( p \) and \( q \) and consider \( E e^{p} \) as a function of the coefficient \( \gamma_b \) of \( z^b \) as long as the resulting polynomial remains in \( P_m(\theta_1, \ldots, \theta_m) \) (note that the set of admissible values of \( |\gamma_b| \) is convex and includes 0). Since \( E e^{p} \neq 0 \) for all \( p \in \)
\( \mathcal{P}_m(\theta_1, \ldots, \theta_m) \), we can choose a continuous branch of \( \ln E e^p \) as a function of \( \gamma_b \). Then we have
\[
\frac{\partial}{\partial \gamma_b} \ln E e^p = \frac{\partial}{\partial \gamma_b} \frac{E e^p}{E} = \frac{E (z^b e^p)}{E e^p}.
\]
Let \( I = \text{supp} b \). Then
\[
|E (z^b e^p)| = |E_I E_T (z^b e^p)| = |E_I (z^b E_T e^p)| = |E_I z^b h_T| \leq E_I |h_T|.
\]
Similarly,
\[
|E e^p| = |E_I E_T e^p| = |E_I h_T| \geq \left( \prod_{i \in I} \cos \frac{\theta_i}{2} \right) E_I |h_T| > 0
\]
by (3.5.1). Therefore,
\[
\left| \frac{\partial}{\partial \gamma_b} \ln E e^p \right| \leq \prod_{i \in \text{supp} b} \frac{1}{\cos(\theta_i/2)}
\]
and hence
\[
|\ln E e^p - \ln E e^q| \leq 2 |\gamma_b| \prod_{i \in \text{supp} b} \frac{1}{\cos(\theta_i/2)}.
\]
Statement 1\(_m\) now follows.

Next, we prove that Statement 1\(_m\) implies Statement 2\(_{m+1}\).
Let \( p \in \mathcal{P}_{m+1}(\theta_1, \ldots, \theta_{m+1}) \) be a polynomial and let us choose an \( 1 \leq i \leq m+1 \). Let \( I = \{1, \ldots, m+1\} \setminus \{i\} \) and let \( h_I (z_i) = E_I e^p \). If we fix \( z_i \), we can consider \( p \) as a Laurent polynomial in \( \mathcal{P}_m(\theta_j : j \neq i) \). Moreover, when \( z_i' \) is changed to \( z_i'' \) only the coefficients \( \gamma_a \) with \( i \in \text{supp} a \) are affected, and each of those coefficients gets multiplied by some \( \zeta_a \in S^1 \). Repeatedly applying Statement 1\(_m\), we conclude that \( h_I (z_i') \neq 0, h_I (z_i'') \neq 0 \) and the angle between the two complex numbers does not exceed
\[
2 \sum_{a \in A: i \in \text{supp} a} |\gamma_a| \prod_{j \in \text{supp} a} \frac{1}{\cos(\theta_j/2)},
\]
which does not exceed \( \theta_i \) by the definition of \( \mathcal{P}_{m+1}(\theta_1, \ldots, \theta_{m+1}) \), and Statement 2\(_{m+1}\) follows.

This concludes the induction and proves that \( E e^p \neq 0 \). \( \square \)

(3.6) **Proof of Corollary 3.3.** Let us choose
\[
\theta_1 = \ldots = \theta_m = \frac{\delta}{\sqrt{c}}
\]
20
for some $0 < \delta < 2\pi/3$ to be determined later. To have the conditions of Theorem 3.2 satisfied, it suffices to have

$$2 \sum_{a \in A: \gamma_a \in \text{supp} \gamma_a} |\gamma_a| \left(\cos \frac{\delta}{2\sqrt{c}}\right)^c \leq \frac{\delta}{\sqrt{c}}$$

for $i = 1, \ldots, m$.

Since

$$\left(\cos \frac{\delta}{2\sqrt{c}}\right)^c \geq \cos \frac{\delta}{2}$$

for $0 \leq \delta \leq \pi$, see [Ba17], it suffices to have

$$\sum_{a \in A: \gamma_a \in \text{supp} \gamma_a} |\gamma_a| \leq \frac{\delta \cos(\delta/2)}{2\sqrt{c}}$$

for $i = 1, \ldots, m$.

Optimizing over $\delta$, we choose $\delta = 1.72$ and

$$\kappa = \frac{\delta \cos(\delta/2)}{2} \approx 0.561,$$

which concludes the proof.

\[4. \text{ Proofs of Theorems 1.2 and 1.6}\]

First, we prove Theorem 1.2.

Let $\mathbb{T}^m = S^1 \times \ldots \times S^1$ be the torus as in Section 3 and let us choose $\mu_i$ to be the rotation invariant (Haar) probability measure on the $i$-th copy of $S^1$. Let $\mu = \mu_1 \times \ldots \times \mu_m$ be the Haar probability measure on $\mathbb{T}^m$.

\[(4.1) \text{Lemma.} \text{ Let } X \subset \mathbb{Z}^n_+ \text{ be a set as in Theorem 1.2. Let } a_j, j = 1, \ldots, n \text{ be the columns of the matrix } A, \text{ considered as integer } m\text{-vectors and let us define a Laurent polynomial } q : \mathbb{T}^m \rightarrow \mathbb{C} \text{ by}

$$q(z_1, \ldots, z_m) = \prod_{j=1}^n \left(1 + w_j z^a_j + \ldots + w_j^{\nu_j} z^a_j\right),$$

where

$$z^a = z_1^{a_1} \cdots z_m^{a_m} \text{ provided } a = (a_1, \ldots, a_m).$$

Then

$$w(X) = EQ.$$
expanding the product that defines \( q \), we get

\[
E_q = \sum_{\xi_1, \ldots, \xi_n: \xi_1 a_1 + \cdots + \xi_n a_n = 0} w_{\xi_1} \cdots w_{\xi_n} = w(X).
\]

(4.2) Proof of Theorem 1.2.

Let \( q(z_1, \ldots, z_m) \) be the Laurent polynomial of Lemma 4.1, so that \( w(X) = E_q \).

Since for \( |w| < 1 \) we have

\[
1 + wz^a + \ldots + w^\nu z^\nu a = \frac{1 - w^{(\nu+1)} z^{(\nu+1)a}}{1 - wz^a},
\]

assuming that \( |w_j| < 1 \) for \( j = 1, \ldots, n \), we can write

\[
\ln q = \sum_{j=1}^n \ln \left(1 + w_j z^{a_j} + \ldots + w_j^{\nu_j} z^{\nu_j a_j}\right)
\]

\[
= \sum_{j=1}^n \ln \left(1 - w_j^{\nu_j+1} z^{(\nu_j+1)a_j}\right) - \sum_{j=1}^n \ln \left(1 - w_j z^{a_j}\right)
\]

\[
= \sum_{j=1}^n \sum_{k=1}^{\infty} \frac{w_j^k z^{ka_j}}{k} - \sum_{j=1}^n \sum_{k=1}^{\infty} \frac{w_j^{(\nu_j+1)k} z^{(\nu_j+1)ka_j}}{k}
\]

\[
= \sum_{j=1}^n \sum_{k=1}^{\infty} \beta_{jk} w_j^k z^{ka_j}
\]

for some real \( \beta_{jk} \) satisfying \( |\beta_{jk}| \leq 1 \).

For a positive integer \( N \), let us define a Laurent polynomial

\[
p_N(z_1, \ldots, z_m) = \sum_{j=1}^n \sum_{k=1}^N \beta_{jk} w_j^k z^{ka_j}
\]

which is just a truncation of the series expansion for \( \ln q \). Now, we observe that for every Laurent monomial \( z^a \) which appears in \( p_N \) with a coefficient \( \gamma_a \neq 0 \), we have

\[
|\text{supp } a| \leq c
\]

and that for \( i = 1, \ldots, m \), we have

\[
\sum_{a: \text{supp } a} |\gamma_a| \leq \sum_{j: \text{a}_{ij} \neq 0} \sum_{k=1}^N \frac{|w_j|^k}{k} \leq r \max_{j=1, \ldots, m} - \ln(1 - |w_j|) \leq 0.56 \frac{\sqrt{c}}{c}
\]

22
as long as
\[(4.2.1) \quad |w_j| \leq \frac{0.46}{r\sqrt{c}} \quad \text{for} \quad j = 1, \ldots, n.\]

(We use that $-\ln(1 - x) \leq 1.2x$ for $0 \leq x \leq 0.3$ and that $r \geq 2$.) Therefore, by Corollary 3.3, $E e^{\rho N} \neq 0$ as long as (4.2.1) holds. On the other hand, $E e^{\rho N}$ is an analytic function of $w_1, \ldots, w_n$ in the open polydisc (4.2.1) and $E e^{\rho N}$ converges to $E q$ uniformly on compact subsets of the polydisc. By the Hurwitz Theorem, see for example, Section 7.5 of [Kr92], we have either $E q \neq 0$ in the polydisc or $E q \equiv 0$ in the polydisc. Since for $w_1 = \ldots = w_n = 0$, we have $E q = 1$, we conclude that $E q \neq 0$ provided (4.2.1) holds. 

\[\square\]

(4.3) Proof of Theorem 1.6. We modify the choice of the probability measure $\mu$ on $\mathbb{T}^m$ as follows: we choose $\mu$ to be the uniform probability measure on the roots of unity of degree $\kappa$ and let $\mu = \mu_1 \times \ldots \times \mu_m$. We note that for $a \in \mathbb{Z}^m$, $a = (\alpha_1, \ldots, \alpha_m)$, we have
\[
\mathbb{E} z^a = \begin{cases} 
1 & \text{if } \alpha_i \equiv 0 \mod \kappa \quad \text{for } i = 1, \ldots, m, \\
0 & \text{otherwise.}
\end{cases}
\]

Given an $m \times n$ integer matrix $A = (a_{ij})$, we define $q(z_1, \ldots, z_m)$ by
\[
q(z_1, \ldots, z_m) = \prod_{j=1}^{n} \left(1 + w_j z^{a_j} + \ldots + w_j z^{(\kappa-1)a_j}\right).
\]

Then
\[
\mathbb{E} q = \sum_{\substack{\xi_1, \ldots, \xi_n: \\
s.t. \xi_1 + \ldots + \xi_n \equiv 0 \mod \kappa \\
\xi_j \neq 0 \quad \text{for } j = 1, \ldots, m, \\
0 \leq \xi_j \leq \kappa - 1 \quad \text{for } j = 1, \ldots, n}} \prod_{j: \xi_j \neq 0} w_j = w(X).
\]

Assuming that $|w_j| < (\kappa - 1)^{-1}$ for $j = 1, \ldots, n$, we expand
\[
\ln q = \sum_{j=1}^{n} \ln \left(1 + w_j z^{a_j} + \ldots + w_j z^{(\kappa-1)a_j}\right)
= \sum_{j=1}^{n} \sum_{s=1}^{\infty} (-1)^s \frac{(w_j z^{a_j} + \ldots + w_j z^{(\kappa-1)a_j})^s}{s}.
\]

For a positive integer $N$, let us define a Laurent polynomial
\[
p_N(z_1, \ldots, z_m) = \sum_{j=1}^{n} \sum_{s=1}^{N} (-1)^s \frac{(w_j z^{a_j} + \ldots + w_j z^{(\kappa-1)a_j})^s}{s}.
\]
For every Laurent monomial $z^a$ which appears in $p_N$ with a coefficient $\gamma_a \neq 0$, we have $|\text{supp } a| \leq c$ and for $i = 1, \ldots, m$, we have

$$\sum_{a: i \in \text{supp } a} |\gamma_a| \leq \sum_{j: a_{ij} \neq 0} \sum_{s=1}^{N} \frac{((\kappa - 1)|w_j|^s}{s} \leq r \max_{j=1, \ldots, n} -\ln(1 - (\kappa - 1)|w_j|) \leq \frac{0.56}{\sqrt{c}}$$

provided

$$|w_j| \leq \frac{0.46}{(\kappa - 1)r\sqrt{c}} \quad \text{for} \quad j = 1, \ldots, n.$$

The proof is then concluded as in Section 4.2. $\square$

5. Approximating $w(X)$ faster

Let $X \subset Z^n_+$ be the set defined in Theorem 1.2. We assume that $A$ has no zero rows or columns, see Section 1.3. Recall that $r \geq 2$ is an upper bound on the number of non-zero entries in a row of $A$ and $c \geq 1$ is an upper bound on the number of non-zero entries in a column of $A$. As in Section 1.3, we define a univariate polynomial $w(X; \zeta)$, that is the weight of the set $X$ under the scaled weights $\zeta w_1, \ldots, \zeta w_n$, so $w(X; \zeta)$ is a polynomial of some degree $d \leq N = \nu_1 + \ldots + \nu_n$. We let $f(\zeta) = \ln w(X; \zeta)$ for $\zeta$ in a neighborhood of 0.

Our goal is to show that the term $f^{(k)}(0)$ in the Taylor expansion (1.3.3) can be computed in $n(r c)^{O(k)}$ time, where we assume the standard RAM machine model with logarithmic-sized words, and additionally we assume that given a column index $j$ of the matrix $A = (a_{ij})$ we can in time $O(c)$ compute the row indices $i$ such that $a_{ij} \neq 0$ (otherwise the running time is bounded by $n m(r c)^{O(k)}$). We note that in this section, all the implied constants in the “$O$” notation are absolute. In particular, if $k = O(\ln N - \ln \epsilon)$ as in Section 1.3, and $r$ and $c$ are fixed beforehand we obtain an algorithm of a polynomial in $N/\epsilon$ complexity.

Our algorithm heavily relies on the ideas of [PR17a], see also [L+17].

(5.1) The idea of the algorithm. Since $w(X; 0) = 1$, we can write

$$w(X; \zeta) = \prod_{i=1}^{d} \left(1 - \frac{\zeta}{\zeta_i}\right),$$

where $\zeta_1, \ldots, \zeta_d \neq 0$ for some $d \leq N$ are the roots if $w(X; \zeta)$, listed with multiplicity. Then

$$f(\zeta) = \sum_{i=1}^{d} \ln \left(\frac{1 - \zeta}{\zeta_i}\right)$$
and
\[ \frac{f^{(k)}(0)}{k!} = -\frac{1}{k} \sum_{i=1}^{d} \zeta_i^{-k}. \]

We introduce the power sums
\[ \sigma_k(A, w, \nu) = \zeta_1^{-k} + \ldots + \zeta_d^{-k}. \]

Hence our goal is to compute \( \sigma_k(A, w, \nu) \) in \( n(rc)^{O(k)} \) time.

The crucial feature of the power sums \( \sigma_k(A, w, \nu) \) is that they are additive functions of \( A \) as is explained below.

In what follows, we consider the set \( M \) of integer matrices \( A \) with rows and columns indexed by non-empty finite subsets of the set \( \mathbb{N} \) of positive integers and without zero rows or columns. For non-empty finite subsets \( R, C \subset \mathbb{N} \), an \( R \times C \) integer-valued matrix \( A \in M \) is a function \( A : R \times C \to \mathbb{Z} \) and we write the \((i, j)\)-th entry of \( A \) as \( A(i, j) \) for \( i \in R \) and \( j \in C \). We fix complex weights \( w_j \) and positive integers \( \nu_j \) for all \( j \in \mathbb{N} \) and define
\[ (5.1.2) \quad X_A = \left\{ (\xi_j : j \in C) : \sum_{j \in C} A(i, j)\xi_j = 0 \quad \text{for } i \in R \right\} \]
\[ \xi_j \in \mathbb{Z}_+ \quad \text{and} \quad 0 \leq \xi_j \leq \nu_j \quad \text{for } j \in C \].

Similarly, we define univariate polynomials
\[ (5.1.3) \quad w(X_A; \zeta) = \sum_{x \in X_A} \prod_{j \in C} (\zeta w_j)^{\xi_j} \]
and define power sums \( \sigma_k(A, w, \nu) \) by \( (5.1.1) \) where \( \zeta_1, \ldots, \zeta_d \) are the roots of \( w(X_A; \zeta) \), listed with multiplicity.

Let \( A_1, A_2 \in M \) be respectively \( R_1 \times C_1 \) and \( R_2 \times C_2 \) matrices. Suppose that \( R_1 \cap R_2 = \emptyset \) and \( C_1 \cap C_2 = \emptyset \). We define the direct sum \( A = A_1 \oplus A_2 \) as the \( R \times C \) matrix, where \( R = R_1 \cup R_2 \), \( C = C_1 \cup C_2 \) and
\[ A(i, j) = \begin{cases} A_1(i, j) & \text{if } i \in R_1 \text{ and } j \in C_1, \\ A_2(i, j) & \text{if } i \in A_2 \text{ and } j \in C_2, \\ 0 & \text{elsewhere}. \end{cases} \]

Clearly, \( A \in M \).

Let \( A_1, A_2 \in M \) be matrices such that \( A = A_1 \oplus A_2 \) is defined. We observe that
\[ w(X_A; \zeta) = w(X_{A_1}; \zeta) w(X_{A_2}; \zeta) \]
and hence
\[ (5.1.4) \quad \sigma_k(A_1 \oplus A_2, w, \nu) = \sigma_k(A_1, w, \nu) + \sigma_k(A_2, w, \nu). \]
Given an \( R \times C \) matrix \( A \in \mathcal{M} \) and an \( R_1 \times C_1 \) matrix \( B \in \mathcal{M} \), we define the index \( \text{ind}(B, A) = 1 \) if \( R_1 \subset R \), \( C_1 \subset C \),

\[
A(i, j) = B(i, j) \quad \text{for all} \quad i \in R_1 \quad \text{and all} \quad j \in C_1
\]

and

\[
A(i, j) = 0 \quad \text{for all} \quad i \in R \setminus R_1 \quad \text{and all} \quad j \in C_1.
\]

Otherwise, we say that \( \text{ind}(B, A) = 0 \).

We define a filtration

\[
\mathcal{M}_1 \subset \mathcal{M}_2 \subset \ldots \subset \mathcal{M}_k \subset \ldots,
\]

where \( \mathcal{M}_k \subset \mathcal{M} \) consists of the matrices with at most \( k \) columns.

In Lemma 5.3 below we show that we can write

\[
(5.1.5) \quad \sigma_k(A; w, \nu) = \sum_{B \in \mathcal{M}_k} \text{ind}(B, A)\mu_k(B, w, \nu) \quad \text{for all} \quad A \in \mathcal{M}
\]

and some complex numbers \( \mu_k(B, w, \nu) \). Although the sum in (5.1.5) contains infinitely many terms, for each \( A \in \mathcal{M} \), only finitely many terms are non-zero, so (5.1.5) is well-defined.

We say that a matrix \( B \in \mathcal{M} \) is connected if it cannot be represented as a direct sum \( B = B_1 \oplus B_2 \) for some matrices \( B_1, B_2 \in \mathcal{M} \) and disconnected otherwise. In Corollary 5.5 below, we deduce from the additivity property (5.1.4) that \( \mu_k(B, w, \nu) = 0 \) in (5.1.5) unless \( B \) is connected. In Section 5.6 we show for any given \( m \times n \) matrix \( A \) with at most \( r \) non-zero entries in each row and at most \( c \) non-zero entries in each column the number of connected matrices \( B \in \mathcal{M}_k \) with \( \text{ind}(B, A) = 1 \) is at most \( n(rc)^{O(k)} \) and that all such matrices \( B \) can be found in \( n(rc)^{O(k)} \) time. Finally, in Section 5.7 we show that for each connected \( B \in \mathcal{M}_k \), one can compute \( \mu_k(B, w, \nu) \) in \( cn2^{O(k)} \) time. This produces an algorithm of \( n(rc)^{O(k)} \) complexity for computing \( \sigma_k(A, w, \nu) \).

Next, we supply the necessary details. We start with a technical result describing how the function \( \text{ind}(B, \cdot) \) behaves under multiplication. Let \( B_1 \in \mathcal{M} \) be an \( R_1 \times C_1 \) matrix and let \( B_2 \in \mathcal{M} \) be an \( R_2 \times C_2 \) matrix. If the restrictions of \( B_1 \) and \( B_2 \) onto \((R_1 \cap R_2) \times (C_1 \cap C_2)\) coincide, we define the connected sum \( B = B_1 \# B_2 \), \( B \in \mathcal{M} \), as the \((R_1 \cup R_2) \times (C_1 \cup C_2)\) matrix such that

\[
B(i, j) = \begin{cases} 
B_1(i, j) & \text{if } i \in R_1 \text{ and } j \in C_1, \\
B_2(i, j) & \text{if } i \in R_2 \text{ and } j \in C_2, \\
0 & \text{otherwise.}
\end{cases}
\]

In particular, if \( R_1 \cap R_2 = \emptyset \) and \( C_1 \cap C_2 = \emptyset \) then \( B_1 \# B_2 = B_1 \oplus B_2 \) is the direct sum of \( B_1 \) and \( B_2 \).
**Lemma 5.2.** Let $B_1 \in \mathcal{M}$ be an $R_1 \times C_1$ matrix and let $B_2 \in \mathcal{M}$ be an $R_2 \times C_2$ matrix.

Suppose that the following conditions (1) – (3) are satisfied:

1. For all $i \in R_1 \cap R_2$ and all $j \in C_1 \cap C_2$ we have $B_1(i, j) = B_2(i, j)$;
2. For all $i \in R_1 \setminus R_2$ and all $j \in C_1 \cap C_2$ we have $B_1(i, j) = 0$;
3. For all $i \in R_2 \setminus R_1$ and all $j \in C_1 \cap C_2$ we have $B_2(i, j) = 0$.

Then $B = B_1 \# B_2$ is defined and

$$\text{ind}(B_1, A) \text{ind}(B_2, A) = \text{ind}(B, A) \quad \text{for all} \quad A \in \mathcal{M}.$$ 

If any of the conditions (1)–(3) is violated then

$$\text{ind}(B_1, A) \text{ind}(B_2, A) = 0 \quad \text{for all} \quad A \in \mathcal{M}.$$ 

**Proof.** Clearly, if (1) is violated then $\text{ind}(B_1, A) \text{ind}(B_2, A) = 0$ for all $A \in \mathcal{M}$.

Suppose that (2) is violated. If $\text{ind}(B_2, A) = 1$ then $A(i, j) = 0$ for all $i \in R_1 \setminus R_2$ and all $j \in C_1 \cap C_2$ and hence $\text{ind}(B_1, A) = 0$ so that $\text{ind}(B_1, A) \text{ind}(B_2, A) = 0$. Similarly, if (3) is violated then $\text{ind}(B_1, A) \text{ind}(B_2, A) = 0$ for all $A \in \mathcal{M}$.

Hence it remains to consider the case when (1)–(3) hold. Without loss of generality we assume that $R_1 \cup R_2$ is a subset of the rows of $A$ and that $C_1 \cup C_2$ is a subset of the columns of $A$.

If $\text{ind}(B_1, A) = 0$ for some $A \in \mathcal{M}$ then either $B_1(i, j) \neq A(i, j)$ for some $i \in R_1$ and some $j \in C_1$ or $A(i, j) \neq 0$ for some $i \notin R_1$ and some $j \in C_1$. In either case $\text{ind}(B, A) = 0$. Similarly, if $\text{ind}(B_2, A) = 0$ then $\text{ind}(B, A) = 0$. If $\text{ind}(B_1, A) = \text{ind}(B_2, A) = 1$ then $B(i, j) = A(i, j)$ for all $i \in R_1 \cup R_2$ and all $j \in C_1 \cup C_2$ while $A(i, j) = 0$ for all $i \notin R_1 \cup R_2$ and all $j \in C_1 \cup C_2$ and hence $\text{ind}(B, A) = 1$ as well. \[\square\]

If the conditions (1)–(3) of Lemma 5.2 are satisfied, we say that the matrices $B_1$ and $B_2$ are **compatible** and denote it $B_1 \sim B_2$. Now we are ready to prove the existence of a decomposition (5.1.5).

**Lemma 5.3.** For a positive integer $k$ and a matrix $B \in \mathcal{M}_k$ one can define complex numbers $\mu_k(B, w, \nu)$ so that (5.1.5) holds for all $A \in \mathcal{M}$.

**Proof.** We write the polynomial (5.1.3) in the monomial basis. Assuming that $A$ is an $R \times C$ matrix, we have

$$w(X_A; \zeta) = 1 + \sum_{k=1}^{N} \pi_k(A, w, \nu)\zeta^k$$

where

$$\pi_k(A, w, \nu) = \sum_{x=(\xi_j: j \in C): j \in C \atop x \in X_A} \prod_{j \in C} w_{\xi_j}^{\xi_j},$$

and

$$\sum_{j \in C} \xi_j = k.$$
where \( X_A \) is defined by (5.1.2) and \( N = \sum_{j \in C} \nu_j \).

We say that a set \( S \subset C \) is the support of a vector \( x \in X_A \), \( x = (\xi_j : j \in C) \), provided \( \xi_j \neq 0 \) if and only if \( j \in S \). Clearly, the support of a vector \( x \) contributing to \( \pi_k(A, w, \nu) \) is a set \( S \subset C \) satisfying \( |S| \leq k \) and the vector \( x_S = (\xi_j : j \in S) \) satisfies \( A_S x_S = 0 \), where \( A_S \) is the \( R \times S \) matrix consisting of the columns of \( A \) with indices in \( S \).

This allows us to write

\[
\pi_k(A, w, \nu) = \sum_{B \in \mathcal{M}_k} \text{ind}(B, A) \lambda_k(B, w, \nu)
\]

where for \( R_1 \times C_1 \) matrix \( B \) we have

\[
\lambda_k(B, w, \nu) = \sum_{x=(\xi_j : j \in C_1)} \prod_{j \in C_1} w_j^{\xi_j},
\]

subject to \( x \in X_B \), support of \( x \) is \( C_1 \), \( \sum_{j \in C_1} \xi_j = k \).

Although formally the sum (5.3.1) is infinite, for each \( A \in \mathcal{M} \) we have \( \text{ind}(B, A) \neq 0 \) for only finitely many \( B \in \mathcal{M} \), so (5.3.1) is well-defined.

We observe that

\[
\pi_k(A, w, \nu) = (-1)^k e_k \left( \zeta_1^{-1}, \ldots, \zeta_d^{-1} \right),
\]

where \( e_k \) is the \( k \)-th elementary symmetric function and \( \zeta_1, \ldots, \zeta_d \) are the roots of \( w(X_A; \zeta) \), listed with their multiplicities (recall that the constant term of \( w(X_A; \zeta) \) is 1). Therefore, the Newton identities imply that

\[
k \pi_k(A, w, \nu) = - \sum_{i=1}^{k} \pi_{k-i}(A, w, \nu) \sigma_i(A, w, \nu) \quad \text{for all } k \geq 1,
\]

where we define

\[
\pi_0(A, w, \nu) = 1.
\]

We define

\[
\mu_1(B, w, \nu) = -\lambda_1(B, w, \nu) \quad \text{for } B \in \mathcal{M}_1.
\]

Assuming that \( \mu_i(B, w, \nu) \) are defined for \( B \in \mathcal{M}_i \) and \( i = 1, \ldots, k-1 \), for \( k \geq 2 \) we define for \( B \in \mathcal{M}_k \)

\[
\mu_k(B, w, \nu) = - k \lambda_k(B, w, \nu)
\]

\[
- \sum_{\substack{B_1 \in \mathcal{M}_{k-1}, B_2 \in \mathcal{M}_i \\text{for } 1 \leq i \leq k-1: \\ B_1 \sim B_2 \text{ and } B_1 \# B_2 = B}} \lambda_{k-i}(B_1, w, \nu) \mu_i(B_2, w, \nu).
\]

28
Here the sum is taken over all distinct ordered pairs of compatible matrices \((B_1, B_2)\) such that \(B_1 \# B_2 = B\) (in particular, we may have \(B_1 = B_2\)). We observe that for each \(B\) the sum contains only finitely many terms, so \(\mu_k(B, w, \nu)\) is well-defined. The identity (5.1.5) now follows from (5.3.1), (5.3.3) and Lemma 5.2.

Our next goal is to show that in (5.1.5) we have \(\mu_k(B, w, \nu) \neq 0\) only for connected matrices \(B\). We start with a general structural result, very similar in spirit to Lemma 4.2 of [CS16], see also [PR17a].

(5.4) Lemma. Let us consider a function \(f : \mathcal{M} \rightarrow \mathbb{C}\) defined by

\[
f(A) = \sum_{B \in \mathcal{S}} \mu_B \text{ind}(B, A),
\]

where \(\mathcal{S} \subset \mathcal{M}\) is a (possibly infinite) set and \(\mu_B \in \mathbb{C} \setminus \{0\}\) for all \(B \in \mathcal{S}\) (for each \(A \in \mathcal{M}\) only finitely many summands are non-zero, so \(f\) is well-defined). Suppose that

\[
f(A_1 \oplus A_2) = f(A_1) + f(A_2)
\]

for any two matrices \(A_1, A_2 \in \mathcal{M}\) such that \(A_1 \oplus A_2\) is defined. Then each \(B \in \mathcal{S}\) is connected, that is, cannot be written as \(B = B_1 \oplus B_2\) for some \(B_1, B_2 \in \mathcal{M}\).

Proof. Seeking a contradiction, assume that there is a disconnected \(B \in \mathcal{S}\). We observe that if \(B \in \mathcal{M}\) is connected, then

\[
\text{ind}(B, A_1 \oplus A_2) = \text{ind}(B, A_1) + \text{ind}(B, A_2)
\]

for any two \(A_1, A_2 \in \mathcal{M}\) such that \(A_1 \oplus A_2\) is defined (since \(B\) is connected, we cannot have \(\text{ind}(B, A_1) = \text{ind}(B, A_2) = 1\) provided \(A_1 \oplus A_2\) is defined). Therefore, without loss of generality, we assume that all \(B \in \mathcal{S}\) are disconnected. Let us choose a \(D \in \mathcal{S}\) that has the smallest number of columns. Hence we have \(D = D_1 \oplus D_2\) for some \(D_1\) and \(D_2\). Then

\[
f(D_i) = \sum_{B \in \mathcal{S}} \mu_B \text{ind}(B, D_i) = 0 \quad \text{for} \quad i = 1, 2
\]

since \(D_i\) has fewer columns than any matrix \(B \in \mathcal{S}\). Therefore,

\[
f(D) = f(D_1) + f(D_2) = 0.
\]

On the other hand, \(\text{ind}(B, D) = 0\) for all \(B \in \mathcal{S} \setminus \{D\}\) and

\[
f(D) = \lambda_D \text{ind}(D, D) = \lambda_D \neq 0,
\]

which is contradiction. \(\square\)
Corollary. In the expansion (5.1.5), we have \( \mu_k(B, w, \nu) = 0 \) whenever \( B \) is disconnected.

Proof. Follows by (5.1.4) and Lemma 5.4.

Enumerating connected matrices. Given an integer \( k \geq 1 \) and an \( R \times C \) matrix \( A \in \mathcal{M} \) with at most \( r \) non-zero entries in each row and at most \( c \) non-zero entries in each column, we want to compile a list of all connected matrices \( B \in \mathcal{M}_k \) such that \( \text{ind}(B, A) = 1 \). First, we observe that an \( R_1 \times C_1 \) matrix \( B \in \mathcal{M} \) such that \( \text{ind}(B, A) = 1 \) is uniquely determined by its set of columns \( C_1 \subset C \) since \( R_1 \subset R \) is then the set of rows of \( A \) whose restriction onto \( C_1 \) are not zero.

We define a graph \( G = (C, E) \). The vertices of \( G \) are the columns of \( A \) and two vertices \( c_1 \) and \( c_2 \) span an edge of \( G \) if and only if there is a row of \( A \) with non-zero entries in columns \( c_1 \) and \( c_2 \). We note that the degree of each vertex of \( G \) does not exceed \( d = rc \). To enumerate connected matrices \( B \in \mathcal{M}_k \) such that \( \text{ind}(B, A) = 1 \) is to enumerate sets of vertices of cardinality at most \( k \) in \( C \) that induce a connected subgraph of \( G \). This is done as in [PR17b]. The crucial observation is that as long as one vertex \( c \) is chosen, there are at most

\[
\binom{k}{d}^{-1} \frac{(ed)^{k-1}}{2},
\]

connected induced subgraphs with \( k \geq 2 \) vertices containing \( c \), see Lemma 2.1 of [B+13]. Consequently, there are \( n d^{O(k)} \) induced connected subgraphs with at most \( k \) vertices in \( G \). Once the vertex \( c \) is chosen, the subgraphs are enumerated with \( d^{O(k)} \) complexity, by successively exploring adjacent vertices, see [PR17b] for details.

Summary of the algorithm. Given an \( m \times n \) matrix \( A \) without zero rows and columns, we interpret it as an \( R \times C \) matrix \( A \in \mathcal{M} \), where \( R = \{1, \ldots , m\} \) and \( C = \{1, \ldots , n\} \). Given a positive integer \( k \), as in Section 5.6 we compile a list \( \mathcal{C} \) of all connected matrices \( B \in \mathcal{M}_k \) such that \( \text{ind}(B, A) = 1 \). We define the filtration

\[
\mathcal{C}_1 \subset \mathcal{C}_2 \subset \ldots \subset \mathcal{C}_{k-1} \subset \mathcal{C}_k = \mathcal{C},
\]

where \( \mathcal{C}_i \) is the set of matrices \( B \in \mathcal{C} \) with at most \( i \) columns.

Given complex numbers \( w_1, \ldots , w_n \) and positive integers \( \nu_1, \ldots , \nu_n \), by (5.3.2) we compute \( \lambda_1(B, w, \nu) \) for all \( B \in \mathcal{C}_1 \) in \( O(cn) \) time (we only need to check the nonzero rows of \( B \)). We let

\[
\mu_1(B, w, \nu) = -\lambda_1(B, w, \nu) \quad \text{for} \quad B \in \mathcal{C}_1,
\]

cf. the proof of Corollary 5.3.

Suppose that we have computed \( \mu_i(B, w, \nu) \) for \( i = 1, \ldots , k-1 \) and all \( B \in \mathcal{C}_{k-1} \) for \( k \geq 2 \). To compute \( \mu_k(B, w, \nu) \) for all \( B \in \mathcal{C}_k \), we use formula (5.3.4). Since every matrix \( B \in \mathcal{C}_k \) has at most \( k \) columns, there are not more than \( 4^k \) pairs
of matrices $B_1 \in \mathcal{M}_{k-i}$ and $B_2 \in \mathcal{M}_i$ such that $B_1 \# B_2 = B$ and all such pairs can be found by inspection in $O(4^k)$ time. We then use (5.3.2) to compute the terms $\lambda_{k-i}(B_1, w, \nu)$ for $i = 0, \ldots, k - 1$. We note that there are $(k-i-1) \leq 2^{k-i}$ non-negative integer vectors with support $S$ and the sum $k-i$ of the coordinates, so each $\lambda_{k-i}(B_1, w, \nu)$ is computed in $c(k-i)n2^{O(k-i)}$ time.

This gives us the list of values $\mu_k(B, w, \nu)$ for all $B \in \mathcal{C}_k$. We then compute

$$\sigma_k(A, w, \nu) = \sum_{B \in \mathcal{C}_k} \mu_k(B, w, \nu),$$

as desired.

ACKNOWLEDGMENTS

The authors are grateful to Martin Dyer for pointing out to possible connections with [Va79] and [VV86], to Alex Samorodnitsky for pointing out to possible connections with low-density parity-check codes and to Prasad Tetali for pointing us to [BS94].

REFERENCES


Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1043, USA
E-mail address: barvinok@umich.edu

Korteweg de Vries Institute for Mathematics, University of Amsterdam
E-mail address: guusregts@gmail.com