Brégman's theorem and extensions

David Galvin*

December 17, 2011

Abstract

Minc conjectured, and Brégman proved, a sharp upper bound on the permanent of an n by n 0-1 matrix with given row sums (equivalently, on the number of perfect matchings in a bipartite graph with each partition class having size n and with fixed degree sequence for one of the two classes). Here we present Radhakrishnan's entropy proof of Brégman's theorem, and Alon and Friedland's proof of an analogous statement for graphs that are not necessarily bipartite. We also discuss progress towards the Upper Matching conjecture of Friedland, Krop and Markström, which extends Brégman's theorem to arbitrary matchings.

1 Introduction

The *permanent* of an n by n matrix $A = (a_{ij})$ is

$$\mathrm{perm}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i\sigma(i)}$$

where S_n is the set of permutations of $[n] := \{1, \ldots, n\}$. This seems superficially quite similar to the determinant, which differs only by the addition of a factor of $(-1)^{\mathrm{sgn}(\sigma)}$ in front of the product. This small difference makes all the difference, however: problems involving the determinant are generally quite tractable algorithmically (because Gaussian elimination can be performed efficiently), but permanent problems seems to be quite intractable (in particular, by a Theorem of Valiant the computation of the permanent of a general n by n matrix in $\#P$-hard).

The permanent of a 0-1 matrix has a nice interpretation in terms of perfect matchings in a graph. There is a natural one-to-one correspondence between 0-1 n by n matrices and bipartite graphs on a fixed bipartition classes each of size n: Given $A = (a_{ij})$ we construct a bipartite graph $G = G(A)$ on bipartition classes $\mathcal{E} = \{v_1, \ldots, v_n\}$ and $\mathcal{O} = \{w_1, \ldots, w_n\}$ by putting $v_i w_j \in E$ if and only if $a_{ij} = 1$. Each $\sigma \in S_n$ that contributes to $\mathrm{perm}(A)$ gives

*dgalvin1@nd.edu; Department of Mathematics, University of Notre Dame, Notre Dame IN 46556. Supported in part by National Security Agency grant H98230-10-1-0364.

rise to the perfect matching (1-regular spanning subgraph) $\{v_i w_{\sigma(i)} : i \in [n]\}$ in G, and this correspondence is bijective. In other words,

$$\text{perm}(A) = |\mathcal{M}_{\text{perfect}}(G)|$$

where $\mathcal{M}_{\text{perfect}}(G)$ is the set of perfect matchings of G.

In 1963 Minc formulated a natural conjecture concerning the permanent of an n by n 0-1 matrix with all row sums fixed. Ten years later Brégman [3] gave the first proof, and the result is now known as Brégman's theorem.

Theorem 1.1 *(Brégman's theorem) Let n non-negative integers $d_1, \ldots, d_n$ by given. Let $A = (a_{ij})$ be an n by n matrix with all entries in $\{0,1\}$ and with $\sum_{j=1}^n a_{ij} = d_i$ for each $i = 1, \ldots, n$ (that is, with the sum of the row i entries of A being d_i, for each i). Then*

$$\text{perm}(A) \leq \prod_{i=1}^{n} (d_i!)^{\frac{1}{d_i}} .$$

Equivalently, let G be a bipartite graph on bipartition classes $\mathcal{E} = \{v_1, \ldots, v_n\}$ and $\mathcal{O} = \{w_1, \ldots, w_n\}$ with each $v_i \in \mathcal{E}$ having degree d_i. Then

$$|\mathcal{M}_{\text{perfect}}(G)| \leq \prod_{i=1}^{n} (d_i!)^{\frac{1}{d_i}} .$$

Notice that the bound is tight: for example, for each fixed d and n with $d|n$, it is achieved by the matrix consisting of n/d blocks down the diagonal with each block being a d by d matrix of all 1's, with zeros everywhere else (or equivalently, by the graph made up of the disjoint union of n/d copies of $K_{d,d}$, the complete bipartite graph with d vertices in each classes).

A short proof of Brégman's theorem was given by Schrijver [12], and a probabilistic reinterpretation of Schrijver's proof was given by Alon and Spencer [2]. In Section 2 we present Radhakrishnan's entropy proof [11].

Brégman's theorem concerns perfect matchings in a bipartite graph. A natural question to ask is what happens in a general (not necessarily bipartite) graph? Kahn and Lovász answered this question.

Theorem 1.2 *(Kahn-Lovász theorem) Let G be a graph on $2n$ vertices $v_1, \ldots, v_{2n}$ with each v_i having degree d_i. Then*

$$|\mathcal{M}_{\text{perfect}}(G)| \leq \prod_{i=1}^{2n} (d_i!)^{\frac{1}{2d_i}} .$$

Notice that this result is also tight: for example, for each fixed d and n with $d|n$, it is achieved by the graph made up of the disjoint union of n/d copies of $K_{d,d}$. Note also that there is no permanent version of this result.

Kahn and Lovász did not publish their proof. Since they first discovered the theorem, it has been rediscovered/reproved a number of times: by Alon and Friedland [1], Cutler and Radcliffe [5], Egorychev [6] and Friedland [7]. Alon and Friedland's is a "book" proof, observing that the theorem is an easy consequence of Brégman's theorem. We present the details in Section 3.

Another direction in which we may consider extending Brégman's theorem is to consider arbitrary matchings in G, rather than perfect matchings. For this discussion, we focus exclusively on the case of d-regular G on $2n$ vertices. Writing $K(n,d)$ for the disjoint union of n/d copies of $K_{d,d}$, we can restate Brégman's theorem as

$$\text{for bipartite } d\text{-regular } G \text{ on } 2n \text{ vertices,} \quad |\mathcal{M}_{\text{perfect}}(G)| \leq d!^{\frac{n}{d}} = |\mathcal{M}_{\text{perfect}}(K(n,d))|$$

and the Kahn-Lovász theorem as

$$\text{for arbitrary } d\text{-regular } G \text{ on } 2n \text{ vertices,} \quad |\mathcal{M}_{\text{perfect}}(G)| \leq d!^{\frac{2n}{2d}} = |\mathcal{M}_{\text{perfect}}(K(n,d))|.$$

Do these inequalities continue to hold if we replace $\mathcal{M}_{\text{perfect}}(G)$ with $\mathcal{M}(G)$, the collection of all matchings (not necessarily perfect) in G?

Conjecture 1.3 *For bipartite d-regular G on $2n$ vertices (or for arbitrary d-regular G on $2n$ vertices),*

$$|\mathcal{M}(G)| \leq |\mathcal{M}(K(n,d))|.$$

Here and in Conjecture 1.4 below, the heart of the matter is the bipartite case: the methods of Alon and Friedland discussed in Section 3 can be modified to show that the bipartite case implies the general case.

Friedland, Krop and Markström [8] have proposed an even stronger conjecture, the Upper Matching conjecture. For each $0 \leq t \leq n$, write $\mathcal{M}_t(G)$ for the number of matchings in G of size t (that is, with t edges).

Conjecture 1.4 *(Upper Matching conjecture) For bipartite d-regular G on $2n$ vertices (or for arbitrary d-regular G on $2n$ vertices), and for all $0 \leq t \leq n$,*

$$|\mathcal{M}_t(G)| \leq |\mathcal{M}_t(K(n,d))|.$$

For $t = n$ this is Brégman's theorem (in the bipartite case) and the Kahn-Lovász theorem (in the general case). For $t = 0, 1$ and 2 it is trivial in both cases. Friedland, Krop and Markström [8] have verified the conjecture (in the bipartite case) for $t = 3$ and 4. For $t = \alpha n$ for $\alpha \in [0,1]$, asymptotic evidence in favor of the conjecture was provided first by Carroll, Galvin and Tetali [4] and then (in a stronger form) by Kahn and Ilinca [10]. We discuss these results in Section 4.

2 Radhakrishnan's proof of Brégman

A perfect matching M in G may be viewed as a permutation σ of $\{1,\ldots,n\}$ via $\sigma(i) = j$ if and only if $v_i w_j \in M$. This is how we will view matchings from now on. Radhakrishnan's first idea is to set up a random variable X which represents the selection of a matching σ from $\mathcal{M}_{\text{perfect}}(G)$, the set of all perfect matchings in G, with all such σ equally likely. By one of the basic properties of entropy, we have

$$H(X) = \log |\mathcal{M}_{\text{perfect}}(G)|,$$

and so to prove the conjecture we need to upper bound

$$H(X) \le \sum_{i=1}^{n} \frac{\log d_i!}{d_i}. \tag{1}$$

As a first attempt, let's view X as a random vector $(\sigma(1), \ldots, \sigma(n))$. By subadditivity we have

$$H(X) \le \sum_{i=1}^{n} H(\sigma(i)).$$

Since there are at most d_i possibilities for the value of $\sigma(i)$, we have $H(\sigma(i)) \le \log d_i$ for all i, and so

$$H(X) \le \sum_{i=1}^{n} \log d_i.$$

This falls somewhat short of (1), since $(\log d_i!)/d_i \approx \log(d_i/e)$ by Stirling.

We might try to improve things by using the sharper chain rule in place of subadditivity:

$$H(X) = \sum_{i=1}^{n} H(\sigma(i)|\sigma(1), \ldots, \sigma(i-1)).$$

Now instead of naively saying that there are d_i possibilities for $\sigma(i)$ for each i, we have a chance to take into account the fact that when it comes time to reveal $\sigma(i)$, some of v_i's neighbors may have already been used (as a match for v_j for some $j < i$), and so there may be a reduced range of choices for $\sigma(i)$.

The problem with this approach is that we have no way of knowing (or controlling) how many neighbors of i have been used at the moment when $\sigma(i)$ is revealed. Radhakrishnan's great idea to deal with this problem is to chose a *random* order in which to examine the vertices of $\mathcal{E}$ (rather than the deterministic order $v_1, \ldots, v_n$). There is a good chance that with a random order, we can say something precise about the average or expected number of neighbors of i that have been used at the moment when $\sigma(i)$ is revealed, and thereby put a better upper bound on the $H(\sigma(i))$ term.

So, let τ be any permutation of $\{1, \ldots, n\}$ (which we will think of as acting on $\mathcal{E}$ in the natural way). We have

$$H(X) = \sum_{j=1}^{n} H(\sigma(\tau(j))|\sigma(\tau(1)), \ldots, \sigma(\tau(j-1))).$$

It will prove convenient to write this sum in a slightly different way. Set $k(\tau, i) = \tau^{-1}(i)$. It will generally be clear from the context which τ and i are being considered, so we will just write "k" for "$k(\tau, i)$". We have

$$H(X) = \sum_{i=1}^{n} H(\sigma(i)|\sigma(\tau(1)), \ldots, \sigma(\tau(k-1))).$$

Since this is true for all τ, we can average over all possible choices of τ to get

$$\begin{aligned} H(X) &= \frac{1}{n!}\sum_{\tau}\left(\sum_{i=1}^{n} H(\sigma(i)|\sigma(\tau(1)),\ldots,\sigma(\tau(k-1)))\right) \\ &= \sum_{i=1}^{n}\frac{1}{n!}\sum_{\tau} H(\sigma(i)|\sigma(\tau(1)),\ldots,\sigma(\tau(k-1))). \end{aligned}$$

We now fix i, and try to understand the inner sum. For each σ and τ, write $D_i(\sigma,\tau)$ for the number of neighbors of v_i that are not among $\sigma(\tau(1)),\ldots,\sigma(\tau(k-1)$. If σ is being revealed to us edge-by-edge, $D_i(\sigma,\tau)$ counts the number of possible values that $\sigma(i)$ can take, given that each of $\sigma(\tau(1)),\ldots,\sigma(\tau(k-1))$ have been revealed. Since σ encodes a matching, we have $1 \le D_i(\sigma,\tau) \le d_i$ always.

For each fixed τ we have

$$H(\sigma(i)|\sigma(\tau(1)),\ldots,\sigma(\tau(k-1))) \le \sum_{j=1}^{d_i}\sum_{\sigma}\frac{\mathbf{1}_{\{D_i(\sigma,\tau)=j\}}}{\#(\sigma)}\log j$$

where $\mathbf{1}$ is an indicator function. Here we are upper bounding entropy by the log of the size of the range. It follows that

$$\begin{aligned} H(X) &\le \sum_{i=1}^{n}\frac{1}{n!}\sum_{\tau}\sum_{j=1}^{d_i}\sum_{\sigma}\frac{\mathbf{1}_{\{D_i(\sigma,\tau)=j\}}}{\#(\sigma)}\log j \\ &= \sum_{i=1}^{n}\frac{1}{n!}\sum_{j=1}^{d_i}\log j\sum_{\sigma}\frac{1}{\#(\sigma)}\sum_{\tau}\mathbf{1}_{\{D_i(\sigma,\tau)=j\}}. \end{aligned} \tag{2}$$

We now reach the point where the power of averaging over τ comes into play. For each fixed σ, $D_i(\sigma,\tau)$ depends only on where $\sigma(i)$ falls in the permutation of the neighbors of i induced by σ and τ (i.e., the permutation $(\sigma(\tau(n_1)),\ldots,\sigma(\tau(n_{d_i})))$, where the n_i's are the neighbors of v_i). If $\sigma(i)$ comes first on this list, then $D_i(\sigma,\tau) = d_i$; if second, then $D_i(\sigma,\tau) = d_i - 1$, and so on. If it comes last, then $D_i(\sigma,\tau) = 1$ since there can be only one possibility for $\sigma(i)$ at that point. By an easy symmetry consideration, $\sigma(i)$ is equally likely to appear in any position in the permutation of the neighbors of i. In other words, for each $1 \le j \le d_i$,

$$\sum_{\tau}\mathbf{1}_{\{D_i(\sigma,\tau)=j\}} = \frac{n!}{d_i}.$$

Plugging into (2) we get

$$\begin{aligned} H(X) &\le \sum_{i=1}^{n}\frac{1}{d_i}\sum_{j=1}^{d_i}\log j \\ &= \sum_{i=1}^{n}\frac{\log d_i!}{d_i} \end{aligned}$$

as required.

3 Alon and Friedland's proof of Kahn-Lovász

Alon and Friedland's idea is to relate $\mathcal{M}_{\text{perfect}}(G)$ to the permanent of the adjacency matrix $\text{Adj}(G) = (a_{ij})$ of G. This is the $2n$ by $2n$ matrix with

$$a_{ij} = \begin{cases} 1 & \text{if } v_i v_j \in E \\ 0 & \text{otherwise.} \end{cases}$$

An element of $\mathcal{M}_{\text{perfect}}(G) \times \mathcal{M}_{\text{perfect}}(G)$ is a pair of perfect matchings. The union of these perfect matchings is a collection of isolated edges (the edges in common to both matchings), together with a collection of disjoint even cycles, that covers the vertex set of the graph. For each such subgraph of G (call it an *even cycle cover*), to reconstruct the pair of matchings from which it arose we have to make an arbitrary choice for each even cycle, since there are two ways of writing an even cycle as an ordered union of matchings. It follows that

$$|\mathcal{M}_{\text{perfect}}(G) \times \mathcal{M}_{\text{perfect}}(G)| = \sum_S 2^{c(S)}$$

where the sum is over all even cycle covers S of G and $c(S)$ counts the number of even cycles in S.

On the other hand, any permutation σ contributing to $\text{perm}(\text{Adj}(G))$ breaks into disjoint cycles each of length at least 2, with the property that for each such cycle $(v_{i_1}, \ldots, v_{i_k})$ we have $v_{i_1}v_{i_2} \in E$, $v_{i_2}v_{i_3} \in E$, …, $v_{i_k}v_{i_1} \in E$. So such σ is naturally associated with a collection of isolated edges (the cycles of length 2), together with a collection of disjoint cycles (some possibly of odd length), that covers the vertex set of the graph. For each such subgraph of G (call it a *cycle cover*), to reconstruct the σ from which it arose we have to make an arbitrary choice for each cycle, since there are two ways of orienting it. It follows that

$$\text{perm}(\text{Adj}(G)) = \sum_S 2^{c(S)}$$

where the sum is over all cycle covers S of G and $c(S)$ counts the number of cycles in S.

It is clear that $|\mathcal{M}_{\text{perfect}}(G) \times \mathcal{M}_{\text{perfect}}(G)| \leq \text{perm}(\text{Adj}(G))$ since there are at least as many S's contributing to the second sum as the first, and the summands are identical for S's contributing to both. Applying Brégman's theorem to the right-hand side, and taking square roots, we get

$$|\mathcal{M}_{\text{perfect}}(G)| \leq \prod_{i=1}^{2n} (d_i!)^{\frac{1}{2d_i}}\ .$$

4 The Upper Matching conjecture

Here we describe the asymptotic evidence in favour of Conjecture 1.4 that has been provided by Carroll, Galvin and Tetali and by Kahn and Ilinca. Both sets of authors focus on the case $t = \alpha n$, where $\alpha \in (0, 1)$ is fixed (and we restrict our attention to those n for which αn is an integer). The first non-trivial task in this range is to determine the asymptotic behavior of

$|\mathcal{M}_{\alpha n}(K(n,d))|$ in n and d. To do this we start from the identity

$$|\mathcal{M}_{\alpha n}(K(n,d))| = \sum_{\substack{a_1,\ldots a_{n/d}:\\ 0\leq a_i \leq d,\ \sum_i a_i = \alpha n}} \prod_{i=1}^{n/d} \binom{d}{a_i}^2 a_i!$$

Here the a_i's are the sizes of the intersections of the matching with each of the components of $K(n,d)$, and the term $\binom{d}{a_i}^2 a_i!$ counts the number of matchings of size a_i in a single copy of $K_{d,d}$. (The binomial term represents the choice of a_i endvertices for the matching from each partition class, and the factorial term tells us how many ways there are to pair the endvertices from each class to form a matching.) Considering only those sequences $(a_1, \ldots a_{n/d})$ in which each a_i is either $\lfloor \alpha d \rfloor$ or $\lceil \alpha d \rceil$, we get

$$\log |\mathcal{M}_{\alpha n}(K(n,d))| = n\left(\alpha \log d + 2H(\alpha) + \alpha \log\left(\frac{\alpha}{e}\right) + \Omega_\alpha\left(\frac{\log d}{d}\right)\right), \tag{3}$$

where $H(\alpha) - \alpha \log \alpha - (1-\alpha)\log(1-\alpha)$ is the binary entropy function. The detailed analysis appears in [4].

In [4], Carroll, Galvin and Tetali prove an upper bound on $\log |\mathcal{M}_{\alpha n}(G)|$ for arbitrary d-regular G on $2n$ vertices that agrees with (3) in the first term, and is off by a constant in the second term:

$$\log |\mathcal{M}_{\alpha n}(G)| \leq n\left(\alpha \log d + H(\alpha)\right). \tag{4}$$

Using a refinement of Radhakrishnan's approach to Brégman's theorem, Kahn and Ilinca [10] have improved this, obtaining a bound that agrees with (3) in the first two terms:

$$\log |\mathcal{M}_{\alpha n}(G)| \leq n\left(\alpha \log d + 2H(\alpha) + \alpha \log\left(\frac{\alpha}{e}\right) + o(d^{-1/4})\right).$$

Here we describe the proof of (4). It is based on finding an upper bound on the *matching polynomial* of G, the polynomial

$$P_{\mathcal{M}}(G,x) = \sum_{t=0}^{n} |\mathcal{M}_t(G)| x^t.$$

For d-regular G on $2n$ vertices, the bound in question is

$$P_{\mathcal{M}}(G,x) \leq (1+dx)^n \tag{5}$$

(we will justify this presently). Using this, we easily obtain (4). Since the matching polynomial has all positive terms, we have $x^{\alpha n}|\mathcal{M}_{\alpha n}| \leq P_{\mathcal{M}}(G,x)$ for any positive x, and so using (5) we get

$$|\mathcal{M}_{\alpha n}(G)| \leq x^{-\alpha n} P_{\mathcal{M}}(G,x) \leq \left(\frac{1+dx}{x^\alpha}\right)^n.$$

Since this is true for all x, we are liberty to choose the best possible x, namely the one that minimizes the right-hand side above. A little calculus shows that $x = \alpha/(d(1-\alpha))$ is the right choice, and a little algebra then gives (4). Essentially what we are trying to do is choose the value of x that maximizes the (percentage) contribution of $x^{\alpha n}|\mathcal{M}_{\alpha n}|$ to $P_{\mathcal{M}}(G,x)$.

The inequality in (4) is trivial for bipartite G ($(1+dx)^n$ is an exact count of subgraphs of G which on one fixed partition class have maximum degree at most one, with a subgraph on t edges counted with weight x^t). For non-bipartite G it is less obvious. The proof given in [4] uses a very general entropy inequality due to Friedgut. A much simpler proof was pointed out by Gurvitz, based on the celebrated Heilmann-Lieb theorem [9]. This asserts that for all graphs G, the polynomial $P_{\mathcal{M}}(G,x) = 0$ has only real, negative roots. It follows that we may write $P_{\mathcal{M}}(G,x) = \prod_{i=1}^{n}(1+\alpha_i x)$ for some positive α_i's with $\sum \alpha_i = (P_{\mathcal{M}}(G,x))'|_{x=0} = |E| = nd$. Applying the arithmetic mean - geometric mean inequality to this expression we obtain

$$P_{\mathcal{M}}(G,x) \le \left(1 + x\frac{\sum \alpha_i}{n}\right)^n = (1+dx)^n\,.$$

References

[1] N. Alon and S. Friedland, The maximum number of perfect matchings in graphs with a given degree sequence, *Electron. J. Combin.* **15** (2008), #N13.

[2] N. Alon and J. Spencer, *The Probabilistic Method*, Wiley, New York, 2000.

[3] L. Brégman, Some properties of nonnegative matrices and their permanents, *Soviet Math. Dokl.* **14** (1973), 945–949.

[4] T. Carroll, D. Galvin and P. Tetali, Matchings and Independent Sets of a Fixed Size in Regular Graphs, *J. Combin. Theory Ser. A* **116** (2009), 1219–1227.

[5] J. Cutler, A. Radcliffe, An entropy proof of the Kahn-Lovász theorem, manuscript.

[6] G. Egorychev, Permanents, Book in Series of Discrete Mathematics (in Russian), Krasnoyarsk, SFU, 2007.

[7] S. Friedland, An upper bound for the number of perfect matchings in graphs, arXiv:0803.0864

[8] S. Friedland, E. Krop and K. Markström, On the Number of Matchings in Regular Graphs, *Electron. J. Combin.* **15** (2008), #R110.

[9] O. Heilmann and E. Lieb, Theory of monomer-dimer systems, *Comm. Math. Phys.* **25** (1972), 190-232.

[10] J. Kahn and L. Ilinca, The number of matchings of a given size. manuscript.

[11] J. Radhakrishnan, An entropy proof of Brégman's theorem, *J. Combin. Theory, Ser. A* **77**, (1997), 161–164.

[12] A. Schrijver, A short proof of Minc's conjecture, *J. Combin. Theory Ser. A* **25** (1978), 80–83.