On counting independent sets in sparse graphs

Martin Dyer*
University of Leeds
United Kingdom

Alan Frieze†
Carnegie Mellon University
U.S.A.

Mark Jerrum‡
University of Edinburgh
United Kingdom

July 16, 1998

Abstract

We prove two results concerning approximate counting of independent sets in graphs with constant maximum degree Δ. The first implies that the Monte Carlo Markov chain technique is likely to fail if $\Delta \geq 6$. The second shows that no fully polynomial randomized approximation scheme can exist if $\Delta \geq 25$, unless RP = NP.

*School of Computer Studies, University of Leeds, LS2 9JT, United Kingdom (dyer@scs.leeds.ac.uk). Supported by ESPRIT Working Group RAND2.

†Department of Mathematics, Carnegie Mellon University, Pittsburgh, PA 15213, U.S.A. (af1p+@andrew.cmu.edu).

‡Department of Computer Science, University of Edinburgh, EH9 3JZ, United Kingdom (mrj@dcs.ed.ac.uk). Supported by ESPRIT Working Group RAND2.

1 Introduction

Counting independent sets in graphs is one of several combinatorial counting problems which have received recent attention. The problem is known to be #P-complete, even for low degree graphs [3]. On the other hand, it has been shown that, for graphs of maximum degree $\Delta = 4$, randomized approximate counting is possible [7, 3]. This success has been achieved using the *Monte Carlo Markov chain* method to construct a *fully polynomial randomized approximation scheme (fpras)*. This has led to a natural question as to how far this success might extend.

Here we consider in more detail this question of counting independent sets in graphs with constant maximum degree. We prove two results. The first, in Section 2, shows that the Monte Carlo Markov chain method is likely to fail for graphs with $\Delta = 6$. This leaves open only the case $\Delta = 5$.

Our second result gives an explicit value of Δ above which approximate counting, using any kind of polynomial-time algorithm, is impossible unless RP = NP. The bound we obtain is $\Delta = 25$, though we suspect that the true value is in single figures, probably 6.

We note that Berman and Karpinski [2] have recently given new explicit bounds for the approximation ratio for the maximum independent set and other problems in low-degree graphs. These directly imply an inapproximability result for counting. (See [7].) However, the bound on Δ obtained this way is larger than ours by at least two orders of magnitude.

2 Monte Carlo Markov chains

For a graph G, let $\mathcal{I}(G)$ denote the collection of independent sets of G. Let $\mathcal{M}(G)$ be any Markov chain, asymptotically uniform on $\mathcal{I}(G)$, with transition matrix P_G. In this section G will be a bipartite graph with a vertex bipartition into classes of equal size n. Let $b(n) \leq n$ be any function of n and suppose we have $P_G(\sigma_1, \sigma_2) = 0$ whenever $|\sigma_1 \oplus \sigma_2| > b(n)$, where $\oplus$ denotes symmetric difference. We will say that $\mathcal{M}(G)$ is $b(n)$-*cautious*. Thus a $b(n)$-cautious chain is not permitted to change the status of more than $b(n)$ of the vertices in G at any step. Ideally, we would wish to have $b(n)$ a constant (as in [7, 3]). However, we will show that no $b(n)$-cautious chain on $\mathcal{I}(G)$ can mix rapidly unless $b(n) = \Omega(n)$. Thus any chain which does mix rapidly on $\mathcal{M}(G)$ must change the status of a sizeable proportion of the vertices at each step.

Specifically, we prove the following

Theorem 1 *Let $\Delta \geq 6$ and $b(n) \leq 0.35n$. Then there exists a bipartite graph G_0, of maximum degree Δ, on $n + n$ vertices (more precisely a sequence of such graphs parameterised by n) with the following property: for any $b(n)$-cautious Markov chain on $\mathcal{I}(G_0)$, and almost every (in the uniform measure) choice of starting state, the mixing time is $\Omega(e^{\gamma n})$, for some constant $\gamma > 0$.*

Since, of course, there does exist an n-cautious chain which mixes rapidly, our result cannot be strengthened much further. The remainder of this section is devoted to the proof of Theorem 1.

We infer the existence of the graph G_0 by random graph methods. Let $K_{n,n}$ denote the complete bipartite graph with vertex bipartition V_1, V_2, where $|V_1| = |V_2| = n$, and let G be the union of Δ perfect matchings selected independently and uniformly at random in $K_{n,n}$. (Since the perfect matchings are independent, they may well share some edges.) Denote by $\mathcal{G}(n,n,\Delta)$ the probability space of bipartite graphs G so defined. Where no confusion can arise, we simply write $\mathcal{G}$ for this class below. Note that $\mathcal{G}$ is a class of graphs with degree bound Δ.

Let $0 < \alpha, \beta < 1$ be chosen values. For $G \in \mathcal{G}$, we consider the collection $\mathcal{I}_G(\alpha,\beta)$ of $\sigma \in \mathcal{I}(G)$ such that $|\sigma \cap V_1| = \alpha n$ and $|\sigma \cap V_2| = \beta n$. (We assume that αn, βn are integers, merely to ease exposition.) We will call $\sigma \in \mathcal{I}_G(\alpha,\beta)$ an (α,β)-*set*. Using linearity of expectation, we may easily compute the expected number $\mathcal{E}(\alpha,\beta) = \mathbf{E}(|\mathcal{I}_G(\alpha,\beta)|)$ of (α,β)-sets in G: it is just the number of ways of choosing an αn-subset from V_1 and a βn-subset from V_2, multiplied by the probability that all Δ perfect matchings avoid connecting the αn-subset to the βn-subset. Thus,

$$\begin{aligned}
\mathcal{E}(\alpha,\beta) &= \binom{n}{\alpha n}\binom{n}{\beta n}\left[\binom{(1-\beta)n}{\alpha n}\Big/\binom{n}{\alpha n}\right]^{\Delta} \\
&= \left(\frac{(1-\beta)^{(\Delta-1)(1-\beta)}(1-\alpha)^{(\Delta-1)(1-\alpha)}}{\alpha^\alpha \beta^\beta (1-\alpha-\beta)^{\Delta(1-\alpha-\beta)}}\right)^{n(1+o(1))} \\
&= e^{\phi(\alpha,\beta)n(1+o(1))}, \qquad (1)
\end{aligned}$$

where

$$\begin{aligned}
\phi(\alpha,\beta) = -\alpha\ln\alpha - \beta\ln\beta - \Delta(1-\alpha-\beta)\ln(1-\alpha-\beta)& \\
+ (\Delta-1)\big((1-\alpha)\ln(1-\alpha) + (1-\beta)\ln(1-\beta)\big)&. \qquad (2)
\end{aligned}$$

It follows that

$$\begin{aligned}
\frac{\partial\phi}{\partial\alpha} &= -\ln\alpha - (\Delta-1)\ln(1-\alpha) + \Delta\ln(1-\alpha-\beta), \qquad (3)\\
\frac{\partial\phi}{\partial\beta} &= -\ln\beta - (\Delta-1)\ln(1-\beta) + \Delta\ln(1-\alpha-\beta), \qquad (4)\\
\frac{\partial^2\phi}{\partial\alpha^2} &= -\frac{1}{\alpha} + \frac{\Delta-1}{1-\alpha} - \frac{\Delta}{1-\alpha-\beta}, \qquad (5)\\
\frac{\partial^2\phi}{\partial\beta^2} &= -\frac{1}{\beta} + \frac{\Delta-1}{1-\beta} - \frac{\Delta}{1-\alpha-\beta}, \qquad (6)\\
\frac{\partial^2\phi}{\partial\alpha\,\partial\beta} &= -\frac{\Delta}{1-\alpha-\beta}. \qquad (7)
\end{aligned}$$

Now ϕ is defined on the triangle

$$\mathcal{T} = \{(\alpha, \beta) : \alpha, \beta \geq 0, \quad \alpha + \beta \leq 1\},$$

and is clearly symmetrical in α, β. (The function ϕ is defined by equation (2) on the interior of $\mathcal{T}$, and can be extended to the boundary by taking limits.) Moreover, using (3)–(7), the following facts are established in Appendix A about the stationary points of ϕ on $\mathcal{T}$:

(i) The function ϕ has no local minima, and no local maxima on the boundary of $\mathcal{T}$.

(ii) All local maxima of ϕ satisfy $\alpha + \beta + \Delta(\Delta - 2)\alpha\beta \leq 1$.

(iii) If $\Delta \leq 5$, ϕ has only a single local maximum, which is on the line $\alpha = \beta$.

(iv) If $\Delta \geq 6$, ϕ has exactly two local maxima, symmetrical in α,β and a single saddle-point on $\alpha = \beta$.

From (iii) and (iv), we see that the distribution of the numbers of (α, β)-sets undergoes a dramatic change from $\Delta = 5$ to $\Delta = 6$. For $\Delta \leq 5$, the "typical" (α, β)-set is "balanced" (i.e. has $\alpha \approx \beta$), whereas for $\Delta \geq 6$ it is "unbalanced". We will examine the first unbalanced case, $\Delta = 6$, and make this precise.

We may determine numerically (see Appendix A) that the maximum of ϕ with smaller α occurs at $(\alpha^*, \beta^*) \approx (0.03546955, 0.40831988)$, and there is a $c \geq 0.71513499$ such that $\phi(\alpha^*, \beta^*) > c$. Therefore let us say (α, β) is a *middle point* if $|\beta - \alpha| \leq 0.35$, and an (α, β)-set is a *middle set* if (α, β) is a middle point. Let ϕ_{mid} denote the maximum of $\phi(\alpha, \beta)$ for (α, β) a middle point. Then, since by (iv) ϕ_{mid} is attained on $|\beta - \alpha| = 0.35$, further numerical computation shows that there is a constant $\gamma \geq 0.00018191$ such that $\phi_{\text{mid}} < c - \gamma$. Hence we have shown that, for large enough n

$$\mathcal{E}^* = \mathcal{E}(\alpha^*, \beta^*) > \exp(cn),$$

and

$$\mathcal{E}_{\text{mid}} = \sum_{|\beta - \alpha| \leq 0.35} \mathcal{E}(\alpha, \beta) < n^2 \exp(\phi_{\text{mid}} n) < \exp((c - \gamma)n).$$

Thus $\mathcal{E}_{\text{mid}} < e^{-\gamma n} \mathcal{E}^*$. For $G \in \mathcal{G}$, let X^* denote the number of (α^*, β^*)-sets G possesses and X_{mid} its number of middle sets. Then

$$\mathbf{E}(X_{\text{mid}} - e^{-\gamma n} X^*) = \mathcal{E}_{\text{mid}} - e^{-\gamma n} \mathcal{E}^* < 0.$$

Thus, for large enough n, there must *exist* a $G_0 \in \mathcal{G}$ such that

$$X_{\text{mid}}(G_0) < e^{-\gamma n} X^*(G_0). \tag{8}$$

This is the graph G_0 of Theorem 1, and the remainder of our argument concerns G_0. Note that we have only shown the *existence* of G_0, We cannot assert from the above that a *random* graph in $\mathcal{G}$ is likely to satisfy (8), though more delicate arguments might well establish this.

Now consider a $0.35n$-cautious chain $\mathcal{M}(G_0) = \mathcal{M}_0$, on $\mathcal{I}(G_0)$. We will examine the *conductance* [6], $\Phi(\mathcal{M}_0)$, of $\mathcal{M}_0$. To this end, let S comprise all (α, β)-sets with $\alpha \geq \beta$, i.e.

$$S = \{\sigma \in \mathcal{I}(G_0) : |\sigma \cap V_1| \geq |\sigma \cap V_2|\}.$$

Then, if P_0 is the transition matrix of $\mathcal{M}_0$,

$$\sum_{\sigma_1 \in S} \sum_{\sigma_2 \notin S} P_0(\sigma_1, \sigma_2) < X_{\text{mid}}(G_0), \qquad \text{and} \qquad |S| > X^*(G_0). \tag{9}$$

It follows from (9) that

$$\Phi(\mathcal{M}_0) < e^{-\gamma n}.$$

Now, suppose the chain is started at a state chosen uniformly in S. By symmetry this is equivalent to choosing a set uniformly from $\mathcal{I}(G_0)$. Then, as shown by Jerrum [5], the probability that the chain is still in S after t steps is at least $(1 - \Phi)^t$. Hence the variation distance of the chain from the uniform distribution is at least $(1 - \Phi)^t - \frac{1}{2}$. It follows, by a simple calculation, that if $n \geq 1/\gamma$, the mixing time (i.e. the time to reach variation distance e^{-1}) is at least $0.0895e^{\gamma n}$.

3 Hardness of approximate counting

The result of the previous section implies that the usual approach to approximating the number of independent sets in a low-degree graph must fail when $\Delta \geq 6$, at least in the worst case. Here we show that, if the degree bound is somewhat larger, then *any* approach to approximating the number of independent sets is doomed to failure, under a reasonable complexity assumption. Precisely, the remainder of this section is devoted to proving the following

Theorem 2 *Unless* NP = RP*, there is no polynomial time algorithm which estimates the* logarithm *of the number of independent sets in a* Δ*-regular graph* ($\Delta \geq 25$) *within relative error at most* $\epsilon = 10^{-6}$.

We give a randomized reduction from the following problem E2LIN2, analysed by Håstad [8]. The input is a system $\mathcal{A}$ of m equations over $\mathbf{Z}_2$ in n variables $x_1, x_2, \ldots, x_n$, such that each equation has exactly two variables. (Thus each equation is of the form $x_i + x_j = 0$ or $x_i + x_j = 1$.) The objective is to find a maximum cardinality consistent subset of equations in $\mathcal{A}$, i.e., to assign values to the variables so as to maximize the number m_C of satisfied equations. Håstad showed, using the powerful theory of PCP, that (unless NP = RP) there is no polynomial time algorithm which can estimate m_C within any constant factor smaller than 12/11.

Therefore consider an instance $\mathcal{A}$ of E2LIN2, as above. We will construct (by a randomized algorithm) a graph $G = (V, E)$, regular of degree Δ. We then show that, if we can approximate the logarithm of the number of independent sets in G to within the required relative error, we can (with high probability) approximate the size of m_C in $\mathcal{A}$, to within a factor $12/11 - \epsilon$. Theorem 2 will then follow.

Let us write $[n] = \{1, 2, \ldots, n\}$. We construct the graph $G = G(\mathcal{A})$ as follows. We assume $m \geq n$, otherwise $\mathcal{A}$ is decomposable or consistent. Let $M = m^6$ and, for each $i \in [n]$, let A_i be the multiset of equations containing x_i, with (multiset) cardinality d_i. We represent x_i by a regular bipartite graph $H_i = (L_i, R_i, F_i)$ of degree $\delta = \Delta - 1$. Here $L_i = \bigcup_{a \in A_i} L_{i,a}$, $R_i = \bigcup_{a \in A_i} R_{i,a}$ where the sets $L_{i,a}$, $R_{i,a}$ partition L_i and R_i respectively, and for all i, a, $|L_{i,a}| = |R_{i,a}| = M$. Thus H_i is bipartite with both its vertex sets of size Md_i. Later, we will associate L_i with the assignment $x_i = 0$, and R_i with $x_i = 1$.

The graph $H_i = (L_i, R_i, F_i)$ will be sampled from $\mathcal{G}(Md_i, Md_i, \delta)$, where $\mathcal{G}$ is the class of random graphs defined in section 2. Here, however, we will reject graphs which are not δ-regular. It is known [1] that a graph in $\mathcal{G}$ is regular of full degree with constant probability (for constant δ). Moreover, the property of being δ-regular can clearly be checked in polynomial time.

The equations a of $\mathcal{A}$ determine the edges connecting the H_i in G, as follows. If a is the equation $x_i + x_j = 1$ ($x_i + x_j = 0$ resp.), we add an arbitrary perfect matching between $L_{i,a}$ and $L_{j,a}$ ($R_{j,a}$ resp.) and another between $R_{i,a}$ and $R_{j,a}$ ($L_{j,a}$ resp.). Thus G is a regular graph of degree Δ. We will show that approximating the logarithm of the number of independent sets in G to within a factor $(1 + 10^{-6})$ will allow us to approximate the E2LIN2 instance within the Håstad bound.

Before returning to the issue of approximation, we will need to establish some crucial properties of the "typical" independent set in G. For this purpose, let I be sampled uniformly from $\mathcal{I}(G)$. First we show that I "occupies about half the available space" in each $L_{i,a}$ or $R_{i,a}$.

Let $\mathcal{L}_{i,a}$ be the set of vertices in $L_{i,a}$ with no neighbour in I and let $\mathcal{L}_i = \bigcup_{a \in A_i} \mathcal{L}_{i,a}$.

Lemma 1 *Then, except for probability* $e^{-\Omega(m^2)}$*, either* $|\mathcal{L}_{i,a}| < m^4$ *or* $|\mathcal{L}_{i,a}| = (2 + O(1/m))|I \cap L_{i,a}|$.

Proof If we condition on $I \cap (V \setminus L_{i,a})$ then $I \cap L_{i,a}$ is a random subset of $\mathcal{L}_{i,a}$. If $|\mathcal{L}_{i,a}| \geq m^4$ then Chernoff's bound implies that

$$\mathbf{Pr}\Big(|I \cap L_{i,a}| \notin \frac{1}{2}\big(1 \pm \frac{1}{m}\big)|\mathcal{L}_{i,a}|\Big) \leq 2\exp\big(-\frac{1}{3}m^2\big),$$

from which the Lemma follows. □

Clearly we may define $\mathcal{R}_{i,a}$ and $\mathcal{R}_i$ symmetrically and prove an analogous result. It is also clear that we may claim the Lemma for all i, a since there are less than m^2 such pairs. We now deduce that at least around half of L_i is "available" to I in H_i, if we fix I outside H_i.

Let $\mathcal{L}'_i$ be the set of vertices in L_i with no neighbour in I outside of H_i.

Lemma 2 *Except for probability $e^{-\Omega(m^2)}$, $|\mathcal{L}'_i| \geq (\frac{1}{2} - O(1/m))|L_i|$.*

Proof We condition on the whole of I outside of H_i. If $L_{i,a}$ is joined by a matching to $V_{j,a}$ ($V \in \{L, R\}$) then, from Lemma 1, $M \geq (2 + O(1/m))|I \cap V_{j,a}|$. Hence

$$|\{v \in L_{i,a} : \{v, w\} \in E \setminus F_i \text{ implies } w \notin I\}| \geq (\tfrac{1}{2} - O(1/m))|L_{i,a}|.$$

Summing this over all $a \in A_i$ gives the Lemma. □

Again, we may define $\mathcal{R}'_i$ and prove a corresponding result. We now show, that for each i either $|\mathcal{L}_i|$ or $|\mathcal{R}_i|$ is "small". We will temporarily drop the suffix i, and write H rather than H_i etc. Let $N = |L| = dM \leq m^7$, $a = |\mathcal{L}'|/N$, $b = |\mathcal{R}'|/N$. Without loss of generality, we assume $a \geq b$ and, from Lemma 2, $b \geq \frac{1}{2} - O(1/m)$. Write $\sigma = I \cap H$, where I is a uniformly chosen independent set in G. We will say that σ is an (α, β)-set if $|\sigma \cap L| = \alpha a N$, $|\sigma \cap R| = \beta b N$. Note that, while we treat α, β as continuous, there are in fact at most N values for each of $a, b, \alpha a, \beta b$ (and hence N^2 values of each of α, β).

Lemma 3 *Let $\delta \geq 24$. If I is a uniformly chosen independent set in G then, except for probability $e^{-\Omega(m^2)}$, $\min(|\mathcal{L}_i|, |\mathcal{R}_i|) \leq \lambda N$, where $\lambda = 0.00943$.*

Proof We argue conditionally on $\mathcal{L}', \mathcal{R}'$. It is easy to see that there are at least 2^{aN} independent sets in H. We will show that, for α, β not satisfying the condition of the Lemma, the number of (α, β)-sets is so much smaller than this that they appear with probability $e^{-\Omega(m^2)}$. It will be sufficient to show that the expected number of (α, β)-sets in such a case is $2^{aN - \Omega(m^2)}$, since Markov's inequality will then imply the inequality for the actual number. Moreover, it will be sufficient to prove this for any particular triple H, α, β, since there can be at most $N^{4n} = e^{O(m \ln m)}$ different combinations of values of α, β in the whole of G. Now the expected number of (α, β)-sets in H is

$$\begin{aligned}
\mathcal{E}(\alpha, \beta) &= \binom{aN}{\alpha a N}\binom{bN}{\beta b N}\left[\binom{(1 - b\beta)N}{\alpha a N}\Big/\binom{N}{\alpha a N}\right]^{\delta} \\
&\leq \binom{aN}{\alpha a N}\binom{bN}{\beta b N}(1 - b\beta)^{\alpha a \delta N} \\
&\leq \left[\left(\alpha^{\alpha}(1-\alpha)^{(1-\alpha)}\right)^{-a}\left(\beta^{\beta}(1-\beta)^{(1-\beta)}\right)^{-b} e^{-\alpha\beta a b \delta}\right]^{N(1+o(1))} \\
&= e^{\psi(\alpha,\beta)N(1+o(1))},
\end{aligned} \tag{10}$$

where

$$\psi(\alpha,\beta) = -a(\alpha\ln\alpha + (1-\alpha)\ln(1-\alpha)) - b(\beta\ln\beta + (1-\beta)\ln(1-\beta)) - \alpha\beta ab\delta. \tag{11}$$

It follows that

$$\frac{\partial\psi}{\partial\alpha} = a(-\ln\alpha + \ln(1-\alpha) - \beta b\delta), \qquad \frac{\partial\psi}{\partial\beta} = b(-\ln\beta + \ln(1-\beta) - \alpha a\delta), \tag{12}$$

$$\frac{\partial^2\psi}{\partial\alpha^2} = \frac{-a}{\alpha(1-\alpha)}, \qquad \frac{\partial^2\psi}{\partial\beta^2} = \frac{-b}{\beta(1-\beta)}, \qquad \frac{\partial^2\psi}{\partial\alpha\,\partial\beta} = -ab\delta. \tag{13}$$

The following two facts about ψ are easily verified.

$$\psi(\alpha,\beta) \geq \psi(1-\alpha,\beta) \qquad \text{if } \alpha \leq \tfrac{1}{2} \tag{14}$$

$$\psi(\alpha,\beta) \geq \psi(\alpha,1-\beta) \qquad \text{if } \beta \leq \tfrac{1}{2} \tag{15}$$

$$\psi(\alpha,\beta) \geq \psi(\beta,\alpha) \qquad \text{if } \beta \leq \alpha \leq 1-\beta \tag{16}$$

We wish to determine the regions where $\psi \geq a\ln 2$. These are connected neighbourhoods of the local maxima of ψ. From (12) we see that ψ has no boundary maxima for α,β in the unit square $\mathcal{U}$. Thus, from (13), ψ has only local maxima or saddle-points in $\mathcal{U}$, and a stationary point is a local maximum if and only if

$$\alpha(1-\alpha)\beta(1-\beta) \leq 1/(ab\delta^2). \tag{17}$$

Thus, at any local maximum, either $\beta(1-\beta) \leq 1/(b\delta)$ or $\alpha(1-\alpha) \leq 1/(a\delta)$. If the former, this and $b\delta \geq 12 - o(1)$ imply $\beta < .1$, and hence $\beta < 1.2/b\delta$. An identical argument holds for α. Let us denote the rectangle $[\ell_\alpha, u_\alpha] \times [\ell_\beta, u_\beta]$ by $[\ell_\alpha, u_\alpha \,|\, \ell_\beta, u_\beta]$. Thus any local maximum of ψ must lie in either $[0,1 \,|\, 0, 1.2/b\delta]$ or $[0, 1.2/a\delta \,|\, 0, 1]$. In $[0, 1.2/b\delta \,|\, 0, 1.2/b\delta]$, $\alpha,\beta < 0.1$ and hence

$$\psi(\alpha,\beta) < 2a(-0.1\ln(0.1) - 0.9\ln(0.9)) < a\ln 2.$$

Then, from (14) and (15), we also have $\psi(\alpha,\beta) < a\ln 2$ in $[1-1.2/b\delta, 1 \,|\, 0, 1.2/b\delta]$ and $[0, 1.2/b\delta \,|\, 1-1.2/b\delta, 1]$. Now, if $\beta \leq 1.2/b\delta$, let $\rho = 1-2\alpha$ and consider the upper bound

$$\psi(\alpha,\beta) \leq \Psi(\rho,\beta) = a(\ln 2 - \tfrac{1}{2}\rho^2) + b\beta(1-\ln\beta) - \tfrac{1}{2}(1-\rho)\beta ab\delta. \tag{18}$$

For fixed β, it is easily shown that Ψ is maximized if $\rho = \frac{1}{2}b\delta\beta < 0.6$. If $b\delta\beta = 1.2$, then $\rho = 0.6$ and

$$\max_{\rho}\ \Psi(\rho,\beta) \leq a(\ln 2 - 0.18) + 0.1a(1-\ln(0.1)) - 0.24a < a\ln 2.$$

Thus $\psi < a\ln 2$ everywhere on the boundary of $[1.2/b\delta, 1-1.2/b\delta \,|\, 0, 1.2/b\delta]$ (but not including the shared boundary with $\mathcal{U}$). Hence, by (16), $\psi < a\ln 2$ everywhere on the boundary of $[0, 1.2/b\delta \,|\, 1.2/b\delta, 1-1.2/b\delta]$, which contains $[0, 1.2/a\delta \,|\, 1.2/b\delta, 1-1.2/b\delta]$. Moreover $\psi(\alpha,\beta) \geq \psi(\beta,\alpha)$ for all points (α,β) in $[1.2/b\delta, 1-1.2/b\delta \,|\, 0, 1.2/b\delta]$. It follows that it is sufficient to determine β^* such that $\psi(\alpha,\beta^*) < a\ln 2$ everywhere in $[1.2/b\delta, 1-1.2/b\delta \,|\, \beta^*, 1.2/b\delta]$. To this end again consider

$$\Psi_0(\beta) = \max_{\rho} \Psi(\rho,\beta) = a\ln 2 + b\beta(1-\ln\beta) - \tfrac{1}{2}ab\beta\delta + \tfrac{1}{8}ab^2\beta^2\delta^2.$$

Now $\Psi_0 < a\ln 2$ if

$$b\beta\delta^2 - 4\delta + 8(1-\ln\beta)/a < 0.$$

This inequality is satisfied provided

$$2\left(1-\sqrt{1-2b\beta(1-\ln\beta)/a}\right) < b\beta\delta < 2\left(1+\sqrt{1-2b\beta(1-\ln\beta)/a}\right).$$

The right hand inequality is clearly irrelevant since we are assuming $\beta \leq 1.2/b\delta$. Thus we need only consider the left hand inequality, i.e. for fixed $\gamma = b\beta < 1.2/\delta$, we require

$$\gamma\delta > \max_{\frac{1}{2}-o(1)\leq b\leq a\leq 1} 2\left(1-\sqrt{1-2\gamma(1-\ln\gamma+\ln b)/a}\right).$$

Considering first b, the maximum occurs when $b = a$. So we have

$$\gamma\delta > \max_{\frac{1}{2}-o(1)\leq a\leq 1} 2\left(1-\sqrt{1-2\gamma(1-\ln\gamma+\ln a)/a}\right).$$

But, since $a \geq \gamma$, the maximum now occurs when $a = \frac{1}{2}$. Thus we require

$$\gamma\delta > 2\left(1-\sqrt{1-4\gamma(1-\ln\gamma-\ln 2)}\right).$$

To achieve $\gamma = 0.004715$, we require $\delta \geq 23.3$.

Thus we have shown that $\min(a\alpha, b\beta) < 0.004715$. The conclusion now follows immediately from Lemma 1. □

We now establish the relationship between the number of independent sets in Γ and the maximum size of a consistent subset of $\mathcal{A}$. Let $\mathcal{I} = \mathcal{I}(G)$. For $\sigma \in \mathcal{I}$ let $S_\sigma \subseteq [n]$ be defined by

$$S_\sigma = \{i : \; |L_i \cap \sigma| > |R_i \cap \sigma|, \; i \in [n]\}.$$

For $S \subseteq [n]$ let $I_S = \{\sigma \in \mathcal{I} : \; S_\sigma = S\}$ and let $\mu_S = |I_S|$. Recall that m is the number of equations in $\mathcal{A}$.

Lemma 4 *For $S \subseteq [n]$, let $\theta(S)$ be the number of equations in $\mathcal{A}$ satisfied by the assignment $x_i = 1$ $(i \in S)$, $x_i = 0$ $(i \notin S)$. Then*

$$4^{M\theta(S)} 3^{M(m-\theta(S))} \leq \mu_S \leq 4^{M\theta(S)} 3^{M(m-\theta(S))} 2^{2\lambda m M}(1 + o(1)), \tag{19}$$

where λ is as in Lemma 3.

Proof Fix $S \subseteq [n]$ and for $\sigma \in I_S$ we let $J_\sigma = \sigma \cap \left(\bigcup_{i \in S} L_i \cup \bigcup_{i \notin S} R_i\right)$. Informally, J_σ restricts σ to the left or right of each subgraph H_i, according to which side contains the larger part of σ. Let

$$\hat{\mu}_S = |\{J_\sigma : \ \sigma \in I_S\}| \leq \mu_S.$$

We show that

$$\hat{\mu}_S = 4^{M\theta(S)} 3^{M(m-\theta(S))}. \tag{20}$$

This immediately proves the lower bound in (19). Furthermore, Lemma 1 implies that for a fixed value J of J_σ there are (up to a factor $(1 + e^{-\Omega(m^2)})$ at most

$$\prod_{i \in [n]} 2^{\lambda d_i M} = 2^{\lambda M \sum_i d_i} = 2^{2\lambda m M}$$

sets $\sigma \in I_S$ with $J_\sigma = J$. The upper bound then follows.

To prove (20) we consider the number of possible choices for $J \cap L_{i,a}$, $J \cap R_{i,a}$, $J \cap L_{j,a}$ and $J \cap R_{j,a}$ for every equation $a : \ x_i + x_j = z_a$ $(z_a \in \{0, 1\})$. For given S, let us define

$$X_{i,a} = \begin{cases} L_{i,a}, & \text{if } i \in S; \\ R_{i,a}, & \text{if } i \notin S. \end{cases}$$

Then there are two cases, determined by the status of a.

(1) Equation a is satisfied by the assignment derived from S. Then there are 2^M choices for each of $J \cap X_{i,a}$, $J \cap X_{j,a}$, giving 4^M in all.

(2) Equation a is not satisfied by the assignment derived from S. Then the subgraph of G induced by $X_{i,a} \cup X_{j,a}$ is a matching of size M, and hence contains 3^M independent sets.

Multiplying the estimates from the two cases over all $a \in \mathcal{A}$ proves (20) and the Lemma. □

We now proceed to the proof of Theorem 2. Let $Z_I = Z_I(G)$ denote the logarithm of the number of independent sets of $G(\mathcal{A})$. Let $Z_C = Z_C(\mathcal{A})$ denote the maximum number of consistent equations in $\mathcal{A}$.

Let Y_I be some estimate of Z_I satisfying $|Y_I/Z_I - 1| \leq \epsilon = 10^{-6}$. Using Y_I we define

$$Y_C = \left(\frac{Y_I}{M} - m \ln 3\right) \frac{1 + 10^{-5}}{\ln(4/3)}.$$

A simple calculation will then show that $1 \leq Y_C/Z_C \leq 12/11 - \epsilon$, so that Y_C determines Z_C with sufficient accuracy to beat the approximability bound for E2LIN2.

From Lemma 4 we see that

$$Y_I \geq (1-\epsilon)M(Z_C \ln(4/3) + m \ln 3).$$

Hence, since $Z_C \geq m/2$,

$$Y_C/1.00001 \geq (1-\epsilon)Z_C - \frac{\epsilon m \ln 3}{\ln(4/3)} \geq Z_C \left(1 - \frac{\epsilon \ln 12}{\ln(4/3)}\right) \geq 0.999991 Z_C,$$

which implies $Y_C \geq Z_C$. On the other hand Lemma 4 also implies that

$$Y_I \leq (1+\epsilon)\big[M\big(Z_C \ln(4/3) + m \ln 3 + 2m\lambda \ln 2\big) + n \ln 2\big],$$

where $\lambda \leq 0.00943$. Hence

$$\begin{aligned} Y_C/1.00001 &\leq (1+\epsilon)Z_C + \frac{\epsilon m \ln 3}{\ln(4/3)} + \frac{(1+\epsilon)2m\lambda \ln 2}{\ln(4/3)} + \frac{(1+\epsilon)\ln 2}{n \ln(4/3)} \\ &\leq Z_C \left(1 + \epsilon + \frac{\epsilon \ln 6}{\ln(4/3)} + \frac{4 \ln 2 (1+\epsilon)\lambda}{\ln(4/3)} + O\left(\frac{1}{n^2}\right)\right) \\ &\leq Z_C \left(1.0908907 + O\left(\frac{1}{n^2}\right)\right), \end{aligned}$$

which implies $Y_C/Z_C \leq 12/11 - \epsilon$ for n large enough.

References

[1] E. A. Bender. The asymptotic number of non-negative integer matrices with given row and column sums. *Discrete Mathematics*, 10:217–223, 1974.

[2] Piotr Berman and Marek Karpinski. On some tighter inapproximability results, 1998. (Preprint.)

[3] Martin Dyer and Catherine Greenhill. On Markov chains for independent sets, 1997. (Preprint.)

[4] Mark Jerrum, Leslie Valiant and Vijay Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 43:169–188, 1986.

[5] Mark Jerrum. Large cliques elude the Metropolis process. *Random Structures and Algorithms*, 3:347–359, 1992.

[6] Mark Jerrum and Alistair Sinclair. The Markov chain Monte Carlo method: an approach to approximate counting and integration. In Dorit Hochbaum, editor, *Approximation Algorithms for NP-Hard Problems*, pages 482–520. PWS Publishing, Boston, 1996.

[7] Michael Luby and Eric Vigoda. Approximately counting up to four. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, pages 682–687. ACM Press, 1997.

[8] Johann Håstad. Some optimal inapproximability results. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, pages 1–10. ACM Press, 1997.

Appendix A

We prove the claims made in Section 2 concerning the function ϕ.

(i) From (5), it can easily be checked that $\partial^2\phi/\partial\alpha^2 < 0$ on the interior of $\mathcal{T}$, and hence ϕ can have no interior local minima. On $\alpha = 0$, ϕ has a maximum at $\beta = \frac{1}{2}$ using (4), but then from (3) we find $\partial\phi/\partial\alpha = +\infty$ at $\alpha = 0, \beta = \frac{1}{2}$. Similarly $\beta = 0$. On $\alpha + \beta = 1$, both $\partial\phi/\partial\alpha, \partial\phi/\partial\beta = -\infty$, so ϕ can have no maximum.

(ii) Since both $\partial^2\phi/\partial\alpha^2, \partial^2\phi/\partial\beta^2 < 0$, ϕ has a maximum if and only if the Hessian of ϕ has positive determinant. The condition for this is $\alpha + \beta + \Delta(\Delta - 2)\alpha\beta \leq 1$, as may be checked from (5)–(7).

(iii) From (3) and (4), the conditions for a stationary point of ϕ may be written

$$\beta = f(\alpha),\ \alpha = f(\beta),$$

where

$$f(x) = 1 - x - x^{1/\Delta}(1-x)^{1-1/\Delta} = (1-x)\left[1 - \left(\frac{x}{1-x}\right)^{1/\Delta}\right] \quad (0 \leq x \leq 1).$$

Thus at any stationary point

$$\alpha = f(f(\alpha)). \tag{21}$$

Clearly $f(x) \leq 0$ for $x \geq \frac{1}{2}$, so $\alpha < \frac{1}{2}$ at any stationary point. Similarly $\beta < \frac{1}{2}$. To study the roots of (21), the change of variable $y = (\alpha/(1-\alpha))^{1/\Delta}$

proves convenient. With a little calculation we may express α, $f(\alpha)$, and $f(f(\alpha))$ in terms of y:

$$\alpha = \frac{y^\Delta}{1+y^\Delta}$$

$$f(\alpha) = (1-\alpha)(1-y) = \frac{1-y}{1+y^\Delta}$$

and

$$\begin{aligned} f(f(\alpha)) &= (1-f(\alpha)) - \left(f(\alpha)(1-f(\alpha))^{\Delta-1}\right)^{1/\Delta} \\ &= \left(\alpha + \frac{y}{1+y^\Delta}\right) - \frac{1}{1+y^\Delta}\left((1-y)(y+y^\Delta)^{\Delta-1}\right)^{1/\Delta} \\ &= \alpha + \frac{y}{1+y^\Delta}\left[1 - \left(\frac{(1-y)(1+y^{\Delta-1})^{\Delta-1}}{y}\right)^{1/\Delta}\right]; \end{aligned}$$

so that (21) is equivalent to

$$(1+y^{\Delta-1})^{\Delta-1} = \frac{y}{1-y} \qquad (0 \le y < 1). \tag{22}$$

Note that the implicit mapping from α to y is a bijection, so we may legitimately study the solution set of (21) through that of (22). Note also that (22) has a root y' satisfying $y + y^\Delta = 1$, and this exists for any $\Delta > 0$. The reader may check that $y + y^\Delta = 1$ is equivalent to $\alpha = f(\alpha)$, and thus y' satisfies $\alpha = \beta$. To analyse (22) in general, let

$$g(y) = (\Delta-1)\ln(1+y^{\Delta-1}) + \ln(1-y) - \ln y,$$

so $g(y) = 0$ has the same roots as (22). Then one may check that $g'(y) = 0$ if and only if

$$h(y) \stackrel{def}{=} \Delta(\Delta-2)y^{\Delta-1} - (\Delta-1)^2 y^\Delta - 1 = 0.$$

But $h(0) = -1$, $h(1) = -2$, and h has a single maximum on $[0,1]$ at $y'' = (\Delta-2)/(\Delta-1)$. Now $h(y'') = (\Delta-2)^\Delta/(\Delta-1)^{\Delta-1} - 1 > 0$ if and only if $\Delta \ge 6$, and $h(y'') < 0$ otherwise. Therefore h has two roots in $[0,1]$ if $\Delta \ge 6$, otherwise no roots.. Thus g has a single root in $[0,1]$ if $\Delta \le 5$, otherwise at most three roots. In the latter case, however, $g(0) = +\infty$, $g(1) = -\infty$, $g(y') = 0$ and a simple calculation shows

$$g'(y') = \frac{(\Delta-1)^2(1-y')^2 - 1}{y'(1-y')} > 0$$

if and only if $\Delta \ge 6$, and $g'(y') < 0$ otherwise. These facts imply that g has exactly three roots if $\Delta \ge 6$.

Now the reader may check that the point (α', α') corresponding to y' (i.e., given by solving $y' = (\alpha'/(1-\alpha'))^{1/\Delta}$) satisfies

$$\alpha + \beta + \Delta(\Delta - 2)\alpha\beta \leq 1, \quad \text{i.e.} \quad \left(\frac{1-\alpha}{\alpha}\right)\left(\frac{1-\beta}{\beta}\right) \geq (\Delta - 1)^2,$$

if and only if $y' \geq y''$. This holds if and only if $\Delta \leq 5$. Thus this point is a maximum for $\Delta \leq 5$, otherwise a saddle-point.

Thus ϕ has one stationary point in $\mathcal{T}$ (on $\alpha = \beta$) if $\Delta \leq 5$, and this is a maximum.

(iv) By the above, if $\Delta \geq 6$, ϕ has no boundary maximum on $\mathcal{T}' = \{(\alpha, \beta) \in \mathcal{T} : \alpha \leq \beta\}$ and therefore has a maximum in the interior of $\mathcal{T}'$ by continuity. By symmetry there is also a maximum in $\mathcal{T} \setminus \mathcal{T}'$. Thus, when $\Delta \geq 6$, ϕ has two symmetrical maxima and a single saddle-point on the line $\alpha = \beta$.