A Spectral Approach to Approximately Counting Independent Sets in Dense Bipartite Graphs

Charlie Carlson
Department of Computer Science, University of California Santa Barbara, CA, USA

Ewan Davies[1]
Department of Computer Science, Colorado State University, Fort Collins, CO, USA

Alexandra Kolla
Computer Science and Engineering, University of California Santa Cruz, CA, USA

Aditya Potukuchi
Department of Electrical Engineering and Computer Science, York University, Toronto, Canada

Abstract

We give a randomized algorithm that approximates the number of independent sets in a dense, regular bipartite graph – in the language of approximate counting, we give an FPRAS for #BIS on the class of dense, regular bipartite graphs. Efficient counting algorithms typically apply to "high-temperature" problems on bounded-degree graphs, and our contribution is a notable exception as it applies to dense graphs in a low-temperature setting. Our methods give a counting-focused complement to the long line of work in combinatorial optimization showing that CSPs such as Max-Cut and Unique Games are easy on dense graphs via spectral arguments.

Our contributions include a novel extension of the method of graph containers that differs considerably from other recent low-temperature algorithms. The additional key insights come from spectral graph theory and have previously been successful in approximation algorithms. As a result, we can overcome some limitations that seem inherent to the aforementioned class of algorithms. In particular, we exploit the fact that dense, regular graphs exhibit a kind of small-set expansion (i.e., bounded threshold rank), which, via subspace enumeration, lets us enumerate small cuts efficiently.

2012 ACM Subject Classification Theory of computation → Approximation algorithms analysis; Mathematics of computing → Approximation algorithms; Theory of computation → Algorithm design techniques

Keywords and phrases approximate counting, independent sets, bipartite graphs, graph containers

Digital Object Identifier 10.4230/LIPIcs.ICALP.2024.35

Category Track A: Algorithms, Complexity and Games

Related Version *Full Version*: `https://arxiv.org/abs/2307.09533`

Funding *Ewan Davies*: Supported in part by NSF grant CCF-2309707.
Aditya Potukuchi: Supported in part by NSERC Discovery grants RGPIN-2023-05087 & DGECR-2023-00408.

1 Introduction

Exactly computing the number $i(G)$ of independent sets in a graph G is #P-hard, even when restricted to bipartite graphs [41]. In the general case, approximating $i(G)$ (to within, say, a constant factor) is NP-hard, even when restricted to d-regular graphs with $d \geq 6$ [20, 46, 45]. Restricted to *bipartite* graphs the problem of counting independent sets is known as #BIS, and the prospect of hardness of approximation is less clear because finding a maximum

[1] Corresponding author

51st International Colloquium on Automata, Languages, and Programming (ICALP 2024).
Editors: Karl Bringmann, Martin Grohe, Gabriele Puppis, and Ola Svensson;
Article No. 35; pp. 35:1–35:18
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

independent set can be done in polynomial time. Under polynomial-time approximation-preserving reductions, many natural counting problems are equivalent to #BIS [17], and the complexity of approximating #BIS has received a lot of attention. Existing approximation algorithms for #BIS include "high-temperature" algorithms that work when degrees on one side of the bipartition are small [38], "low-temperature" algorithms that require additional assumptions such as expansion [11, 29] or unbalanced degrees [8], and exponential-time algorithms that are nonetheless faster than algorithms for the general, non-bipartite case [24]. The description of these methods in terms of temperature is due to a common generalization in terms of weighted counting and strong connections to statistical physics, where counting (weighted) independent sets corresponds to computing the partition function of the hard-core model.

The idea that Max-CSP optimization problems such as Max-Cut and Unique Games should be easy to approximate on dense graphs – perhaps because they have good expansion properties – is well-established [3, 18, 19]. Many of the techniques that apply to dense or expanding graphs have been generalized in interesting directions. In particular, spectral methods give good results in both dense graphs and expanders, and in many cases can be extended to more refined structural properties such as small-set expansion and threshold rank to great effect. Most of the prominent approaches to Max-CSPs relevant to this work fall into three categories: algorithmic regularity lemmas which began with Frieze and Kannan [19] and were extended to threshold rank by Oveis Gharan and Trevisan [39]; convex hierarchies and correlation rounding [4, 6, 25]; and the spectral technique of subspace enumeration due to Kolla and Tulsiani [36, 37]. Prior to these developments were several algorithms demonstrating that counting problems on dense graphs admit efficient approximation algorithms [1, 16, 33], though these results do not apply to counting independent sets.

An analogous theme in approximate counting is to obtain algorithms on expander graphs or random graphs [7, 10, 21, 26, 29]. Despite superficial similarity to the aforementioned work on Max-CSPs in the sense that these works give algorithms for dense or expanding instances, there is relatively little work establishing any common underlying phenomenon that makes Max-CSP problems and counting problems easy on dense or expanding graphs. A notable exception is due to Risteski [42], who connected the work on correlation rounding and convex hierarchies [6] to the broad and well-studied problem of approximating partition functions. His approach is also known as the variational method. Regularity methods and correlation rounding do provide some evidence of structure common to these problems; for example, Coja–Oghlan and various coauthors have developed a range of regularity lemmas and applied them to both Max-CSPs and spin models on random graphs [5, 13, 14], and Coja–Oghlan and Perkins independently discovered correlation rounding in the context of Gibbs measures and partition functions [15]. Counting independent sets is not typically one of the examples studied, though occasionally this is more for convenience than for fundamental reasons.

In the specific context of #BIS, connections to Max-CSP research are even more scarce. The polymer approach of Jenssen, Keevash and Perkins [29] is a major algorithmic breakthrough for #BIS which shows that several prominent #BIS-hard problems can be approximated in polynomial time on bounded-degree expander graphs (and thus random d-regular graphs for $d = O(1)$). Further refinements of the method broaden the range of problems covered [21, 26], provide faster algorithms based on rapid mixing of Markov chains known as polymer dynamics [11], or weaken the structural properties required by applying container theorems to combinatorial enumeration problems that arise in the method [10, 32]. None of these developments give polynomial-time algorithms in dense graphs, however. Carlson, Davies, and Kolla [9] applied the polymer method to approximate the Potts model partition function on (bounded-degree) graphs with bounded threshold rank, but the conditions their analysis requires are prohibitively restrictive, and it is unclear whether their techniques can be applied to #BIS. While Risteski's approach has been extended and improved [28, 35], results are stated for spin models with soft constraints such as the Ising and Potts models, and the approximation guarantees degrade in the presence of the hard constraints that are inherent to independent sets.

1.1 Main result

We specifically address the superficial similarities between algorithms for Max-CSPs and counting independent sets by giving an algorithm for approximately counting independent sets in dense, regular bipartite graphs which combines the highly successful techniques of polymer models, subspace enumeration, and container theorems for the enumeration of independent sets in bipartite graphs. Our approximation guarantee is of the strong type typically sought in approximate counting. We say that a relative ϵ-approximation of a real number x is a real number y such that $e^{-\epsilon} \leq x/y \leq e^{\epsilon}$, and a fully polynomial randomized approximation scheme (FPRAS) for a counting problem is an algorithm that with probability at least $3/4$ outputs a relative ϵ-approximation to the solution in time polynomial in the instance size and $1/\epsilon$.

▶ **Theorem 1.** *For each $\delta \in (0, 1)$ there is an FPRAS for #BIS on the class of d-regular bipartite graphs G with $d = \lfloor \delta |V(G)|/2 \rfloor$.*

We use spectral methods and subspace enumeration to enumerate small cuts in d-regular bipartite graphs via an ϵ-net of the vector space spanned by small eigenvalues of the Laplacian matrix of the graph, influenced by the use of these methods in combinatorial optimization [2, 36, 37] and approximate counting. Some of our analysis builds upon the perturbative approach of [27, 29] and an important refinement of this method due to Jenssen and Perkins [30] (and with Potukuchi [31]) that uses graph container lemmas of the type developed by Sapozhenko [43, 44]. While container theorems for independent sets have been used to control enumeration problems that arise in establishing the convergence of the cluster expansion [30, 31, 32], and these have inspired container-like theorems for controlling analogous enumeration problems [10], our addition of subspace enumeration here has a different purpose.

In terms of running time, our result improves upon the dense case of an algorithm of Jenssen, Perkins, and Potukuchi [32] which runs in subexponential time on d-regular bipartite graphs for all $d \geq \omega(1)$. In the case $d = \Theta(n)$ their algorithm takes time $\exp(\Omega(\log^4 n))$, and our contribution works for any accuracy parameter ϵ, which is not given by the methods in [32]. The improvement stems from incorporating the spectral techniques mentioned above, which lets us sidestep algorithmic cluster expansion. That is, our spectral techniques overcome an obstacle in the algorithm of [32] related to polynomial accuracy: we can achieve arbitrary accuracy without resorting to a naive enumeration of polymers (which in this setting are connected subgraphs of the square of the instance).

An interesting question posed in [32] is whether #BIS admits a general subexponential-time algorithm. One of our technical contributions is to show that a perspective on graph spectra involving higher-order eigenvalues and eigenvectors advances our understanding of #BIS.

2 Overview

Fix $\delta > 0$ and for $d = \lfloor \delta n \rfloor$, a bipartite graph $G = (X \cup Y, E)$ on $2n$ vertices. Let $\epsilon > 0$ and note that we allow ϵ to depend on n.

Our proof begins with the well-known observation that to enumerate independent sets in a bipartite graph it suffices to enumerate deviations from the "ideal" independent set X. That is, we have the identity

$$i(G) = \sum_{A \subseteq X} 2^{|Y \setminus N(A)|} \tag{1}$$

because, for a fixed $A \subseteq X$, any vertex of $Y \setminus N(A)$ can be added to A without spanning an edge. There is a similar formula for $i(G)$ based on enumerating deviations from Y. An important achievement of [29] is to give a rigorous proof that in bipartite graphs with strong expansion, typical independent sets are small deviations from either X or Y. Intuitively, we see a hint of this idea in equation (1) as when G is an expander we expect that $N(A) \gg |A|$ and so the terms on the right-hand side are small unless $|A|$ is small. Given this, one might hope to obtain an algorithm provided one can solve the problem of efficiently enumerate the small deviations and quantifing their contributions to $i(G)$. This is done in [29] by approximating $i(G)$ with the sum of two polymer models, and brute force enumeration of terms in the cluster expansion for these models.

If the bipartite graph is not an expander, then large deviations from X and Y must be handled. For example, in a $2n$-vertex disjoint union of complete d-regular bipartite graphs, a significant number of independent sets intersect both X and Y on $\Omega(n)$ vertices. To extend the algorithm to all bipartite graphs, using an idea from [32] we can separate contributions from expanding and non-expanding pieces of the deviation A. The first step is to break $A \subseteq X$ in the sum in (1) into pieces with disjoint neighborhoods. We say that a subset $A \subseteq X$ is *polymer*[2] if it is connected in the square G^2 of G, and note that any $A \subseteq X$ admits a unique partition into polymers which have disjoint neighborhoods. We call the polymers in this partition the *components* of A and denote the set of components of A by $\mathcal{K}(A)$. We say that two polymers are *compatible* if their neighborhoods are disjoint, and that a set or tuple of polymers is *compatible* if the polymers in it are pairwise compatible. Thus, subsets $A \subseteq X$ correspond to compatible sets of polymers via the unique partition into polymers with disjoint neighborhoods.

▷ Claim 2.

$$i(G) = \sum_{k \geq 0} \frac{1}{k!} \sum_{\substack{(A_1,\ldots,A_k) \text{ s.t.} \\ \text{each } A_i \text{ is a polymer and} \\ (A_1,\ldots,A_k) \text{ compatible}}} 2^{\left|Y \setminus \bigcup_{j=1}^{k} N(A_j)\right|}. \tag{2}$$

Proof. The claim follows from the correspondence between subsets $A \subseteq X$ and sets of compatible polymers given by $A \mapsto \mathcal{K}(A)$. By convention, we sum over compatible tuples of polymers which leads to the term $1/k!$ to account for the permutations of each tuple. We use the fact that compatible polymers have disjoint neighborhoods for the correspondence of the summands. ◁

The *closure* $[A]$ of a subset $A \subseteq X$ is $[A] := \{x \in X : N(x) \subseteq N(A)\}$, and we say that A is *closed* if $A = [A]$. Note that A is closed if and only if each component of A is closed. A subset $A \subseteq X$ is called *t-expanding* if $|N(A)| = |[A]| + t$, and (in a slight abuse of terminology

[2] In related works the term "2-linked" is used for the property of being connected in G^2.

that we hope the reader permits) t-contracting if $|N(A)| < |[A]| + t$. For a fixed t_0 that we determine later, we split the sum over polymers in (2) according to t_0-contraction. To do this, for each subset $A \subseteq X$, let $X_A = X \setminus N(N(A))$ and $Y_A = Y \setminus N(A)$. Let $\mathcal{P}_A$ be the set of polymers which are subsets of X_A and define

$$\Xi(A) := \sum_{k\geq 0} \frac{1}{k!} \sum_{\substack{(B_1,\ldots,B_k) \text{ s.t.} \\ B_i \in \mathcal{P}_A \text{ is not } t_0\text{-contracting} \\ \text{and } (B_1,\ldots,B_k) \text{ compatible}}} 2^{-\sum_{i=1}^k |N(B_i)|},$$

where the inner sum is over compatible k-tuples of polymers, each of which is not t_0-contracting (equivalently, t-expanding for some $t \geq t_0$).

▷ Claim 3.

$$i(G) = \sum_{k\geq 0} \frac{1}{k!} \sum_{\substack{(A_1,\ldots,A_k) \text{ s.t.} \\ \text{each } A_i \text{ is a } t_0\text{-contracting polymer} \\ \text{and } (A_1,\ldots,A_k) \text{ compatible}}} 2^{\left|Y \setminus \bigcup_{j=1}^k N(A_j)\right|} \cdot \Xi\left(\bigcup_{j=1}^k A_j\right). \tag{3}$$

Proof. From Claim 2 we can split the sum over tuples of polymers into a sum over tuples of t_0-contracting polymers and tuples of non-t_0-contracting polymers. The idea is to first sum over tuples $(A_1, \ldots, A_k)$ of t_0-contracting polymers and then use the fact that for $A = \bigcup_{j=1}^k A_j$ the quantity $\Xi(A)$ contains a sum over the ways to extend this tuple to one containing non-t_0-contracting polymers and the summand is the additional contribution that each such extension makes. The definition of $\mathcal{P}_A$ means that any $B_i \in \mathcal{P}_A$ is compatible with each component of A. With a little care, one can check that the permutaions of the tuples are correctly taken into account and the claim follows. ◁

A further refinement of the expression for $i(G)$ groups t_0-contracting polymers according to their neighborhoods. The motivation for this is that two subsets $A \subseteq X$ and $B \subseteq X$ (each corresponding to the union of some compatible tuple of t_0-contracting polymers) have the same contribution in the sum if $N(A) = N(B)$ because this implies that $\Xi(A) = \Xi(B)$. For a subset $A \subseteq X$ write

$$\mathcal{D}(A) := \prod_{A' \in \mathcal{K}(\mathcal{A})} \left|\{B' \subseteq A' : B' \text{ is a polymer and } N(B') = N(A')\}\right|.$$

The quantity $\mathcal{D}(A)$ counts the number of subsets B of A such that $N(B) = N(A)$ and which are formed by choosing for each component A' of A, a subset $B' \subseteq A'$ which is a polymer. Note that if A' is t_0-contracting then so is any polymer $B' \subseteq A'$ with $N(B') = N(A')$. For convenience, we define $\mathcal{A}$ to be the set of all $A \subseteq X$ with closed, t_0-contracting components.

▷ Claim 4.

$$i(G) = \sum_{A\in\mathcal{A}} \mathcal{D}(A) \cdot 2^{|Y \setminus N(A)|} \cdot \Xi(A), \tag{4}$$

Proof. From Claim 3 we can restrict the sum over tuples of t_0-contracting polymers to closed t_0-contracting polymers provided, for each compatible tuple $(A_1, \ldots, A_k)$ of closed t_0-contracting polymers, we multiply their contribution to the sum by a term counting the number of ways of getting that contribution with polymers that are not necessarily closed. Identifying compatible tuples of closed polymers with their union, i.e. setting $A = \bigcup_{j=1}^k A_j$, the contribution to the sum from A is $2^{|Y \setminus N(A)|} \cdot \Xi(A)$. The term $\mathcal{D}(A)$ is exactly the number of ways of getting this contribution. The claim follows from the conversion of the sum back into one over suitable subsets of X, namely those in $\mathcal{A}$, instead of a sum over compatible tuples of polymers. ◁

Now that we have a suitable expression (4) for $i(G)$, we can describe how our algorithm approximates $i(G)$. Our algorithm simply enumerates the sets $A \in \mathcal{A}$, approximates each $\mathcal{D}(A)$ term, and uses the fact (which we must prove) that 1 is a good approximation of each $\Xi(A)$ to approximate $i(G)$. Given these subroutines, computing the sum (4) is straightforward. The analysis of our algorithm thus splits into three separate components. Recall that the input is a d-regular bipartite graph G on $2n$ vertices such that for some constant $\delta > 0$ we have $d = \lfloor \delta n \rfloor$, and an approximation error ϵ. We set $t_0 = C \log(n/\epsilon)$, where $C = C(\delta)$ is large enough, and the correctness and running time of our algorithm follows from the results below. Note that for this choice of t_0 an exponential such as 4^{t_0} is polynomial in n and $1/\epsilon$.

▶ **Lemma 5.** *For $t_0 \le 2^{-8}d$, the set $\mathcal{A} = \{A \subseteq X : A$ closed and t_0-contracting$\}$ has size at most $n^{O(1/\delta)} \cdot 4^{t_0}$ and can be enumerated in the same time.*

The proof of this lemma uses subspace enumeration to find small cuts in G, and then for each such small cut enumerates the sets $A \in \mathcal{A}$ which are close to the cut. See Section 4.

▶ **Lemma 6.** *Let $A \subseteq X$ be a closed t_0-contracting polymer. Then for $\epsilon', \rho' > 0$ there is a randomized algorithm running in time polynomial in n, $1/\epsilon'$ and $\log(1/\rho')$ that with probability at least $1-\rho'$ outputs a relative ϵ'-approximation to the number of polymers $B \subseteq A$ such that $N(B) = N(A)$.*

This lemma uses straightforward estimation of an expectation by repeated sampling, and is very similar to the analogous result in [32]. The proof is in Section 5. We use the lemma in each component of the sets $A \in \mathcal{A}$ in the claim below. This claim requires an upper bound on t_0, but this is a small technical detail as the only way to violate this bound is to choose an error parameter ϵ so small that one has time for brute force because an FPRAS can take time polynomial in $1/\epsilon$, see Section 3 where we use the claim.

▷ Claim 7. Suppose that $t_0 \le d/2$. Then each set $A \in \mathcal{A}$ has at most $2/\delta = O(1)$ components, and for $\epsilon', \rho' > 0$ there is a randomized algorithm running in time polynomial in n, $1/\epsilon'$ and $\log(1/\rho')$ that, given a set $A \in \mathcal{A}$ as input, with probability $1 - \rho'$ obtains an ϵ'-approximation of $\mathcal{D}(A)$.

Proof. Any t_0-contracting set must have size at least $d - t_0$, and in the case $t_0 \le d/2$ we have $d - t_0 \ge d/2$ and hence each $A \in \mathcal{A}$ has at least $2n/d = 2/\delta$ components.

Observe that if $A \in \mathcal{A}$ has ℓ components then running the algorithm of Lemma 6 on each component with error parameter ϵ'/ℓ and probability parameter ρ'/ℓ yields, with probability at least $1-\rho'$, a relative ϵ-approximation to $\mathcal{D}(A)$ in time polynomial in n, ℓ/ϵ' and $\log(\ell/\rho')$. When $t_0 \le d/2$ we have the upper bound $\ell = O(1)$ from above and the claim follows. ◁

▶ **Lemma 8.** *Let $A \in \mathcal{A}$, then $1 \le \Xi(A) \le e^{\epsilon/2}$.*

This result means that 1 is a relative $\epsilon/2$-approximation for each of the $\Xi(A)$ terms appearing in (4). The proof is based on graph container methods due to Sapozhenko [43, 44], which have since been refined, [23, 22, 34, 40], and their application to algorithmic counting [30, 31, 32]. We give the proof in Section 6.

3 The algorithm and proof of Theorem 1

Input A $\lfloor \delta n \rfloor$-regular bipartite graph $G = (X \cup Y, E)$ on $2n$ vertices and an approximation error $\epsilon > 0$.
Output A relative ϵ-approximation i' of $i(G)$.

Recall that $C = C(\delta)$ is a large enough constant, and that $t_0 = C \log(n/\epsilon)$. In the following proof, implicit constants in the $O(\cdot)$ notation and implicit polynomials are allowed to depend on δ but not ϵ. If $\epsilon \le n \exp(-d/(2^8 C))$ then we can afford to run a brute force algorithm that computes $i(G)$ exactly in time $e^{O(n)}$ and the running time is still polynomial in $1/\epsilon$. Otherwise, we note that for all large enough n we have $d - 2^7 t_0 \ge d/2$ and run the following algorithm. For convenience, we assume that $\epsilon \le 1$ and simply run the algorithm for $\epsilon = 1$ if the given ϵ is larger.

First, construct the set $\mathcal{A}$, which can be done in time $(n/\epsilon)^{O(1)}$ by Lemma 5. Note also that $|\mathcal{A}|$ is polynomial in n and $1/\epsilon$. Then, for each $A \in \mathcal{A}$ compute an $\epsilon/2$-approximation $\tilde{\mathcal{D}}_A$ of $\mathcal{D}_A$ with the algorithm of Claim 7 and probability parameter $\rho' = 3/(4|\mathcal{A}|)$. Then $\log(1/\rho')$ is polynomial in $\log n$ and $\log(1/\epsilon)$ so the running time of this step is polynomial in n and $1/\epsilon$. By a union bound, with probability at least $3/4$ we get the desired approximation in each application of the claim, and thus a valid relative $\epsilon/2$-approximation $\tilde{\mathcal{D}}_A$ of each $\mathcal{D}_A$. Then output $i' = \sum_{A \in \mathcal{A}} \tilde{\mathcal{D}}_A 2^{|Y \setminus N(A)|}$. By Lemma 8 and the analysis above, the output is a valid ϵ-approximation of $i(G)$ obtained in time $(n/\epsilon)^{O(1)}$, thus proving Theorem 1.

4 Subspace enumeration and contracting sets

The proof of Lemma 5 has two parts. First, we show how to enumerate small cuts using subspace enumeration. For related results see [2, 36, 37]. We use the term *cut* to mean a subset of $V = X \cup Y$, and the *value* $|\nabla(C)|$ of a cut C is the number of edges with precisely one endpoint in C. Subspace enumeration involves what is commonly called an ϵ-net of a subset U' of a vector space, which is a collection of points such that U' is contained in the union of the balls of radius ϵ around each point. Since we reserve ϵ for the error parameter in our algorithm, our nets are ξ-nets.

▶ **Lemma 9.** *Let $G = (V, E)$ be a d-regular bipartite graph on $N = 2n$ vertices. There is a set $\mathcal{C}^{\text{cut}} \subseteq 2^V$ such that $|\mathcal{C}^{\text{cut}}| \le n^{O(n/d)}$ and $\mathcal{C}^{\text{cut}}$ has the following property. For all $t \ge 1$ and cuts $S \subseteq V$ with value $|\nabla(S)| \le td$, there is some $C \in \mathcal{C}^{\text{cut}}$ such that $|S \triangle C| \le 32t$ and $|\nabla(C)| \le 33td$. Moreover, the set $\mathcal{C}^{\text{cut}}$ can be constructed in time $n^{O(n/d)}$.*

Proof. Let $d = \lambda_1 \ge \cdots \ge \lambda_N = -d$ be the spectrum of the adjacency matrix A of G. The facts that $\lambda_1 = d = -\lambda_N$ and that the spectrum of A is symmetric about zero are standard, see e.g. [12]. Let k be such that A has precisely $2k$ eigenvalues of absolute value at least $d/2$. Counting closed walks of length two gives

$$\mathrm{Tr}(A^2) = Nd = \sum_{i=1}^{N} \lambda_i^2 \ge kd^2/2,$$

and hence $k \le 4n/d$.

Let $L = dI - A$ be the Laplacian matrix of G and let $\mathbf{e}_1, \ldots, \mathbf{e}_N$ be an orthonormal basis of eigenvectors of L such that $\mathbf{e}_i$ has eigenvalue μ_i with $0 = \mu_1 \le \cdots \le \mu_N = 2d$. By the definition of k, it must be the case that $\mu_{k+1} > d/2$. Let U be the span of $\mathbf{e}_1, \ldots \mathbf{e}_k$, and $U^\perp$ be the orthogonal complement of U. For $\xi = \sqrt{2}$, we require an efficient construction of a ξ-net $\mathcal{E} \subseteq U$ covering all vectors of L^2-norm at most $\sqrt{n}$ in U. For example, we can take

$$\mathcal{E} := \left\{ \mathbf{p} = \sum_{i=1}^{k} x_i \mathbf{e}_i : x_1, \ldots, x_k \in (\xi/\sqrt{k}) \cdot \mathbb{Z},\ \|\mathbf{p}\| \le \sqrt{n} \right\},$$

yielding $|\mathcal{E}| \le (2\sqrt{nk}/\xi)^k$. Then every vector in U with L^2-norm at most $\sqrt{n}$ lies at most distance ξ from a vector in $\mathcal{E}$.

The algorithm to construct $\mathcal{C}^{\mathrm{cut}}$ is as follows. Start with $\mathcal{C}^{\mathrm{cut}} = \emptyset$ and for each point $\mathbf{p} \in \mathcal{E}$, form $\mathbf{p}'$ by rounding each coordinate of $\mathbf{p}$ to $\{0,1\}$ (breaking ties with $1/2 \mapsto 1$) and add the vertex subset with indicator vector $\mathbf{p}'$ to $\mathcal{C}^{\mathrm{cut}}$.

We now show that $\mathcal{C}^{\mathrm{cut}}$ has the desired properties. By the construction of $\mathcal{C}^{\mathrm{cut}}$ and $\mathcal{E}$ we have $|\mathcal{C}^{\mathrm{cut}}| \le |\mathcal{E}| \le n^{O(n/d)}$. To establish the other property of $\mathcal{C}^{\mathrm{cut}}$, let $t \ge 1$ and consider an arbitrary subset $S \subseteq V$ with $|\nabla(S)| \le td$. Let $\mathbf{s}$ be the indicator vector of the set S and write this vector in the eigenbasis of L as $\mathbf{s} = \sum_{i=1}^{N} s_i \mathbf{e}_i$. Let $\mathbf{u} = \sum_{i=1}^{k} s_i \mathbf{e}_i$ be the projection of $\mathbf{s}$ onto U and let $\mathbf{p}$ be the point in $\mathcal{E}$ closest to $\mathbf{u}$. Indicator vectors of subsets of V have L^2-norm at most $\sqrt{n}$, and hence $\|\mathbf{u} - \mathbf{p}\| \le \xi$.

Without considering our need for an efficient construction, the idea is that because $\nabla(S)$ is small we know that $\mathbf{s}$ is an indicator vector close to its projection $\mathbf{u}$ onto U. Thus, if we form $\mathcal{C}^{\mathrm{cut}}$ as the union of all sets whose indicator vectors are close to vectors in U, each set S of interest has an indicator vector that lies within a distance twice the definition of "close" to a set in C.

To make the above sketch efficient, we replace U with the ξ-net $\mathcal{E}$. Note that

$$td \ge |\nabla(S)| = \mathbf{s}^{\mathrm{T}} L \mathbf{s} = \sum_{i=1}^{N} \mu_i s_i^2 \ge \frac{d}{2} \sum_{i=k+1}^{N} s_i^2.$$

But $\sum_{i=k+1}^{N} s_i^2 = \|\mathbf{s} - \mathbf{u}\|^2$, so we have the bound $\|\mathbf{s} - \mathbf{u}\| \le \sqrt{2t}$. Then we immediately have $\|\mathbf{s} - \mathbf{p}\| \le \sqrt{2t} + \xi$ from the triangle inequality. Let $\mathbf{p}'$ be obtained from $\mathbf{p}$ by rounding each coordinate to $\{0,1\}$, breaking ties with $1/2 \mapsto 1$, and let $C \subseteq V$ be the set whose indicator vector is $\mathbf{p}'$. We have $|S \triangle C| = \|\mathbf{s} - \mathbf{p}'\|^2$ and we bound the latter with the triangle inequality. In particular, $\mathbf{s}$ is an indicator vector of distance at most $\sqrt{2t} + \xi$ from $\mathbf{p}$ and $\mathbf{p}'$ must be the closest indicator vector to $\mathbf{p}$, hence $\|\mathbf{p} - \mathbf{p}'\| \le \sqrt{2t} + \xi$. Then $\|\mathbf{s} - \mathbf{p}\| \le 2(\sqrt{2t} + \xi)$, and because $t \ge 1$ and $\xi = \sqrt{2}$ we have

$$|S \triangle C| \le 4\left(\sqrt{t} + \xi\right)^2 \le 32t.$$

It remains to bound the value of the cut $|\nabla(C)|$, and the desired bound follows from the observation that

$$|\nabla(C)| \le |\nabla(S)| + d|S \triangle C| \le td + 32td = 33td. \qquad ◀$$

Lemma 9 tells us that there is an efficient construction of a collection $\mathcal{C}^{\mathrm{cut}}$ of cuts such that any small cut S must be close to a cut in $\mathcal{C}^{\mathrm{cut}}$ in Hamming distance. We now show that given a small cut S we can enumerate the sets $A \in \mathcal{A}$ which are close to S. For this to be useful, it must be that each $A \in \mathcal{A}$ is close to some small cut, and we give the details of this later.

▶ **Lemma 10.** *Fix any $c \ge 1$ and let $t \le \frac{d}{8c}$. Given a cut C with value at most td, there are at most 4^t closed t-contracting subsets $A \subseteq X$ such that $|A \triangle (C \cap X)| \le ct$ and $|N(A) \triangle (C \cap Y)| \le ct$. Moreover, these sets A can be enumerated in time $4^t \cdot n^{O(1)}$.*

Proof. Let $A' := C \cap X$ and $W' := C \cap Y$. By the fact that G is d-regular, $|E(A', W')| \leq d \min\{|A'|, |W'|\}$ and hence $|\nabla(C)| \geq d \max\{|W'| - |A'|, |A'| - |W'|\}$. By assumption, we have $|\nabla(C)| \leq td$ and therefore $\big||W'| - |A'|\big| \leq t$.

Set

$$S_X := \{v \in X \setminus A' : |N(v) \setminus W'| \leq 3ct\}, \text{ and}$$
$$S_Y := \{v \in W' : |N(v) \cap A'| \leq ct\},$$

so that $S_X \subseteq X$ consists of vertices in $X \setminus A'$ with almost all of their neighbors in W' and $S_Y \subseteq Y$ consists of vertices in W' with almost all of their neighbors in $X \setminus A'$. We have the following claims.

▷ Claim 11. For any closed t-contracting subset $A \subseteq X$ such that $|A \triangle A'| \leq ct$, $A \setminus A' \subseteq S_X$.

Proof. Suppose for contradiction that there is a vertex $v \in A \setminus A'$ such that

$$|N(v) \setminus W'| > 3ct.$$

We derive the contradiction using the facts that $|\nabla(A \cap A')| = d|A \cap A'|$ and that any of the edges in $\nabla(A \cap A')$ not incident to W' contribute to the value of the cut C. These facts imply that $|E(A \cap A', W')| \geq d|A \cap A'| - t \cdot d$, and hence

$$|N(A \cap A') \cap W'| \geq |A \cap A'| - t.$$

Then because A is closed and non-expanding,

$$\begin{aligned} |A| + t \geq |N(A)| &\geq |N((A \cap A') \cup \{v\})| \\ &> |N((A \cap A') \cup \{v\}) \cap W'| + 3ct \\ &\geq |A \cap A'| + 2ct \geq |A| + ct, \end{aligned}$$

which is a contradiction because there is a strict inequality in the chain and $c \geq 1$. ◁

▷ Claim 12. For any t-contracting subset $A \subseteq X$ such that $|A \triangle A'| \leq ct$, $W' \setminus N(A \cap A') \subseteq S_Y$.

Proof. We note that for each vertex v in $W' \setminus N(A \cap A')$, we have that

$$N(v) \cap A' \subseteq A' \setminus A.$$

Since $|A' \setminus A| \leq ct$, it follows that $|N(v) \cap A'| \leq ct$. ◁

We can now complete the proof of the lemma. Using the degree constraints in the definitions of S_X and S_Y, we have

$$\begin{aligned} td &\geq |\nabla(C)| \\ &\geq |S_X|(d - 3ct) + |S_Y|(d - ct) \\ &\geq (d/2) \cdot (|S_X| + |S_Y|) \end{aligned}$$

where the last inequality uses $t < \frac{d}{8c}$. As a result, we have

$$|S_X| + |S_Y| \leq 2t.$$

Putting Claim 11 and Claim 12 together, we have that each closed t-contracting sets A with $|A \triangle A'|$, $|N(A) \triangle W'| \leq ct$ must be of the form

$$A = [(A' \setminus N(S'_Y)) \cup S'_X]$$

for some subsets $S'_Y \subseteq S_Y$ and $S'_X \subseteq S_X$. Thus, the total number of such A is at most $2^{|S_X| + |S_Y|} \leq 4^t$. Since we are given the cut C, S_X and S_Y can be found in time polynomial in n as required. ◀

With these ingredients we can proof Lemma 5, which we recall states that $\mathcal{A}$ can be enumerated in time $n^{O(1/\delta)}4^{t_0}$.

Proof of Lemma 5. Since $d = \lfloor \delta n \rfloor$, we construct $\mathcal{C}^{\text{cut}}$ as in Lemma 9 in time $n^{O(1/\delta)}$. We then choose $c = 32$ and enumerate for each $C \in \mathcal{C}^{\text{cut}}$, every closed t_0-contracting subset A with $|A \triangle (C \cap X)| \leq 32t_0$ and $|N(A) \triangle (C \cap Y)| \leq 32t_0$ using Lemma 10. We are done if every $A \in \mathcal{A}$ appears in this enumeration process, as the running times combine to give the required $n^{O(1/\delta)}4^{t_0}$. This holds because each $A \in \mathcal{A}$ is closed and t_0-contracting and hence setting $S_A = A \cup N(A)$, by a double counting argument, we have $|\nabla(S_A)| = d|N(A)| - d|A| \leq t_0 d$. So each $A \in \mathcal{A}$ corresponds to a cut of value at most $t_0 d$ and hence some $C \in \mathcal{C}^{\text{cut}}$ has $|S_A \triangle C| \leq 32t_0$ by Lemma 9. ◀

5 Approximating the number of covers

For convenience, we restate Lemma 6 here.

▶ **Lemma 6.** *Let $A \subseteq X$ be a closed t_0-contracting polymer. Then for $\epsilon', \rho' > 0$ there is a randomized algorithm running in time polynomial in n, $1/\epsilon'$ and $\log(1/\rho')$ that with probability at least $1 - \rho'$ outputs a relative ϵ'-approximation to the number of polymers $B \subseteq A$ such that $N(B) = N(A)$.*

Proof. The method is exactly the same as [32, Lem. 17], but in our setting with $d = \lfloor \delta n \rfloor$ the resulting algorithm runs in time polynomial in n.

Let $|A| = a$, $N(A) = W$ have size $|W| = w$, and let $W' = \{v \in W : |N(v) \cap A| \leq d/2\}$ have size $|W'| = w'$. Let

$$\mathcal{D} = \{B \subseteq A : N(B) = W \text{ and } B \text{ is a polymer}\}$$

be the set whose size we wish to estimate.

By [32, Cor. 10], there is a polymer $A' \subseteq A$ of size at most

$$\frac{2a}{d} \log d + \frac{2w}{d} + 2(w - a) \leq \frac{2}{\delta}(1 + \log n) + 2t_0$$

such that $N(A') = W$. Then $|\mathcal{D}| \geq 2^{a - \left(\frac{2}{\delta}(1+\log n) + 2t_0\right)}$, because any subset of A which contains A' is a polymer. Now $|\mathcal{D}|$ can be estimated to relative error ϵ' with probability at least $1 - \rho$ by sampling

$$\frac{1}{(\epsilon')^2} \log(1/\rho) n^{O(1/\delta)} 4^{t_0}$$

subsets of A uniformly at random, and this can be proved with a suitable application of the Chernoff bound. ◀

6 Enumerative lemmas

In this section we prove Lemma 8 which states that for $A \in \mathcal{A}$ we have $1 \leq \Xi_A \leq e^{\epsilon/2}$.

Proof of Lemma 8. For the proof, we fix an arbitrary $A \in \mathcal{A}$. The terms in the sum giving Ξ_A are non-negative, and the lower bound comes from the term $k = 0$ which contributes 1. For the upper bound, we use recent results on graph containers and adapt them to our purposes.

Recall that a polymer is a 2-linked subset $B \subseteq X$ and that the function Ξ_A involves a sum over tuples of non-t_0-contracting polymers. For convenience, we define $\mathcal{G}(w,t)$ to be the set of t-expanding polymers with neighborhood size w,

$$\mathcal{G}(w,t) = \{B \subseteq X, \text{ polymer} : |N(B)| = w, |N(B)| - |[B]| = t\}.$$

In terms of this notation, we have

$$\Xi_A = \sum_{k\geq 0} \sum_{\substack{\{B_1,\ldots,B_k\}\in\mathcal{P}_A \text{ compatible} \\ \text{s.t. each } B_i \text{ not } t_0\text{-contracting}}} 2^{-\sum_{i=1}^k |N(B_i)|} \tag{5}$$

$$\leq \sum_{k\geq 0} \frac{1}{k!} \left(\sum_{t\geq t_0} \sum_{w\geq 0} |\mathcal{G}(w,t)| 2^{-w} \right)^k, \tag{6}$$

where we drop the requirement on the tuples of being compatible and relax the requirement that the B_i are subsets of X_A to being subsets of X, and hence have an upper bound. To proceed, we require upper bounds on $|\mathcal{G}(w,t)|$ and split into two cases according to t. The following result is proved in the rest of this section and Appendix A.

▶ **Lemma 13.** *There is an absolute constant $\gamma > 0$ such that for $t_0 \leq t$, and any integer w,*

$$|\mathcal{G}(w,t)| \leq 2^{w-\gamma t}.$$

With this lemma in hand, and because each neighborhood size w that we see is in $[1,n]$, there is an absolute constant $\gamma > 0$ such that

$$\Xi_A \leq \sum_{k\geq 0} \frac{1}{k!} \left(\sum_{t\geq t_0} n 2^{-\gamma t} \right)^k \tag{7}$$

$$= \sum_{k\geq 0} \frac{1}{k!} \left(n \frac{2^{-\gamma t_0}}{1-2^{-\gamma}} \right)^k = \exp\left(n \frac{2^{-\gamma t_0}}{1-2^{-\gamma}} \right). \tag{8}$$

This at most the required $e^{\epsilon/2}$ provided that

$$t_0 \geq \frac{1}{\gamma} \log_2 \left(\frac{2}{1-2^{-\gamma}} \frac{n}{\epsilon} \right),$$

which our choice $t_0 = C\log(n/\epsilon)$ satisfies for all large enough constants $C = C(\delta)$. ◀

Before we proceed with the proof of Lemma 13, we would like to remark out that one of the main contributions of this paper is to handle the case when t is small.

Proof of Lemma 13. We first take care of the case when $t \geq \log^4 n$. For each $v \in V$, let us define

$$\mathcal{G}'(v,w,t) = \{A \in \mathcal{G}(w,t) : v \in A\}.$$

First, we observe that $\log^2 d \cdot \frac{t}{d} \leq \log^2 n \cdot \frac{n}{\delta n} \ll \log^4 n$. Lemma 4 in [32] gives us that there is a constant c such that for each v, $\mathcal{G}'(v,w,t) \leq 2^{w-ct}$. Thus, we have

$$|\mathcal{G}(w,t)| \leq \sum_v |\mathcal{G}'(v,w,t)| \leq n \cdot 2^{n-ct} \leq 2^{n-ct/2}$$

for n large enough.

Before we address the case when $t_0 \leq t < \log^4 n$, let us set up some additional notation. Given a vertex $v \in V$ and a subset $S \subseteq V$, we write $d_S(v)$ for the number of neighbors of v in S.

▶ **Definition 14** (Essential subset). *For a subset $A \subseteq X$, we write $W = N(A)$ and $W_s = \{y \in W : d_A(y) \geq s\}$. We say that F is an* essential set *for A if $W \supseteq F \supseteq W_{d/2}$ and $N(F) \supseteq [A]$.*

It may be useful to consider such an F an approximation for the neighborhood $W = N(A)$.

▶ **Definition 15** (Container). *We call a tuple $(S', T') \in 2^X \times 2^Y$ a γ'-container for a subset $A \subseteq X$ that is t-contracting, with neighborhood $W = N(A)$ if*
1. *$S' \supseteq [A]$ and $W_{d/2} \subseteq T' \subseteq W$,*
2. *$d_{Y \setminus T'}(v) \leq \gamma' t$ for each $v \in S'$, and*
3. *$d_{S'}(v) \leq \gamma' t$ for each $v \in Y \setminus T'$.*

The following two results show the existence of containers and bound the number of sets for which a given container is a γ'-container.

▶ **Lemma 16.** *For any $\gamma' > 0$ and any set $F \subseteq Y$, there is a set $\mathcal{C}^{\text{ind}} \subseteq 2^X \times 2^Y$ of size at most $n^{O(1/\gamma')}$ such that any $A \subseteq X$ for which F is an essential set, has a γ'-container in $\mathcal{C}^{\text{ind}}$.*

▶ **Lemma 17.** *There is an absolute constant $\gamma'' > 0$ such that the following holds.*

For any $\gamma' > 0$, $w < n$, $t < \log^4 n$, and tuple $(S', T') \in \mathcal{C}^{\text{ind}}$, there are at most $2^{w - \gamma'' t}$ sets $A \in \mathcal{G}(w, t)$ such that (S', T'), and is a γ'-container for A.

Since the proofs of these results are small modifications of existing container results, e.g. [40], we defer their proofs to Appendix A. We are now ready to handle the case of small t as follows. Consider an integer $t \in [t_0, \log^4 n]$ and a set $A \in \mathcal{G}(w, t)$. Define $S_A := [A] \cup N(A)$. As in the proof of Lemma 5, we have that $|\nabla(S_A)| = d|N(A)| - d|[A]| = td$. By Lemma 9, there is a cut $C \in \mathcal{C}^{\text{cut}}$ such that $g := |S_A \triangle C| \leq O(t)$. Let $A' := C \cap X$ and $W' := C \cap Y$.

Consider the set $W'_g = \{u \in Y : d_{A'}(u) > g\}$. We have the following two claims.

▷ Claim 18. $W \supseteq W'_g \supseteq W_{d/2}$.

Proof. Consider a vertex $u \in W_{d/2}$. We have

$$d_{A'}(u) \geq d_A(u) - |A \setminus A'| \geq d/2 - |L \triangle L'| \geq d/2 - g > g,$$

where the last inequality holds since $d = \lfloor \delta n \rfloor$ and $g = O(\log^4 n)$. Therefore $u \in W'_g$. Moreover, consider a vertex $u \in W'_g$. We have

$$d_A(u) \geq d_{A'}(u) - |A' \setminus A| > g - |L \triangle L'| > 0,$$

and hence $u \in W$. ◁

▷ Claim 19. $A \subseteq N(W'_g)$.

Proof. Suppose otherwise, i.e. there is a vertex $u \in A$ such that for each vertex $v \in N(u)$ we have $d_{A'}(v) \leq g$. For any such v, we have

$$d_A(v) \leq d_{A'}(v) + |A \setminus A'| \leq d_{A'}(v) + |L \triangle L'| \leq 2g.$$

This gives us that

$$t \cdot d = |E(W, X \setminus A)| \geq |E(N(u), W \setminus A)| \geq d(d - 2g),$$

contradicting the assumptions that $d = \lfloor \delta n \rfloor$ and t and g are both $O(\log^4 n)$. ◁

Claims 18 and 19 show that W'_g is an essential set for A. The set $A \in \mathcal{G}(w,t)$ may be constructed by

1. choosing the appropriate cut C in the set $\mathcal{C}^{\text{cut}}$ constructed in Lemma 9,
2. constructing the essential subset W'_g for it as above,
3. using Lemma 16 to obtain a γ'-container of A, where γ' is the absolute constant of Lemma 17, and finally
4. reconstructing A from the γ'-container with Lemma 17.

There are $n^{O(1/\delta)}$ choices for L' in the first step, a unique construction of W'_g for the second, $n^{O(1/\gamma')}$ possible containers in the third step, and $2^{w-\gamma''t}$ ways for the final step. In total there are

$$2^{w-\gamma''t+O(1/\gamma'+1/\delta)\log n} \leq 2^{w-\gamma''t/2}$$

such sets $A \in \mathcal{G}(w,t)$. The last inequality comes from our assumption that $t \geq t_0$ for our choice of $t_0 = C(\delta)\log(n/\epsilon) \geq C\log(n)$ (because wlog $\epsilon \leq 1$) satisfying

$$t_0 \geq \Omega\left(\frac{\log n}{\gamma''}\left(\frac{1}{\gamma'}+\frac{1}{\delta}\right)\right). \blacktriangleleft$$

7 Concluding remarks and future directions

1. Naturally, a next goal is to understand the power and limitations of the methods presented, especially in conjunction with existing cluster expansion methods. More specifically, we are curious about the following two questions:
   a. Can this spectral point of view help with our understanding of independent sets in a larger class of bipartite graphs?
   b. To what extent do these methods help in reducing the computation needed to implement algorithmic cluster expansion?

   In this context, the problem of approximating the number of independent sets in *small-set expanders* feels within striking distance.
2. Our next remark concerns Lemma 9. As mentioned before, similar results have had other applications in optimization and Unique Games [2, 36, 37], though we take a subtly different viewpoint worth noting: we seek to approximate *all* cuts in the graph, not just small ones. In any case, we find the lemma interesting in its own right and conjecture something stronger.

   ▶ **Conjecture 20.** *Lemma 9 holds with* $|\mathcal{C}^{\text{cut}}| \leq 2^{O(n/d)}$.

   If true, this would be best possible, as evidenced by a disjoint union of $1/\delta$ components. Setting $t = 0$ in this case gives exactly $2^{1/\delta}$ cuts of size 0.
3. Finally, we leave open the problem of making our algorithm deterministic. At the moment, the only step where randomness is used is Lemma 6.

References

1 J. D. Annan. A Randomised Approximation Algorithm for Counting the Number of Forests in Dense Graphs. *Combinatorics, Probability and Computing*, 3(3):273–283, 1994. doi:10.1017/S0963548300001188.

2 Sanjeev Arora, Boaz Barak, and David Steurer. Subexponential Algorithms for Unique Games and Related Problems. *Journal of the ACM*, 62(5):1–25, 2015. doi:10.1145/2775105.

3 Sanjeev Arora, David Karger, and Marek Karpinski. Polynomial time approximation schemes for dense instances of NP-hard problems. In *Proceedings of the Twenty-Seventh Annual ACM Symposium on Theory of Computing*, STOC '95, pages 284–293, New York, NY, USA, 1995. Association for Computing Machinery. doi:10.1145/225058.225140.

4 Sanjeev Arora, Subhash A. Khot, Alexandra Kolla, David Steurer, Madhur Tulsiani, and Nisheeth K. Vishnoi. Unique games on expanding constraint graphs are easy: Extended abstract. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, pages 21–28, Victoria British Columbia Canada, 2008. ACM. doi:10.1145/1374376.1374380.

5 Victor Bapst and Amin Coja-Oghlan. Harnessing the Bethe free energy. *Random Structures & Algorithms*, 49(4):694–741, 2016. doi:10.1002/rsa.20692.

6 Boaz Barak, Prasad Raghavendra, and David Steurer. Rounding Semidefinite Programming Hierarchies via Global Correlation. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pages 472–481, 2011. doi:10.1109/FOCS.2011.95.

7 Antonio Blanca, Andreas Galanis, Leslie Ann Goldberg, Daniel Stefankovic, Eric Vigoda, and Kuan Yang. Sampling in Uniqueness from the Potts and Random-Cluster Models on Random Regular Graphs. In Eric Blais, Klaus Jansen, José D. P. Rolim, and David Steurer, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2018)*, volume 116 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 33:1–33:15, Dagstuhl, Germany, 2018. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.APPROX-RANDOM.2018.33.

8 Sarah Cannon and Will Perkins. Counting independent sets in unbalanced bipartite graphs. In *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1456–1466, 2020. doi:10.1137/1.9781611975994.88.

9 Charles Carlson, Ewan Davies, and Alexandra Kolla. Efficient algorithms for the Potts model on small-set expanders. *To appear in Chicago Journal of Theoretical Computer Science*, March 2020. arXiv:2003.01154.

10 Charlie Carlson, Ewan Davies, Nicolas Fraiman, Alexandra Kolla, Aditya Potukuchi, and Corrine Yap. Algorithms for the ferromagnetic Potts model on expanders. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 344–355, 2022. doi:10.1109/FOCS54457.2022.00040.

11 Zongchen Chen, Andreas Galanis, Leslie A. Goldberg, Will Perkins, James Stewart, and Eric Vigoda. Fast algorithms at low temperatures via Markov chains. *Random Structures & Algorithms*, 58(2):294–321, 2021. doi:10.1002/rsa.20968.

12 Fan Chung. *Spectral Graph Theory*, volume 92 of *CBMS Regional Conference Series in Mathematics*. American Mathematical Society, December 1996. doi:10.1090/cbms/092.

13 Amin Coja-Oghlan, Colin Cooper, and Alan Frieze. An Efficient Sparse Regularity Concept. *SIAM Journal on Discrete Mathematics*, 23(4):2000–2034, 2010. doi:10.1137/080730160.

14 Amin Coja-Oghlan and Will Perkins. Belief Propagation on Replica Symmetric Random Factor Graph Models. In Klaus Jansen, Claire Mathieu, José D. P. Rolim, and Chris Umans, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2016)*, volume 60 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 27:1–27:15, Dagstuhl, Germany, 2016. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.APPROX-RANDOM.2016.27.

15 Amin Coja-Oghlan and Will Perkins. Bethe States of Random Factor Graphs. *Communications in Mathematical Physics*, 366(1):173–201, February 2019. doi:10.1007/s00220-019-03387-7.

16 Martin Dyer, Alan Frieze, and Mark Jerrum. Approximately counting Hamilton cycles in dense graphs. In *Proceedings of the Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '94, pages 336–343, USA, 1994. Society for Industrial and Applied Mathematics. URL: https://dl.acm.org/doi/abs/10.5555/314464.314557.

17 Martin Dyer, Leslie Ann Goldberg, Catherine Greenhill, and Mark Jerrum. The Relative Complexity of Approximate Counting Problems. *Algorithmica*, 38(3):471–500, 2004. doi:10.1007/s00453-003-1073-y.

18 A. Frieze. A new rounding procedure for the assignment problem with applications to dense graph arrangement problems. In *Proceedings of the 37th Annual Symposium on Foundations of Computer Science*, FOCS '96, page 21, USA, 1996. IEEE Computer Society.

19 A. Frieze and R. Kannan. The regularity lemma and approximation schemes for dense problems. In *Proceedings of 37th Conference on Foundations of Computer Science*, pages 12–20, 1996. `doi:10.1109/SFCS.1996.548459`.

20 Andreas Galanis, Qi Ge, Daniel Štefankovič, Eric Vigoda, and Linji Yang. Improved inapproximability results for counting independent sets in the hard-core model. *Random Structures & Algorithms*, 45(1):78–110, 2014. `doi:10.1002/rsa.20479`.

21 Andreas Galanis, Leslie Ann Goldberg, and James Stewart. Fast Algorithms for General Spin Systems on Bipartite Expanders. *ACM Transactions on Computation Theory*, 13(4):25:1–25:18, 2021. `doi:10.1145/3470865`.

22 David Galvin. A Threshold Phenomenon for Random Independent Sets in the Discrete Hypercube. *Combinatorics, Probability and Computing*, 20(1):27–51, 2011. `doi:10.1017/S0963548310000155`.

23 David Galvin and Prasad Tetali. Slow mixing of Glauber dynamics for the hard-core model on regular bipartite graphs. *Random Structures & Algorithms*, 28(4):427–443, 2006. `doi:10.1002/rsa.20094`.

24 Leslie Ann Goldberg, John Lapinskas, and David Richerby. Faster exponential-time algorithms for approximately counting independent sets. *Theoretical Computer Science*, 892:48–84, November 2021. `doi:10.1016/j.tcs.2021.09.009`.

25 Venkatesan Guruswami and Ali Kemal Sinop. Lasserre Hierarchy, Higher Eigenvalues, and Approximation Schemes for Graph Partitioning and Quadratic Integer Programming with PSD Objectives. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pages 482–491. IEEE Computer Society, 2011. `doi:10.1109/FOCS.2011.36`.

26 Tyler Helmuth, Matthew Jenssen, and Will Perkins. Finite-size scaling, phase coexistence, and algorithms for the random cluster model on random graphs. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 59(2):817–848, 2023. `doi:10.1214/22-AIHP1263`.

27 Tyler Helmuth, Will Perkins, and Guus Regts. Algorithmic Pirogov–Sinai theory. *Probability Theory and Related Fields*, 2019. `doi:10.1007/s00440-019-00928-y`.

28 Vishesh Jain, Frederic Koehler, and Andrej Risteski. Mean-field approximation, convex hierarchies, and the optimality of correlation rounding: A unified perspective. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, pages 1226–1236, Phoenix, AZ, USA, 2019. Association for Computing Machinery. `doi:10.1145/3313276.3316299`.

29 Matthew Jenssen, Peter Keevash, and Will Perkins. Algorithms for #BIS-Hard Problems on Expander Graphs. *SIAM Journal on Computing*, 49(4):681–710, 2020. `doi:10.1137/19M1286669`.

30 Matthew Jenssen and Will Perkins. Independent sets in the hypercube revisited. *Journal of the London Mathematical Society*, 102(2):645–669, 2020. `doi:10.1112/jlms.12331`.

31 Matthew Jenssen, Will Perkins, and Aditya Potukuchi. Independent sets of a given size and structure in the hypercube. *Combinatorics, Probability and Computing*, 31(4):702–720, July 2022. `doi:10.1017/S0963548321000559`.

32 Matthew Jenssen, Will Perkins, and Aditya Potukuchi. Approximately counting independent sets in bipartite graphs via graph containers. *Random Structures & Algorithms*, 63(1):215–241, 2023. `doi:10.1002/rsa.21145`.

33 Mark Jerrum and Alistair Sinclair. Approximating the Permanent. *SIAM Journal on Computing*, 18(6):1149–1178, 1989. `doi:10.1137/0218077`.

34 Jeff Kahn and Jinyoung Park. The Number of Maximal Independent Sets in the Hamming Cube. *Combinatorica*, 42(6):853–880, December 2022. `doi:10.1007/s00493-021-4729-9`.

35 Frederic Koehler, Holden Lee, and Andrej Risteski. Sampling Approximately Low-Rank Ising Models: MCMC meets Variational Methods. In *Proceedings of Thirty Fifth Conference on Learning Theory*, pages 4945–4988. PMLR, June 2022. URL: `https://proceedings.mlr.press/v178/koehler22a.html`.

36 Alexandra Kolla. Spectral Algorithms for Unique Games. In *2010 IEEE 25th Annual Conference on Computational Complexity*, pages 122–130, Cambridge, MA, USA, June 2010. IEEE. `doi:10.1109/CCC.2010.20`.

37 Alexandra Kolla and Madhur Tulsiani. Playing random and expanding unique games. Unpublished, 2007. URL: `https://home.cs.colorado.edu/~alko5368/UGspec.pdf`.

38 Jingcheng Liu and Pinyan Lu. FPTAS for #BIS with Degree Bounds on One Side. In *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing*, STOC '15, pages 549–556, New York, NY, USA, June 2015. Association for Computing Machinery. `doi:10.1145/2746539.2746598`.

39 Shayan Oveis Gharan and Luca Trevisan. A New Regularity Lemma and Faster Approximation Algorithms for Low Threshold Rank Graphs. *Theory of Computing*, 11(1):241–256, 2015. `doi:10.4086/toc.2015.v011a009`.

40 Jinyoung Park. Note on the Number of Balanced Independent Sets in the Hamming Cube. *The Electronic Journal of Combinatorics*, page P2.34, 2022. `doi:10.37236/10471`.

41 J. Scott Provan and Michael O. Ball. The Complexity of Counting Cuts and of Computing the Probability that a Graph is Connected. *SIAM Journal on Computing*, 12(4):777–788, 1983. `doi:10.1137/0212053`.

42 Andrej Risteski. How to calculate partition functions using convex programming hierarchies: Provable bounds for variational methods. In *Conference on Learning Theory*, pages 1402–1416. PMLR, 2016. URL: `https://proceedings.mlr.press/v49/risteski16.html`.

43 A. A. Sapozhenko. On the number of connected subsets with given cardinality of the boundary in bipartite graphs. *Metody Diskretnogo Analiza*, (45):42–70, 96, 1987.

44 A. A. Sapozhenko. On the number of independent sets in extenders. *Diskretnaya Matematika*, 13(1):56–62, 2001. `doi:10.1515/dma.2001.11.2.155`.

45 A. Sly and N. Sun. The Computational Hardness of Counting in Two-Spin Models on d-Regular Graphs. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 361–369, 2012. `doi:10.1109/FOCS.2012.56`.

46 Allan Sly. Computational Transition at the Uniqueness Threshold. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pages 287–296, Las Vegas, NV, USA, 2010. IEEE. `doi:10.1109/FOCS.2010.34`.

A Deferred proofs

A.1 Proof of Lemma 16

We restate the result for convenience.

▶ **Lemma 16.** *For any $\gamma' > 0$ and any set $F \subseteq Y$, there is a set $\mathcal{C}^{\text{ind}} \subseteq 2^X \times 2^Y$ of size at most $n^{O(1/\gamma')}$ such that any $A \subseteq X$ for which F is an essential set, has a γ'-container in $\mathcal{C}^{\text{ind}}$.*

Proof. Let $A \subseteq X$ be a subset for which F is an essential set and let $W = N(A)$, $t := |N(A)| - |[A]|$. Consider the following algorithm

---

initialize $T' \leftarrow F$
while $\exists\, v \in [A]$ s.t. $d_{W \setminus T'}(v) > \gamma' t$, pick such a v **do**
  $T' \leftarrow T' \cup N(v)$
end while
initialize $S' \leftarrow \{v \in X : d_{Y \setminus T'}(v) \leq \gamma' t\}$
while $\exists\, v \in Y \setminus W$ s.t. $d_{S'}(v) > \gamma' t$, pick such a v **do**
  $S' \leftarrow S' \setminus N(v)$
end while
$T' \leftarrow T' \cup \{v \in Y : d_S(v) > \gamma' t\}$
return (S', T')

The lemma follows provided we can show that (S', T') as given by the algorithm above is a γ'-container for A by establishing properties 1–3, and provided we can show a good enough bound on the total number of outputs (S', T') which can occur for a fixed F as A varies.

To prove that the output (S', T') is a γ'-container of A, we first show that $S' \supseteq [A]$ and $W_{d/2} \subseteq T' \subseteq W$, establishing 1. Since F is an essential subset for A, we initialize $T' \leftarrow F$, and T' can then only grow, we have $W_{d/2} \subseteq T'$. Clearly, $T' \subseteq W$ at the end of the first while loop. After the second initialize statement, we have that each vertex $v \in [A]$ satisfies $d_{W \setminus T'}(v) \leq d_{Y \setminus T'}(v) \leq \gamma' t$. Therefore, $S' \supseteq A$ at the end of this line. This property is maintained during the second while loop since we only delete $N(v)$ from S for $v \notin W$. This also means that in the penultimate line, all vertices added to T' are from W. Thus $T' \subseteq W$ is also maintained at the end of the algorithm. Next, we prove 3. At the beginning of the second loop, every $v \in S'$ satisfies $d_{Y \setminus T'}(v) \leq \gamma' t$. Since vertices are only removed from S and added to T' after this point, this property is preserved till the end. Finally, to prove 2 note that the penultimate line of the algorithm ensures that every $v \in Y \setminus T$ satisfies $d_S(v) \leq \gamma' t$.

To bound the number of possible outputs for a fixed F, note that before the start of the first loop we have $|W \setminus T'| \leq O(t)$. Each step in the first loop of the algorithm removes γt vertices from $W \setminus T'$. Therefore, this loop runs at most $O(1/\gamma')$ times. Next, each step in the second loop removes at least $\gamma' t$ vertices from $S \setminus [A]$. Immediately after the second initialize statement, we have

$$dt \geq |E(S' \setminus [A], T')| \geq (d - \gamma' t)|S' \setminus [A]|.$$

As a result, $|S' \setminus [A]| = O(t)$. So the second loop runs for at most $1/\gamma'$ steps. The output is determined by the set of $O(1/\gamma')$ vertices chosen in both loops, so the number of possible outputs for the algorithm for a given F is at most $n^{O(1/\gamma')}$. ◀

A.2 Proof of Lemma 17

We restate the result for convenience.

▶ **Lemma 17.** *There is an absolute constant $\gamma'' > 0$ such that the following holds.*
For any $\gamma' > 0$, $w < n$, $t < \log^4 n$, and tuple $(S', T') \in \mathcal{C}^{\text{ind}}$, there are at most $2^{w - \gamma'' t}$ sets $A \in \mathcal{G}(w, t)$ such that (S', T'), and is a γ'-container for A.

We need the following lemma

▶ **Lemma 21.** *Let (S', T') be a γ'-container for a set $A \in \mathcal{G}(w, t)$. Then $|S'| \leq |T'|$.*

Proof. Let us denote $W = N(A)$. First, we observe that $|E(S', W)| \leq d|T'| + \gamma' t|W \setminus T'|$ by 3. We also have that $|E(S', W)| \geq d|[A]| + |S' \setminus [A]|(d - \gamma' t) = d|S'| - \gamma' t|S' \setminus [A]|$ by 1 and 2. Combining these inequalities, we have

$$|S'| \leq |T'| + \frac{\gamma' t(|S' \setminus [A]| + |W \setminus |T'||)}{d}. \tag{9}$$

Since $T' \supseteq W_{d/2}$, we have that $|W \setminus T'| \leq O(t)$ and

$$td = |E(W, X \setminus [A])| \geq \sum_{v \in S' \setminus [A]} d_{T'}(v) \geq |S' \setminus [A]|(d - \gamma' t)$$

which gives $|S' \setminus [A]| = O(t)$. So (9) implies

$$|S'| \leq |T'| + O\left(\frac{\gamma' t^2}{d}\right).$$

Since $t \leq \log^4 n$, $d = \lfloor \delta n \rfloor$, and $|S'|$ and $|T'|$ are both integers, we have that $|S'| \leq |T'|$. ◀

We finish the proof using the following lemma from [40], whose proof we reproduce for clarity.

▶ **Lemma 22** ([40], Lemma 11). *There is an absolute constant $\gamma'' > 0$ such that the following holds.*

For any tuple $(S', T') \in 2^X \times 2^Y$ such that $|S'| \leq |T'|$, there are at most $2^{w-\gamma'' t}$ sets $A \in \mathcal{G}(w,t)$ such that $[A] \subseteq S'$ and $T' \subseteq N(A)$.

To be precise, in [40] the graph in question is the d-dimensional hypercube and additional hypotheses are stated, namely $w - t < n/4$ and $w > d^4$. These play no role in the proof, however, and it extends verbatim to the result stated above.

Proof. Throughout, we denote $W = N(A)$, and let $\alpha > 0$ be a constant that will be determined later.

If $|S'| < w - \alpha t$, then A is among the possible $2^{w-\alpha t}$ subsets of S'. Suppose otherwise, that $|S'| > w - \alpha t$. Let $A^* \in \mathcal{G}(w,t)$ such that (S', T') is a γ'-container for A^* and let $W^* = N(A^*)$. We have that $[A]$ is completely determined by $W \setminus W^*$ and $W^* \setminus W$. Since $W^* \setminus W \subseteq W^* \setminus T$, and

$$|W^* \setminus T'| \leq |W^*| - |T'| = |W| - |T'| \leq |W| - |S'| \leq \alpha t,$$

there are at most $2^{\alpha t}$ choices for $W^* \setminus W$. Next, for each vertex in $W \setminus W^*$, we choose a neighbor in $A \setminus A^* \subseteq S' \setminus A^*$. Observe that $W \setminus W^* = N(A \setminus A^*) \setminus W^*$. Since

$$|W \setminus W^*| \leq |W \setminus F| = |W| - |F| \leq |W| - |S'| \leq \alpha t,$$

and

$$|S' \setminus A^*| \leq |S'| - |A^*| = |S'| - |A| \leq |T'| - |A| \leq |W| - |A| = t.$$

Therefore, the number of choices for $W \setminus W^*$ is at most

$$\binom{t}{\alpha t} \leq 2^{H(\alpha)t}.$$

Once we have $[A]$, there are at most 2^{w-t} possibilities for A. Thus the total number of choices is at most

$$2^{w-t+t(\alpha+H(\alpha))}.$$

Choosing e.g., $\alpha = 0.17$ allows one to choose $\gamma'' = 0.17$. ◀