Hardness of Identity Testing for Restricted Boltzmann Machines and Potts models

Antonio Blanca ABLANCA@CSE.PSU.EDU
Pennsylvania State University

Zongchen Chen CHENZONGCHEN@GATECH.EDU
Georgia Institute of Technology

Daniel Štefankovič STEFANKO@CS.ROCHESTER.EDU
University of Rochester

Eric Vigoda VIGODA@GATECH.EDU
Georgia Institute of Technology

Abstract

We study identity testing for restricted Boltzmann machines (RBMs), and more generally for undirected graphical models. Given sample access to the Gibbs distribution corresponding to an unknown or hidden model M^* and given an explicit model M, can we distinguish if either $M = M^*$ or if they are (statistically) far apart? Daskalakis et al. (2018) presented a polynomial-time algorithm for identity testing for the ferromagnetic (attractive) Ising model. In contrast, for the antiferromagnetic (repulsive) Ising model, Bezáková et al. (2019) proved that unless RP $=$ NP there is no identity testing algorithm when $\beta d = \omega(\log n)$, where d is the maximum degree of the visible graph and β is the largest edge weight (in absolute value).

We prove analogous hardness results for RBMs (i.e., mixed Ising models on bipartite graphs), even when there are no latent variables or an external field. Specifically, we show that if RP $\neq$ NP, then when $\beta d = \omega(\log n)$ there is no polynomial-time algorithm for identity testing for RBMs; when $\beta d = O(\log n)$ there is an efficient identity testing algorithm that utilizes the structure learning algorithm of Klivans and Meka (2017). In addition, we prove similar lower bounds for purely ferromagnetic RBMs with inconsistent external fields, and for the ferromagnetic Potts model. Previous hardness results for identity testing of Bezáková et al. (2019) utilized the hardness of finding the maximum cuts, which corresponds to the ground states of the antiferromagnetic Ising model. Since RBMs are on bipartite graphs such an approach is not feasible. We instead introduce a novel methodology to reduce from the corresponding approximate counting problem and utilize the phase transition that is exhibited by RBMs and the mean-field Potts model. We believe that our method is general, and that it can be used to establish the hardness of identity testing for other spin systems.

1. Introduction

For graphical models, there are several fundamental computational tasks which are essential for utilizing these models. These computational problems can be broadly labeled as follows: sampling, counting, structure learning, and testing. Our big picture aim is to understand the relationship between these problems. The specific focus in this paper is on the computational complexity of the *identity testing* problem for *undirected* graphical models and its connections to the hardness of the counting problem.

Identity testing is a basic question in statistics for testing whether a given model fits a dataset. Roughly speaking, given data $\mathcal{D}$ sampled from the posterior or likelihood distribution of an unknown/hidden model M^* and given an explicit model M, can we distinguish whether $M = M^*$?

We study identity testing in the context of undirected graphical models [38], which correspond to (pairwise) Markov random fields in probability theory and computer vision [22] and to spin systems in statistical physics [23]. We focus attention on examples of graphical models of particular interest: the Ising model, the Potts model, and Restricted Boltzmann Machines. The Ising model is the simplest example of an undirected graphical model, and, in fact, it is one of the most well-studied models in statistical physics where it is used to study phase transitions. The Potts model is the generalization of the Ising model from a two state system to an integer $q \geq 3$ state system. It is also well-studied in statistical physics as the nature of the phase transition changes as q increases [15, 16].

Restricted Boltzmann Machines (RBMs) are a simple class of undirected graphical models corresponding to the Ising model on bipartite graphs. Originally introduced by Smolensky in 1986 [46], they have played an important role in the history of computational learning theory. They have two layers of variables: one layer corresponding to the observed variables and another layer corresponding to the hidden/latent variables, and no intralayer connections so that the underlying graph is bipartite. Learning was shown to be practical in these restricted models [30, 32] and henceforth played a seminal role in the development of deep learning [41, 39, 42, 31].

We define first the Potts model, as both the Ising model and RBMs may be viewed as special cases of this model. The Potts model is specified by a graph $G = (V, E)$, a set of vertex labels or spins $[q] = \{1, \ldots, q\}$, a set of edge weights defined by $\beta : E \to \mathbb{R}$ and a set of vertex weights $h : V \times [q] \to \mathbb{R}$. Configurations of the Potts model are the collection of vertex labelings $\Omega = \{1, \ldots, q\}^V$. The *Gibbs distribution* associated with the Potts model is a distribution over all configurations $\sigma \in \Omega$ such that:

$$\mu(\sigma) = \mu_{G,\beta,h}(\sigma) := \frac{1}{Z} \exp\left(\sum_{\{u,v\} \in E} \beta(\{u,v\}) \mathbb{1}(\sigma(u) = \sigma(v)) + \sum_{v \in V} h(v, \sigma(v)) \right),$$

where $Z = Z_{G,\beta,h}$ is corresponding the normalizing factor or *partition function*.

When $\beta(e) > 0$ for every $e \in E$, the model is called *ferromagnetic* and neighboring vertices prefer to align to the same spin. Conversely, when $\beta(e) < 0$ for every $e \in E$ the model is called *antiferromagnetic*. Models where β is allowed to be both positive or negative for distinct edges are called *mixed* models.

The Ising model corresponds to the special case where there are only two spins; i.e., $q = 2$. RBMs are mixed Ising models restricted to bipartite graphs; that is, G is bipartite with bipartition $V = L \cup R$. Since the focus in this paper is on lower bounds, we often consider the case of no external field ($h = 0$) in order to obtain stronger hardness results.

Given a model specification, that is, a graph $G = (V, E)$, an edge weight function β and an external field h, the goal in the *sampling problem* is to generate samples from the Gibbs distribution $\mu = \mu_{G,\beta,h}$ (or from a distribution close to μ in total variation distance). The corresponding *counting problem* is to compute the partition function $Z = Z_{G,\beta,h}$. The (exact) counting problem is #P-hard [50] even for restricted classes of graphs [26, 49], and hence the focus on the approximate

counting problem of obtaining an FPRAS (fully-polynomial randomized approximation scheme[1]) for Z. For a general class of models, the approximate counting and the approximate sampling problems are equivalent, i.e., there are polynomial-time reductions between them [34, 47, 36]. A seminal result of Jerrum and Sinclair [33] (see also [40, 11, 28]) presented an FPRAS for the partition function of the ferromagnetic Ising model.

Another two fundamental problems for undirected graphical models are *structure learning* and *identity testing*. The structure learning problem is as follows: given oracle access to samples from the Gibbs distribution μ_{M^*} for an unknown (i.e., "hidden") model $M^* = (G^*, \beta^*, h^*)$, can we learn G^* (i.e., the structure of the model) in polynomial-time with probability at least $2/3$? In the case of no latent variables (so the samples from the Gibbs distribution reveal the label of all vertices V of G) recent work of Klivans and Meka [35] (see also [6, 52, 29, 51, 53]) learns n-vertex graphs with $O(\log n) \times \exp(O(\beta d))$ samples and $O(n^2 \log n) \times \exp(O(\beta d))$ time where d is the maximum degree of G and $\beta := \max_{e \in E} |\beta(e)|$ is the maximum edge weight in absolute value; this bound has nearly-optimal sample complexity from an information-theory perspective [43].

For RBMs with latent variables (thus samples only reveal the labels for vertices on one side R), structure learning can be done in time $O(n^{d_L+1})$ where d_L is the maximum degree of the latent variables. Recent work of Bresler, Koehler and Moitra [7] proves that there is no algorithm with running time $n^{o(d_L)}$ assuming k-sparse noisy parity on n bits is hard to learn in time $n^{o(k)}$; they also show that for the special case of ferromagnetic RBMs with hidden variables there is a structure learning algorithm with $O(\log n) \times \exp(O(\beta d^2))$ sample complexity and $O(n^2 \log n) \times \exp(O(\beta d^2))$ running time.

In the *identity testing* problem we are given oracle access to samples from the Gibbs distribution μ_{M^*} for an unknown model $M^* = (G^*, \beta^*, h^*)$ (as in structure learning) and we are also given an explicit model $M = (G, \beta, h)$. Our goal is to determine, with probability $\geq 2/3$, if either $M = M^*$ or if the models are $(1-\varepsilon)$-far apart; specifically, if the total variation distance between their Gibbs distributions is at least $1-\varepsilon$ for a given $\varepsilon > 0$. (Previous works assumed separation $\geq \varepsilon$ in the later case, whereas we prove hardness even when we assume separation $\geq 1-\varepsilon$.)

It is known that identity testing cannot be solved in polynomial time for general graphical models in the presence of hidden variables unless RP = NP [4] . In this paper we assume there are no hidden variables and hence the samples from μ_{M^*} reveal the label of every vertex in the graph G; this setting is more interesting for hardness results. We explore a more refined picture of hardness of identity testing vs. polynomial-time algorithms.

It is known that identity testing can be reduced to sampling [14] or structure learning [1]: given an efficient algorithm for the associated sampling problem or an efficient algorithm for structure learning, then one can efficiently solve the identity testing problem. Hence, identity testing is (computationally) easier than sampling and structure learning. (To be precise, one needs to solve both the structure learning and the parameter estimation problems to solve identity testing; the algorithm of Klivans and Meka [35] does in fact provide this.) This raises the question of whether identity testing can be efficiently solved in cases where sampling and structure learning are known to be hard. We prove (for the models studied here) that when sampling and structure learning are hard, then identity testing is also hard.

1. A fully polynomial-time randomized approximation scheme (FPRAS) for an optimization problem with optimal solution Z produces an approximate solution $\hat{Z}$ such that, with probability at least $1-\delta$, $(1-\varepsilon)\hat{Z} \leq Z \leq (1+\varepsilon)\hat{Z}$ with running time polynomial in the instance size, ε^{-1} and $\log(\delta^{-1})$.

1.1. Our results

The ε-identity testing problem for the Ising and Potts models is formally defined as follows. For positive integers n and d, and positive real numbers β and h, let $\mathcal{M}_{\text{RBM}}(n, d, \beta, h)$ denote the family of RBMs on n-vertex bipartite graphs $G = (V, E)$ of maximum degree at most d, where the absolute value of all edge interactions is at most β and the field $|h(v, i)| \leq h$ for all $v \in V$ and $i \in [q]$; see Definition 4. We define $\mathcal{M}_{\text{POTTS}}(n, d, \beta, h)$ analogously for the family of Potts models, without the restriction of G being bipartite.

Given an RBM $M \in \mathcal{M}_{\text{RBM}}(n, d, \beta, h)$, and sample access to a distribution μ_{M^*} for an unknown RBM $M^* \in \mathcal{M}_{\text{RBM}}(n, d, \beta, h)$, distinguish with probability $\geq 3/4$ between the cases:

$$1.\ \mu_M = \mu_{M^*}; \qquad 2.\ \|\mu_M - \mu_{M^*}\|_{\text{TV}} \geq 1 - \varepsilon.$$

The choice of $3/4$ for the probability of success is arbitrary, and it can be replaced by any constant in the interval $(\frac{1}{2}, 1)$ at the expense of a constant factor in the running time of the algorithm. The ε-identity testing problem for the Potts model is defined in the same manner, but assuming that both M and M^* belong to $\mathcal{M}_{\text{POTTS}}(n, d, \beta, h)$ instead.

Our first result concerns the identity testing problem on $\mathcal{M}_{\text{RBM}}(n, d, \beta, 0)$; that is, (mixed) RBMs without external fields: $h(v, i) = 0$ for all $v \in V$, $i \in [q]$. We show that for RBMs the approach utilizing structure learning is essentially best possible. In particular we prove that when $\beta d = \omega(\log n)$ there is no poly-time identity testing algorithm, unless RP = NP. Note that when $\beta d = O(\log n)$, the algorithm of Klivans and Meka [35] for structure learning and parameter estimation provides an identity testing algorithm with $\text{poly}(n)$ sample complexity and running time.

Theorem 1 *Suppose n, d are positive integers such that $3 \leq d \leq n^\theta$ for constant $\theta \in (0, 1)$ and let $\varepsilon \in (0, 1)$. If* RP $\neq$ NP*, then for all real $\beta > 0$ satisfying $\beta d = \omega(\log n)$ there is no polynomial running time algorithm to solve the ε-identity testing problem for the class $\mathcal{M}_{\text{RBM}}(n, d, \beta, 0)$ of mixed RBMs without external fields.*

In contrast to the above result, Daskalakis, Dikkala and Kamath [14] provided a poly-time identity testing algorithm for all *ferromagnetic* Ising model with *consistent* fields (the external field is consistent if it only favors the same unique spin at every vertex; otherwise it is called inconsistent). Their algorithm crucially utilizes the known poly-time sampling methods for the ferromagnetic Ising model [33, 40, 11, 28]. On the hardness side, super-polynomial lower bounds were recently established for identity testing for the *antiferromagnetic* Ising model on general (not necessarily bipartite) graphs when $\beta d = \omega(\log n)$ [1]. This previous result utilizes the hardness of the maximum cut problem, since maximum cuts correspond to the "ground states" (maximum likelihood configurations) of the antiferromagnetic model; this is not the case for RBMs, and new insights are required (see Section 1.2 for a more detailed discussion). In particular we show a new approach to reduce from the counting problem.

Ferromagnetic and antiferromagnetic RBMs are equivalent models; that is, there is a one-to-one correspondence between configurations with the same weight. Hence, the results in [14] solve the identity testing problem for both ferromagnetic and antiferromagnetic RBMs with no latent variables, even in the presence of a consistent external field. Moreover, Klivans and Meka's algorithm from [35] together with the hardness results of Theorem 1 provides a fairly complete picture of the computational complexity of identity testing for (mixed) RBMs with no external field ($h = 0$).

Our next result concerns the hardness of identity testing for purely *ferromagnetic* RBMs with an *inconsistent* magnetic field; that is, a field that favors one spin for some of the vertices and the other spin for the rest. For this we utilize the complexity of #BIS, which is the problem of counting the independent sets in a bipartite graph. #BIS is believed not to have an FPRAS, and it has achieved considerable interest in approximate counting as a tool for proving relative complexity hardness [17, 25, 18, 8, 10, 9, 20]. Let $\mathcal{M}^{+}_{\mathrm{RBM}}(n, d, \beta, h)$ be set of all ferromagnetic RBMs in $\mathcal{M}_{\mathrm{RBM}}(n, d, \beta, h)$.

Theorem 2 *Suppose n, d are positive integers such that $3 \leq d \leq n^{\theta}$ for constant $\theta \in (0, 1)$ and let $\varepsilon \in (0, 1)$. If* #BIS *does not admit an* FPRAS*, there exists $h = O(1)$ such that when $\beta d = \omega(\log n)$ there is no polynomial running time algorithm that solves the ε-identity testing problem for the class $\mathcal{M}^{+}_{\mathrm{RBM}}(n, d, \beta, h)$ of* ferromagnetic *RBMs with* inconsistent *external fields.*

Given the efficient identity testing algorithm for ferromagnetic Ising models [14, 33], we may ask whether there are other (ferromagnetic) models that allow efficient testing algorithms. A prime candidate is the ferromagnetic Potts model. Both the ferromagnetic Ising and Potts models have a rich structure; for instance, their random-cluster representation [27] enables sophisticated (and widely-used) sampling algorithms such as the Swendsen-Wang algorithm [48]. However, while there are efficient samplers for the ferromagnetic Ising model for all graphs G and all edge interactions β [33, 11, 28], the case of the ferromagnetic Potts model (i.e., $q > 2$ spins) looks less promising. In fact, it is unlikely that there is an efficient sampling/counting algorithm for general ferromagnetic Potts models since this is a known #BIS-hard problem [25, 21]; this is due to a phenomena called *phase co-existence*, which we will also exploit; see Section 2.2.1. Given the weaker hardness of sampling and approximate counting for the ferromagnetic Potts model, the hardness of the identity problem was less clear.

We prove that identity testing for the ferromagnetic Potts model is in fact hard in the same regime of parameters where sampling and structure learning are known to be hard. Specifically, we observe that the structure learning algorithm from [35] applies to the Potts model, and hence implies a testing algorithm when $\beta d = O(\log n)$; we establish lower bounds when $\beta d = \omega(\log n)$ that hold even for the simpler case of models with no external field.

Theorem 3 *Suppose n, d, $q \geq 3$ are positive integers such that $3 \leq d \leq n^{\theta}$ for constant $\theta \in (0, 1)$ and let $\varepsilon \in (0, 1)$. If* #BIS *does not admit an* FPRAS*, then there is no polynomial running time algorithm that solves the ε-identity testing problem for the class $\mathcal{M}^{+}_{\mathrm{POTTS}}(n, d, \beta, 0)$ of ferromagnetic q-state Potts models without an external field. Moreover, our lower bound applies restricted to the class of ferromagnetic Potts models on bipartite graphs in $\mathcal{M}^{+}_{\mathrm{POTTS}}(n, d, \beta, 0)$.*

1.2. Our techniques

Our proof is a general approach that allows us to obtain hardness results for several models of interest. Specifically, we devise a novel methodology to reduce the problem of approximate counting (i.e., approximating partition functions) to identity testing. For this we consider a decision version of approximate counting and prove that this variant is as hard as the standard approximation problem; this first step of our reduction applies to many other models of interest (see Theorem 7).

In the second step of our reduction, given a hard counting instance, we use insights about the phase transition of the models to construct a testing instance whose output allows us to solve the

decision version of approximate counting. The actual reduction is generic (see Theorem 9), but the insights about each model are needed to build a suitable testing instance; this construction is the only part of our proof that is model specific, whereas every other step in the proof applies to more general spin systems. Our approach is nicely illustrated in the context of the ferromagnetic Potts model; that is, in the proof of Theorem 3 in Section 2. There, we utilize the phase transition phenomenon in the associated mean-field Potts model which corresponds to the complete graph. In particular, there is a phase co-existence corresponding to a first-order phase transition which we utilize to approximate the partition function of the input graph; see Section 2.

In the third and final step of the reduction, we reduce the maximum degree of the graph in the testing instance by using random bipartite graphs as gadgets, as has been done in seminal hardness results for approximate counting [44, 45], and more recently in [1] for the hardness of testing for the antiferromagnetic Ising model. This step is also generic and applies to a large class of models; see [3]. One interesting implication of our approach is that our gadget and reduction yields always bipartite graphs, and hence we immediately get hardness results for bipartite graphs for all of the models studied in this paper.

We pause to briefly contrast the above proof approach with that in [1], where it was established hardness of identity testing for the antiferromagnetic Ising model. As mentioned earlier, in the antiferromagnetic Ising model, the configurations with the highest weight or likelihood (i.e., the ground states) correspond to the maximum cuts of the original graph. Hence, it is natural to prove hardness of identity testing for the antiferromagnetic Ising model using a reduction from the maximum cut problem. The ground states of ferromagnetic systems, on the other hand, correspond to the monochromatic configurations, so there is no hard optimization problem in the background to utilize in the reduction. (The similar obstacle for RBMs is that the maximum cut problem is trivial in bipartite graphs, so we cannot hope to use it to prove hardness.) We use the hardness of approximating the partition function instead and utilize the unique nature of the phase transition in these models in an essential way.

To reduce the degree of the graphs in our construction we do utilize insights and certain technical lemmas from [1]. Specifically, those concerning the expansion of random near-regular bipartite graphs. We note that the models we consider on these random graphs are different than those in [1]; in particular, we consider mixed models and allowed external fields, whereas in [1] these gadgets are purely antiferromagnetic and there is no external field.

We present our proof approach in the context of the ferromagnetic Potts model, specifically in Section 2 we sketch the prove Theorem 3. The complete proof of Theorem 3 and the proofs for RBMs, i.e., Theorems 1 and 2 which follow the same approach, are provided in the full version [3].

2. Testing ferromagnetic Potts models

In this section we prove Theorem 3, our lower bound for identity testing for the ferromagnetic Potts model. To prove this theorem, we introduce a new methodology to reduce approximate counting (i.e., the problem of finding an FPRAS for the partition function of a model), to identity testing. We believe our methods could be used to establish the hardness of identity testing for other models.

We introduce some useful notation next. Recall that in the introduction we define the families of models $\mathcal{M}_{\text{RBM}}$, $\mathcal{M}^{+}_{\text{RBM}}$, $\mathcal{M}_{\text{POTTS}}$ and $\mathcal{M}^{+}_{\text{POTTS}}$. We formalize and extend this notation as follows.

Definition 4 *For integers $n, d \geq 3$ and $\beta, h \in \mathbb{R}$, let $\mathcal{M}_{\text{POTTS}}(n, d, \beta, h, q)$ denote the family of q-state Potts models on n-vertex graphs $G = (V_G, E_G)$ of maximum degree at most d with edge*

interactions $\beta_G : E_G \to \mathbb{R}$ *such that for every edge* $\{u, v\} \in E_G$, $|\beta_G(\{u, v\})| \leq \beta$, *and external field* $h_G : V_G \times [q] \to \mathbb{R}$ *such that for every vertex* $v \in V_G$ *and spin* $i \in [q]$, $|h_G(v, i)| \leq h$.

Remark 5 *We omit* q *from the notation above as it is usually clear from context. For the special case of* $q = 2$, *i.e., the Ising model, we use* $\mathcal{M}_{\text{ISING}}$; *when* $q = 2$ *and the underlying graph is bipartite we use* $\mathcal{M}_{\text{RBM}}$. *In addition, we add "+" or "−" as a superscript to the notation to denote the corresponding ferromagnetic or antiferromagnetic subfamilies; e.g.,* $\mathcal{M}^+_{\text{POTTS}}(n, d, \beta, h)$ *denotes the subset of ferromagnetic Potts models in* $\mathcal{M}_{\text{POTTS}}(n, d, \beta, h)$. *Finally, we add a circumflex, e.g.,* $\hat{\mathcal{M}}^+_{\text{POTTS}}(n, d, \beta, h)$, *for the subfamily of models where every edge weight is* exactly *equal to* β.

2.1. Step 1: Decision version of approximate counting

Our starting point is always a known hard approximate counting instance. For the ferromagnetic Potts model, we consider the problem of approximating its partition function on a graph G. As mentioned, this problem is known to be #BIS-hard, even under the additional assumptions that all edges have the same interaction parameter $0 < \beta_G = \Theta(1)$ and that there is no external field (i.e., $h = 0$) [25, 21]. Our goal is to design an FPRAS for the partition function $Z_{G,\beta_G} := Z_{G,\beta_G,0}$ using a poly-time algorithm for identity testing, thus establishing the #BIS-hardness of this problem.

Our first step is to reduce the problem of approximating Z_{G,β_G} to a natural decision variant of the problem. This decision version will be more naturally solved by the testing algorithm and is more generally defined as follows:

Definition 6 (Decision r-approximate counting) *Given a Potts model* (G, β_G, h_G), *an approximation ratio* $r > 1$ *and an input* $\hat{Z} \in \mathbb{R}$, *distinguish with probability at least* $5/8$ *between the cases: (i)* $Z_{G,\beta_G,h_G} \leq \frac{1}{r}\hat{Z}$ *and (ii)* $Z_{G,\beta_G,h_G} \geq r\hat{Z}$.

We show that the decision version of approximate counting is as hard as the standard problem of approximating Z_{G,β_G,h_G}. Our proof of this theorem is provided in the full version [3].

Theorem 7 *Let* $n, d \geq 1$ *be integers and let* $\beta, h \geq 0$ *be real numbers. Suppose that there is no* FPRAS *for the counting problem for a family of Potts models* $\mathcal{M}$, *where* $\mathcal{M}$ *is any of the models* $\hat{\mathcal{M}}^+_{\text{POTTS}}(n, d, \beta, h)$, $\hat{\mathcal{M}}^-_{\text{ISING}}(n, d, \beta, h)$ *or* $\hat{\mathcal{M}}^+_{\text{ISING}}(n, d, \beta, h)$. *Then, for any* $c > 0$ *there is no polynomial-time algorithm for the decision version of* n^c*-approximate counting for* $\mathcal{M}$.

2.2. Step 2: Testing instance construction

We first construct a hard instance for the identity testing problem for the ferromagnetic Potts model on general graphs, with no restriction on the maximum degree and with a constant upper bound on the edge interactions. We prove first that identity testing is #BIS-hard in this setting.

Theorem 8 *Consider a ferromagnetic Potts model with no external field* ($h = 0$) *where the interaction on every edge is ferromagnetic and bounded from above by a constant* $\beta_0 > 0$. *Then, there is no polynomial-time identity testing algorithm for the model unless there is an* FPRAS *for #BIS.*

To establish this theorem, we construct an identity testing instance that allows us to solve the decision variant of approximate counting. This theorem does not immediately imply Theorem 3 from

the introduction because we allow the degree to be unbounded; specifically, Theorem 8 establishes hardness for $\mathcal{M}^{+}_{\text{POTTS}}(n, n, \beta, 0)$. The final step of the proof uses this result and a degree-reducing gadget to establish Theorem 3 (see Section 2.3). Our main gadget in the proof of Theorem 8 will be a complete graph; this is known as the *mean-field* case in statistical physics.

2.2.1. THE FERROMAGNETIC MEAN-FIELD q-STATE POTTS MODEL

Let $H = K_m$ be a complete graph on m vertices and let β_H be the interaction strength on the edges of H. By symmetry, the q-state Potts configurations on a complete graph can be described by their "signature"—by "signature" we mean the vector $(\sigma_1, \ldots, \sigma_q) \in \mathbb{Z}^q$ where $\sigma_i \geq 0$ is the number of vertices that have spin i; note that $\sum_{i=1}^{q} \sigma_i = m$.

In the complete graph, the ferromagnetic Potts model is known to undergo an "order-disorder" phase transition. Specifically, there exists a critical value $\beta_H = \mathfrak{B}_o/m$ such that when $\beta_H < \mathfrak{B}_o/m$, long-range correlations do not exist; the system is then said to be in a "disordered" state as the typical configurations have signature $\approx (m/q, \ldots, m/q)$ where each spin has roughly the same density (up to lower order terms). In contrast, when $\beta_H > \mathfrak{B}_o/m$, typical configurations have a dominant spin and the remaining spins are uniformly distributed. These configurations are thus referred to as "majority" configurations. More precisely there exists a constant $\alpha = \alpha(\beta_H) > 1/q$ and, with high probability, configurations from the Gibbs distribution have signature $\approx \left(\alpha m, \frac{(1-\alpha)m}{q-1}, \ldots, \frac{(1-\alpha)m}{q-1}\right)$ up to permutations and lower order terms.

When $q \geq 3$, the phase transition is known to be of *first-order*, which means that at the critical point $\beta_H = \mathfrak{B}_o/m$ both disordered and majority configurations occur with constant probability. This phenomena is referred to as *phase co-existence*, and it is known (or conjectured) to be present in a variety of graphs, being the root reason for the hardness of sampling and counting for the ferromagnetic Potts model. In contrast, in the Ising model (i.e., when $q = 2$), there is a *second-order phase transition* and the majority density $\alpha(\mathfrak{B}_o)$ is $1/q$ at the critical point; hence these two phases – disordered and majority – coincide at this point.

We now formalize the notion of the majority phase M, the disordered phase D, and the remaining configurations S with their corresponding partition functions Z_H^{M}, Z_H^{D}, and Z_H^{S}. The majority phase is defined with respect to a fixed constant $\hat{\alpha} = \hat{\alpha}(\mathfrak{B}_o)$ which is the density of the dominant color at the phase coexistence point $\mathfrak{B}_o/m$. Let Ω_H denote the set of Potts configurations on H and for $\sigma \in \Omega_H$, let $(\sigma_1, \ldots, \sigma_q) \in \mathbb{Z}^q$ denote its signature. Consider the following sets:

$$M := \left\{ \sigma \in \Omega_H \,\middle|\, \exists j \in [q] : \ |\sigma_j - \hat{\alpha} m| \leq m^{3/4} \text{ and } \left|\sigma_i - \frac{1-\hat{\alpha}}{q-1} m\right| \leq m^{3/4} \text{ for } i \in [q] \setminus \{j\} \right\},$$

$D := \{\sigma \in \Omega_H \,\big|\, \forall i \in [q] : \ |\sigma_i - m/q| \leq m^{3/4}\}$, and $S := \Omega_H \setminus (M \cup D)$.

For a configuration σ on H, let $w_H^{\sigma}(\beta_H) = \exp\left(\sum_{\{u,v\} \in E(H)} \beta_H \mathbb{1}(\sigma(u) = \sigma(v))\right)$ denote the weight of σ in the mean-field model (H, β_H). Consider the contributions of each type of configuration to the partition function. That is, $Z_H^{\text{M}}(\beta_H) := \sum_{\sigma \in M} w_H^{\sigma}(\beta_H)$, $Z_H^{\text{D}}(\beta_H) := \sum_{\sigma \in D} w_H^{\sigma}(\beta_H)$, and $Z_H^{\text{S}}(\beta_H) := \sum_{\sigma \in S} w_H^{\sigma}(\beta_H)$. Hence, the partition function of (H, β_H) is given by $Z_H(\beta_H) = Z_H^{\text{M}}(\beta_H) + Z_H^{\text{D}}(\beta_H) + Z_H^{\text{S}}(\beta_H)$. We note that in our reduction later, we will choose a specific $\beta_H > 0$ depending on the instance of the approximate counting problem and the parameters of the identity testing algorithm; hence, to emphasize the effect of β_H, we parameterize Z_H^{M} (and other functions in this section) in terms of β_H.

The next two lemmas concern the mean-field Potts model near the critical point $\mathfrak{B}_o/m$. We note that as a consequence of the first-order phase transition, there is a critical window around $\mathfrak{B}_o/m$ where the non-dominant phase (i.e., disorder or majority) is still much more likely than any other type configurations; this phenomena is known as *metastability* and will also be crucial for us.

First we establish that in the critical window around $\mathfrak{B}_o/m$ the majority M and disordered D configurations are exponentially more likely than the remaining configurations S. Several variants of this result have been proved in some fashion before, e.g., [5, 37, 25, 13, 24, 19, 2]; however, the precise bound we require in our proofs does not seem to be available in the literature.

Lemma 1 *There exists constants $c, c' > 0$ such that for any β_H satisfying $|\beta_H - \mathfrak{B}_o/m| \le c' m^{-3/2}$, we have $Z_H^{\mathrm{S}}(\beta_H) \le \min\{Z_H^{\mathrm{M}}(\beta_H), Z_H^{\mathrm{D}}(\beta_H)\} \exp(-c\sqrt{m})$.*

In addition, we show that we can find in $\mathrm{poly}(m)$ time a value for the parameter β_H in the critical window to achieve a specified ratio R of the majority partition function $Z_H^{\mathrm{M}}(\beta_H)$ to the disordered partition function $Z_H^{\mathrm{D}}(\beta_H)$.

Lemma 2 *There exist constants $c, c' > 0$ such that for any $R \in [e^{-c\sqrt{m}}, e^{c\sqrt{m}}]$ and any constant $\delta \in (0,1)$, we can efficiently find $\beta_H > 0$ in $\mathrm{poly}(m)$ time such that $|\beta_H - \mathfrak{B}_o/m| \le c' m^{-3/2}$ and $(1-\delta)R \le \frac{Z_H^{\mathrm{M}}(\beta_H)}{Z_H^{\mathrm{D}}(\beta_H)} \le R$.*

The proof of these two lemmas is provided in the full version [3].

2.2.2. IDENTITY TESTING REDUCTION

Visible Model Construction. Let (G, β_G) be the instance of the ferromagnetic Potts model with no external field (i.e., $h = 0$) for which we are trying to approximate the partition function Z_{G,β_G}; we shall assume $G = (V_G, E_G)$ is an N-vertex graph and that every edge has interaction strength $0 < \beta_G = \Theta(1)$. Let $H = (V_H, E_H)$ be a complete graph on $m = N^{10}$ vertices. The graph $F = (V_F, E_F)$ is the result of connecting the vertices of H and G with a complete bipartite graph $K_{m,N}$ with edges $E_{m,N}$; that is, $V_F = V_G \cup V_H$ and $E_F = E_H \cup E_G \cup E_{m,N}$. We consider the Potts model on the graph F with edge interactions $\beta_F : E_F \to \mathbb{R}$ such that $\beta_F(e) = \beta_H$ when $e \in E_H$; $\beta_F(e) = \beta_G$ when $e \in E_G$; and $\beta_F(e) = \beta$ when $e \in E_{m,N}$, where $\beta_H, \beta > 0$ will be chosen later.

We use $n := N + m$ for the number of vertices of F, and, with a slight abuse of notation, we use F for the Potts model (F, β_F) which will play the role of the visible model in our reduction; μ_F denotes the corresponding Gibbs distribution.

We study first the properties of "typical" configurations on G conditional on a configuration σ on the complete graph H. Let Ω_F, Ω_H and Ω_G be the set of Potts configuration on the graph F, H and G respectively; note that $\Omega_F = \Omega_H \times \Omega_G$. For $\sigma \in \Omega_H$, define

$$Z_F^{\sigma}(\beta_H) := \sum_{\eta \in \Omega_F : \eta(V_H) = \sigma} w_F^{\eta}(\beta_H)$$

where $w_F^{\eta}(\beta_H) = \exp\left(\sum_{\{u,v\} \in E_F} \beta_F(\{u,v\}) \mathbb{1}(\eta(u) = \eta(v))\right)$; that is, $Z_F^{\sigma}(\beta_H)$ is the total contribution to the partition $Z_F(\beta_H)$ of F of the configurations that agree with σ on H.

If we fix a configuration σ on H and look at the configuration on G (under the Gibbs distribution on F conditional on σ) then σ will act as an external field on the vertices of G. We show that if σ is in

the majority phase (i.e., in the set M), then the configuration on G will be monochromatic with high probability as these configurations will maximize the number of monochromatic edges between G and H. In contrast, when σ is in the disordered phase (i.e., in D), then every configuration on G will have (roughly) the same number of monochromatic edges between G and H; hence, the partition function $Z_F^\sigma(\beta_H)$ in this case will be proportional to Z_{G,β_G}.

To formalize this, we split the partition function of F into three parts depending on the signature on the complete graph H. Let $Z_F^{\mathrm{M}}(\beta_H) = \sum_{\sigma\in M} Z_F^\sigma(\beta_H)$, $Z_F^{\mathrm{D}}(\beta_H) = \sum_{\sigma\in D} Z_F^\sigma(\beta_H)$ and $Z_F^{\mathrm{S}}(\beta_H) = \sum_{\sigma\in S} Z_F^\sigma(\beta_H)$. Then, $Z_F(\beta_H) = Z_F^{\mathrm{M}}(\beta_H) + Z_F^{\mathrm{D}}(\beta_H) + Z_F^{\mathrm{S}}(\beta_H)$.

The following lemma details the above description of the properties of configurations on the original instance G conditional on the configuration on the complete graph H.

Lemma 3 *For any constants $\delta \in (0,1)$ and $c > 0$, and any β_H such that $|\beta_H - \mathfrak{B}_o/m| \le cm^{-3/2}$, there exists constants $c_1, c_2 > 0$ such that for any $\beta \in [c_1 N/m, c_2/Nm^{3/4}]$:*

1. *When the configuration on H is in the majority phase, the configuration on G is likely to be monochromatic; more precisely,*

$$e^{-\delta} \cdot Z_H^{\mathrm{M}} \cdot \exp\left(\hat{\alpha}\beta Nm + \beta_G|E_G|\right) \le Z_F^{\mathrm{M}}(\beta_H) \le e^{\delta} \cdot Z_H^{\mathrm{M}} \cdot \exp\left(\hat{\alpha}\beta Nm + \beta_G|E_G|\right). \quad (1)$$

2. *When the configuration on H is in the disordered phase, the configuration on G will have very limited influence from the configuration on H; more precisely,*

$$e^{-\delta} \cdot Z_H^{\mathrm{D}} \cdot Z_G \cdot \exp\left(\beta Nm/q\right) \le Z_F^{\mathrm{D}}(\beta_H) \le e^{\delta} \cdot Z_H^{\mathrm{D}} \cdot Z_G \cdot \exp\left(\beta Nm/q\right). \quad (2)$$

3. *The remaining configurations on H have a small contribution to the partition function of the model F; more precisely, $Z_F^{\mathrm{S}}(\beta_H) \le Z_F(\beta_H)\exp\left(-\Omega(\sqrt{m})\right)$.*

The proof of this lemma, which uses Lemma 1, can be found in the full version [3].

Hidden Model Construction. We now construct our hidden model and show that we can efficiently generate samples from its Gibbs distribution. Let F^* be the graph obtained by our construction above where we replace the graph G by a complete graph on N vertices. More precisely, let $K = K_N$ be a complete graph on N vertices and let F^* be the graph that results from connecting the vertices of K and H with a complete bipartite graph $K_{N,m}$.

The edges of K have parameter $\beta_K = \beta_G + 4\log q$, whereas the remaining edges have the same interaction strength as in F; that is, edges between K and H will have parameter β and those in H parameter β_H. This Potts model on F^*, which again with a slight abuse of notation we denote by F^*, will act as the hidden model. We choose $\beta_K = \beta_G + 4\log q$, so that K is more likely to be monochromatic than G. Let μ_{F^*} the corresponding Gibbs distribution on F^*. We establish next that we can efficiently generate samples from μ_{F^*}; see [3] for the proof.

Lemma 4 *There is an exact sampling algorithm for the distribution μ_{F^*} with running time* $\mathrm{poly}(n)$.

Proof Overview. We provide the high-level idea of the reduction next. Recall that our goal is to provide a polynomial-time algorithm for the decision version of the r-approximate counting problem for the ferromagnetic Potts model (G, β_G). That is, for a real number $\hat{Z}$ we want to check whether $Z_G \le \frac{1}{r}\hat{Z}$ or $Z_G \ge r\hat{Z}$, where $Z_G := Z_{G,\beta_G}$ is the partition function of (G, β_G).

For any "reasonable" $\hat{Z} \in \mathbb{R}$ (i.e., $\hat{Z}$ not too small or too large, in which case the approximate counting problem is trivial), we can find a value of the parameter β_H for our construction such that

$$\frac{Z_F^{\mathrm{D}}(\beta_H)}{Z_F^{\mathrm{M}}(\beta_H)} \approx \frac{1}{\sqrt{\varepsilon L}} \frac{Z_G}{\hat{Z}},$$

where $L = L(n)$ and $\varepsilon = \varepsilon(n)$ are the sample complexity and accuracy parameter of the testing algorithm, respectively. This is possible because of the first-order phase transition of the ferromagnetic mean-field q-state Potts model for $q \geq 3$, and the associated phase coexistence and metastability phenomena discusses earlier; see Section 2.2.1. (Specifically, by Lemma 2 we can find β_H so that $Z_H^{\mathrm{M}}(\beta_H)/Z_H^{\mathrm{D}}(\beta_H) \approx R$ for any target R, and then we can use Lemma 3 to translate this value to a value for $Z_G \cdot Z_F^{\mathrm{M}}(\beta_H)/Z_F^{\mathrm{D}}(\beta_H)$.)

Now, for this choice of β_H and setting $r \approx \sqrt{L/\varepsilon}$, note that if $Z_G \leq \frac{1}{r}\hat{Z}$, then the ratio $Z_F^{\mathrm{D}}(\beta_H)/Z_F^{\mathrm{M}}(\beta_H)$ is small ($\lesssim 1/L$). Conversely, when $Z_G \geq r\hat{Z}$, the ratio is large ($\gtrsim 1/\varepsilon$). Therefore, to distinguish whether $Z_G \leq \frac{1}{r}\hat{Z}$ or $Z_G \geq r\hat{Z}$ it is sufficient to determine whether the ratio $Z_F^{\mathrm{D}}(\beta_H)/Z_F^{\mathrm{M}}(\beta_H)$ is small or large. For this we can use the identity testing algorithm. In particular, when the ratio is small ($\lesssim 1/L$), the majority phase of H is dominant in F, and G will likely be monochromatic. Since this is also the case in F^* (i.e., K is monochromatic with high probability), then the models F and F^* will be close in total variation distance ($\lesssim 1/L$), and the testing algorithm using only L samples would output YES. Otherwise, when $Z_F^{\mathrm{D}}(\beta_H)/Z_F^{\mathrm{M}}(\beta_H)$ is large ($\gtrsim 1/\varepsilon$), the disorder phase is dominant, so F and F^* are likely to disagree on the spins of G and K; this implies that their total variation distance is large ($\gtrsim 1-\varepsilon$), and so the tester would output NO.

We start with a useful technical lemma which is proved in [3] using Lemmas 2 and 3.

Lemma 5 *Let $\varepsilon \in (0,1)$ be a constant, $L = L(n) = \mathrm{poly}(n)$ and $r = 96\varepsilon^{-1}\sqrt{\varepsilon L + 1}$. Suppose $\hat{Z} \in \mathbb{R}$ is such that $rq\exp(\beta_G|E_G|) \leq \hat{Z} \leq \frac{1}{r}q^N\exp(\beta_G|E_G|)$. Then, there exists constants $c, c_1, c_2 > 0$ such that the following holds. For any $\beta \in \left[\frac{c_1 N}{m}, \frac{c_2}{Nm^{3/4}}\right]$, we can find $\beta_H > 0$ in the range $|\beta_H - \mathfrak{B}_o/m| \leq cm^{-3/2}$ in $\mathrm{poly}(n)$ time such that all of the following holds:*

(i) $\frac{1}{4\sqrt{\varepsilon L+1}}\frac{Z_G}{\hat{Z}} \leq \frac{Z_F^{\mathrm{D}}(\beta_H)}{Z_F^{\mathrm{M}}(\beta_H)} \leq \frac{1}{\sqrt{\varepsilon L+1}}\frac{Z_G}{\hat{Z}}$, *and* $\frac{Z_F^{\mathrm{S}}(\beta_H)}{Z_F(\beta_H)} \leq e^{-c_3\sqrt{m}}$;

(ii) $\frac{Z_{F^*}^{\mathrm{D}}(\beta_H)}{Z_{F^*}^{\mathrm{M}}(\beta_H)} \leq \frac{2}{r\sqrt{\varepsilon L+1}}$, *and* $\frac{Z_{F^*}^{\mathrm{S}}(\beta_H)}{Z_{F^*}(\beta_H)} \leq e^{-c_3\sqrt{m}}$;

(iii) If $Z_G \leq \frac{1}{r}\hat{Z}$, then $\|\mu_F - \mu_{F^*}\|_{\mathrm{TV}} \leq \frac{1}{16L}$;

(iv) If $Z_G \geq r\hat{Z}$, then $\|\mu_F - \mu_{F^*}\|_{\mathrm{TV}} \geq 1-\varepsilon$.

2.2.3. A GENERIC REDUCTION FROM COUNTING TO TESTING

Theorem 8 will follow from Lemmas 4 and 5 using the following general reduction from the decision version of r-approximate counting to testing, which we prove in the full version [3].

Theorem 9 *Let (G, β_G, h_G) be a Potts model on an N-vertex graph G with partition function Z_G and let $\hat{Z} \in \mathbb{R}$. Let $\varepsilon \in (0,1)$ be a constant, $n = \mathrm{poly}(N)$ and suppose there exists an ε-identity testing algorithm for a family of Potts models $\mathcal{M}$ on n-vertex graphs with sample complexity $L = L(n) = \mathrm{poly}(n)$ and $\mathrm{poly}(n)$ running time. Suppose that given (G, β_G, h_G), $\hat{Z}$, ε and L, there is $r = \mathrm{poly}(L, \varepsilon^{-1})$ such that we can construct two models $F, F^* \in \mathcal{M}$ in $\mathrm{poly}(n)$ time satisfying:*

(i) If $Z_G \le \frac{1}{r}\hat{Z}$*, then* $\|\mu_F - \mu_{F^*}\|_{\mathrm{TV}} \le \frac{1}{16L}$*;*

(ii) If $Z_G \ge r\hat{Z}$*, then* $\|\mu_F - \mu_{F^*}\|_{\mathrm{TV}} \ge 1 - \varepsilon$*; and*

(iii) We can sample from a distribution $\mu_{F^*}^{\mathrm{ALG}}$ *such that* $\left\|\mu_{F^*} - \mu_{F^*}^{\mathrm{ALG}}\right\|_{\mathrm{TV}} \le \delta$ *in time* $\mathrm{poly}(n, \delta^{-1})$.

Then, there is a $\mathrm{poly}(N)$ *running time algorithm for the decision version of* r*-approximate counting for* (G, β_G, h_G) *that succeeds with probability at least* $5/8$.

2.2.4. PROOF OF THEOREM 8

We can now establish the hardness of testing for the ferromagnetic Potts model on general graphs.

Proof of Theorem 8 Consider the ferromagnetic Potts model on an N-vertex graph $G=(V_G, E_G)$ with constant edge weight β_G in every edge and no external field. Let $\hat{Z} > 0$ be a real number and let $n = N^{10} + N$. Suppose there is an ε-identity testing algorithm for $\mathcal{M}^{+}_{\mathrm{POTTS}}(n, n, \beta_G, 0)$ with sample complexity $L = L(n) = \mathrm{poly}(n)$ and running time $\mathrm{poly}(n)$. Let $r = 96\varepsilon^{-1}\sqrt{\varepsilon L + 1}$; our goal is to determine whether $Z_G \le \frac{1}{r}\hat{Z}$ or $Z_G \ge r\hat{Z}$ where $Z_G := Z_{G,\beta_G}$.

We construct the Potts models F and F^* as describe in Section 2.2.2 with corresponding Gibbs distributions μ_F and μ_{F^*} using the values of β and β_H supplied by Lemma 5; hence the models F and F^* belong to $\mathcal{M}^{+}_{\mathrm{POTTS}}(n, n, \beta_G, 0)$, since $\beta_G > \max\{\beta, \beta_H\}$.

Lemmas 5 ensures that when $rq\exp(\beta_G|E_G|) \le \hat{Z} \le \frac{q^N}{r}\exp(\beta_G|E_G|)$ conditions (i) and (ii) in Theorem 9 are satisfied. Moreover, Lemma 4 gives condition (iii). Thus, Theorem 9 implies that we have an algorithm for the decision version of r-approximate counting for the Potts model on G when $rq\exp(\beta_G|E_G|) \le \hat{Z} \le \frac{q^N}{r}\exp(\beta_G|E_G|)$. Meanwhile, we can bound Z_G crudely by $qe^{\beta_G|E_G|} \le Z_G \le q^N e^{\beta_G|E_G|}$. Thus, if $\hat{Z} < rq\exp(\beta_G|E_G|) \le rZ_G$, we can output $\hat{Z} \le \frac{1}{r}Z_G$. Similarly, when $\hat{Z} > \frac{1}{r}q^N\exp(\beta_G|E_G|) \ge \frac{1}{r}Z_G$ we can output $\hat{Z} \ge rZ_G$. Therefore, we have a $\mathrm{poly}(N)$ algorithm for the decision version of r-approximate counting for $\hat{\mathcal{M}}^{+}_{\mathrm{POTTS}}(N, N, \beta_G, 0)$ where $N = \Theta(n^{1/10})$, $r = \mathrm{poly}(N)$ and $\beta_G = \Theta(1)$. The result then follows from Theorem 7 and the fact that there is no FPRAS for $\hat{\mathcal{M}}^{+}_{\mathrm{POTTS}}(N, N, \beta_G, 0)$ unless there is one for #BIS [25, 21]. ■

2.3. Step 3: Degree reduction

The following result is a special case of a more general theorem we prove in the full version [3], and it provides a reduction from identity testing in the family $\mathcal{M}_{\mathrm{POTTS}}(\hat{n}, d, \hat{\beta}, \hat{h})$ to identity testing in $\mathcal{M}_{\mathrm{POTTS}}(n, n, \beta, h)$, under some mild assumptions on the model parameters; this allows us to deduce the hardness of testing on graphs of bounded degree as stated in Theorem 3.

Theorem 10 *Let* $\hat{n}, d \in \mathbb{N}^+$ *be such that* $3 \le d \le \hat{n}^{1-\rho}$ *for some constant* $\rho \in (0, 1)$*. Suppose that for some constants* $\beta, h \ge 0$ *there is no* $\mathrm{poly}(n)$ *running time* ε*-identity testing algorithm for* $\mathcal{M}_{\mathrm{POTTS}}(n, n, \beta, h)$*. Then there exists a constant* $c \in (0, 1)$ *such that, for any constant* $\hat{\varepsilon} > \varepsilon$ *there is no* $\mathrm{poly}(\hat{n})$ *running time* $\hat{\varepsilon}$*-identity testing algorithm for* $\mathcal{M}_{\mathrm{POTTS}}(\hat{n}, d, \hat{\beta}, \hat{h})$ *provided* $\hat{\beta}d = \omega(\log \hat{n})$ *and* $\hat{h} \le h\hat{n}^{-c}$.

We conclude with the proof of Theorem 3.

Proof of Theorem 3 Follows from Theorems 8 and 10. ■

References

[1] I. Bezáková, A. Blanca, Z. Chen, D. Štefankovič, and E. Vigoda. Lower bounds for testing graphical models: Colorings and antiferromagnetic ising models. *Journal of Machine Learning Research*, 21(25):1–62, 2020.

[2] A. Blanca and A. Sinclair. Dynamics for the mean-field random-cluster model. *Proceedings of the 19th International Workshop on Randomization and Computation*, pages 528–543, 2015.

[3] A. Blanca, Z. Chen, D. Štefankovič, and E. Vigoda. Hardness of Identity Testing for Restricted Boltzmann Machines and Potts models. *arXiv preprint arXiv:2004.10805*, 2020.

[4] A. Bogdanov, E. Mossel, and S. Vadhan. The Complexity of Distinguishing Markov Random Fields. In A. Goel, K. Jansen, J.D.P Rolim, and R. Rubinfeld, editors, *Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques*, pages 331–342, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg. ISBN 978-3-540-85363-3.

[5] B. Bollobás, G. Grimmett, and S. Janson. The random-cluster model on the complete graph. *Probability Theory and Related Fields*, 104(3):283–317, 1996. ISSN 0178-8051. doi: 10.1007/BF01213683.

[6] G. Bresler. Efficiently learning Ising models on arbitrary graphs. In *Proceedings of the 47th Annual ACM Symposium on Theory of Computing (STOC)*, pages 771–782, 2015.

[7] G. Bresler, F. Koehler, and A. Moitra. Learning restricted Boltzmann machines via influence maximization. In *Proceedings of the 51st Annual ACM Symposium on Theory of Computing (STOC)*, pages 828–839, 2019.

[8] A.A. Bulatov, M. Dyer, L.A. Goldberg, M. Jerrum, and C. McQuillan. The expressibility of functions on the Boolean domain, with applications to Counting CSPs. *Journal of the ACM (JACM)*, 60(5):32, 2013.

[9] J.-Y. Cai, A. Galanis, L.A. Goldberg, H. Guo, M. Jerrum, D. Štefankovič, and E. Vigoda. $\#BIS$-hardness for 2-spin systems on bipartite bounded degree graphs in the tree non-uniqueness region. *Journal of Computer and System Sciences*, 82(5):690–711, 2016.

[10] X. Chen, M. Dyer, L.A. Goldberg, M. Jerrum, P. Lu, C. McQuillan, and D. Richerby. The complexity of approximating conservative counting CSPs. *Journal of Computer and System Sciences*, 81(1):311–329, 2015.

[11] A. Collevecchio, T.M. Garoni, T. Hyndman, and D. Tokarev. The Worm process for the Ising model is rapidly mixing. *Journal of Statistical Physics*, 164(5):1082–1102, 2016.

[12] I. Csiszár and P.C. Shields. Information theory and statistics: A tutorial. *Foundations and Trends® in Communications and Information Theory*, 1(4):417–528, 2004.

[13] P. Cuff, J. Ding, O. Louidor, E. Lubetzky, Y. Peres, and A. Sly. Glauber dynamics for the mean-field Potts model. *Journal of Statistical Physics*, 149(3):432–477, 2012. ISSN 0022-4715. doi: 10.1007/s10955-012-0599-2.

[14] C. Daskalakis, N. Dikkala, and G. Kamath. Testing Ising models. In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1989–2007, 2018.

[15] H. Duminil-Copin, M. Gagnebin, M. Harel, I. Manolescu, and V. Tassion. Discontinuity of the phase transition for the planar random-cluster and Potts models with $q > 4$. *arXiv preprint arXiv:1611.09877*, 2016.

[16] H. Duminil-Copin, V. Sidoravicius, and V. Tassion. Continuity of the Phase Transition for Planar Random-Cluster and Potts Models with $1 \leq q \leq 4$. *Communications in Mathematical Physics*, 349(1):47–107, 2017.

[17] M. Dyer, L.A. Goldberg, C. Greenhill, and M. Jerrum. The relative complexity of approximate counting problems. *Algorithmica*, 38(3):471–500, 2004.

[18] M. Dyer, L.A. Goldberg, and M. Jerrum. An approximation trichotomy for Boolean #CSP. *Journal of Computer and System Sciences*, 76(3-4):267–277, 2010.

[19] A. Galanis, D. Štefankovič, and E. Vigoda. Swendsen-Wang algorithm on the mean-field Potts model. *Proceedings of the 19th International Workshop on Randomization and Computation*, pages 815–828, 2015.

[20] A. Galanis, L.A. Goldberg, and M. Jerrum. Approximately Counting H-Colourings is #BIS-Hard. *SIAM Journal on Computing*, 45(3):680–711, 2016.

[21] A. Galanis, D. Štefankovič, E. Vigoda, and L. Yang. Ferromagnetic Potts model: Refined #BIS-hardness and related results. *SIAM Journal on Computing*, 45(6):2004–2065, 2016.

[22] S. Geman and C. Graffigne. Markov random field image models and their applications to computer vision. In *Proceedings of the International Congress of Mathematicians*, volume 1, pages 1496–1517. Berkeley, CA, 1986.

[23] H.-O. Georgii. *Gibbs measures and phase transitions*, volume 9. Walter de Gruyter, 2011.

[24] R. Gheissari, E. Lubetzky, and Y. Peres. Exponentially slow mixing in the mean-field Swendsen-Wang dynamics. In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1981–1988. SIAM, 2018.

[25] L.A. Goldberg and M. Jerrum. Approximating the partition function of the ferromagnetic Potts model. *Journal of the ACM*, 59(5):25, 2012.

[26] C. Greenhill. The complexity of counting colourings and independent sets in sparse graphs and hypergraphs. *Computational Complexity*, 9(1):52–72, 2000.

[27] G.R. Grimmett. *The Random-Cluster Model*, volume 333 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006.

[28] H. Guo and M. Jerrum. Random cluster dynamics for the Ising model is rapidly mixing. In *Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1818–1827, 2017.

[29] L. Hamilton, F. Koehler, and A. Moitra. Information theoretic properties of Markov random fields, and their algorithmic applications. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 2460–2469, 2017.

[30] G.E. Hinton. Training products of experts by minimizing contrastive divergence. *Neural computation*, 14(8):1771–1800, 2002.

[31] G.E. Hinton and R.R. Salakhutdinov. Replicated softmax: an undirected topic model. In *Advances in neural information processing systems*, pages 1607–1614, 2009.

[32] G.E. Hinton, S. Osindero, and Y-W. Teh. A fast learning algorithm for deep belief nets. *Neural computation*, 18(7):1527–1554, 2006.

[33] M. Jerrum and A. Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM Journal on computing*, 22(5):1087–1116, 1993.

[34] M.R. Jerrum, L.G. Valiant, and V.V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 43:169–188, 1986.

[35] A. Klivans and R. Meka. Learning graphical models using multiplicative weights. In *Proceedings of the 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 343–354. IEEE, 2017.

[36] V. Kolmogorov. A faster approximation algorithm for the Gibbs partition function. In *Proceedings of the 31st Conference On Learning Theory*, volume 75 of *Proceedings of Machine Learning Research*, pages 228–249. PMLR, 06–09 Jul 2018.

[37] M.J. Luczak and T. Łuczak. The phase transition in the cluster-scaled model of a random graph. *Random Structures & Algorithms*, 28(2):215–246, 2006.

[38] K.P. Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.

[39] S. Osindero and G.E. Hinton. Modeling image patches with a directed hierarchy of Markov random fields. In *Advances in neural information processing systems*, pages 1121–1128, 2008.

[40] D. Randall and D. Wilson. Sampling spin configurations of an Ising system. In *Proceedings of the 10th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, volume 17, pages 959–960, 1999.

[41] R. Salakhutdinov and G.E. Hinton. Deep Boltzmann machines. In *Artificial intelligence and statistics*, pages 448–455, 2009.

[42] R. Salakhutdinov, A. Mnih, and G.E. Hinton. Restricted Boltzmann machines for collaborative filtering. In *Proceedings of the 24th international conference on Machine learning*, pages 791–798, 2007.

[43] N.P. Santhanam and M.J. Wainwright. Information-theoretic limits of selecting binary graphical models in high dimensions. *IEEE Trans. Information Theory*, 58(7):4117–4134, 2012.

[44] A. Sly. Computational transition at the uniqueness threshold. In *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 287–296, 2010.

[45] A. Sly and N. Sun. The computational hardness of counting in two-spin models on d-regular graphs. In *Proceedings of the 53rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 361–369, 2012.

[46] P. Smolensky. Information processing in dynamical systems: Foundations of harmony theory. Technical report, Colorado Univ at Boulder Dept of Computer Science, 1986.

[47] D. Štefankovič, S. Vempala, and E. Vigoda. Adaptive simulated annealing: A near-optimal connection between sampling and counting. *Journal of the ACM (JACM)*, 56(3):1–36, 2009.

[48] R.H. Swendsen and J.S. Wang. Nonuniversal critical dynamics in Monte Carlo simulations. *Physical Review Letters*, 58:86–88, 1987.

[49] S.P Vadhan. The complexity of counting in sparse, regular, and planar graphs. *SIAM Journal on Computing*, 31(2):398–427, 2001.

[50] L.G. Valiant. The complexity of enumeration and reliability problems. *SIAM Journal on Computing*, 8(3):410–421, 1979.

[51] M. Vuffray, S. Misra, A. Lokhov, and M. Chertkov. Interaction screening: Efficient and sample-optimal learning of Ising models. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 2595–2603, 2016.

[52] M. Vuffray, S. Misra, and A.Y. Lokhov. Efficient learning of discrete graphical models. *arXiv preprint arXiv:1902.00600*, 2019.

[53] S. Wu, S. Sanghavi, and A.G. Dimakis. Sparse logistic regression learns all discrete pairwise graphical models. In *Advances in Neural Information Processing Systems*, pages 8069–8079, 2019.