Received: 21 April 2021 | Revised: 20 July 2022 | Accepted: 8 August 2022 | Published on: 23 December 2022

DOI: 10.1002/rsa.21131

RESEARCH ARTICLE

WILEY

Efficient sampling and counting algorithms for the Potts model on $\mathbb{Z}^d$ at all temperatures

Christian Borgs[1] | **Jennifer Chayes**[1] | **Tyler Helmuth**[2] | **Will Perkins**[3] | **Prasad Tetali**[4]

[1]University of California Berkeley, Berkeley, California, USA
[2]University of Durham, Durham, UK
[3]Georgia Institute of Technology, Atlanta, Georgia, USA
[4]Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

Correspondence
Will Perkins, Georgia Institute of Technology, Atlanta, GA, USA.
Email: william.perkins@gmail.com

Funding information
EPSRC, Grant/Award Number: EP/P003656/1; NSF, Grant/Award Numbers: DMS-1847451, CCF-1934915, DMS-1811935.

Abstract

For $d \geq 2$ and all $q \geq q_0(d)$ we give an efficient algorithm to approximately sample from the q-state ferromagnetic Potts and random cluster models on finite tori $(\mathbb{Z}/n\mathbb{Z})^d$ for any inverse temperature $\beta \geq 0$. This shows that the physical phase transition of the Potts model presents no algorithmic barrier to efficient sampling, and stands in contrast to Markov chain mixing time results: the Glauber dynamics mix slowly at and below the critical temperature, and the Swendsen–Wang dynamics mix slowly at the critical temperature. We also provide an efficient algorithm (an FPRAS) for approximating the partition functions of these models at all temperatures. Our algorithms are based on representing the random cluster model as a contour model using Pirogov–Sinai theory. The main innovation of our approach is an algorithmic treatment of unstable ground states, which is essential for our algorithms to apply to all inverse temperatures β.

KEYWORDS

approximate counting and sampling, phase transition, Pirogov–Sinai theory, Potts model, random cluster model

An extended abstract of this article appeared at STOC 2020 [6].

1 | INTRODUCTION

The Potts model is a probability distribution on assignments of q colors to the vertices of a finite graph G. For $\sigma \in [q]^{V(G)} := \{1, 2, \dots, q\}^{V(G)}$ let

$$H_G(\sigma) := \sum_{(i,j)\in E(G)} \delta_{\sigma_i \neq \sigma_j} \tag{1}$$

be the the number of bichromatic edges of G under the coloring σ. The *q-state ferromagnetic Potts model* at inverse temperature $\beta \geq 0$ is the probability distribution μ_G^{Potts} on $[q]^{V(G)}$ defined by

$$\mu_G^{\text{Potts}}(\sigma) := \frac{e^{-\beta H_G(\sigma)}}{Z_G^{\text{Potts}}(\beta)}, \qquad Z_G^{\text{Potts}}(\beta) := \sum_{\sigma\in[q]^{V(G)}} e^{-\beta H_G(\sigma)}. \tag{2}$$

The normalizing constant $Z_G^{\text{Potts}}(\beta)$ is the Potts model partition function. Since $\beta \geq 0$, monochromatic edges are preferred. This is often referred to as the *ferromagnetic* Potts model.

In this article, we are interested in computational aspects of the Potts model. To this end, we view Z_G^{Potts} and μ_G^{Potts} as functions and probability measures indexed by finite graphs G, and consider two computational tasks associated to these objects. The first is the *approximate counting* problem: for a partition function Z_G and error tolerance $\epsilon > 0$, compute a number $\hat{Z}$ so that $e^{-\epsilon}\hat{Z} \leq Z_G \leq e^{\epsilon}\hat{Z}$. We say that such a $\hat{Z}$ is an *ϵ-relative approximation* to Z_G. The second is the *approximate sampling* problem: for a probability measure μ_G and error tolerance $\epsilon > 0$, output a random configuration $\hat{\sigma}$ with distribution $\hat{\mu}$ so that $\|\hat{\mu} - \mu_G\|_{TV} < \epsilon$. We say $\hat{\sigma}$ is an *ϵ-approximate sample* from μ_G.

Approximate counting and sampling algorithms can always be obtained by brute force in time exponential in the size of the graph, and the interesting question is if more efficient algorithms exist. To formalize this, a *fully polynomial-time approximation scheme* (FPTAS) is an algorithm that given G and $\epsilon > 0$ returns an ϵ-relative approximation to Z_G and runs in time polynomial in $|V(G)|$ and $1/\epsilon$. If the algorithm uses randomness it is a *fully polynomial-time randomized approximation scheme* (FPRAS). A randomized algorithm that given G and $\epsilon > 0$ outputs an ϵ-approximate sample from μ_G and runs in time polynomial in both $|V(G)|$ and $1/\epsilon$ is an *efficient sampling scheme*. These notions are standard in the study of the computational complexity of approximate sampling and counting, see Section 1.1 below.

One of the main result of this article is the development of an FPRAS and an efficient sampling scheme for the q-state Potts model on the discrete tori $\mathbb{T}_n^d = (\mathbb{Z}/n\mathbb{Z})^d$ for *all* inverse temperatures $\beta \geq 0$, provided q is large enough as a function of d.

Theorem 1.1. *For all $d \geq 2$ there exists $q_0 = q_0(d)$ such that for $q \geq q_0$ and all inverse temperatures $\beta \geq 0$ there is an FPRAS and efficient sampling scheme for the q-state Potts model at inverse temperature β on the torus $\mathbb{T}_n^d$.*

If ϵ is not too small, meaning $\epsilon \geq \exp(-O(n^{d-1}))$, our approximate counting algorithm is deterministic. We will comment on this further in what follows, see below Theorem 1.2.

In the next section we discuss the context and motivation behind Theorem 1.1. The remainder of the introduction then turns to some relevant facts about the Potts model (Section 1.2), the closely related random cluster model (Section 1.3), and a discussion of our proof strategy and the main challenges in proving these results (Section 1.4).

1.1 | Context and motivation: Approximation algorithms and computational phase transitions

For many statistical mechanics models like the Potts model *exact* computation of the partition function has long been known to be #P-hard, even for restricted classes of graphs and parameters. In particular, if P≠NP, this task cannot be performed in polynomial time. Current research therefore is focused on approximate counting.

For some special models (the ferromagnetic Ising model [28], the monomer-dimer model [27]), there is an FPRAS for all graphs and all parameters. For other models, the computational complexity of approximate counting and sampling depends on the class of graphs and on the parameters of the model. These models exhibit *computational phase transitions*. We now briefly introduce a well-known example of such a transition. Recall that a subset $I \subset V$ of vertices of a graph $G = (V, E)$ is independent if no two vertices in I are joined by an edge. Given $\lambda > 0$, the independent set (or hard-core) model with fugacity λ is the probability distribution on independent sets that chooses I with probability proportional to $\lambda^{|I|}$. An important series of results in the field of approximate counting has established the existence of a computational phase transition for the independent set model. More precisely, restrict the set of input graphs to be those of maximum degree Δ, and let $\lambda_c := \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^{\Delta}}$. Then there exists an FPTAS and an efficient sampling scheme if $\lambda < \lambda_c$ [42], while there does not if $\lambda > \lambda_c$ unless NP=RP [18, 38, 39]. The parameter λ_c also appears in statistical physics. Namely, it is the point where the independent set model has a phase transition on the Δ-regular tree in the sense of uniqueness ($\lambda < \lambda_c$) and non-uniqueness ($\lambda > \lambda_c$) of Gibbs measures. The hardness result is obtained by a reduction to MAX-CUT, an NP-hard problem.

A third class of model lies between the other two: those for which no FPRAS is known in general, but no computational hardness is known either. An important example of such a model is the independent set model when one restricts the inputs to *bipartite* graphs. Counting independent sets in bipartite graphs is called #BIS, and many approximate counting problems of interest turn out to be equivalent to the existence of an FPRAS for #BIS, see [16]. It has been conjectured that no FPRAS exists for #BIS. The connection to the Potts models, and hence the present work, is as follows. Fix $q \geq 3$. The existence of an FPRAS for the q-state Potts model on graphs of maximum degree Δ at large enough inverse temperature β would imply the existence of an FPRAS for #BIS [19, Theorem 2]. The conjecture, therefore, is that no such FPRAS for the q-state Potts model exists.

Until recently the construction of efficient approximate counting and sampling schemes for statistical physics models was largely restricted to the uniqueness regime of the respective models, for example, via Markov-chain mixing or correlation decay arguments. Notable exceptions include the Ising and monomer-dimers models and special classes of graphs with dualities, for example, planar duality [5, 20, 21, 41]. Recently, efficient algorithms in non-uniqueness regimes have been developed. These algorithms are primarily based on the observation that classical tools from mathematical physics, the cluster expansion and Pirogov–Sinai theory, can be used to obtain efficient algorithms deep inside the non-uniqueness regime on lattices [25]. Further works [11, 13, 26] extended the use of the cluster expansion to obtain algorithms for other classes of graphs and models for parameters, again deep inside non-uniqueness regimes.

As will be discussed in Section 1.4 below, our proof of Theorem 1.2 relies on a significant extension of the Pirogov–Sinai methodology of [25]. Note that Theorem 1.2 completely rules out the existence of a computational phase transition for the q-state Potts model on tori $(\mathbb{Z}/n\mathbb{Z})^d$ when $q \gg 1$. The Potts model on tori has a uniqueness/non-uniqueness phase transition in the infinite volume limit and hence our result shows that any relation between computational and physical phase transitions for the Potts model is subtle, in that it sensitive to the class of graphs being considered. It is important to note

that this potential sensitivity is not new, as it also follows from algorithmic results [28] concerning the Ising model. Our contribution, therefore, is a proof of this subtlety in a less specialized context, by fairly robust methods, and for a problem directly related to #BIS. While the tori we consider are rather special graphs, we view them as a starting point for understanding potential barriers to the existence of an FPRAS for the Potts model, and hence to understanding the existence or non-existence of an FPRAS for #BIS.

After the appearance of the extended abstract of this work in [6], results concerning all-temperature algorithms for the Potts model on expander graphs appeared [24]. The methods of [24] are different than those of the present paper, as the geometry of expander graphs allows one to avoid the use of Pirogov–Sinai theory and work with simpler polymer models instead.

1.2 | The Potts model on $\mathbb{Z}^d$

The Potts model is known to exhibit a phase transition on $\mathbb{Z}^d$ when $d \geq 2$, and when q is sufficiently large the phase diagram has been completely understood for some time [31, 32]. For large q there is a critical temperature $\beta_c = \beta_c(d, q)$ satisfying

$$\beta_c = \frac{\log q}{d} + O(q^{-1/d}), \tag{3}$$

such that for $\beta < \beta_c$ there is a unique infinite-volume Gibbs measure, while if $\beta > \beta_c$ there are q extremal translation-invariant Gibbs measures. Each of these low-temperature measures favor one of the q colors. At the transition point $\beta = \beta_c$ there are $q+1$ extremal translation-invariant Gibbs measures; q of these measures favor one of the q colors, and the additional measure is the "disordered" measure from $\beta < \beta_c$. We note that the phenomenology of the model is q-dependent [14]. The preceding results require q large as they use q^{-1} as a small parameter in proofs.

The existence of multiple measures in the low-temperature phase is reflected in the dynamical aspects of the model. While Glauber dynamics for the Potts model mix rapidly at sufficiently high temperatures, they mix in time $\exp(\Theta(n^{d-1}))$ when $\beta \geq \beta_c$ [8, 10]. Even the global-move Swendsen–Wang dynamics take time $\exp(\Theta(n^{d-1}))$ to mix when $\beta = \beta_c$ [8].

1.3 | Random cluster model

Given a finite graph $G = (V(G), E(G))$ the *random cluster model* is a probability distribution on edge sets of G given by

$$\mu_G^{\mathrm{RC}}(A) := \frac{p^{|A|}(1-p)^{|E(G)|-|A|}q^{c(G_A)}}{Z_G^{\mathrm{RC}}(p,q)}, \qquad A \subseteq E(G), \tag{4}$$

where $c(G_A)$ is the number of connected components of the graph $G_A = (V(G), A)$ and

$$Z_G^{\mathrm{RC}}(p,q) := \sum_{A \subseteq E(G)} p^{|A|}(1-p)^{|E(G)|-|A|}q^{c(G_A)} \tag{5}$$

is the random cluster model partition function.

The Potts model and the random cluster model can be put onto the same probability space via the Edwards–Sokal coupling (see, e.g., [14]). We recall this coupling in Appendix A; one consequence is the relation, for $\beta \geq 0$ and integer $q \geq 2$,

$$Z_G^{\mathrm{Potts}}(\beta) = e^{\beta|E(G)|}Z_G^{\mathrm{RC}}(1-e^{-\beta}, q). \tag{6}$$

With the parameterization $p = 1 - e^{-\beta}$ the random cluster model on $\mathbb{Z}^d$, $d \geq 2$, also has a critical inverse temperature $\beta_c = \beta_c(q, d)$ that satisfies (3) and that coincides with the Potts critical inverse temperature for integer q. For $\beta < \beta_c$ the random cluster model has a unique infinite volume measure (the *disordered* measure), while for $\beta > \beta_c$ the *ordered* measure is the unique infinite volume measure. For $\beta = \beta_c$ the two measures coexist, in the sense that there are multiple infinite-volume Gibbs measures, with one corresponding to the ordered and one corresponding to the disordered measure.

Our counting and sampling algorithms for the Potts model extend to the random cluster model on finite subgraphs of $\mathbb{Z}^d$ with two different types of boundary conditions. To make this precise requires a few definitions. Let Λ be a finite set of vertices of $\mathbb{Z}^d$ and let G_Λ be the subgraph of $\mathbb{Z}^d$ induced by Λ. We say G_Λ is *simply connected* if G_Λ is connected and the subgraph induced by $\Lambda^c = \mathbb{Z}^d \setminus \Lambda$ is connected. The random cluster model with *free boundary conditions* on G_Λ is just the random cluster model on the induced subgraph G_Λ as defined by (4). The random cluster model with *wired boundary conditions* on G_Λ is the random cluster model on the (multi-)graph G'_Λ obtained from G_Λ by identifying all of the vertices on the boundary of Λ to be one vertex; see [14, Section 1.2.2] for a formal definition. We refer to the Gibbs measures and partition functions with free and wired boundary conditions as $\mu^f_\Lambda, \mu^w_\Theta, Z^f_\Theta, Z^w_\Theta$. Explicitly,

$$Z^f_\Lambda := \sum_{A \subset E(G_\Lambda)} p^{|A|}(1-p)^{|E(G_\Lambda)|-|A|} q^{c(G_A)}, \qquad \text{and} \tag{7}$$

$$Z^w_\Lambda := \sum_{A \subset E(G'_\Lambda)} p^{|A|}(1-p)^{|E(G'_\Lambda)|-|A|} q^{c(G'_A)}, \tag{8}$$

where $c(G_A)$ is the number of connected components of the graph (Λ, A) and $c(G'_A)$ is the number of components of the graph (Λ', A) in which we identify all vertices on the boundary of Λ.

Theorem 1.2. *For $d \geq 2$ there exists $q_0 = q_0(d)$ so that for $q \geq q_0$ the following is true.*

For $\beta \geq \beta_c$ there is an FPTAS and efficient sampling scheme for the random cluster model on all finite, simply connected induced subgraphs of $\mathbb{Z}^d$ with wired boundary conditions.

For $\beta \leq \beta_c$ there is an FPTAS and efficient sampling scheme for the random cluster model on all finite, simply connected induced subgraphs of $\mathbb{Z}^d$ with free boundary conditions.

Theorem 1.2 yields an FPTAS, while Theorem 1.1 gave an FPRAS for the torus. The reason for this is that our Pirogov–Sinai based methods become more difficult to implement on the torus if the error parameter ϵ is smaller than $\exp(-O(n^{d-1}))$. The algorithm for Theorem 1.1 circumvents this by making use of the Glauber dynamics for this range of ϵ. This is possible because, despite being slow mixing, the Glauber dynamics are fast enough when given time $O(\epsilon^{-1})$ for ϵ this small by [8]. By using Glauber dynamics in a similar manner we could obtain an FPRAS for the random cluster model on $\mathbb{T}^d_n$.

We note that our methods are likely capable of handling boundary conditions other than those described above, but we leave an investigation of the full scope of their applicability for the future.

1.4 | Proof overview

The results of this article are based on non-trivial extensions of the recent work [25]. To discuss the new ingredients, we first recall two key ideas from [25]. The first, which has since gone on to be used in many subsequent works [11–13, 26, 33], is the notion of a *polymer model algorithm*. We discuss this method in a self-contained way in Section 2 below; it is based on the well-developed ideas of polymer models and cluster expansion from mathematical physics [22, 30]. In [25] this was combined with Barvinok's interpolation method [2] to devise efficient algorithms. Polymer model

algorithms are efficient algorithms for estimating the partition function of low-density independent set models. The power of the method is that it can handle independent sets on very general graphs with vertex-dependent activities. Many problems of interest can be rephrased in terms of independent set models of this type.

The second key idea from [25] for this work is the algorithmic use of Pirogov–Sinai theory. An important ingredient for this is the notion of a *ground state*. Formally, for the Potts models, the ground states are the colorings σ of $\mathbb{Z}^d$ that minimize $H_{\mathbb{Z}^d}(\sigma)$. Rigorously, the ground states are colorings σ for which any finite perturbation σ' satisfies $H_G(\sigma')-H_G(\sigma) > 0$ as $G \uparrow \mathbb{Z}^d$; since σ' is a finite perturbation this sequence is constant for large enough volumes G. The ground states of the ferromagnetic Potts model are the q monochromatic colorings.

This notion of a ground state is meant to capture the intuition that when $\beta \gg 1$, one expects a typical configuration of the Potts model to look essentially like one the ground states, with some small local deviations. Rigorously verifying this picture is non-trivial, and is part of the subject of Pirogov–Sinai theory. The key output of the theory is a convergent expansion for the logarithm of the partition function of the model with monochromatic boundary conditions, where the terms of the expansion correspond to local deviations from the given ground state. The expansion has a recursive flavor: the terms of the expansion are themselves given by ratios of partition functions with different boundary conditions. This recursion can be traced back to the fact that local deviations can have internal structures: there could be a red island inside of a blue lake inside of a red sea. See Figure 1. The algorithms of [25] made use of the symmetry of the ground states of the Potts model in handling this recursion, the key point being that symmetry implies (when $\beta \gg 1$) the deviations are rare enough that their contribution to the relevant partition functions can be controlled by a convergent cluster expansion.

Theorems 1.1 and 1.2 concern not just low temperatures, but all temperatures. Pirogov–Sinai theory has been developed for the Potts model at all temperatures when $q \gg 1$, and for doing this it is very helpful to use the random cluster representation [32]. Our algorithms rely on this, and we follow the sophisticated approach from [8]. Algorithmically, however, the reliance on the random cluster model creates a key difficulty. As discussed above, in the Potts model representation, the q ground states

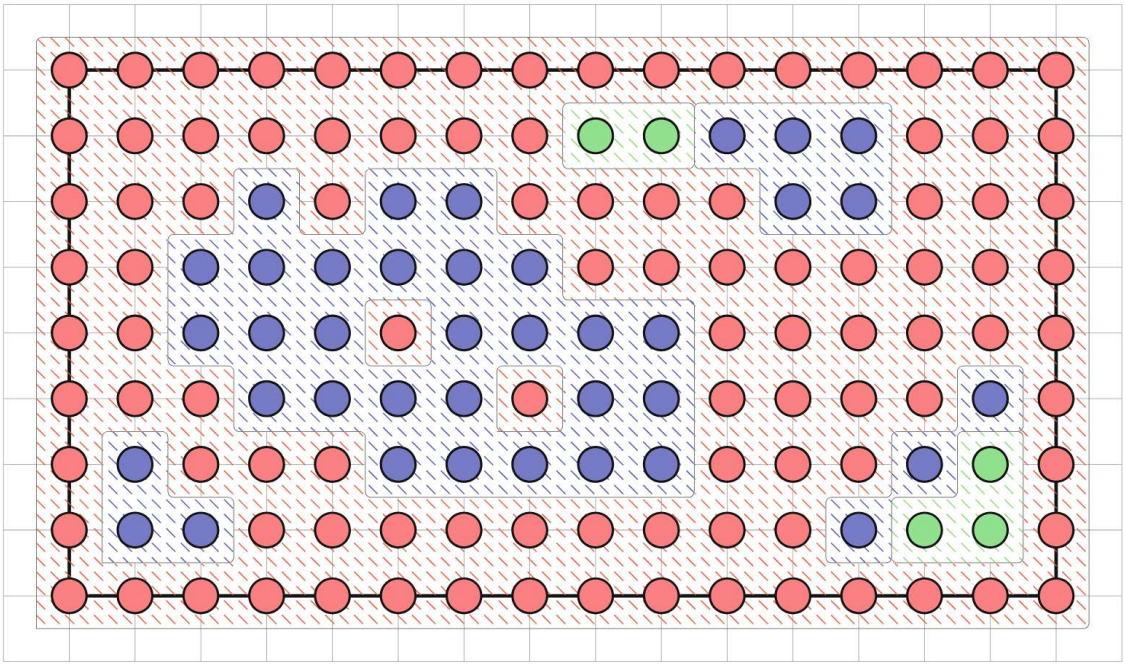

FIGURE 1 A $q = 3$ Potts model configuration depicting nested regions of constant color.

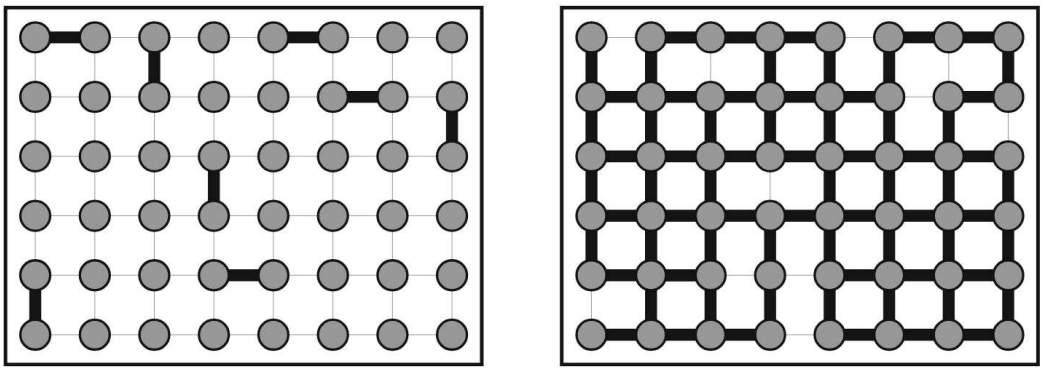

FIGURE 2 Two random cluster model configurations. The configuration on the left is a perturbation of the empty set of edges, and the configuration on the right a perturbation of the full set of edges.

(one for each color) are completely symmetric. These ground states were defined with $\beta \gg 1$ in mind, while applying Pirogov–Sinai theory at all temperatures requires having a ground state corresponding to the typical behavior when $\beta \ll 1$ as well. The random cluster representation achieves this naturally: it has two ground states, the ordered (full set of edges) and disordered (empty set of edges) ground states. These ground states are not symmetric; the former captures the low-temperature behavior and the latter the high-temperature behavior. See Figure 2.

In Pirogov–Sinai theory ground states are categorized based on free energies of truncated models, as is discussed in [29, Section 1.5]. For a given choice of parameters, ground states minimizing the truncated free energy are *stable* while other ground states are *unstable*. In the Potts representation all ground states are stable by symmetry, and this was exploited in the low temperature algorithms in [25]. In the random cluster representation, one of the ground states may well be unstable (in fact only at $\beta = \beta_c$ are both ground states stable). Thus while working with the random cluster representation gives us a convergent cluster expansion at all temperatures, it also necessitates an algorithmic approach that accommodates unstable ground states. The next paragraph discusses our algorithmic approach. We believe this approach could be adapted to other models with unstable ground states, but for the sake of concreteness we restrict our discussion to the setting of the random cluster model.

To see the issue that unstable ground states create for algorithms, recall that when $\beta \gg 1$ the intuition is that most configurations look like the ordered ground state, with local deviations that look like the disordered ground state. Since the disordered ground state is not stable at low temperatures, it does not suppress local deviations that flip back to the ordered ground state. This prevents us from analyzing the recursive structure of the Pirogov–Sinai expansion by using polymer model methods: the polymer model expansion in an unstable ground state does not have a convergent expansion. To circumvent this, we use tools from [8] to establish that inside of any unstable deviation there will be a further deviation back to the stable ground state. This flip back to the stable ground state happens rapidly enough that we can use brute-force methods. Since there may be many unstable deviations, it is also important for us to control their total volume, and again we use tools from [8] to do this.

As is clear from this discussion, this articles makes significant use of the methods developed in [8, 25]. For the ease of the reader who wishes to see the proofs of results we use from [8] we have largely stuck to the definitions presented in that article, and have made careful note of the situations in which we have chosen alternative definitions that facilitate our algorithms. To complement the discussion above, we conclude this section with an outline of our arguments along with pointers to the technical content of the article.

1. In Section 2, we briefly recall the notion of a polymer model and convergence criteria for the cluster expansion, and recall from [25] how this can be used for approximation algorithms. A key improvement upon [25] is that we work directly with the cluster expansion rather than using Barvinok's method [3]. This is essential, as Barvinok's method relies on the existence of a zero-free region. In the Potts model there cannot be a zero-free region uniformly in the volume near β_c, precisely because this is the point at which a phase transition occurs. In this section, we also apply the polymer model algorithm to the random cluster model at very high temperatures, meaning $\beta \leq \beta_h := \frac{3 \log q}{4d}$.
2. In Section 3, we first recall the tools from Pirogov–Sinai theory developed in [8] for the random cluster model. We then use these tools to establish the necessary ingredients for an algorithmic implementation of the method.
3. Section 4 contains estimates for the contour model representation derived in Section 3. We prove some consequences of estimates from [8] that are needed for our algorithms. As discussed above, the key additional estimates concern how unstable contours rapidly "flip" to stable contours, which are essential for our algorithms to be efficient.
   This section focuses on the most interesting case of $\beta \geq \beta_c$. The case $\beta_h < \beta < \beta_c$, which is very similar to $\beta > \beta_c$ and again uses estimates from [8], is discussed in Appendix B.
4. In Section 5, we present our approximate counting algorithms. The broad idea is to use the inductive Pirogov–Sinai method of [25], but with significant refinements to deal with the presence of an unstable ground state. Similar refinements are then used in Section 6 to develop sampling algorithms.

We remark that it may be possible to combine results and proof techniques from [1, 15, 35] to prove that the Glauber dynamics mix rapidly on the torus and sufficiently regular subsets of $\mathbb{Z}^d$ for all $\beta < \beta_c$, which would yield a much faster sampling algorithm than the one we have given here. We are not aware, however, of any existing statement in the literature which would directly imply rapid mixing in the whole range $\beta < \beta_c$, and leave this as an open problem. Further open problems can be found in the conclusion of this article, Section 7.

2 | POLYMER MODELS, CLUSTER EXPANSIONS, AND ALGORITHMS

This section describes how two related tools from statistical physics, abstract polymer models and the cluster expansion, can be used to design efficient algorithms to approximate partition functions.

An *abstract polymer model* [22, 30] consists of a set $\mathcal{C}$ of *polymers*, with each polymer $\gamma \in \mathcal{C}$ equipped with a complex-valued *weight* w_γ and a non-negative *size* $\|\gamma\|$. The set $\mathcal{C}$ also comes equipped with a symmetric compatibility relation $\sim$ such that each polymer is incompatible with itself, denoted $\gamma \nsim \gamma$. Let $\mathcal{G}$ denote the collection of all sets of pairwise compatible polymers from $\mathcal{C}$, including the empty set of polymers. The polymer model partition function is defined to be

$$Z(\mathcal{C}, w) := \sum_{\Gamma \in \mathcal{G}} \prod_{\gamma \in \Gamma} w_\gamma. \tag{9}$$

In (9) w is shorthand for the collection of polymer weights.

Let Γ be a non-empty tuple of polymers. The *incompatibility graph* H_Γ of Γ has vertex set Γ and edges linking any two incompatible polymers, that is, $\{\gamma, \gamma'\}$ is an edge if and only if $\gamma \nsim \gamma'$. A non-empty ordered tuple Γ of polymers is a *cluster* if its incompatibility graph H_Γ is connected. Let

$\mathcal{G}^c$ be the set of all clusters of polymers from $\mathcal{C}$. The cluster expansion is the following formal power series for $\log Z(\mathcal{C}, w)$ in the variables w_γ:

$$\log Z(\mathcal{C}, w) = \sum_{\Gamma \in \mathcal{G}^c} \phi(H_\Gamma) \prod_{\gamma \in \Gamma} w_\gamma. \tag{10}$$

In (10) $\phi(H)$ denotes the *Ursell function* of the graph $H = (V(H), E(H))$, that is,

$$\phi(H) := \frac{1}{|V(H)|!} \sum_{\substack{A \subseteq E(H) \\ (V(H), A) \text{ connected}}} (-1)^{|A|}.$$

For a proof of (10) see, for example, [17, 30]. Define $\|\Gamma\| := \sum_{\gamma \in \Gamma} \|\gamma\|$, and define the truncated cluster expansion by

$$T_m(\mathcal{C}, w) := \sum_{\substack{\Gamma \in \mathcal{G}^c \\ \|\Gamma\| < m}} \phi(H_\Gamma) \prod_{\gamma \in \Gamma} w_\gamma.$$

Henceforth we will restrict our attention to a special class of polymer models defined in terms of a graph G with maximum degree Δ on N vertices. Namely, we will assume that each polymer is a connected subgraph $\gamma = (V(\gamma), E(\gamma))$ of G. The compatibility relation is defined by disjointness in G: $\gamma \sim \gamma'$ iff $V(\gamma) \cap V(\gamma') = \emptyset$. We write $|\gamma|$ for $|V(\gamma)|$, the number of vertices in the polymer γ.

A useful criteria for convergence of the formal power series in (10) is given by the following adaptation of a theorem of Kotecký and Preiss [30].

Lemma 2.1. *Let G be a graph of maximum degree $\Delta \geq 2$ on N vertices. Suppose that polymers are connected subgraphs of G that contain at least two vertices. Suppose further that for some $b > 0$ and all $\gamma \in \mathcal{C}$,*

$$\|\gamma\| \geq b|E(\gamma)|, \tag{11}$$

$$|w_\gamma| \leq \exp\left(-\left(\frac{4 + \log \Delta}{b} + 3\right) \|\gamma\|\right). \tag{12}$$

Then the cluster expansion (10) converges absolutely, and for $m \in \mathbb{N}$,

$$|T_m(\mathcal{C}, w) - \log Z(\mathcal{C}, w)| \leq N e^{-3m}. \tag{13}$$

Moreover, if instead all polymers are connected, induced subgraphs of G, and for some $b > 0$ and all $\gamma \in \mathcal{C}$,

$$\|\gamma\| \geq b|\gamma|, \tag{14}$$

$$|w_\gamma| \leq \exp\left(-\left(\frac{3 + \log \Delta}{b} + 3\right) \|\gamma\|\right), \tag{15}$$

then the same conclusion holds.

This lemma implies that if conditions (11) and (12) hold, then $\exp(T_m(\mathcal{C}, w))$ is an ϵ-relative approximation to $Z(\mathcal{C}, w)$ for $m \geq \log(N/\epsilon)/3$.

Proof. We append to C a polymer γ_v for each $v \in V(G)$ consisting only of that vertex, with size $\|\gamma_v\| = 1$ and $w_{\gamma_v} = 0$. By definition, γ_v is incompatible with every other polymer that contains v. Then

$$\begin{aligned}\sum_{\gamma \not\sim \gamma_v} |w_\gamma| e^{|E(\gamma)|+3\|\gamma\|} &\le \sum_{\gamma \not\sim \gamma_v} e^{|E(\gamma)|} e^{-(\frac{4+\log\Delta}{b})\|\gamma\|} \\ &\le \sum_{\gamma \not\sim \gamma_v} e^{|E(\gamma)|} e^{-(4+\log\Delta)|E(\gamma)|} \\ &\le \sum_{k\ge 1} (e\Delta)^k e^{-(3+\log\Delta)k},\end{aligned}$$

where the first inequality is by (12), the second by (11), and the third is by bounding the number of connected subgraphs of G with k edges that contain v by $(e\Delta)^k$ [7]. This yields

$$\sum_{\gamma \not\sim \gamma_v} |w_\gamma| e^{|E(\gamma)|+3\|\gamma\|} \le \sum_{k\ge 1} e^{-2k} < 1/2. \tag{16}$$

Moreover, under the second assumption, that all polymers are connected induced subgraphs, we have a similar bound, with $|\gamma|$ in place of $|E(\gamma)|$:

$$\sum_{\gamma \not\sim \gamma_v} |w_\gamma| e^{|\gamma|+3\|\gamma\|} \le \sum_{k\ge 1} e^{-k} < 1, \tag{17}$$

where we have used that the number of connected induced subgraphs on k vertices that contain v is at most $(e\Delta)^k$ [7].

Now fix a polymer γ. By summing (16) over all $v \in \gamma$ we obtain

$$\sum_{\gamma' \not\sim \gamma} |w_{\gamma'}| e^{|E(\gamma')|+3\|\gamma'\|} < |\gamma|/2 \le |E(\gamma)|. \tag{18}$$

By applying the main theorem of [30] with $a(\gamma) = |E(\gamma)|$, $d(\gamma) = 3\|\gamma\|$ we obtain that the cluster expansion converges absolutely. Moreover, we also obtain that

$$\sum_{\substack{\Gamma \in \mathcal{G}^c \\ \Gamma \ni v}} \left| \phi(H_\Gamma) \prod_{\gamma \in \Gamma} w_\gamma \right| e^{3\|\Gamma\|} \le 1, \tag{19}$$

where the sum is over all clusters that contain a polymer containing the vertex v. By using this estimate, restricting to $\|\Gamma\| \ge m$, and summing over all $v \in V(G)$ one obtains

$$\sum_{\substack{\Gamma \in \mathcal{G}^c \\ \|\Gamma\| \ge m}} \left| \phi(H_\Gamma) \prod_{\gamma \in \Gamma} w_\gamma \right| \le N e^{-3m}, \tag{20}$$

which is (13).

The same argument works under the second assumption by taking $a(\gamma) = |\gamma|$. ∎

Because clusters are connected objects arising from a bounded-degree graph, the truncated cluster expansion can be computed efficiently. Recall that $N = |V(G)|$.

Lemma 2.2. *Suppose the conditions of Lemma 2.1 hold. Then given a list of all polymers γ of size at most m along with the weights w_γ of these polymers, the truncated cluster expansion $T_m(C, w)$ can be computed in time $O(N \exp(O(m)))$.*

Proof. This is [25, Theorem 6]. ∎

The next lemma says that, for the purposes of approximating a polymer partition function, it is sufficient to have approximate evaluations $\tilde{w}_\gamma$ of the weights w_γ.

Lemma 2.3. *Let $v : \mathcal{C} \to [0, \infty)$ be a non-negative function on polymers such that $v(\gamma) \le \|\gamma\|^2$. Suppose $0 < \epsilon < N^{-1}$, and let $m = \log(8/\epsilon)/3$. Suppose the conditions of Lemma 2.1 hold and that for all $\gamma \in \mathcal{C}$ with $\|\gamma\| \le m$, $\tilde{w}_\gamma$ is an $\epsilon v(\gamma)$-relative approximation to w_γ. Then $\exp(T_m(C, \tilde{w}))$ is an $N\epsilon/4$-relative approximation to $Z(C, w)$.*

Proof. Using the definition of m and applying Lemma 2.1, we have

$$|\log Z_G(C, w) - T_m(C, w)| \le N\epsilon/8,$$

so by the triangle inequality it is enough to show that

$$|T_m(C, \tilde{w}) - T_m(C, w)| \le N\epsilon/8. \tag{21}$$

Define r_γ by $\log \tilde{w}_\gamma = \log w_\gamma + r_\gamma$. To prove (21), note the identity

$$T_m(C, \tilde{w}) - T_m(C, w) = \sum_{\substack{\Gamma \in \mathcal{C}^c(G) \\ \|\Gamma\| < m}} \phi(H_\Gamma) \prod_{\gamma \in \Gamma} w_\gamma \cdot \left[\exp\left(\sum_{\gamma \in \Gamma} r_\gamma \right) - 1 \right].$$

Our hypotheses imply $|r_\gamma| \le \epsilon v(\gamma)$, and hence by the triangle inequality we obtain

$$|T_m(C, \tilde{w}) - T_m(C, w)| \le \sum_{\substack{\Gamma \in \mathcal{C}^c(G) \\ \|\Gamma\| < m}} \left(\exp\left(\sum_{\gamma \in \Gamma} \epsilon v(\gamma) \right) - 1 \right) \left| \phi(H_\Gamma) \prod_{\gamma \in \Gamma} w_\gamma \right|,$$

where we have used the elementary inequality $|e^a - 1| \le e^b - 1$ when $|a| \le b$ to bound the term in square brackets. Since $v(\gamma) \le \|\gamma\|^2$ this yields, after ordering the sum over clusters according to their size k,

$$|T_m(C, \tilde{w}) - T_m(C, w)| \le \sum_{k=1}^{m-1} (\exp(\epsilon k^2) - 1) \sum_{\substack{\Gamma \in \mathcal{C}^c(G) \\ \|\Gamma\| = k}} \left| \phi(H_\Gamma) \prod_{\gamma \in \Gamma} w_\gamma \right| \le \sum_{k=1}^{m-1} (\exp(\epsilon k^2) - 1) N e^{-3k}.$$

The last inequality follows from the convergence of the cluster expansion (see (20) in the proof of Lemma 2.1). Since $\epsilon < N^{-1}$ we can bound $e^{\epsilon k^2} - 1$ by $2\epsilon k^2$, and (21) follows since $\sum_{k \ge 1} k^2 e^{-3k} < 1/16$. ∎

Putting Lemmas 2.1–2.3 together we see that the partition function $Z(C, w)$ can be approximated efficiently if

1. conditions (11) and (12) hold,
2. polymers of size at most $m = O(\log N/\epsilon)$ can be enumerated efficiently, that is, in time polynomial in N and exponential in m, and
3. the polymer weights w_γ can be approximated efficiently, that is, in time polynomial in the size of γ.

2.1 | High temperature expansion

This section explains how the polymer model algorithm of the previous section yields efficient counting and sampling algorithms for the random cluster model when q is sufficiently large and $\beta \le \beta_h = \frac{3\log q}{4d}$. This use of the polymer model algorithm also serves as a warm-up for the more sophisticated contour-based algorithms we will use in later sections when $\beta > \beta_h$.

In fact, the simpler setting of $\beta \le \beta_h$ allows for greater generality: we will derive an algorithm that applies to the random cluster model on *any* graph G of maximum degree at most $2d$.

Theorem 2.4. *Suppose $d \ge 2$ and $q = q(d)$ is sufficiently large. Then for $\beta \le \beta_h$ there is an FPTAS and efficient sampling scheme for the Potts model and the random cluster model with $p = 1 - e^{-\beta}$ on all graphs of maximum degree at most $2d$.*

Proof. Let $G = (V(G), E(G))$ be such a graph. We define polymers to be connected subgraphs of G with at least two vertices. As per our convention, polymers are compatible if they are vertex disjoint, and $|\gamma| = |V(\gamma)|$. We set $\|\gamma\| = |E(\gamma)|$, and define the weight of a polymer γ to be

$$w_\gamma := \left(\frac{p}{1-p}\right)^{\|\gamma\|} q^{1-|\gamma|} = (e^\beta - 1)^{\|\gamma\|} q^{1-|\gamma|} . \tag{22}$$

Let $\mathcal{C}(G)$ be the set of all polymers on G, $\mathcal{G}(G)$ be the collection of all sets of pairwise compatible polymers from $\mathcal{C}(G)$, and let

$$\Xi(G) := \sum_{\Gamma \in \mathcal{G}(G)} \prod_{\gamma \in \Gamma} w_\gamma \tag{23}$$

be the corresponding polymer model partition function. Then we have the identity

$$Z_G^{\mathrm{RC}}(p, q) = (1-p)^{|E(G)|} q^{|V(G)|}\, \Xi(G). \tag{24}$$

The relation (24) follows by extracting a common prefactor of $(1-p)^{|E(G)|} q^{|V(G)|}$ from the random cluster partition function. To see this relation it may help to temporarily allow polymers that consist of a single vertex; since these receive weight one by (22) it is equivalent to remove these from the set of polymers.

Condition (11) holds with $b = 1$ since $\|\gamma\| = |E(\gamma)|$. We will show that condition (12) holds if $\beta \le \beta_h$ and q is sufficiently large as a function of d. Suppose there is a q_0 such that for all γ, all $\beta \le \beta_h$, and all $q \ge q_0$

$$w_\gamma \le C^{-\|\gamma\|}. \tag{25}$$

Then if $C = C(d) > 0$ is large enough, (12) holds. Since $b = 1$, $C = \exp(7 + \log 2d)$ suffices, and we fix C to be this value hereon. We now verify (25) in three steps, by considering polymers grouped according to the value of $k = \|\gamma\|$.

1. For $k > 5d$ we will use the fact that $|\gamma| \geq \|\gamma\|/d$ since every edge is incident to two vertices and every vertex is incident to at most $2d$ edges. Then using the fact that $\beta \leq \beta_h$ we have

$$w_\gamma \leq q(e^\beta - 1)^k q^{-k/d} \leq q^{1-\frac{k}{4d}} \leq q^{-\frac{k}{20d}}, \tag{26}$$

which is at most $C^{-\|\gamma\|}$ if $q \geq C^{20d}$.

2. For $d < k \leq 5d$, we will use the fact that $|\gamma| \geq \frac{1}{2} + \sqrt{2\|\gamma\|}$ since the number of edges in a graph on r vertices is at most $\binom{r}{2}$. Then we have

$$w_\gamma \leq qq^{\frac{3k}{4d}} q^{-\frac{1}{2}-\sqrt{2k}} \leq q^{\frac{1}{2}+\frac{3c}{4}-2\sqrt{c}}, \tag{27}$$

where $c = k/d$ and where we use the fact that $d \geq 2$ and $q \geq 1$. Then since $\frac{1}{2} + \frac{3c}{4} - 2\sqrt{c} \leq -\frac{1}{5}$ for $c \in [1, 5]$, we have

$$w_\gamma \leq q^{-1/5}, \tag{28}$$

which is at most $C^{-\|\gamma\|}$ if $q \geq C^{25d}$.

3. For $1 \leq k \leq d$, since $|\gamma| \geq 2$, we have

$$w_\gamma \leq q^{-1}(e^\beta - 1)^k \leq q^{-1}e^{\beta k} \leq q^{-1/4}, \tag{29}$$

which is at most $C^{-\|\gamma\|}$ provided $q \geq C^{4d}$.

Thus taking $q_0 \geq \exp(25d(7 + \log 2d))$ suffices. Lemmas 2.1 and 2.2 then give an FPTAS for computing the random cluster partition function $Z_G^{\mathrm{RC}}(1 - e^{-\beta}, q)$ for all graphs of maximum degree $2d$, as enumerating subgraphs of size m in a bounded degree graph takes time $\exp(O(m))$, and computing the weight functions only requires counting the number of edges and vertices in each subgraph.

The efficient sampling scheme follows from [25, Theorem 10]. Counting and sampling algorithms for the random cluster model can be converted into algorithms for the Potts model via the Edwards–Sokal coupling described in Appendix A. ∎

Proof of Theorems 1.1 and 1.2 for $\beta \leq \beta_h$. Theorem 1.1 follows immediately from Theorem 2.4 since $\mathbb{T}_n^d$ is $2d$-regular.

By (3), $\beta_h < \beta_c$ when q is large enough. Thus Theorem 1.2 requires we provide approximate counting and sampling algorithms for free boundary conditions. Since induced subgraphs of $\mathbb{Z}^d$ have degree bounded by $2d$, the result follows by Theorem 2.4. ∎

3 | CONTOUR MODEL REPRESENTATIONS

Contour models refer to a class of polymer models that arise in Pirogov–Sinai theory [37]. For a given spin configuration, contours represent geometric boundaries between regions dominated by different ground states; the precise definition for the purposes of this article will be given below. This section describes an important contour model representation for the random cluster model on the torus $\mathbb{T}_n^d$ that is the basic combinatorial object in our algorithms. This contour representation was originally developed for obtaining optimal lower bounds on the mixing time for Glauber and Swendsen–Wang dynamics [8]. In addition to recalling the construction from [8] this section also develops the additional ingredients necessary for algorithmic applications of the representation.

3.1 | Continuum embedding

The contour model representation from [8] is based on the natural embedding of the discrete torus $\mathbb{T}_n^d = (\mathbb{Z}/n\mathbb{Z})^d$ of side-length $n \in \mathbb{N}$ into the continuum torus $\boldsymbol{T}_n^d := (\mathbb{R}/n\mathbb{R})^d$. This subsection recalls the basic definitions, and explains how they can be rephrased in terms of discrete graph-theoretic notions.[1]

In what follows we abuse notation slightly and write $\mathbb{T}_n^d$ for the graph $(\mathbb{T}_n^d, E)$, where E is the edge set of the discrete torus. We will follow the convention that bold symbols, for example, $\boldsymbol{V}$, denote subsets of $\boldsymbol{T}_n^d$, while objects denoted by non-bold symbols like V reside in $\mathbb{T}_n^d$. Thus each vertex $v \in \mathbb{T}_n^d$ is identified with a point $\boldsymbol{v} \in \boldsymbol{T}_n^d$, and we will identify each edge $e = \{u, v\} \in E$ with the unit line segment $\boldsymbol{e} \subset \boldsymbol{T}_n^d$ that joins $\boldsymbol{u}$ to $\boldsymbol{v}$. We will also drop $\mathbb{T}_n^d$ from the notation when possible, for example, E for $E(\mathbb{T}_n^d)$.

Let $\Omega = 2^E$ denote the set of configurations of the random cluster model on $\mathbb{T}_n^d$. Let $\boldsymbol{c} \subset \boldsymbol{T}_n^d$ denote a closed k-dimensional hypercube with vertices in $\mathbb{T}_n^d$ for some $k = 1, \ldots, d$. We say a hypercube $\boldsymbol{c}$ is *occupied* with respect to $A \in \Omega$ if for all edges e with $\boldsymbol{e} \subset \boldsymbol{c}$, e is in A. Define

$$\boldsymbol{A} := \left\{ \boldsymbol{x} \in \boldsymbol{T}_n^d \,|\, \text{there exists } \boldsymbol{c} \text{ occupied s.t. } d_\infty(\boldsymbol{x}, \boldsymbol{c}) \leq \frac{1}{4} \right\}, \tag{30}$$

where d_∞ is the ℓ_∞-distance, and the distance from a point to a set is defined in the standard way: $d_\infty(\boldsymbol{x}, \boldsymbol{c}) = \inf_{\boldsymbol{y} \in \boldsymbol{c}} d_\infty(\boldsymbol{x}, \boldsymbol{y})$. Thus $\boldsymbol{A}$ is the closed $1/4$-neighborhood of the occupied hypercubes of A. The connected components of the (topological) boundary $\partial \boldsymbol{A}$ of the set $\boldsymbol{A}$ are the crucial objects in what follows. Since each connected component arises from an edge configuration in Ω, it is clear that the set of possible connected components is a finite set. As the connected components of $\partial \boldsymbol{A}$ are continuum objects, it may not be immediately apparent how to represent them in a discrete manner. We briefly describe how to do this now.

Let $\frac{1}{2}\mathbb{T}_n^d$ denote the graph $(\frac{1}{2}\mathbb{Z}/n\mathbb{Z})^d$; as a graph this is equivalent to the discrete torus $(\mathbb{Z}/(2n)\mathbb{Z})^d$. The notation $\frac{1}{2}\mathbb{T}_n^d$ is better because we will embed $\frac{1}{2}\mathbb{T}_n^d$ in $\boldsymbol{T}_n^d$ such that (i) $\boldsymbol{0}$ coincides in $\mathbb{T}_n^d$ and $\frac{1}{2}\mathbb{T}_n^d$, and (ii) the nearest neighbors of 0 in $\frac{1}{2}\mathbb{T}_n^d$ are the midpoints of the edges e containing 0 in $\mathbb{T}_n^d$.[2]

An important observation is that $\boldsymbol{A}$ can be written as a union of collections of adjacent closed d-dimensional hypercubes of side-length $1/2$ centered at vertices in $\frac{1}{2}\mathbb{T}_n^d$, where two hypercubes are called *adjacent* if they share a $(d-1)$-dimensional face. Adjacency of a set of hypercubes means the set of hypercubes is connected under the binary relation of being adjacent. By construction the connected components of $\boldsymbol{A}$ correspond to the connected components of the edge configuration A.

The boundary $\partial \boldsymbol{A}$ of $\boldsymbol{A}$ is just the sum, modulo two, of the boundaries of the hypercubes whose union gives A. These boundaries are $(d-1)$-dimensional hypercubes dual to edges in $\frac{1}{2}\mathbb{T}_n^d$; here dual means that the barycenter of the $(d-1)$-dimensional hypercube is the same as barycenter of the edge in $\frac{1}{2}\mathbb{T}_n^d$. The $(d-1)$-dimensional hypercubes that arise from this duality are the vertices in $\left(\frac{1}{2}\mathbb{T}_n^d\right)^\star$, the graph dual to $\frac{1}{2}\mathbb{T}_n^d$; two vertices in $\left(\frac{1}{2}\mathbb{T}_n^d\right)^\star$ are connected by an edge if and only if the corresponding $(d-1)$-dimensional hypercubes intersect in one $(d-2)$-dimensional hypercube. The preceding discussion implies $\partial \boldsymbol{A}$ can be identified with a subgraph of $\left(\frac{1}{2}\mathbb{T}_n^d\right)^\star$.

In the sequel we will discuss components of $\partial \boldsymbol{A}$ as continuum objects; by the preceding discussion this could be reformulated in terms of subgraphs of $\left(\frac{1}{2}\mathbb{T}_n^d\right)^\star$. In Appendix C, we show that the

[1]This continuum construction allows for tools from algebraic topology to be used. We have chosen to follow the continuum terminology to allow the interested reader to easily consult [8].

[2]More formally, since $\mathbb{Z}^d \subset \frac{1}{2}\mathbb{Z}^d \subset \mathbb{R}^d$, we obtain a common embedding of $\frac{1}{2}\mathbb{T}_n^d$ and $\mathbb{T}_n^d$ in $\boldsymbol{T}_n^d$.

computations we perform involving components of $\partial \boldsymbol{A}$ can be efficiently computed using their representations as subgraphs of $(\frac{1}{2}\mathbb{T}_n^d)^\star$.

3.2 | Contours and interfaces

An important aspect of the analysis in [8] is that it distinguishes topologically trivial and non-trivial components of $\partial \boldsymbol{A}$. To make this precise, for $i = 1, \dots, d$ we define the *ith fundamental loop* $\boldsymbol{L}_i$ to be the set $\{\boldsymbol{y} \in \boldsymbol{T}_n^d | \boldsymbol{y}_j = 1 \text{ for all } j \neq i\}$. The *winding vector* $N(\boldsymbol{\gamma}) \in \{0,1\}^d$ of a connected component $\boldsymbol{\gamma} \in \partial \boldsymbol{A}$ is the vector whose ith component is the number of intersections (mod 2) of $\boldsymbol{\gamma}$ with $\boldsymbol{L}_i$.

Definition 1. Let $A \in \Omega$ be an edge configuration.

1. The set of contours $\Gamma(A)$ associated to A is the set of connected components of $\partial \boldsymbol{A}$ with winding vector 0.
2. The interface network $\mathcal{S}(A)$ associated to A is the set of connected components of $\partial \boldsymbol{A}$ with non-zero winding vector. Each connected component of an interface network is an interface.

Without reference to any particular edge configuration, a subset $\boldsymbol{\gamma} \subset \boldsymbol{T}_n^d$ is a contour if there is an $A \in \Omega$ such that $\boldsymbol{\gamma} \in \Gamma(A)$. Interfaces and interface networks are defined analogously.

Since each fundamental loop intersects each $(d-1)$-dimensional face of a hypercube centered on $\frac{1}{2}\mathbb{T}_n^d$ exactly zero or one times, we have the following lemma, which ensures contours can be efficiently distinguished from interfaces.

Lemma 3.1. *Suppose $\boldsymbol{\gamma} \in \partial \boldsymbol{A}$ is comprised of K $(d-1)$-dimensional faces. Then the winding vector of $\boldsymbol{\gamma}$ can be computed in time $O(nK)$.*

Proof. Fix $i \in \{1, 2, \dots, d\}$. Each fundamental loop L_i has length $O(n)$, and hence the set F_i of faces that have non-trivial intersection with L_i has cardinality $|F_i| = O(n)$. Given the list of faces in $\boldsymbol{\gamma}$ we can compute the ith component of the winding vector by (i) iterating through the list of faces of $\boldsymbol{\gamma}$ and adding one each time we find a face in F_i, and (ii) taking the result modulo two. ∎

The connected components of $\boldsymbol{T}_n^d \setminus \partial \boldsymbol{A}$ are subsets of either $\boldsymbol{A}$ or $\boldsymbol{T}_n^d \setminus \boldsymbol{A}$. In the former case we call a component *ordered* and in the latter case *disordered*. We write $\boldsymbol{A}_{\text{ord}}$ (resp. $\boldsymbol{A}_{\text{dis}}$) for the union of the ordered (resp. disordered) components associated to A.

Definition 2. The labeling ℓ_A associated to A is the map from the connected components of $\boldsymbol{T}_n^d \setminus \partial \boldsymbol{A}$ to the set $\{\text{dis}, \text{ord}\}$ that assigns ord to components in $\boldsymbol{A}_{\text{ord}}$ and dis to components in $\boldsymbol{A}_{\text{dis}}$.

Definition 3. Two contours $\boldsymbol{\gamma}_i$, $i = 1, 2$ are compatible if $d_\infty(\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2) \geq \frac{1}{2}$. We extend this definition analogously to two interfaces, or one interface and one contour.

Definition 4. A matching collection of contours Γ and interfaces $\mathcal{S}$ is a triple $(\Gamma, \mathcal{S}, \ell)$ such that $\mathcal{S}$ is an interface network and

1. The contours and interfaces in $\Gamma \cup \mathcal{S}$ are pairwise compatible, and
2. ℓ is a map from the set of connected components of $\boldsymbol{T}_n^d \setminus \cup_{\boldsymbol{\gamma} \in \Gamma \cup \mathcal{S}} \boldsymbol{\gamma}$ to the set $\{\text{dis}, \text{ord}\}$ such that for every $\boldsymbol{\gamma} \in \Gamma \cup \mathcal{S}$, distinct components adjacent to $\boldsymbol{\gamma}$ are assigned different labels.

Lemma 3.2. *The map from edge configurations $A \in \Omega$ to triples (Γ, S, ℓ) of matching contours and interfaces is a bijection.*

Proof. See [8, p. 15]. ∎

3.3 | Contour and interface formulation of Z

By Lemma 3.2 we can rewrite the partition function in terms of matching collections of contours and interfaces by re-writing the weight $w(A)$ of a configuration A in terms of its contours and interfaces. By weight $w(A)$ we mean the numerator of (4), that is, $w(A) = p^{|A|}(1-p)^{|E\setminus A|}q^{c(V,A)}$. To this end, define

$$e_{\text{ord}} := -d\log(1-e^{-\beta}), \quad e_{\text{dis}} := d\beta - \log q, \quad \kappa := \frac{1}{2}\log(e^{\beta}-1). \tag{31}$$

Further, define the *size* $\|\boldsymbol{\gamma}\|$ of a contour $\boldsymbol{\gamma}$ (resp. *size* $\|\boldsymbol{S}\|$ of an interface $\boldsymbol{S}$) by

$$\|\boldsymbol{\gamma}\| := \left|\boldsymbol{\gamma} \cap \bigcup_{e\in E} \boldsymbol{e}\right|, \quad \|\boldsymbol{S}\| := \left|\boldsymbol{S} \cap \bigcup_{e\in E} \boldsymbol{e}\right|. \tag{32}$$

This is the number of intersections of $\boldsymbol{\gamma}$ (resp. $\boldsymbol{S}$) with $\bigcup_{e\in E}\boldsymbol{e}$. For a continuum set $\boldsymbol{\Lambda}$ we write $|\boldsymbol{\Lambda}|$ for $|\boldsymbol{\Lambda} \cap \mathbb{T}_n^d|$, that is, the number of vertices of $\mathbb{T}_n^d$ in $\boldsymbol{\Lambda}$ in the embedding of $\mathbb{T}_n^d$ into $\boldsymbol{T}_n^d$. This will cause no confusion as we never need to measure the volume of a continuum set.

Using these definitions, $w(A)$ can be written as

$$w(A) = q^{c(\boldsymbol{A}_{\text{ord}})}e^{-e_{\text{dis}}|\boldsymbol{A}_{\text{dis}}|}e^{-e_{\text{ord}}|\boldsymbol{A}_{\text{ord}}|}\prod_{\boldsymbol{S}\in S} e^{-\kappa\|\boldsymbol{S}\|}\prod_{\boldsymbol{\gamma}\in\Gamma} e^{-\kappa\|\boldsymbol{\gamma}\|}, \tag{33}$$

where $c(\boldsymbol{A}_{\text{ord}})$ is the number of connected components of $\boldsymbol{A}_{\text{ord}}$. The products run over the sets of interfaces and contours associated to the edge configuration A, respectively. We indicate the derivation of (33) in Section 3.3.1 below; see also [8, p. 13–15]. Since (recall $\Omega = 2^E$)

$$Z = Z^{\text{RC}}_{\mathbb{T}_n^d}(1-e^{-\beta}, q) = \sum_{A\in\Omega} w(A), \tag{34}$$

it follows from (33) and Lemma 3.2 that

$$Z = \sum_{(S,\Gamma)} q^{c(\boldsymbol{A}_{\text{ord}})}e^{-e_{\text{dis}}|\boldsymbol{A}_{\text{dis}}|}e^{-e_{\text{ord}}|\boldsymbol{A}_{\text{ord}}|}\prod_{\boldsymbol{S}\in S} e^{-\kappa\|\boldsymbol{S}\|}\prod_{\boldsymbol{\gamma}\in\Gamma} e^{-\kappa\|\boldsymbol{\gamma}\|}, \tag{35}$$

where the sum runs over matching collections of contours and interfaces. This is the contour and interface network representation of the random cluster model partition function.

In what follows it will be necessary to separate different contributions to Z. To this end, let

$$\Omega_{\text{tunnel}} := \{A \in \Omega | S(A) \neq \emptyset\}, \quad \Omega_{\text{rest}} := \Omega \setminus \Omega_{\text{tunnel}}, \tag{36}$$

and define the corresponding partition functions

$$Z_{\text{tunnel}} := \sum_{A\in\Omega_{\text{tunnel}}} w(A), \quad Z_{\text{rest}} := \sum_{A\in\Omega_{\text{rest}}} w(A). \tag{37}$$

By (35) Z_{rest} can be expressed in terms of contours alone. We will see later that Z_{tunnel} is very small compared to Z_{rest}, and so the task of approximating Z is essentially the task of approximating Z_{rest}.

3.3.1 | Derivation of contour representation

We briefly indicate how to obtain (33). Recall that G_A denotes the graph $(V(G), A)$. Let $\|\delta A\| = |\delta_1 A| + 2|\delta_2 A|$, where $\delta_k A$ is the set of edges in $E \setminus A$ that contain k vertices in $V(A)$. Observe

$$c(V, A) = c(G_A) + |V \setminus V(A)|, \tag{38}$$

$$2|A| = 2d|V(A)| - \|\delta A\|. \tag{39}$$

The first of these relations follows since every vertex not contained in an edge of A belongs to a singleton connected component, and the second is a counting argument. Using these relations one obtains

$$w(A) = q^{c(G_A)} e^{-e_{\text{dis}}|V \setminus V(A)|} e^{-e_{\text{ord}}|V(A)|} e^{-\kappa\|\delta A\|}. \tag{40}$$

To pass from (40) to (33) requires just a few observations. First, $c(G_A)$ equals the number of components of $\boldsymbol{A}$, which is the number of connected components of $\boldsymbol{A}_{\text{ord}}$. Second, $|V(A)| = |\boldsymbol{A}_{\text{ord}}|$, and similarly $|V \setminus V(A)| = |\boldsymbol{A}_{\text{dis}}|$. Lastly, $\|\delta A\|$ is precisely the sum of sizes of the contours and interfaces, as each contribution to $\|\delta A\|$ is given by a transverse intersection of an edge $\boldsymbol{e}$ with the boundary of $\boldsymbol{A}$.

3.4 | External contour representations

Next we will take the first steps to construct a representation of Z_{rest} as the sum of polymer model partition functions. We begin with some basic results and definitions. Fix an arbitrary point $\boldsymbol{x}_0 \in \boldsymbol{T}_n^d$ that cannot be contained in any contour, and let $\sqcup$ denote disjoint union.

Lemma 3.3 ([8, Lemma 4.3]). *For any contour $\boldsymbol{\gamma}$, $\boldsymbol{T}_n^d \setminus \boldsymbol{\gamma}$ has exactly two components.*

Definition 5. Let $\boldsymbol{\gamma}$ be a contour, and suppose $\boldsymbol{T}_n^d \setminus \boldsymbol{\gamma} = \boldsymbol{C} \sqcup \boldsymbol{D}$. Then the exterior Ext $\boldsymbol{\gamma}$ of $\boldsymbol{\gamma}$ is $\boldsymbol{C}$ if $|\boldsymbol{C}| > |\boldsymbol{D}|$, and is $\boldsymbol{D}$ if the inequality is reversed. In the case of equality the exterior is the component containing $\boldsymbol{x}_0$. The interior Int $\boldsymbol{\gamma}$ of $\boldsymbol{\gamma}$ is the component of $\boldsymbol{T}_n^d \setminus \boldsymbol{\gamma}$ that is not Ext $\boldsymbol{\gamma}$.

Note that the notion of exterior is defined relative to $\boldsymbol{T}_n^d$, though we omit this from the notation.

Remark. This is a different definition of exterior than is used in [8]; our definition is more convenient for algorithmic purposes. Most of the results of [8] concerning the interiors/exteriors of contours apply verbatim with this change, and whenever we use these results we will remark on why they apply.

If two contours $\boldsymbol{\gamma}$ and $\boldsymbol{\gamma}'$ are compatible, then we write (i) $\boldsymbol{\gamma} < \boldsymbol{\gamma}'$ if Int $\boldsymbol{\gamma} \subset$ Int $\boldsymbol{\gamma}'$ and (ii) $\boldsymbol{\gamma} \perp \boldsymbol{\gamma}'$ if Int $\boldsymbol{\gamma} \cap$ Int $\boldsymbol{\gamma}' = \emptyset$. Given a matching collection of contours Γ, $\boldsymbol{\gamma} \in \Gamma$ is an *external contour* if there does not exist $\boldsymbol{\gamma}' \in \Gamma$ such that $\boldsymbol{\gamma}' < \boldsymbol{\gamma}$. The *exterior* of a matching collection of contours Γ is

$$\text{Ext } \Gamma := \bigcap_{\boldsymbol{\gamma} \in \Gamma} \text{Ext } \boldsymbol{\gamma}. \tag{41}$$

If Γ is matching, then Ext Γ is a connected subset of $\mathbb{T}_n^d$. This follows by noting that [8, Lemma 5.5] holds with Definition 5 of the interior and exterior, and given this, the connectedness of Ext Γ follows by the argument in [8, Lemma 5.6]. Note that since Ext Γ is contained in $\mathbb{T}_n^d \setminus \bigcup_{\boldsymbol{\gamma} \in \Gamma} \boldsymbol{\gamma}$, this implies that Ext Γ is labeled either ord or dis.

As usual in Pirogov–Sinai theory, see, for example, [8, Section 6.2], it is useful to resum the matching compatible contours that contribute to (35) according to the external contours of the configuration. To make this precise, we require several definitions. A matching collection of contours Γ is *mutually external* if $\boldsymbol{\gamma} \perp \boldsymbol{\gamma}'$ for all $\boldsymbol{\gamma} \neq \boldsymbol{\gamma}' \in \Gamma$. For a continuum set $\boldsymbol{\Lambda} \subseteq \boldsymbol{T}_n^d$, we say a contour $\boldsymbol{\gamma}$ is *a contour in* $\boldsymbol{\Lambda}$ if $d_\infty(\boldsymbol{\gamma}, \boldsymbol{T}_n^d \setminus \boldsymbol{\Lambda}) \geq 1/2$. The distance to the empty set is infinite by convention.

Write $\mathcal{C}(\boldsymbol{\Lambda})$ for the set of contours in $\boldsymbol{\Lambda}$, and $\mathcal{C} = \mathcal{C}(\boldsymbol{T}_n^d)$ for the set of all contours. For $\boldsymbol{\Lambda} \subseteq \boldsymbol{T}_n^d$ define $\mathcal{G}^{\text{ext}}(\boldsymbol{\Lambda})$ to be the set of matching mutually external contours in $\boldsymbol{\Lambda}$, and then define

$$Z_{\text{ord}}(\boldsymbol{\Lambda}) := \sum_{\Gamma \in \mathcal{G}_{\text{ord}}^{\text{ext}}(\boldsymbol{\Lambda})} e^{-e_{\text{ord}}|\boldsymbol{\Lambda} \cap \text{Ext } \Gamma|} \prod_{\boldsymbol{\gamma} \in \Gamma} e^{-\kappa \|\boldsymbol{\gamma}\|} Z_{\text{dis}}(\text{Int } \boldsymbol{\gamma}), \tag{42}$$

$$Z_{\text{dis}}(\boldsymbol{\Lambda}) := \sum_{\Gamma \in \mathcal{G}_{\text{dis}}^{\text{ext}}(\boldsymbol{\Lambda})} e^{-e_{\text{dis}}|\boldsymbol{\Lambda} \cap \text{Ext } \Gamma|} \prod_{\boldsymbol{\gamma} \in \Gamma} e^{-\kappa \|\boldsymbol{\gamma}\|} q Z_{\text{ord}}(\text{Int } \boldsymbol{\gamma}), \tag{43}$$

where the sums in (42) and (43) run over sets of matching mutually external contours in which Ext Γ is labeled ord and dis, respectively. This is the desired resummation. In the special case $\boldsymbol{\Lambda} = \boldsymbol{T}_n^d$ these partition functions represent the sums of $w(A)$ over

$$\Omega_{\text{ord}} := \{A \in \Omega \setminus \Omega_{\text{tunnel}} | \text{Ext } \Gamma(A) \text{ is labeled ord}\}, \tag{44}$$

$$\Omega_{\text{dis}} := \{A \in \Omega \setminus \Omega_{\text{tunnel}} | \text{Ext } \Gamma(A) \text{ is labeled dis}\}. \tag{45}$$

That is, we get a decomposition $Z_{\text{rest}} = qZ_{\text{ord}} + Z_{\text{dis}}$, where

$$Z_{\text{ord}} = q^{-1} \sum_{A \in \Omega_{\text{ord}}} w(A), \quad Z_{\text{dis}} = \sum_{A \in \Omega_{\text{dis}}} w(A). \tag{46}$$

Subsection 3.8 will give interpretations of these quantities in terms of random cluster model partition functions for many other choices of $\boldsymbol{\Lambda}$.

3.5 | Labeled contours

This subsection introduces labeled contours and establishes some basic properties of these objects. These properties will ensure that we can efficient enumerate labeled contours.

In Definition 2, we associated a labeling to an entire collection of matching and compatible contours and interfaces. For collections of contours, since each contour splits $\mathbb{T}_n^d$ into two pieces, it is more convenient to associate the labeling to individual contours. We do this by assigning a label to Int $\boldsymbol{\gamma}$ (resp. Ext $\boldsymbol{\gamma}$) according to the label of the region of $\mathbb{T}_n^d \setminus \cup_{\boldsymbol{\gamma} \in \Gamma} \boldsymbol{\gamma}$ adjacent to $\boldsymbol{\gamma}$ contained in Int $\boldsymbol{\gamma}$ (resp. Ext $\boldsymbol{\gamma}$).

A *compatible set of labeled contours* Γ is a set of compatible contours Γ such that the connected components of $\mathbb{T}_n^d \setminus \cup_{\boldsymbol{\gamma} \in \Gamma} \boldsymbol{\gamma}$ are assigned the same labels by the labeled contours. More precisely, for a component $\boldsymbol{B}$ of $\mathbb{T}_n^d \setminus \cup_{\boldsymbol{\gamma} \in \Gamma} \boldsymbol{\gamma}$, $\partial \boldsymbol{B}$ is a union of compatible contours $\boldsymbol{\gamma}_0, \ldots, \boldsymbol{\gamma}_k$ for some $k \geq 0$, and (up to relabeling) either (i) $\boldsymbol{\gamma}_i < \boldsymbol{\gamma}_0$ for $i = 1, \ldots, k$ or (ii) $\boldsymbol{\gamma}_i \perp \boldsymbol{\gamma}_j$ for $i \neq j$. The condition of compatibility of the labels in the first case is that the interior label of $\boldsymbol{\gamma}_0$ is the same as the exterior label of $\boldsymbol{\gamma}_i$ for all $i = 1, \ldots k$, and in the second case is that all exterior labels agree.

By construction, the set of collections of matching and compatible contours is the same as the set of collections of compatible labeled contours. The advantage of the latter is that it enables us to define a labeled contour $\boldsymbol{\gamma}$ to be *ordered* if its exterior label is ord, and *disordered* if its exterior label is dis.

We let $C_{\mathrm{ord}}(\boldsymbol{\Lambda})$ and $C_{\mathrm{dis}}(\boldsymbol{\Lambda})$ denote the sets of labeled contours in $\boldsymbol{\Lambda}$ with external labels ord and dis, respectively, with $C_{\mathrm{ord}} = C_{\mathrm{ord}}(\boldsymbol{T}_n^d)$ and $C_{\mathrm{dis}} = C_{\mathrm{dis}}(\boldsymbol{T}_n^d)$. The next lemma gives a way to construct a labeled contour $\boldsymbol{\gamma}$ from an edge configuration.

Lemma 3.4. *Let $\ell \in \{\mathrm{ord}, \mathrm{dis}\}$, let $\boldsymbol{\gamma} \in C_\ell$, and $\boldsymbol{\Lambda} = \mathrm{Int}\ \boldsymbol{\gamma}$. Then*

- *If $\ell = \mathrm{dis}$, let $E'(\boldsymbol{\Lambda})$ be set of edges contained in $\boldsymbol{\Lambda}$. Then $\boldsymbol{\gamma}$ is the unique component of $\partial \boldsymbol{A}$ where $A = E'(\boldsymbol{\Lambda}) \subset E$.*
- *If $\ell = \mathrm{ord}$, let $E'(\boldsymbol{\Lambda})$ be the set of edges whose midpoints are contained in $\boldsymbol{\Lambda}$. Then $\boldsymbol{\gamma}$ is the unique component of $\partial \boldsymbol{A}$ where $A = E \setminus E'(\boldsymbol{\Lambda})$.*

Proof. These claims follows from [8, Lemma 5.1]; see the proof of [8, Lemma 5.11].[3] ∎

Lemma 3.4 gives a way to construct a given contour from some set of edges A. For our algorithms it will be important to be able to generate contours from a relatively small set of edges. We first explain how to do this for disordered contours.

Suppose $\boldsymbol{\gamma} \in C_{\mathrm{dis}}$ and let $\Lambda = \mathrm{Int}\ \boldsymbol{\gamma} \cap \mathbb{T}_n^d$. Define

$$\mathcal{E}_\gamma := \{e = \{i,j\} | i,j \in \Lambda, \quad d_\infty(\mathrm{mid}(\boldsymbol{e}), \boldsymbol{\gamma}) \geq 3/4\}, \tag{47}$$

where $\mathrm{mid}(\boldsymbol{e})$ denotes the midpoint of the edge $\boldsymbol{e}$; this is the vertex of $\frac{1}{2}\mathbb{T}_n^d$ on the two-step path from i to j in $\frac{1}{2}\mathbb{T}_n^d$.

Lemma 3.5. *Suppose $\boldsymbol{\gamma} \in C_{\mathrm{dis}}$ and let $\boldsymbol{\Lambda} = \mathrm{Int}\ \boldsymbol{\gamma}$. Suppose $F \subseteq \mathcal{E}_\gamma$ and let $A = E' \setminus F$, where $E' = E'(\Lambda)$ is defined as in Lemma 3.4. Let Γ be the set of contours in $\partial \boldsymbol{A}$. Then $\boldsymbol{\gamma} \in \Gamma$, and for all $\boldsymbol{\gamma}' \in \Gamma$ with $\boldsymbol{\gamma}' \neq \boldsymbol{\gamma}$ we have $\boldsymbol{\gamma}' < \boldsymbol{\gamma}$. Moreover, all sets of matching contours consisting of γ and contours in $\mathrm{Int}\ \gamma$ arise from such F.*

Proof. We begin by recalling an alternate construction of $\boldsymbol{A}$ from [8]. Let $E \subset E(\mathbb{T}_n^d)$, and let $D \subset E$. Set $D^\star$ to be the set of $(d-1)$-dimensional unit hypercubes dual to the edges of D, and set

$$V_-(D) = \left\{x \in V(\mathbb{T}_n^d) | \{x,y\} \in D \text{ if } \{x,y\} \in E\right\}.$$

Set $\boldsymbol{D}_{\mathrm{dis}}$ to be the union of the open $3/4$-neighborhood of $V_-(D)$ and the open $1/4$-neighborhood of $D^\star$. Then by [8, Lemma 5.1, (iv)], if $D = E \setminus A$, $\boldsymbol{E} \setminus \boldsymbol{A} = \boldsymbol{D}_{\mathrm{dis}}$. That is, $\boldsymbol{D}_{\mathrm{dis}}$ is the disordered region associated to A (relative to the region $\boldsymbol{E}$).

To prove the lemma, we apply this construction with $E = E'(\Lambda)$ and $D = F$. The definition of $\mathcal{E}_\gamma$ ensures that both the open $3/4$-neighborhoods of the included vertices and the open $1/4$-neighborhoods of the included dual facets are at distance at least $1/2$ from $\boldsymbol{\gamma}$. This implies that $\boldsymbol{\gamma}$ is a boundary component of $\boldsymbol{E} \setminus F$, and the first claim follows as all other boundary components are adjacent to $\boldsymbol{D}_{\mathrm{dis}}$. The second claim follows from the bijection of Lemma 3.2, which restricts to a bijection in this setting. ∎

Lemma 3.6. *Suppose $\boldsymbol{\gamma} \in C_{\mathrm{dis}}$. Then there is a connected graph with edge set A such that (i) $|A| \leq 2d\|\boldsymbol{\gamma}\|$ and (ii) $\boldsymbol{\gamma}$ is the outermost contour in $\partial \boldsymbol{A}$.*

[3] These results rely only on the geometry of hypercubes and not on the definitions of interior/exterior.

Proof. Choose $F = \mathcal{E}_\gamma$ in Lemma 3.5. Then the subgraph of $\mathbb{T}_n^d$ induced by $E'' = E'(\Lambda) \setminus F$ is connected: if not $\partial \boldsymbol{E}''$ would contain two compatible exterior contours as the boundaries of the thickenings of the connected components of E'' are compatible. This would contradict the conclusion of Lemma 3.5 that there is a unique exterior contour.

The bound on the size of A is crude; it can be obtained by noting that the included edges all contain a vertex from which there is an edge outgoing from Λ, and the count of these vertices is a lower bound for $\|\boldsymbol{\gamma}\|$. Each of the vertices is contained in at most $2d$ edges. ∎

We now establish a similar way to construct an ordered contour from a small edge set. The situation is slightly different due to the differences between ordered and disordered contours in Lemma 3.4. Define, for $\boldsymbol{\gamma} \in \mathcal{C}_{\text{ord}}$, $\Lambda = \text{Int } \boldsymbol{\gamma} \cap \mathbb{T}_n^d$,

$$\mathcal{E}_\gamma := \{\{i,j\} | i,j \in \Lambda\}. \tag{48}$$

Lemma 3.7. *Suppose $\boldsymbol{\gamma} \in \mathcal{C}_{\text{ord}}$ and $F \subseteq \mathcal{E}_\gamma$. Let $A = (E \setminus E'(\Lambda)) \cup F$, where $E'(\Lambda)$ is defined as in Lemma 3.4. Let Γ be the set of contours in $\partial \boldsymbol{A}$. Then $\boldsymbol{\gamma} \in \Gamma$, and for all $\boldsymbol{\gamma}' \in \Gamma$ with $\boldsymbol{\gamma}' \neq \boldsymbol{\gamma}$ we have $\boldsymbol{\gamma}' < \boldsymbol{\gamma}$. Moreover, all sets of matching contours consisting of γ and contours in* Int γ *arise from such F.*

Proof. The proof is essentially the same as for Lemma 3.5. Let $A' = E \setminus E'(\Lambda)$. The set $\boldsymbol{F}$ is disjoint from $\boldsymbol{A}'$ as every vertex i interior to $\boldsymbol{\gamma}$ is at distance at least $3/4$ from $\boldsymbol{\gamma}$. This implies $\partial \boldsymbol{A}$ is the union of $\partial \boldsymbol{A}'$ and $\partial \boldsymbol{F}$, which implies the first claim. The second claim follows from the bijection of Lemma 3.2, which restricts to a bijection in this setting. ∎

Two edges $e,f \in E$ are called *1-adjacent* if $d_\infty(\boldsymbol{e},\boldsymbol{f}) \leq 1$. A set of edges A is *1-connected* if for any $e,f \in A$, there is a sequence of 1-adjacent edges in A from e to f. In the next lemma, $\partial \boldsymbol{A}^{\boldsymbol{c}}$ is the boundary of the thickening of the edge set $A^c = E \setminus A$.

Lemma 3.8. *Suppose $\boldsymbol{\gamma} \in \mathcal{C}_{\text{ord}}$. Then there is a 1-connected set of edges A of size at most $\|\boldsymbol{\gamma}\|$ such that $\boldsymbol{\gamma}$ is the outermost contour in $\partial \boldsymbol{A}^{\boldsymbol{c}}$.*

Proof. Let A be the set of all edges that intersect $\boldsymbol{\gamma}$. By the definition of $\|\cdot\|$, $|A| \leq \|\boldsymbol{\gamma}\|$. By Lemma 3.7 $\boldsymbol{\gamma}$ is the outermost contour in $\boldsymbol{A}^{\boldsymbol{c}}$, as $A^c = E'(\Lambda) \cup \mathcal{E}_{\text{dis}}(\Lambda)$. The 1-connectedness of A follows from the connectedness of $\boldsymbol{\gamma}$ and the observation that every point of $\boldsymbol{\gamma}$ is at most d_∞ distance $1/2$ from an edge in A. ∎

3.6 | Contour enumeration

This section uses the results of the previous subsection to guarantee the existence of an efficient algorithm for enumerating contours. This requires a few additional lemmas.

Lemma 3.9. *For all $\boldsymbol{\gamma} \in \mathcal{C}$, $|\text{Int } \boldsymbol{\gamma}| \leq \|\boldsymbol{\gamma}\|^2$, and $|\text{Int } \boldsymbol{\gamma}| \leq (n/2)\|\boldsymbol{\gamma}\|$.*

Proof. This follows by [8, Lemma 5.7], as the interior of a contour as defined by Definition 5 is always smaller than the definition of the interior of a contour in [8]. ∎

Lemma 3.10. *There is an algorithm that determines the vertex set* Int $\boldsymbol{\gamma} \cap \mathbb{T}_n^d$ *in time $O(\|\boldsymbol{\gamma}\|^3)$.*

Proof. Let $m = \|\boldsymbol{\gamma}\|$. Let G be the subgraph of $\mathbb{T}_n^d$ that arises after removing all edges that intersect some $(d-1)$-dimensional face in $\boldsymbol{\gamma}$. Consider the following greedy algorithm to determine the two connected components of G. The algorithm starts at $C_0 = x$, where x is chosen such that it is contained in an edge not present in G. The algorithm determines the connected component containing x in G by adding at step $i+1$ the first vertex (with respect to lexicographic order) in $\mathbb{T}_n^d \setminus C_i$ that neighbors C_i; if no neighbors exist the component has been determined. The kth step takes time at most $(2d)k$, so performing N steps of this algorithm takes time $O(N^3)$.

Since $|\mathrm{Int}\,\boldsymbol{\gamma}| \le \|\boldsymbol{\gamma}\|^2$ by Lemma 3.9, we can stop the greedy procedure after m^2+1 steps. If the algorithm terminates due to this condition, the component being explored is the exterior component. The interior component can then be determined in at most $O(m^3)$ additional steps by running the greedy algorithm from the neighbor of x that is in the interior component. Otherwise the algorithm will have already terminated and determined the interior component. ∎

Lemma 3.11. *Fix an edge $e \in E$. There is an algorithm to construct all contours $\boldsymbol{\gamma} \in \mathcal{C}_{\mathrm{ord}}$ that (i) can arise from a connected edge set A that contains e and (ii) have $\|\boldsymbol{\gamma}\| \le m$. The algorithm runs in time* $\exp(O(m))$.

Similarly, there is an $\exp(O(m))$*-time algorithm to construct all contours $\boldsymbol{\gamma} \in \mathcal{C}_{\mathrm{dis}}$ that (i) can arise from an edge set A such that A^c is 1-connected and contains e and (ii) have $\|\boldsymbol{\gamma}\| \le m$.*

Proof. We first consider disordered contours, and begin by enumerating all connected sets A of edges that contain e that are of size at most $2dm$. This can be done in time $\exp(O(m))$. If $2m \le n$ then we consider the enumerated edge sets as subsets of $E(\mathbb{T}_{2m}^d)$; otherwise we consider them as subsets of $E(\mathbb{T}_n^d)$.

For each edge set A, construct $\partial \boldsymbol{A}$ and take the outermost contour (if there is not a single outermost contour, discard the result). By Lemma 3.6 this generates all disordered contours of size at most m that arise from connected edge sets containing e. We obtain the desired list of contours by removing any duplicates, which takes time at most $\exp(O(m))$. The remainder of the proof shows that the operations in this paragraph can be done in time polynomial in m.

The constructions of $\partial \boldsymbol{A}$ takes time at most $O(m)$ as it is a $\mathbb{Z}_2$ sum of $(d-1)$-dimensional facets, and determining these facets takes a constant amount of time (depending only on the dimension d) for each edge. Determining if a component of $\partial \boldsymbol{A}$ is a contour can be done by computing the winding number of the component; this takes time $O((2dm \wedge n)K)$ for a component with K facets by Lemma 3.1. Determining the interior of a given contour takes time at most $O(m^3)$ by Lemma 3.10, and hence we can check if $\mathrm{Int}\,\boldsymbol{\gamma}' \subset \mathrm{Int}\,\boldsymbol{\gamma}$ for all pairs in time $O(m^4)$ since there are at most m^2 contours. This completes the proof for disordered contours.

For ordered contours the argument applies nearly verbatim. The changes are as follows. First, enumerate 1-connected sets A^c that contain e. Second, to see that we get the desired contours, appeal to Lemma 3.8. Lastly, computing $\partial \boldsymbol{A}$ takes time $O((m \wedge n)^d)$ which is polynomial in m; this is by our choice of torus in the first paragraph of the proof. ∎

The next definition is useful for inductive arguments involving contours.

Definition 6. *The level $\mathcal{L}(\boldsymbol{\gamma})$ of a contour $\boldsymbol{\gamma}$ is defined inductively as follows. If $\boldsymbol{\gamma}$ is thin, meaning $C(\mathrm{Int}\,\boldsymbol{\gamma}) = \emptyset$, then $\mathcal{L}(\boldsymbol{\gamma}) = 0$. Otherwise, $\mathcal{L}(\boldsymbol{\gamma}) = 1 + \max\{\mathcal{L}(\boldsymbol{\gamma}') | \boldsymbol{\gamma}' < \boldsymbol{\gamma}\}$.*

Call a set $\boldsymbol{\Lambda} \subseteq \boldsymbol{T}_n^d$ a *region* if $\boldsymbol{\Lambda} = \boldsymbol{T}_n^d$ or if $\boldsymbol{\Lambda}$ is a connected component of $\boldsymbol{T}_n^d \setminus \partial \boldsymbol{A}$ for some $A \subset E$. In the former case set $\partial \boldsymbol{\Lambda} = \emptyset$, and in the latter case set $\partial \boldsymbol{\Lambda}$ to be the union of all connected components

of ∂A incident to Λ. In particular if $\Lambda = \operatorname{Int} \gamma$ for some contour γ, then Λ is a region and $\partial\Lambda = \gamma$. Finally, for compatible contours $\gamma_1, \dots, \gamma_t$, define $\|\gamma_1 \cup \cdots \cup \gamma_t\| = \|\gamma_1\| + \cdots + \|\gamma_t\|$. We conclude this subsection by stating our main algorithmic result on efficiently computing sets of contours.

Proposition 3.12. *There is an $O((|\Lambda| + \|\partial\Lambda\|)\exp(O(m)))$-time algorithm that, for all regions Λ, (i) enumerates all contours in $C_{\mathrm{ord}}(\Lambda) \cup C_{\mathrm{dis}}(\Lambda)$ with size at most m and (ii) sorts this list consistent with the level assignments.*

Proof. We begin by proving the first item. Apply Lemma 3.11 for each edge contained in Λ. This takes time $O(|\Lambda|\exp(O(m)))$ as there are at most $2d$ edges in Λ for each vertex of $\mathbb{T}_n^d$ in Λ. The output is a (multi-)set of contours of size at most m contained in $\mathbb{T}_n^d$. Trim the resulting list of contours to remove duplicates.

By Lemma 3.10 in time $\exp(O(m))$ we can determine $\operatorname{Int}\gamma$ for every γ from the list obtained in the first paragraph. We determine the list of level zero contours by iterating through the list, checking for each γ if $\gamma' < \gamma$ for some other $\gamma' \neq \gamma$ in the list. If not, assign γ level 0. This takes time at most $\exp(O(m))$. We continue by running the same operation on the sublist of all contours of level at least one, that is, the sublist of contours not assigned level 0. If γ has level at least one and there is no $\gamma' < \gamma$, γ' also of level at least one, then γ is assigned level one. By repeating this we assign a level to every contour. The maximal level of a contour is m^2, the maximal size of the interior of a contour of size m, and hence the total running time is at most $m^2\exp(O(m)) = \exp(O(m))$.

To conclude, trim the list to retain only contours γ' contained in Λ. This can be done by removing contours at distance less than $1/2$ from γ. Computing this distance takes time $O(\|\gamma\|\|\gamma'\|)$, which is at most $O(\|\gamma\|m)$. ∎

3.7 | Polymer representations for Z_{ord} and Z_{dis}

To obtain polymer model representations of Z_{ord} and Z_{dis}, define $\tilde{\Omega}_{\mathrm{ord}}(\Lambda)$ and $\tilde{\Omega}_{\mathrm{dis}}(\Lambda)$ to be the sets of compatible collections of contours in Λ that are labeled ord and dis, respectively. Define

$$K_{\mathrm{ord}}(\gamma) = e^{-\kappa\|\gamma\|}\frac{Z_{\mathrm{dis}}(\operatorname{Int}\gamma)}{Z_{\mathrm{ord}}(\operatorname{Int}\gamma)}, \qquad K_{\mathrm{dis}}(\gamma) = e^{-\kappa\|\gamma\|}\frac{qZ_{\mathrm{ord}}(\operatorname{Int}\gamma)}{Z_{\mathrm{dis}}(\operatorname{Int}\gamma)}. \tag{49}$$

By following a well trodden path in Pirogov–Sinai theory (see, e.g., [8, p. 28] or [25, p. 28]), these definitions give the following representations for Z_{ord} and Z_{dis} as partition functions of abstract polymer models:

$$Z_{\mathrm{ord}}(\Lambda) = e^{-e_{\mathrm{ord}}|\Lambda|}\sum_{\Gamma\in\tilde{\Omega}_{\mathrm{ord}}(\Lambda)}\prod_{\gamma\in\Gamma}K_{\mathrm{ord}}(\gamma), \tag{50}$$

$$Z_{\mathrm{dis}}(\Lambda) = e^{-e_{\mathrm{dis}}|\Lambda|}\sum_{\Gamma\in\tilde{\Omega}_{\mathrm{dis}}(\Lambda)}\prod_{\gamma\in\Gamma}K_{\mathrm{dis}}(\gamma), \tag{51}$$

where the sums run over collections of compatible labeled contours in Λ with external label ord and dis, respectively.

In fact, for $\ell \in \{\mathrm{ord}, \mathrm{dis}\}$, the above formulas represent $Z_\ell(\Lambda)$ as the partition function of a polymer model in the form discussed in Section 2, that is, where polymers are subgraphs of a fixed graph G with bounded degree. In detail, recalling the discussion in Section 3.1, we consider contours as induced subgraphs of (a subgraph of) the bounded-degree graph $(\frac{1}{2}\mathbb{T}_n^d)^\star$. Thus $|\gamma|$ is the number of vertices in

a contour when represented as a subgraph. Condition (14) holds with $b = 1$ since $\|\boldsymbol{\gamma}\| \geq |\gamma|$ by (32). The more substantial hypothesis (15) will be verified in later sections for appropriate choices of the label and of β.

In the sequel we will write $|\boldsymbol{\Lambda}|_{(\frac{1}{2}\mathbb{T}_n^d)^\star}$ for the size of set of vertices of $\left(\frac{1}{2}\mathbb{T}_n^d\right)^\star$ that are part of some contour $\boldsymbol{\gamma}$ in $C_\ell(\boldsymbol{\Lambda})$ for some ℓ. The next technical lemma shows it is enough to find algorithms that are polynomial time in $|\boldsymbol{\Lambda}|_{(\frac{1}{2}\mathbb{T}_n^d)^\star}$.

Lemma 3.13. *For $\boldsymbol{\Lambda}$ a continuum set, $|\boldsymbol{\Lambda}|_{(\frac{1}{2}\mathbb{T}_n^d)^\star}$ is polynomial in $|\Lambda|$.*

Proof. By construction, contours inside $\boldsymbol{\Lambda}$ arise from edge configurations of edges inside $\boldsymbol{\Lambda}$. The number of such edges is at most $2d$ times the number of vertices inside. Since contours are boundaries of unions of $(d-1)$-dimensional hypercubes centered at vertices in $(\frac{1}{2}\mathbb{T}_n^d)^\star$ that lie on edges, this proves the claim, since there are a bounded number of such hypercubes associated to each edge. ∎

3.8 | Random cluster model formulations of contour partition functions

The definitions of the partition functions $Z_{\text{ord}}(\boldsymbol{\Lambda})$ and $Z_{\text{dis}}(\boldsymbol{\Lambda})$ in (42) and (43) only involve contours. In general, these contour partition functions do not correspond to random cluster model partition functions due to the exclusion of interfaces. However, we will show that when $\Lambda = \text{Int}\,\boldsymbol{\gamma} \cap \mathbb{T}_n^d$ can be embedded as a subgraph of $\mathbb{Z}^d$, there is such an interpretation.

To make this precise, recall the definitions (7) and (8) of Z_Λ^f and Z_Λ^w for $\Lambda \subset \mathbb{Z}^d$ such that the subgraph G_Λ induced by Λ is simply connected. Recall that $p = 1 - e^{-\beta}$.

Proposition 3.14. *Suppose $\Lambda \subset \mathbb{Z}^d$ is simply connected, and let $n = 3|\Lambda|$. Then there are contours $\boldsymbol{\gamma}_{\text{dis}} \in C_{\text{dis}}(\mathbb{T}_n^d)$ and $\boldsymbol{\gamma}_{\text{ord}} \in C_{\text{ord}}(\mathbb{T}_n^d)$ determined by Λ such that*

$$Z_{\text{dis}}(\text{Int}\,\boldsymbol{\gamma}_{\text{ord}}) = (1-p)^{-\frac{1}{2}\|\boldsymbol{\gamma}_{\text{ord}}\|} Z_\Lambda^f, \quad Z_{\text{ord}}(\text{Int}\,\boldsymbol{\gamma}_{\text{dis}}) = q^{-1} p^d \left|\text{Int}\,\boldsymbol{\gamma}_{\text{ord}} - |\{E(\Lambda)\}|\right| Z_\Lambda^w.$$

Proof. Since $n = 3|\Lambda|$, we can embed $\Lambda \subset \mathbb{T}_n^d$. Moreover, the set of boundary vertices $\partial\Lambda := \{i \in \Lambda : \exists j \in \mathbb{Z}^d \setminus \Lambda, (i,j) \in E(\mathbb{Z}^d)\}$ can be identified with $\{i \in \Lambda : \exists j \in \Lambda^c, (i,j) \in E\}$. Thus the graphs G_Λ and G'_Λ used in the definitions of Z_Λ^f and Z_Λ^w are the same whether defined by considering Λ as a subset of $\mathbb{Z}^d$ or $\mathbb{T}_n^d$. Note that by our choice of n we know that any component of $\partial\boldsymbol{A}$ will be a contour if A is a subset of edges that are at graph distance at most two from Λ. To see this in an elementary way, note that we can further consider Λ as a subset of $\mathbb{T}_n^d$ such that the fundamental loops of $\mathbb{T}_n^d$ are at distance at least (say) ten from $\boldsymbol{\Lambda}$.

We first consider the case of Z_Λ^f. To do this, let $A_0 \subset E$ be the set of edges with both endpoints in Λ^c. Let $\boldsymbol{\gamma}_{\text{ord}}$ be the unique contour in $\partial\boldsymbol{A_0}$; the fact that there is a unique contour follows from the fact that Λ is simply connected. By Lemma 3.7, for any subset A of edges in $E(G_\Lambda) = \mathcal{E}_{\boldsymbol{\gamma}_{\text{ord}}}$, the contours of $\partial\boldsymbol{A}$ are contained in Int $\boldsymbol{\gamma}_{\text{ord}}$. Moreover, this lemma ensures that by carrying out the contour construction of Section 3.3.1 for subsets of edges $A' = A_0 \cup A$ where all edges of A are from $E(G_\Lambda)$, we obtain all contour configurations $\Gamma = \{\gamma_{\text{ord}}\} \cup \Gamma'$ where the contours of Γ' are contained in Int $\boldsymbol{\gamma}_{\text{ord}}$.

To obtain the conclusion, note that (i) $\sum_{A'} w(A')$ is proportional to Z_Λ^f, where the sum runs over these $A' = A_0 \cup A$ described above, and (ii) $\sum_{A'} w(A')$ is proportional to $Z_{\text{dis}}(\text{Int}\,\boldsymbol{\gamma}_{\text{ord}})$. To obtain the proportionality constant we compare the contributions of the empty edge configuration (empty contour

configuration), see (43). These are, respectively, $q^{|\Lambda|}(1-p)^{|E(\Lambda)|}$ and $q^{|\Lambda|}(1-p)^{d|\Lambda|}$. The ratio of these terms is $(1-p)^{-\frac{1}{2}\|\boldsymbol{\gamma}_{\text{ord}}\|}$ since $\|\boldsymbol{\gamma}_{\text{ord}}\|$ is exactly the number of edges between Λ and Λ^c.

We now consider the case of Z^w_Λ. Let $A = E(G_\Lambda)$, and consider the ordered contour $\boldsymbol{\gamma}'$ that arises from the edge set $E \setminus A$. Define

$$\tilde{A} := A \cup \{e \in E | d_\infty(\text{mid}(\boldsymbol{e}), \boldsymbol{\gamma}') \leq 1/2\}, \tag{52}$$

the set of edges whose midpoints are either in the interior of $\boldsymbol{\gamma}'$ or within distance $1/2$ of $\boldsymbol{\gamma}'$. Then set $\boldsymbol{\gamma}_{\text{dis}}$ to be the single contour in $\partial\tilde{\boldsymbol{A}}$; there is only one contour in this set by the assumption Λ is simply connected. Note that A is precisely $\mathcal{E}_{\boldsymbol{\gamma}_{\text{dis}}}$ as defined above Lemma 3.5, and hence there is a bijection between contour configurations in Int $\boldsymbol{\gamma}_{\text{dis}}$ and subsets of $\tilde{A}$ in which each edge not in A is occupied. As for the case of Z^f_Λ we can now conclude, as summing over such edge sets is proportional to both $Z_{\text{ord}}(\text{Int } \boldsymbol{\gamma}_{\text{dis}})$ (recall (46)) and Z^w_Λ. To compute the proportionality constant, we compare the all occupied configuration to the empty contour configuration, see (42). This gives, respectively, $qp^{|E(\Lambda)|}$ and $e^{-e_{\text{ord}}|\text{Int } \boldsymbol{\gamma}_{\text{ord}}|}$, and hence

$$Z_{\text{ord}}(\text{Int } \boldsymbol{\gamma}_{\text{dis}}) = q^{-1}p^{d|\text{Int } \boldsymbol{\gamma}_{\text{ord}}|-|\{E(\Lambda)\}|}Z^w_\Theta. \tag{53}$$

■

4 | CONTOUR MODEL ESTIMATES

In this section, we state several estimates related to the contour representations from the previous section. Recall the definition (34) of Z.

Lemma 4.1 ([8, Lemma 6.1(a)]). *There are constants $c > 0$, $q_0 = q_0(d) < \infty$, and $n_0 < \infty$ such that if $q \geq q_0$, $n \geq n_0$, and $\beta \geq \beta_c$,*

$$\frac{Z_{\text{tunnel}}}{Z} \leq \exp(-c\beta n^{d-1}). \tag{54}$$

In what follows c will always denote the constant from Lemma 4.1, and q_0 and n_0 will always be at least as large as the constants in the lemma. More precisely, several lemmas will require q_0 to be chosen large enough, and we implicitly take q_0 to be the maximum of these requirements. We also choose n_0 large enough so that via (3) we have $\beta_c > \beta_h$. Lemma 4.1 ensures that Z_{tunnel} is neglectable when approximating Z up to relative errors $\epsilon \gg \exp(-c\beta n^{d-1})$. We will also need to know that Z_{dis} is neglectable when $\beta > \beta_c$. This requires two lemmas.

Lemma 4.2. *If $q \geq q_0$, $n \geq n_0$, and $\beta > \beta_c$ there exist constants $a_{\text{dis}} > 0$ and $f > 0$ so that if $\epsilon_n := 2\exp(-c\beta n)$, then*

$$Z_{\text{ord}} \geq \exp(-(f+\epsilon_n)n^d), \quad Z_{\text{dis}} \leq \exp((-f+\epsilon_n)n^d)\max_{\Gamma \in \mathcal{G}^{\text{ext}}_{\text{dis}}} e^{-\frac{a_{\text{dis}}}{2}|\text{Ext } \Gamma|}\prod_{\boldsymbol{\gamma}\in\Gamma} e^{-\frac{c}{2}\beta\|\boldsymbol{\gamma}\|}. \tag{55}$$

Proof. With $a_{\text{dis}} \geq 0$ this follows from [8, Lemma 6.3] provided $f = f_{\text{ord}}$ for $\beta \geq \beta_c$, and that $f = f_{\text{ord}}$ follows from [8, Lemma A.3]. What remains is to prove $a_{\text{dis}} > 0$ when $\beta > \beta_c$. The results of [32] imply that there is a unique Gibbs measure for the random cluster model when $\beta > \beta_c$. If a_{dis} was 0 for some $\beta > \beta_c$, then the argument establishing [8, Lemma 6.1 (b)] implies the existence of multiple Gibbs measures, a contradiction. ■

Lemma 4.3. *If $q \geq q_0$, $n \geq n_0$, and $\beta > \beta_c$, then there exists a constant $b_{\text{dis}} > 0$ so that*

$$\frac{Z_{\text{dis}}}{Z} \leq 2\exp(-b_{\text{dis}} n^{d-1}) \,. \tag{56}$$

Proof. Suppose $\Gamma \in \mathcal{G}_{\text{dis}}^{\text{ext}}$. Then we claim that

$$|\text{Ext } \Gamma| + \sum_{\gamma \in \Gamma} \|\gamma\| \geq 2n^{d-1}. \tag{57}$$

To see this, note that

$$|\text{Ext } \Gamma| + \sum_{\gamma \in \Gamma} |\text{Int } \gamma| = n^d, \tag{58}$$

which combined with Lemma 3.9 implies

$$|\text{Ext } \Gamma| + \frac{n}{2} \sum_{\gamma \in \Gamma} \|\gamma\| \geq n^d, \tag{59}$$

which implies (57) when $n \geq 2$.

By Lemma 4.2, if n is large enough,

$$\frac{Z_{\text{dis}}}{Z_{\text{ord}}} \leq 2 \max_{\Gamma \in \mathcal{G}_{\text{dis}}^{\text{ext}}} e^{-\frac{a_{\text{dis}}}{2}|\text{Ext } \Gamma|} \prod_{\gamma \in \Gamma} e^{-\frac{c}{2}\beta\|\gamma\|}. \tag{60}$$

Set $b_{\text{dis}} := \min\{a_{\text{dis}}, c\beta\} > 0$. By (57),

$$e^{-\frac{a_{\text{dis}}}{2}|\text{Ext } \Gamma|} \prod_{\gamma \in \Gamma} e^{-\frac{c}{2}\beta\|\gamma\|} \leq \exp(-b_{\text{dis}} n^{d-1}) \tag{61}$$

for all $\Gamma \in \mathcal{G}_{\text{dis}}^{\text{ext}}$. The lemma now follows from (60). ∎

The next two lemmas will allow us to verify the Kotecký–Preiss condition for the contour models defining Z_{dis} and Z_{ord} from the previous section.

Lemma 4.4 ([8, Lemma 6.3])**.** *If $q \geq q_0$ and $\beta = \beta_c$, then*

$$K_{\text{ord}}(\boldsymbol{\gamma}) \leq e^{-c\beta\|\gamma\|}, \quad \text{and} \quad K_{\text{dis}}(\boldsymbol{\gamma}) \leq e^{-c\beta\|\gamma\|},$$

for all $\boldsymbol{\gamma}$ in $\mathcal{C}_{\text{ord}}$ and $\mathcal{C}_{\text{dis}}$, respectively.

Lemma 4.5 ([8, Lemma 6.3])**.** *If $q \geq q_0$ and $\beta > \beta_c$, then*

$$K_{\text{ord}}(\boldsymbol{\gamma}) \leq e^{-c\beta\|\gamma\|} \quad \text{for all} \quad \boldsymbol{\gamma} \in \mathcal{C}_{\text{ord}}.$$

In particular, since $\beta > \beta_h = \frac{3\log q}{4d}$, then for sufficiently large q the contour weights $w_\gamma = K_{\text{ord}}(\gamma)$ (for $\beta \geq \beta_c$) and $w_\gamma = K_{\text{dis}}(\gamma)$ (for $\beta = \beta_c$) will satisfy condition (15). Condition (14) is satisfied with $b = 1$ by the discussion in Section 3.7.

Next we will show that when $\beta > \beta_c$ and the disordered ground state is unstable, that regions with disordered boundary conditions "flip" quickly to ordered regions by way of a large contour; more precisely, the dominant contribution to $Z_{\text{dis}}(\mathbf{\Lambda})$ comes from collections of contours with small external volume.

For a region $\mathbf{\Lambda}$ and $M > 0$ we define

$$\mathcal{H}^{\text{flip}}_{\text{dis}}(\mathbf{\Lambda}, M) := \{\Gamma \in \mathcal{G}^{\text{ext}}_{\text{dis}}(\mathbf{\Lambda}) | \, |\{\text{Ext } \Gamma \cap \mathbf{\Lambda}\}| \leq M\},$$

and

$$Z^{\text{flip}}_{\text{dis}}(\mathbf{\Lambda}, M) := \sum_{\Gamma \in \mathcal{H}^{\text{flip}}_{\text{dis}}(\mathbf{\Lambda}, M)} e^{-e_{\text{dis}}|\text{Ext } \Gamma \cap \mathbf{\Lambda}|} \prod_{\gamma \in \Gamma} e^{-\kappa \|\gamma\|} q Z_{\text{ord}}(\text{Int } \boldsymbol{\gamma}). \tag{62}$$

Thus, compare (43), $Z^{\text{flip}}_{\text{dis}}(\mathbf{\Lambda}, M)$ is the contribution to $Z_{\text{dis}}(\mathbf{\Lambda})$ from contour configurations with small exterior volume.

Lemma 4.6. *Suppose $q \geq q_0$ and $\beta > \beta_c$. Then there exists $a_{\text{dis}} > 0$ so that the following holds for all $n \geq n_0$. Suppose $\boldsymbol{\gamma} \in C_{\text{ord}}$. For any $\epsilon > 0$, if*

$$M \geq \frac{2}{a_{\text{dis}}}(\kappa + 3)\|\boldsymbol{\gamma}\|, \tag{63}$$

then $Z^{\text{flip}}_{\text{dis}}(\text{Int } \boldsymbol{\gamma}, M)$ is an ϵ-relative approximation to $Z_{\text{dis}}(\text{Int } \boldsymbol{\gamma})$.

Proof. Let $\mathbf{\Lambda} = \text{Int } \boldsymbol{\gamma}$. Note that the lemma is immediate if Int $\boldsymbol{\gamma}$ does not contain any contours. Let

$$Z^{\text{err}}_{\text{dis}}(\mathbf{\Lambda}) := Z_{\text{dis}}(\Lambda) - Z^{\text{flip}}_{\text{dis}}(\mathbf{\Lambda}, M).$$

To prove the lemma it suffices to show that

$$0 \leq Z^{\text{err}}_{\text{dis}}(\mathbf{\Lambda}) / Z^{\text{flip}}_{\text{dis}}(\mathbf{\Lambda}, M) \leq \epsilon/2. \tag{64}$$

The lower bound is immediate since Z_{dis} is a sum of non-negative terms and $Z^{\text{flip}}_{\text{dis}}(\mathbf{\Lambda}, M)$ is at least one. Thus the proof of (64) has two parts: lower bounding $Z^{\text{flip}}_{\text{dis}}(\mathbf{\Lambda}, M)$ and upper bounding $Z^{\text{err}}_{\text{dis}}(\mathbf{\Lambda})$. The combination of these bounds will prove (64).

We begin with the lower bound on $Z^{\text{flip}}_{\text{dis}}(\mathbf{\Lambda}, M)$. Recall the definition (48) of $\mathcal{E}_\gamma$. Let $\boldsymbol{\gamma}' \in \mathcal{C}_{\text{dis}}(\mathbf{\Lambda})$ be the contour obtained by thickening $\mathcal{E}_\gamma$ and taking the boundary, that is, $\partial \mathcal{E}_\gamma$. Let $\Gamma = \{\boldsymbol{\gamma}'\}$. Note that Ext Γ contains no vertices, because $\mathbf{\Lambda}$ is connected and all edges inside $\mathbf{\Lambda}$ are in $\mathcal{E}_\gamma$.

Next observe that $\|\boldsymbol{\gamma}'\| \leq \|\boldsymbol{\gamma}\|$. This is because by construction any edge contributing to $\|\boldsymbol{\gamma}'\|$ must have one vertex outside of $\mathbf{\Lambda}$, and such an edge also contributes to $\|\boldsymbol{\gamma}\|$. In particular, $\Gamma \in \mathcal{H}^{\text{flip}}_{\text{dis}}(\Lambda, M)$, and hence

$$\begin{aligned} Z^{\text{flip}}_{\text{dis}}(\mathbf{\Lambda}, M) &\geq e^{-e_{\text{dis}}|\{\text{Ext } \Gamma \cap \mathbf{\Lambda}\}|} e^{-\kappa \|\boldsymbol{\gamma}'\|} q Z_{\text{ord}}(\text{Int } \boldsymbol{\gamma}') \\ &\geq e^{-\kappa \|\boldsymbol{\gamma}\|} q Z_{\text{ord}}(\text{Int } \boldsymbol{\gamma}') \\ &\geq e^{-(\kappa+1)\|\boldsymbol{\gamma}\|} q e^{-(f+\epsilon_n)\left|\{\text{Int } \boldsymbol{\gamma}'\}\right|} \\ &\geq \frac{1}{2} e^{-(\kappa+1)\|\boldsymbol{\gamma}\|} q e^{-f|\{\mathbf{\Lambda}\}|}, \end{aligned}$$

where $\epsilon_n = 2e^{-c\beta n}$ as above and f is the constant from Lemma 4.2. The second inequality used that Ext Γ contains no vertices. The second-to-last inequality follows from Lemma 4.2, and the last inequality follows since (i) $|\text{Int } \boldsymbol{\gamma}| = |\text{Int } \boldsymbol{\gamma}'|$ and (ii) for n large enough we have $e^{\epsilon_n|\text{Int } \gamma|} \leq 2$ for all $\boldsymbol{\gamma} \in \mathcal{C}$.

Next we prove an upper bound on $Z^{\text{flip}}_{\text{dis}}(\boldsymbol{\Lambda}, M)$. In fact, the upper bound is essentially contained in [8, Appendices A.2 and A.3], and we explain it here. Some further notation will be helpful. Let $a_{\text{dis}} > 0$ be the constant from Lemma 4.2. We call a contour $\boldsymbol{\gamma} \in \mathcal{C}_{\text{dis}}$ "small" if $\text{diam}(\boldsymbol{\gamma}) \leq \frac{c\beta}{a_{\text{dis}}}$ and "large' otherwise. Here $\text{diam}(\boldsymbol{\gamma})$ denotes the diameter of $\boldsymbol{\gamma}$, the maximum over $i = 1, \dots, n$ of $|\{I_i(\boldsymbol{\gamma})\}|$, where $I_i(\boldsymbol{\gamma}) = \{k \in \mathbb{Z}/n\mathbb{Z} | \boldsymbol{S}^{(i)}_k \cap \boldsymbol{\gamma} \neq \emptyset\}$, where $\boldsymbol{S}^i_k$ is the set $\{\boldsymbol{x} \in \boldsymbol{T}^d_n | \boldsymbol{x}_i = k\}$. See [8, p. 22].

For a region $\boldsymbol{\Lambda}'$, let

$$\mathcal{G}^{\text{ext,small}}_{\text{dis}}(\boldsymbol{\Lambda}') := \{\Gamma \in \mathcal{G}^{\text{ext}}_{\text{dis}}(\boldsymbol{\Lambda}') | \boldsymbol{\gamma}' \text{ is small } \forall \boldsymbol{\gamma}' \in \Gamma\},$$
$$\mathcal{G}^{\text{ext,large}}_{\text{dis}}(\boldsymbol{\Lambda}') := \{\Gamma \in \mathcal{G}^{\text{ext}}_{\text{dis}}(\boldsymbol{\Lambda}') | \boldsymbol{\gamma}' \text{ is large } \forall \boldsymbol{\gamma}' \in \Gamma\},$$

and

$$\begin{aligned} Z^{\text{small}}_{\text{dis}}(\boldsymbol{\Lambda}') &:= \sum_{\Gamma \in \mathcal{G}^{\text{ext}}_{\text{dis,small}}(\boldsymbol{\Lambda}')} e^{-e_{\text{dis}}|\text{Ext } \Gamma \cap \boldsymbol{\Lambda}'|} \prod_{\boldsymbol{\gamma} \in \Gamma} e^{-\kappa\|\boldsymbol{\gamma}\|} q Z_{\text{ord}}(\text{Int } \boldsymbol{\gamma}) \\ &= e^{-e_{\text{dis}}|\boldsymbol{\Lambda}'|} \sum_{\Gamma \in \mathcal{G}^{\text{ext,small}}_{\text{dis}}(\boldsymbol{\Lambda}')} \prod_{\boldsymbol{\gamma}' \in \Gamma} K_{\text{dis}}(\boldsymbol{\gamma}'). \end{aligned}$$

Moreover, let

$$\mathcal{H}^{\text{err}}_{\text{dis}}(\boldsymbol{\Lambda}) := \{\Gamma \in \mathcal{G}^{\text{ext}}_{\text{dis}}(\boldsymbol{\Lambda}) | \, |\{\text{Ext } \Gamma \cap \boldsymbol{\Lambda}\}| > M\}, \quad \text{and}$$
$$\mathcal{H}^{\text{err,large}}_{\text{dis}}(\boldsymbol{\Lambda}) := \{\Gamma \in \mathcal{G}^{\text{ext,large}}_{\text{dis}}(\boldsymbol{\Lambda}) | \, |\{\text{Ext } \Gamma \cap \boldsymbol{\Lambda}\}| > M\}.$$

Following the proof of [8, Lemma A.1], we have that

$$\begin{aligned} Z^{\text{err}}_{\text{dis}}(\boldsymbol{\Lambda}, M) &= \sum_{\Gamma \in \mathcal{H}^{\text{err}}_{\text{dis}}(\boldsymbol{\Lambda})} e^{-e_{\text{dis}}|\text{Ext } \Gamma \cap \boldsymbol{\Lambda}|} \prod_{\boldsymbol{\gamma}' \in \Gamma} e^{-\kappa\|\boldsymbol{\gamma}'\|} q Z_{\text{ord}}(\text{Int } \boldsymbol{\gamma}') \\ &\leq \sum_{\Gamma \in \mathcal{H}^{\text{err,large}}_{\text{dis}}(\boldsymbol{\Lambda})} Z^{\text{small}}_{\text{dis}}(\text{Ext } \Gamma \cap \boldsymbol{\Lambda}) \prod_{\boldsymbol{\gamma}' \in \Gamma} q e^{-\kappa\|\boldsymbol{\gamma}'\|} Z_{\text{ord}}(\text{Int } \boldsymbol{\gamma}') \\ &\leq e^{(\epsilon_n - f)|\boldsymbol{\Lambda}| + \|\boldsymbol{\gamma}\|} e^{-\frac{a_{\text{dis}}}{2} M} \sum_{\Gamma \in \mathcal{H}^{\text{err,large}}_{\text{dis}}(\boldsymbol{\Lambda})} e^{-\frac{a_{\text{dis}}}{2}|\text{Ext } \Gamma \cap \boldsymbol{\Lambda}|} \prod_{\boldsymbol{\gamma}' \in \Gamma} e^{-(\frac{\beta}{8} - 3)\|\boldsymbol{\gamma}'\|} \\ &\leq 2e^{-f|\Lambda| + 2\|\boldsymbol{\gamma}\|} e^{-\frac{a_{\text{dis}}}{2} M}. \end{aligned}$$

The first inequality follows since for each $\Gamma \in \mathcal{H}^{\text{err}}_{\text{dis}}(\boldsymbol{\Lambda})$, the set of large contours in Γ appear in $\mathcal{H}^{\text{err,large}}_{\text{dis}}(\boldsymbol{\Lambda})$. The second inequality follows from the proof of [8, Lemma A.1]; as above we are using that $f = f_{\text{ord}}$ when $\beta > \beta_c$. The last inequality follows from [8, (A.12)] and the fact that $e^{\epsilon_n|\boldsymbol{\Lambda}|} \leq 2$ for large enough n.

We can now conclude and prove (64): putting the bounds together and using (63) we get

$$\frac{Z^{\mathrm{err}}_{\mathrm{dis}}(\mathbf{\Lambda})}{Z^{\mathrm{flip}}_{\mathrm{dis}}(\mathbf{\Lambda}, M)} \leq 4q^{-1}e^{(\kappa+3)\|\gamma\|-\frac{a_{\mathrm{dis}}}{2}M} \leq \epsilon/2.$$

■

We conclude this section with an enumerative lemma concerning $\mathcal{H}^{\mathrm{flip}}_{\mathrm{dis}}$.

Proposition 4.7. *There is an algorithm that given $\boldsymbol{\gamma} \in \mathcal{C}_{\mathrm{ord}}$ and $M \in \mathbb{N}$ outputs $\mathcal{H}^{\mathrm{flip}}_{\mathrm{dis}}(\mathrm{Int}\ \boldsymbol{\gamma}, M)$ in time $\|\boldsymbol{\gamma}\|e^{O(\|\boldsymbol{\gamma}\|+M)}$.*

Proof. This follows from a variation on the proof of Proposition 3.12. To determine $\mathcal{H}^{\mathrm{flip}}_{\mathrm{dis}}(\mathrm{Int}\ \boldsymbol{\gamma})$ we will consider $\boldsymbol{\gamma}$ to be a contour in a torus of side-length $\|\boldsymbol{\gamma}\| \wedge n$; this torus has volume polynomial in $\|\boldsymbol{\gamma}\|$.

$\mathcal{H}^{\mathrm{flip}}_{\mathrm{dis}}(\mathrm{Int}\ \boldsymbol{\gamma})$ is the set of mutually external contour configurations $\Gamma \setminus \boldsymbol{\gamma}$ obtained as F ranges over the possibilities listed in Lemma 3.7. As in Lemma 3.8 we can determine $E' \cup F$ by considering it as the complement of 1-connected set of edges $A = A' \sqcup B$, where A' is the set of edges that intersect $\boldsymbol{\gamma}$. For any choice of such an A, $\mathrm{Ext}\ \Gamma \cap \mathbb{T}^d_n$ is of size at least $O(|\{B\}|)$, so to determine $\mathcal{H}^{\mathrm{flip}}_{\mathrm{dis}}(\mathrm{Int}\ \boldsymbol{\gamma}, M)$ it is enough to consider all possible sets B of size at most M. The claim now follows by arguing as in the proof of Proposition 3.12. ■

5 | APPROXIMATE COUNTING ALGORITHMS

This section describes our approximate counting algorithms for $\beta > \beta_h$. The algorithms differ depending on whether $\beta = \beta_c$, $\beta > \beta_c$, or $\beta_h < \beta < \beta_c$. Recall that for $\ell \in \{\mathrm{dis}, \mathrm{ord}\}$, $Z_\ell(\mathbf{\Lambda})$ was defined for all regions $\mathbf{\Lambda}$ in (42)–(43). The heart of this section is the following lemma.

Lemma 5.1. *For $d \geq 2$ and $q \geq q_0$ the following hold.*

1. *If $\beta = \beta_c$ there is an FPTAS to approximate $Z_{\mathrm{ord}}(\mathbf{\Lambda})$ and $Z_{\mathrm{dis}}(\mathbf{\Lambda})$.*
2. *If $\beta > \beta_c$ there is an FPTAS to approximate $Z_{\mathrm{ord}}(\mathbf{\Lambda})$.*
3. *If $\beta_h < \beta < \beta_c$ there is an FPTAS to approximate $Z_{\mathrm{dis}}(\mathbf{\Lambda})$.*

In each case the FPTAS applies to any region $\mathbf{\Lambda}$, with running time polynomial in $|\mathbf{\Lambda}|$, the number of vertices of $\mathbb{T}^d_n$ in $\mathbf{\Lambda}$.

Sections 5.1 and 5.2 prove the first two cases of Lemma 5.1. The case $\beta_h < \beta < \beta_c$ is very similar to $\beta > \beta_c$, and we defer the details to Appendix B. In Section 5.3, we show how these results, together with a result from [8], suffice to give an FPRAS for Z on the torus.

5.1 | Proof of Lemma 5.1 when $\beta = \beta_c$

We begin by defining a useful variant of the truncated cluster expansion for $Z_{\mathrm{ord}}(\mathbf{\Lambda})$ and $Z_{\mathrm{dis}}(\mathbf{\Lambda})$. Let K be a function from contours to positive real numbers. For $\ell \in \{\mathrm{ord}, \mathrm{dis}\}$ define

$$T_{\ell,m}(\mathbf{\Lambda}, K) := \sum_{\substack{\Gamma \in \mathcal{C}^c_\ell(\mathbf{\Lambda}) \\ \|\Gamma\| < m}} \phi(\Gamma) \prod_{\boldsymbol{\gamma} \in \Gamma} K(\boldsymbol{\gamma}),$$

so that by (50) and (51) $Z_\ell(\Lambda) = \exp(-e_\ell|\Lambda|)T_{\ell,\infty}(\mathbf{\Lambda}, K_\ell)$ provided the cluster expansion for the polymer models converge.

Recall that the level of a contour was defined in Definition 6, and that $|\mathbf{\Lambda}|_{(\frac{1}{2}\mathbb{T}_n^d)^\star}$ was defined immediately prior to Lemma 3.13.

Lemma 5.2. *Suppose $d \geq 2$, $q \geq q_0$ and $\beta = \beta_c$. Given $\mathbf{\Lambda}$ with $|\mathbf{\Lambda}|_{(\frac{1}{2}\mathbb{T}_n^d)^\star} = N$, and an error parameter $\epsilon > 0$, let $m = \log(8N^2/\epsilon)/3$. Inductively (by level) define weights $\tilde{K}_{\text{ord}}(\boldsymbol{\gamma})$ and $\tilde{K}_{\text{dis}}(\boldsymbol{\gamma})$ for all contours $\boldsymbol{\gamma}$ in $C_{\text{ord}}(\mathbf{\Lambda})$ and $C_{\text{dis}}(\mathbf{\Lambda})$ with size $\|\boldsymbol{\gamma}\| \leq m$ by:*

1. *If $\boldsymbol{\gamma}$ is thin, then set*

$$\tilde{K}_{\text{ord}}(\boldsymbol{\gamma}) = e^{-\kappa\|\boldsymbol{\gamma}\|-(e_{\text{dis}}-e_{\text{ord}})|\text{Int }\boldsymbol{\gamma}|}, \quad \tilde{K}_{\text{dis}}(\boldsymbol{\gamma}) = qe^{-\kappa\|\boldsymbol{\gamma}\|-(e_{\text{ord}}-e_{\text{dis}})|\text{Int }\boldsymbol{\gamma}|}.$$

2. *If $\boldsymbol{\gamma}$ is not thin, then set*

$$\begin{aligned}\tilde{K}_{\text{ord}}(\boldsymbol{\gamma}) &= e^{-\kappa\|\boldsymbol{\gamma}\|-(e_{\text{dis}}-e_{\text{ord}})|\text{Int }\boldsymbol{\gamma}|}\exp\left[T_{m,\text{dis}}(\text{Int }\boldsymbol{\gamma},\tilde{K}) - T_{m,\text{ord}}(\text{Int }\boldsymbol{\gamma},\tilde{K})\right],\\ \tilde{K}_{\text{dis}}(\boldsymbol{\gamma}) &= qe^{-\kappa\|\boldsymbol{\gamma}\|-(e_{\text{ord}}-e_{\text{dis}})|\text{Int }\boldsymbol{\gamma}|}\exp\left[T_{m,\text{ord}}(\text{Int }\boldsymbol{\gamma},\tilde{K}) - T_{m,\text{dis}}(\text{Int }\boldsymbol{\gamma},\tilde{K})\right].\end{aligned}$$

Then for N sufficiently large $e^{-e_\ell|\Lambda|}\exp(T_{\ell,m}(\mathbf{\Lambda},\tilde{K}_\ell))$ is an ϵ-relative approximation to $Z_\ell(\mathbf{\Lambda})$ for $\ell \in \{\text{ord},\text{dis}\}$.

Proof. Suppose $\ell \in \{\text{dis},\text{ord}\}$. First note that the inductive definition of the weights $\tilde{K}_\ell(\boldsymbol{\gamma})$ makes sense: to compute $\tilde{K}_\ell(\boldsymbol{\gamma})$ for a contour $\boldsymbol{\gamma}$ of level $t+1$ only requires knowing $\tilde{K}_\ell(\boldsymbol{\gamma}')$ for contours $\boldsymbol{\gamma}'$ of level t and smaller.

Since $\beta = \beta_c$ and $q \geq q_0$, Lemma 4.4 tells us that

$$K_\ell(\boldsymbol{\gamma}) \leq e^{-c\beta\|\boldsymbol{\gamma}\|} \tag{65}$$

for $\ell \in \{\text{dis},\text{ord}\}$ and for all $\boldsymbol{\gamma} \in C_\ell(\Lambda)$. If q_0 is large enough then (65) implies condition (15) holds since β_c grows like $\log q$ by (3). Thus by Section 3.7 the hypotheses of Lemma 2.1 are satisfied and the cluster expansion for $Z_\ell(\mathbf{\Lambda})$ converges for $\ell \in \{\text{ord},\text{dis}\}$.

Now let $\epsilon' = \epsilon/N$, so that $m = \log(8N/\epsilon')/3$. We will apply Lemma 2.3 with $v(\boldsymbol{\gamma}) = |\text{Int }\boldsymbol{\gamma}|$. This is a valid choice of $v(\boldsymbol{\gamma})$ by Lemma 3.9. Lemma 2.3 says that

$$e^{-e_{\text{ord}}|\Lambda|}\exp\left(T_{\text{ord},m}(\mathbf{\Lambda},\tilde{K}_{\text{ord}})\right) \quad \text{and} \quad e^{-e_{\text{dis}}|\Lambda|}\exp\left(T_{\text{dis},m}(\mathbf{\Lambda},\tilde{K}_{\text{dis}})\right),$$

are ϵ-relative approximations to $Z_{\text{ord}}(\mathbf{\Lambda})$ and $Z_{\text{dis}}(\mathbf{\Lambda})$ if for all $\boldsymbol{\gamma} \in C_\ell(\mathbf{\Lambda})$ of size at most m, $\tilde{K}_\ell(\boldsymbol{\gamma})$ is an $\epsilon'|\text{Int }\boldsymbol{\gamma}|$-relative approximation to $K_\ell(\boldsymbol{\gamma})$. We will prove this by induction on the level of $\boldsymbol{\gamma}$.

For a thin contour, $\tilde{K}_\ell(\boldsymbol{\gamma}) = K_\ell(\boldsymbol{\gamma})$. Now suppose that for all contours $\boldsymbol{\gamma}$ of level at most t and size at most m, $\tilde{K}_\ell(\boldsymbol{\gamma})$ is an $\epsilon'|\text{Int }\boldsymbol{\gamma}|$-relative approximation of $K_\ell(\boldsymbol{\gamma})$. Consider a contour $\boldsymbol{\gamma}$ of level $t+1$ and size at most m. Then all contours $\boldsymbol{\gamma}'$ that appear in the expansions

$$T_{m,\text{dis}}(\text{Int }\boldsymbol{\gamma},\tilde{K}_{\text{dis}}) \quad \text{and} \quad T_{m,\text{ord}}(\text{Int }\boldsymbol{\gamma},\tilde{K}_{\text{ord}}),$$

are of level at most t and size at most m, and so for each such $\boldsymbol{\gamma}'$, by the inductive hypothesis $\tilde{K}_\ell(\boldsymbol{\gamma}')$ is an $\epsilon'|\text{Int }\boldsymbol{\gamma}'|$-relative approximation to $K_\ell(\boldsymbol{\gamma}')$. Then by Lemma 2.3, we have that

$$e^{-(e_{\text{dis}}-e_{\text{ord}})|\text{Int }\boldsymbol{\gamma}|}\exp\left[T_{m,\text{dis}}(\text{Int }\boldsymbol{\gamma},\tilde{K}_{\text{dis}}) - T_{m,\text{ord}}(\text{Int }\boldsymbol{\gamma},\tilde{K}_{\text{ord}})\right]$$

is an $|\text{Int }\boldsymbol{\gamma}|\epsilon'$-relative approximation to $\frac{Z_{\text{dis}}(\text{Int }\boldsymbol{\gamma})}{Z_{\text{ord}}(\text{Int }\boldsymbol{\gamma})}$ (and likewise for dis and ord swapped). Multiplying by the prefactor $e^{-\kappa\|\boldsymbol{\gamma}\|}$ for ord and by $qe^{-\kappa\|\boldsymbol{\gamma}\|}$ for dis shows that $\tilde{K}_\ell(\boldsymbol{\gamma})$ is an $\epsilon'|\text{Int }\boldsymbol{\gamma}|$-relative approximation to $K_\ell(\boldsymbol{\gamma})$ as desired. ∎

With this, we can prove the $\beta = \beta_c$ case of Lemma 5.1.

Proof of Lemma 5.1 when $\beta = \beta_c$. Let $N = |\boldsymbol{\Lambda}|_{(\frac{1}{2}\mathbb{T}_n^d)^\star}$ and let $m = \log(8N^2/\epsilon)/3$. We need to show that the expansion $T_{\ell,m}(\boldsymbol{\Lambda}, \tilde{K}_\ell)$ and the weights $\tilde{K}_\ell(\boldsymbol{\gamma})$ for all $\boldsymbol{\gamma}$ of size at most m in $\mathcal{C}_\ell(\boldsymbol{\Lambda})$ can be computed in time polynomial in N and $1/\epsilon$ for $\ell \in \{\text{dis}, \text{ord}\}$. We can list the sets of contours in $\mathcal{C}_{\text{ord}}(\boldsymbol{\Lambda})$ and $\mathcal{C}_{\text{dis}}(\boldsymbol{\Lambda})$ of size at most m, together with their labels and levels, in time $O(N\exp(O(m))$ by Proposition 3.12. Since $m = \log(8N^2/\epsilon)/3$, $O(N\exp(O(m))$ is polynomial in N and $1/\epsilon$. The number N itself is polynomial in $|\Lambda|$ by Lemma 3.13.

To prove the lemma we must compute the weights $\tilde{K}_\ell(\boldsymbol{\gamma})$ and the truncated cluster expansions $T_{m,\ell}(\text{Int }\boldsymbol{\gamma}, \tilde{K}_\ell)$ for each contour in the list. We do this inductively by level. For level zero contours $\tilde{K}_\ell(\boldsymbol{\gamma}) = K_\ell(\boldsymbol{\gamma})$ only depends on $\|\boldsymbol{\gamma}\|$ and $|\{\text{Int }\boldsymbol{\gamma}\}|$, so $\tilde{K}_\ell(\boldsymbol{\gamma})$ can be computed in time $O(\|\boldsymbol{\gamma}\|^3)$ by computing these quantities by using Lemma 3.10. We then continue inductively; each $\tilde{K}_\ell(\boldsymbol{\gamma})$ can be computed efficiently since the truncated cluster expansions can be computed in time polynomial in N and $1/\epsilon$ using Lemma 2.2. ∎

5.2 | Proof of Lemma 5.1 when $\beta > \beta_c$

When $\beta > \beta_c(q, d)$ the ordered ground state is stable, but the disordered state is unstable. For a definition of stability of ground states, see, for example, [9]; the upshot for this article is that we cannot use the cluster expansion to approximate $Z_{\text{dis}}(\boldsymbol{\Lambda})$ for a region $\boldsymbol{\Lambda}$.

To deal with this complication we will appeal to Lemma 4.6. In words, this lemma says that for $\beta > \beta_c$, a typical contour configuration in a region with disordered boundary conditions will have very few external vertices. We will exploit this fact to enumerate all sets of typical external contours in the region. This is possible since the number of external vertices is small. Once we have fixed a set of external contours we are back to the task of approximating partition functions with ordered boundary conditions.

We now make the preceding discussion precise. Given $K : \mathcal{C}_{\text{ord}}(\boldsymbol{\Lambda}) \to [0, \infty)$ and $M > 0$, define

$$\Xi^M_{\text{dis}}(\boldsymbol{\Lambda}, K) := e^{e_{\text{dis}}|\Lambda|} \sum_{\Gamma \in \mathcal{H}^{\text{flip}}_{\text{dis}}(\boldsymbol{\Lambda}, M)} e^{-e_{\text{dis}}|\text{Ext }\Gamma|} \prod_{\boldsymbol{\gamma}\in\Gamma} e^{-\kappa\|\boldsymbol{\gamma}\|} q \exp\left[T_{m,\text{ord}}(\text{Int }\boldsymbol{\gamma}, K)\right].$$

Lemma 5.3. *Suppose $d \geq 2$, $q \geq q_0$ and $\beta > \beta_c$. Let $\boldsymbol{\Lambda}$ be a region with $|\boldsymbol{\Lambda}|_{(\frac{1}{2}\mathbb{T}_n^d)^\star} = N$, fix $\epsilon > 0$, and let $m = \log(8N^2/\epsilon)/3$. Inductively (by level) define $\tilde{K}_{\text{ord}}(\boldsymbol{\gamma})$ for $\boldsymbol{\gamma} \in \mathcal{C}_{\text{ord}}(\boldsymbol{\Lambda})$ with size $\|\boldsymbol{\gamma}\|$ at most m by*

1. *If $\boldsymbol{\gamma}$ is thin, then*

$$\tilde{K}_{\text{ord}}(\boldsymbol{\gamma}) = e^{-\kappa\|\boldsymbol{\gamma}\| - (e_{\text{dis}} - e_{\text{ord}})|\text{Int }\boldsymbol{\gamma}|}.$$

2. *If $\boldsymbol{\gamma}$ is not thin, define*

$$\tilde{K}_{\text{ord}}(\boldsymbol{\gamma}) = e^{-\kappa\|\boldsymbol{\gamma}\| - (e_{\text{dis}} - e_{\text{ord}})|\text{Int }\boldsymbol{\gamma}|} \exp\left[-T_{m,\text{ord}}(\text{Int }\boldsymbol{\gamma}, \tilde{K})\right] \Xi^M_{\text{dis}}(\text{Int }\boldsymbol{\gamma}, \tilde{K}_{\text{ord}}),$$

with $M = \frac{2}{a_{\text{dis}}}\left(\log(\frac{32q}{\epsilon'}) + (\kappa + 3)m\right)$.

Then for all N large enough, $e^{-e_{\text{ord}}|\Lambda|}\exp\big(T_{\text{ord},m}(\Lambda,\tilde{K}_{\text{ord}})\big)$ is an ϵ-relative approximation to $Z_{\text{ord}}(\Lambda)$.

Proof. Let $\epsilon' = \epsilon/N$ so that $m = \log(8N/\epsilon')/3$.

If q_0 is large enough then we have $K_{\text{ord}}(\gamma) \le e^{-c\beta\|\gamma\|}$ by Lemma 4.5 since $\beta > \beta_c$. This along with (3) implies condition (15) holds for ordered contours, and thus by Section 3.7 the hypotheses of Lemma 2.1 are satisfied and the cluster expansion for $Z_{\text{ord}}(\Lambda)$ converges. Applying Lemma 2.3 with $v(\gamma) = |\text{Int }\gamma|$ then tells us that

$$e^{-e_{\text{ord}}|\Lambda|}\exp\big(T_{\text{ord},m}(\Lambda,\tilde{K}_{\text{ord}})\big)$$

is an ϵ-relative approximation to $Z_{\text{ord}}(\Lambda)$ if for all $\gamma \in C_{\text{ord}}(\Lambda)$ of size at most m, $\tilde{K}_{\text{ord}}(\gamma)$ is an $\epsilon'|\text{Int }\gamma|$-relative approximation to $K_{\text{ord}}(\gamma)$. We will prove this is the case by induction. The base case of the induction (thin contours) holds since $\tilde{K}_{\text{ord}}(\gamma) = K_{\text{ord}}(\gamma)$. Now suppose that the statement holds for all contours of level at most t and size at most m, and consider a contour γ of level $t+1$ and size at most m.

The inductive hypothesis and Lemma 2.3 imply that

$$e^{-e_{\text{ord}}|\Lambda|}\exp\left[T_{m,\text{ord}}(\text{Int }\gamma,\tilde{K})\right]$$

is an $\epsilon'|\text{Int }\gamma|/2$-relative approximation to $Z_{\text{ord}}(\text{Int }\gamma)$, and so it suffices to show that $e^{-e_{dis}|\Lambda|}\Xi^M_{\text{dis}}(\text{Int }\gamma,\tilde{K}_{\text{ord}})$ is an $\epsilon'|\text{Int }\gamma|/2$-relative approximation to $Z_{\text{dis}}(\text{Int }\gamma)$.

By Lemma 4.6, $Z^{\text{flip}}_{\text{dis}}(\text{Int }\gamma,M)$ is an $\epsilon'/4$-relative approximation to $Z_{\text{dis}}(\text{Int }\gamma)$ for $M = \frac{2}{a_{\text{dis}}}\left(\log(\frac{32q}{\epsilon'}) + (\kappa+3)m\right)$, and so it suffices to show that $e^{-e_{dis}|\Lambda|}\Xi^M_{\text{dis}}(\text{Int }\gamma,\tilde{K}_{\text{ord}})$ is an $\epsilon'|\text{Int }\gamma|/4$-relative approximation to $Z^{\text{flip}}_{\text{dis}}(\text{Int }\gamma,M)$. We will accomplish this by showing, for each $\Gamma \in \mathcal{H}^{\text{flip}}(\text{Int }\gamma,M)$, that

$$e^{-e_{\text{dis}}|\text{Ext }\Gamma|}\prod_{\gamma'\in\Gamma}e^{-\kappa\|\gamma'\|}q\exp\left[T_{m,\text{ord}}(\text{Int }\gamma',\tilde{K})\right]$$

is an $\epsilon'|\text{Int }\gamma|/4$-relative approximation to

$$e^{-e_{\text{dis}}|\text{Ext }\Gamma|}\prod_{\gamma'\in\Gamma}e^{-\kappa\|\gamma'\|}qZ_{\text{ord}}(\text{Int }\gamma'),$$

and then summing over Γ. The prefactors are identical, and so it comes down to comparing $\prod_{\gamma'\in\Gamma}\exp\left[T_{m,\text{ord}}(\text{Int }\gamma',\tilde{K})\right]$ to $\prod_{\gamma'\in\Gamma}Z_{\text{ord}}(\text{Int }\gamma')$. Since the contours in Γ are mutually external,

$$\sum_{\gamma'\in\Gamma}|\text{Int }\gamma'| \le |\text{Int }\gamma|,$$

and hence it suffices to show that for each γ', $\exp\left[T_{m,\text{ord}}(\text{Int }\gamma',\tilde{K})\right]$ is an $\epsilon'|\text{Int }\gamma'|/4$-relative approximation to $Z_{\text{ord}}(\text{Int }\gamma')$. This follows from Lemma 2.3 since $m = \log(8N/\epsilon')/3$ and by induction we have that $\tilde{K}_{\text{ord}}(\gamma'')$ is an $\epsilon'|\text{Int }\gamma''|$-relative approximation to $K_{\text{ord}}(\gamma'')$ for all contours γ'' that contribute to $T_{m,\text{ord}}(\text{Int }\gamma',\tilde{K})$. ∎

With this, we can prove the $\beta > \beta_c$ case of Lemma 5.1.

Proof of Lemma 5.1 when $\beta > \beta_c$. Given Lemma 5.3, we need to show that we can compute $\tilde{K}_{\text{ord}}(\boldsymbol{\gamma})$ for all $\boldsymbol{\gamma}$ of size at most $m = \log(8N^2/\epsilon)/3$ in time polynomial in N and $1/\epsilon$. The proof of this is the same as the proof of the $\beta = \beta_c$ case of the lemma except that now we have to account for the computation of $\Xi^M_{\text{dis}}(\text{Int }\boldsymbol{\gamma}, \tilde{K})$ for all $\boldsymbol{\gamma} \in \mathcal{C}_{\text{dis}}(\boldsymbol{\Lambda})$ of size at most m, with $M = \frac{2}{a_{\text{dis}}}\left(\log(\frac{32q}{\epsilon'}) + (\kappa+3)m\right)$.

For a given $\Gamma \in \mathcal{H}^{\text{flip}}_{\text{dis}}(\text{Int }\boldsymbol{\gamma}, M)$, the computation of

$$e^{-e_{\text{dis}}|\text{Ext }\Gamma|}\prod_{\boldsymbol{\gamma}'\in\Gamma} e^{-\kappa\|\boldsymbol{\gamma}'\|}q\exp\left[T_{m,\text{ord}}(\text{Int }\boldsymbol{\gamma}', \tilde{K})\right] \tag{66}$$

can be done in time polynomial in N and $1/\epsilon$ since it just involves computing the truncated cluster expansions $T_{m,\text{ord}}(\text{Int }\boldsymbol{\gamma}', \tilde{K})$ for at most m^2 contours $\boldsymbol{\gamma}'$, and since we compute $\tilde{K}_{\text{ord}}(\boldsymbol{\gamma}')$ in order of the level of $\boldsymbol{\gamma}'$, we will have already computed all the weight functions needed in the expansion.

To conclude, note the set $\mathcal{H}^{\text{flip}}_{\text{dis}}(\text{Int }\boldsymbol{\gamma}, M)$ can be enumerated in polynomial time by Proposition 4.7 since both $\|\boldsymbol{\gamma}\|$ and M are $O(\log(N^2/\epsilon))$. Since N is polynomial in $|\boldsymbol{\Lambda}|$ by Lemma 3.13, the proof is complete. ∎

Note that Lemma 5.3 used the value of $a_{\text{dis}} > 0$ to determine the value of M in the definitions of the weights $\tilde{K}$. It is desirable to avoid using a_{dis} as an input of the algorithm, and hence we close this section with a lemma that shows how to bound M without knowing a_{dis} precisely.

Lemma 5.4. *Suppose* $d \geq 2$, $q \geq q_0$, *and* $\beta > \beta_c$. *There is an* $O(1)$*-time algorithm to determine a constant* $a^\star_{\text{dis}} > 0$ *such that* $a_{\text{dis}} > a^\star_{\text{dis}}$. *The constants in the* $O(1)$ *term may depend on* q, β, d.

Proof. We follow the notation from [8, Appendix A.1]. As discussed below [8, (A.7)], we have $\left|\{f_\ell - f^{(n)}_\ell\}\right| \leq \epsilon_n$ for $\ell \in \{\text{ord}, \text{dis}\}$, where $\epsilon_n = 2e^{-c\beta n}$, where n is the side-length of the torus $\mathbb{T}^d_n$, and $f_\ell = \lim_{n\to\infty} f^{(n)}_\ell$.

Compute $f^{(n)}_\ell$ for $\ell \in \{\text{ord}, \text{dis}\}$ until $\left|\{f^{(n)}_{\text{ord}} - f^{(n)}_{\text{dis}}\}\right|$ is at least $3\epsilon_n$. Let n_0 be the first such n that is found. Then by the triangle inequality, a_{dis} is at least $a^\star_{\text{dis}} = \epsilon_{n_0}$.

Note that n_0 can be bounded above in terms of the value of $a_{\text{dis}} = a_{\text{dis}}(\beta, d, q)$ and ϵ_n, so the above procedure terminates in a finite time (depending on β, d, q). ∎

5.3 | Proof of Theorem 1.1

To prove Theorem 1.1 we will need the following result from [8] about the mixing time of the Glauber dynamics.

Theorem 5.5 ([8, Theorem 1.1])**.** *The mixing time of the Glauber dynamics for the q-state ferromagnetic Potts model satisfies*

$$\tau_{q,\beta}(\mathbb{T}^d_n) = e^{O(n^{d-1})}, \tag{67}$$

where the $O(\cdot)$ *in the exponent hides constants that depend on* q, β.

We will use this result to give an approximation algorithm when the approximation parameter ϵ is extremely small. The reason we are able to combine the Glauber dynamics with our contour-based algorithm to give an FPRAS is that [8] proves *optimal* slow mixing results for the Glauber and Swendsen–Wang dynamics. That is, up to a constant in the exponent, the upper bound of the mixing

time of the Glauber dynamics (or Swendsen–Wang dynamics) is the inverse of the bound on Z_{tunnel}/Z from Lemma 4.1. Thus when ϵ is too small for the contour algorithms to work, the Glauber dynamics can take over.

Proof of Theorem 1.1. Let $N = n^d$ be the number of vertices of $\mathbb{T}_n^d$. We will use a simple fact several times below: if $\epsilon \in (0,1)$, $Z, Z^* > 0$, and $Z^*/Z < \epsilon/2$, then $(Z - Z^*)$ is an ϵ-relative approximation to Z.

We first consider the case $\beta = \beta_c$. To give an FPRAS for $Z = Z_{\mathbb{T}_n^d}$ we consider two subcases. Let c be the constant from Lemma 4.1.

Suppose $\epsilon < 4e^{-c\beta n^{d-1}}$. Since $e^{O(n^{d-1})}$ is polynomial in N and $1/\epsilon$, we can use Glauber dynamics to obtain an ϵ-approximate sample in polynomial time. By using simulated annealing (e.g., [40]) we can also approximate the partition function in time polynomial in N and $1/\epsilon$.

If $\epsilon \geq 4e^{-c\beta n^{d-1}}$, then by Lemma 4.1, $Z_{\text{rest}} = Z_{\text{dis}} + Z_{\text{ord}}$ is an $\epsilon/2$-relative approximation to Z, so it suffices to find an $\epsilon/4$-relative approximation to both Z_{dis} and Z_{ord}. This can be done in time polynomial in N and $1/\epsilon$ by Lemma 5.1.

Next we consider the case $\beta > \beta_c$. Again there are two subcases. Let c be the constant from Lemma 4.1 as before, and let b_{dis} be the constant from Lemma 4.3. If $\epsilon < 4e^{-c\beta n^{d-1}} + 4e^{-b_{\text{dis}} n^{d-1}}$, then again $e^{O(n^{d-1})}$ is polynomial in N and $1/\epsilon$ and we can approximately count and sample by using the Glauber dynamics.

If $\epsilon \geq 4e^{-c\beta n^{d-1}} + 4e^{-b_{\text{dis}} n^{d-1}}$, then by Lemmas 4.1 and 4.3, Z_{ord} is an $\epsilon/2$-relative approximation to Z and so it suffices to give an $\epsilon/2$-relative approximation to Z_{ord}. This can be done in time polynomial in N and $1/\epsilon$ by Lemma 5.1.

Lastly, consider $\beta < \beta_c$. The case $\beta \leq \beta_h$ was completed in Section 2. The case $\beta_h < \beta < \beta_c$ is done exactly as the case $\beta > \beta_c$ with the roles of ord and dis reversed; see Appendix B for details. ∎

Proof of Theorem 1.2 for counting. Let $\Lambda \subset \mathbb{Z}^d$ be such that the induced subgraph G_Λ is finite and simply connected. By Proposition 3.14, we can construct an ordered contour $\boldsymbol{\gamma}_{\text{ord}}$ and a disordered contour $\boldsymbol{\gamma}_{\text{dis}}$ so that

$$Z_{\text{dis}}(\text{Int } \boldsymbol{\gamma}_{\text{ord}}) = (1-p)^{-\frac{1}{2}\|\boldsymbol{\gamma}_{\text{ord}}\|} Z_\Lambda^f, \qquad Z_{\text{ord}}(\text{Int } \boldsymbol{\gamma}_{\text{dis}}) = q^{-1} p^{d\left|\{\text{Int } \boldsymbol{\gamma}_{\text{ord}}\}\right| - |E(\Lambda)|} Z_\Lambda^w .$$

The FPTAS for Z_Λ^w for $\beta \geq \beta_c$ then follows from Lemma 5.1, as does the FPTAS for Z_Λ^f for $\beta_h < \beta \leq \beta_c$. The case $\beta \leq \beta_h$ was covered in Section 2. ∎

6 | SAMPLING

In this section, we present efficient approximate sampling algorithms for the random cluster and Potts models when $\beta > \beta_h$. By the Edwards–Sokal coupling, see Appendix A, it suffices to obtain algorithms for the random cluster model. Describing the strategy, which is based on that of [25, Sections 5 and 6], requires a few definitions.

Recall the definition (4) of the random cluster measure μ^{RC} on $\mathbb{T}_n^d$. Thus μ^{RC} is a measure on subsets of edges $A \in \Omega$. Recalling the definitions (44) and (45) of the sets Ω_{ord} and Ω_{dis} of ordered and disordered edge configurations, we analogously define

$$\mu_\ell(A) := \frac{w(A)}{Z_\ell}, \quad A \in \Omega_\ell \text{ with } \ell \in \{\text{ord}, \text{dis}\}.$$

For a region $\mathbf{\Lambda}$, define measures $\nu_\ell^{\mathbf{\Lambda}}$ on the sets of external contours $\mathcal{C}_\ell^{\text{ext}}(\mathbf{\Lambda})$ as follows.

$$\nu_{\text{dis}}^{\mathbf{\Lambda}}(\Gamma) := \frac{e^{-e_{\text{dis}}|\Lambda \cap \text{Ext } \Gamma|} \prod_{\gamma \in \Gamma} e^{-\kappa\|\gamma\|} q Z_{\text{ord}}(\text{Int } \boldsymbol{\gamma})}{Z_{\text{dis}}(\mathbf{\Lambda})}, \qquad \Gamma \in \mathcal{C}_{\text{dis}}^{\text{ext}}(\mathbf{\Lambda}), \tag{68}$$

$$\nu_{\text{ord}}^{\mathbf{\Lambda}}(\Gamma) := \frac{e^{-e_{\text{ord}}|\Lambda \cap \text{Ext } \Gamma|} \prod_{\gamma \in \Gamma} e^{-\kappa\|\gamma\|} Z_{\text{dis}}(\text{Int } \boldsymbol{\gamma})}{Z_{\text{ord}}(\mathbf{\Lambda})}, \qquad \Gamma \in \mathcal{C}_{\text{ord}}^{\text{ext}}(\mathbf{\Lambda}), \tag{69}$$

where $|\Lambda \cap \text{Ext } \Gamma|$ is the number of vertices contained in the continuum set $\mathbf{\Lambda} \cap \text{Ext } \Gamma$.

We now outline our strategy for approximately sampling from μ_{ord} and μ_{dis}; a small modification will also apply to sampling from μ^{RC} on the torus. The key idea is that the inductive representations of the partition functions in (42) and (43) yield a procedure to sample from μ_{dis} and μ_{ord} if we can sample from the measures $\nu_\ell^{\mathbf{\Lambda}}$ for $\ell \in \{\text{ord}, \text{dis}\}$ and for all regions $\mathbf{\Lambda}$. The procedure, which we call the *inductive contour sampling algorithm*, is as follows. Consider μ_{ord}. To sample a set of compatible, matching contours with ordered external contours, we first sample Γ from $\nu_{\text{ord}}^{\boldsymbol{T}_n^d}$, then for each $\boldsymbol{\gamma} \in \Gamma$ we sample from $\nu_{\text{dis}}^{\text{Int } \boldsymbol{\gamma}}$ and repeat inductively until there are no interiors left to sample from. The union of all contours sampled is a set of matching and compatible contours, and these contours are distributed as the restriction of (35) to contour configurations that arise from ordered edge configurations. This set of contours can then be mapped to an edge set via the bijection of Lemma 3.2, and the distribution of this edge set is μ_{ord}. The procedure for sampling from μ_{dis} is analogous. For a more detailed discussion of the validity of this algorithm, see [25, Section 5].

By using the same procedure it is possible to efficiently approximately sample from μ_{ord} and μ_{dis} provided one can efficiently approximately sample from the external contour measures $\nu_{\text{dis}}^{\mathbf{\Lambda}}$ and $\nu_{\text{ord}}^{\mathbf{\Lambda}}$. Again, we refer to [25, Section 5] for further details.

The next lemma is an essential input for developing efficient approximate samplers for $\nu_\ell^{\mathbf{\Lambda}}$ as it tells us we need only consider "small" contours. For $\ell \in \{\text{ord}, \text{dis}\}$ let $\nu_\ell^{\mathbf{\Lambda},m}$ be the probability measure defined as in (68) and (69), but restricted to Γ with $\|\Gamma\| < m$. The normalization factor for $\nu_\ell^{\mathbf{\Lambda},m}$ is thus the contour partition function restricted to Γ with $\|\Gamma\| < m$.

Lemma 6.1. *Suppose $d \geq 2$, $q \geq q_0$, and $\epsilon > 0$. Then, letting $N = |\{\mathbf{\Lambda}\}|_{(\frac{1}{2}\mathbb{T}_n^d)^\star}$, for $m \geq C' \log(N/\epsilon)$ with C' a large enough absolute constant,*

1. *If $\beta \geq \beta_c$, then $\|\nu_{\text{ord}}^{\mathbf{\Lambda},m} - \nu_{\text{ord}}^{\mathbf{\Lambda}}\|_{TV} < \epsilon$.*
2. *If $\beta_h < \beta \leq \beta_c$, then $\|\nu_{\text{dis}}^{\mathbf{\Lambda},m} - \nu_{\text{dis}}^{\mathbf{\Lambda}}\|_{TV} < \epsilon$.*

for all regions $\mathbf{\Lambda}$.[4]

Proof. This follows from the convergence of the cluster expansion for $Z_\ell(\mathbf{\Lambda})$ for the specified choices of ℓ and β. For details see, for example, [Proof of Lemma 13] 4 [25]. ∎

Lemma 6.2. *Suppose $d \geq 2$ and $q \geq q_0$. Then*

1. *For $\beta = \beta_c$, there are efficient sampling schemes for $\nu_{\text{ord}}^{\mathbf{\Lambda}}$ and $\nu_{\text{dis}}^{\mathbf{\Lambda}}$.*
2. *For $\beta > \beta_c$ there is an efficient sampling scheme for $\nu_{\text{ord}}^{\mathbf{\Lambda}}$.*
3. *For $\beta_h < \beta < \beta_c$ there is an efficient sampling scheme for $\nu_{\text{dis}}^{\mathbf{\Lambda}}$.*

In each case these algorithms apply for all regions $\mathbf{\Lambda}$.

[4]The constant C' depends only on the constants c in the bounds on $K_\ell(\boldsymbol{\gamma}) \leq \exp(-c\beta\|\boldsymbol{\gamma}\|)$. These bounds are given by Lemmas 4.4 and 4.5 for $\beta \geq \beta_c$, and in Appendix B for $\beta_h < \beta < \beta_c$.

Proof. First we consider $\beta = \beta_c$. By Lemma 5.1 there are efficient algorithms to approximate $Z_{\text{dis}}(\mathbf{\Lambda})$ and $Z_{\text{ord}}(\mathbf{\Lambda})$ for all regions $\mathbf{\Lambda}$. With this, we can apply the approximate sampling algorithms given in [25, Theorems 11 and 13]. We summarize the algorithm here, assuming that we want to sample a collection of ordered contours (the disordered case is identical).

By Lemma 6.1 it is enough to obtain an ϵ-approximate sample from $\nu_{\ell}^{\mathbf{\Lambda},m}$ with $m = O(\log(N/\epsilon))$. List all contours of size at most m in $\mathcal{C}_{\text{ord}}(\mathbf{\Lambda})$, and call this collection $\mathcal{C}$. Order the vertices of $\mathbf{\Lambda}$ arbitrarily as $v_1, \ldots, v_N$. We will form a random collection $\Gamma = \Gamma_N$ of mutually external ordered contours step by step. Begin with $\Gamma_0 = \emptyset$. At step i, let $\mathcal{C}_i$ be the subset of contours $\boldsymbol{\gamma}$ in $\mathcal{C}$ such that (i) $v_i \in \text{Int } \boldsymbol{\gamma}$ (ii) $\boldsymbol{\gamma}$ is external to Γ_{i-1}, and (iii) $\text{Int } \boldsymbol{\gamma} \cap \{v_1, \ldots, v_{i-1}\} = \emptyset$. We can efficiently approximate the conditional probability of each contour in $\mathcal{C}_i$, or of adding no contour at step i, by using Lemma 5.1 to approximate the relevant polymer partition functions. The result of this procedure is the desired approximate sampling algorithm.

Sampling from $\nu_{\text{ord}}^{\mathbf{\Lambda}}$ for $\beta > \beta_c$ also follows from the algorithm described above since we have an FPTAS for computing $Z_{\text{ord}}(\mathbf{\Lambda})$, and similarly for $\nu_{\text{dis}}^{\mathbf{\Lambda}}$ when $\beta_h < \beta < \beta_c$. ∎

Our strategy for efficiently approximately sampling from μ_{ord} and μ_{dis} requires that we can also efficiently approximately sample from $\nu_{\text{dis}}^{\mathbf{\Lambda}}$ for small regions $\mathbf{\Lambda}$ when $\beta > \beta_c$ (and likewise from $\nu_{\text{ord}}^{\mathbf{\Lambda}}$ when $\beta < \beta_c$). We cannot use the cluster expansion for this task since the disordered (resp. ordered) ground state is unstable, and so instead our approach is based on the intuition from Lemma 4.6 that a disordered region will quickly "flip" to being ordered when $\beta > \beta_c$.

Lemma 6.3. *Suppose $d \geq 2$ and $q \geq q_0$. Then*

1. *For $\beta > \beta_c$ there is an ϵ-approximate sampling algorithm for $\nu_{\text{dis}}^{\mathbf{\Lambda}}$ that runs in time polynomial in $1/\epsilon$ and exponential in $\|\partial\mathbf{\Lambda}\|$.*
2. *For $\beta_h < \beta < \beta_c$ there is an ϵ-approximate sampling algorithm for $\nu_{\text{ord}}^{\mathbf{\Lambda}}$ that runs in time polynomial in $1/\epsilon$ and exponential in $\|\partial\mathbf{\Lambda}\|$.*

In each case these algorithms apply for all regions $\mathbf{\Lambda}$.

In our sampling algorithms we can allow exponential dependence on $\|\partial\mathbf{\Lambda}\|$ since by Lemma 6.1 we need only consider contours γ with $\|\gamma\| = O(\log(N/\epsilon))$.

Proof of Lemma 6.3. Consider the case $\beta > \beta_c$ and suppose $\mathbf{\Lambda} = \text{Int } \boldsymbol{\gamma}$. The lemma follows from Proposition 4.7 and Lemma 4.6. More precisely, set M according to Lemma 4.6, and then compute $\mathcal{H}_{\text{dis}}^{\text{flip}}(\text{Int } \boldsymbol{\gamma}, M)$ by Proposition 4.7. As in the proof of Lemma 5.1, compute accurate approximations to the weight of each summand in $Z_{\text{dis}}^{\text{flip}}(\text{Int } \boldsymbol{\gamma}, M)$. These approximations determine the probabilities according to which we sample $\Gamma \in \mathcal{H}_{\text{dis}}^{\text{flip}}(\text{Int } \boldsymbol{\gamma}, M)$. By Lemma 4.6 the result is an ϵ-approximation to $\nu_{\text{ord}}^{\text{Int } \boldsymbol{\gamma}}$.

For $\beta_h < \beta < \beta_c$ the proof is essentially the same given the inputs discussed in Appendix B. ∎

Proof of Theorems 1.1 and 1.2, sampling. We first consider the sampling part of Theorem 1.2, which follows similarly to the proof of the approximate counting algorithm given in the previous section. Given (i) $\Lambda \subset \mathbb{Z}^d$ such that G_Λ is simply connected and (ii) a choice of wired or free boundary conditions, Proposition 3.14 gives a contour $\boldsymbol{\gamma}$ such that the partition function associated to Int $\boldsymbol{\gamma}$ is Z_Λ^w or Z_Λ^f up to an efficiently computable prefactor. Thus if $\beta = \beta_c$ we can use Lemma 6.2 to implement the inductive contour algorithm, but using ϵ'-approximations to $\nu_{\text{ord}}^{\mathbf{\Lambda}}$ and $\nu_{\text{dis}}^{\mathbf{\Lambda}}$ in place of the true measures. If $\epsilon' = \epsilon^2/(9N^2)$ where $N = |\{\mathbf{\Lambda}\}|_{(\frac{1}{2}\mathbb{T}_n^d)^\star}$, the result is an ϵ-approximate sample by [25, Lemma 12]. Here we are using N as a crude bound for the depth of the inductive contour algorithm.

If $\beta > \beta_c$, then Lemma 6.1 tells us that it suffices to sample from $\nu_{\text{ord}}^{\Lambda,m}$ with $m = O(\log(N/\epsilon))$. The consequence of this fact is that we can use the algorithm described above for $\beta = \beta_c$, as each call for an ϵ-approximate sample of $\nu_{\text{dis}}^{\Lambda}$ takes time $\exp(O(\log N/\epsilon))$ by Lemma 6.3 since each contour is of size at most $O(\log(N/\epsilon))$. For $\beta_h < \beta < \beta_c$ an analogous argument applies with the roles of ord and dis reversed.

For Theorem 1.1 the situation is similar to what we have just discussed, except for the fact that μ^{RC} is not an ordered or a disordered measure: it includes configurations with ordered and disordered external contours and includes the configurations with interfaces. If $\beta > \beta_c$, however, we have (see Lemmas 4.1 and 4.3) $\|\mu^{\text{RC}} - \mu_{\text{ord}}\|_{TV} = \exp(-\Omega(n^{d-1}))$, and hence if ϵ is not too small, we can sample from μ_{ord} as above. *Mutatis mutandis* the same argument applies for μ_{dis} if $\beta_h < \beta < \beta_c$. On the other hand if $\epsilon = \exp(-\Omega(n^{d-1}))$, then we can use the Glauber dynamics to sample efficiently by Theorem 5.5.

For $\beta = \beta_c$ the situation is slightly different as the probability of both the ordered and disordered configurations are both of constant order, while the probability of configurations with interfaces is still $\exp(-\Omega(n^{d-1}))$. The solution is to use the approximate counting algorithm of Lemma 5.1 to approximate the relative probabilities of Ω_{ord} and Ω_{dis} under μ^{RC} and then to sample from each using the procedure above. Again if $\epsilon = \exp(-\Omega(n^{d-1}))$ we can use the Glauber dynamics.

Note that our sampling algorithm will not return any configurations with interfaces if $\epsilon \geq 4e^{-c\beta n^{d-1}}$, but such configurations have probability smaller than ϵ. On the other hand, if $\epsilon < 4e^{-c\beta n^{d-1}}$, then running Glauber dynamics may indeed return a configuration with interfaces. ∎

7 | CONCLUSIONS

In this article, we have given efficient approximate counting and sampling algorithms for the random cluster and q-state Potts models on $\mathbb{Z}^d$ at all inverse temperatures $\beta \geq 0$, provided $q \geq q_0(d)$ and $d \geq 2$. We believe the ideas of this article will, however, allow for approximate counting and sampling algorithms to be developed for a much broader class of statistical mechanics models. The necessary conditions for the development of algorithms for a given model is that there are only finitely many ground states, and that there is "sufficient τ-functionality." These are the necessary ingredients for the implementation of Pirogov–Sinai theory, see [9]. Our methods allow for the presence of unstable ground states, a significant improvement compared to the algorithms in [25].

Our results suggest that the algorithmic tasks of counting and sampling may be performed efficiently for a fairly broad class of statistical mechanics models with first-order phase transitions, but we leave a fuller investigation of this for future work. A related interesting questions is the existence of efficient algorithms for all $\beta \geq \beta_c$ in the presence of a second-order phase transition; we are not aware of any results in this direction with the exception of the Ising model, that is, the $q = 2$ state Potts model [23, 28]. To conclude we list some further open questions related to this article.

1. Our algorithms are restricted to $q \geq q_0(d)$ with $q_0(d)$ more than exponentially large in d. Do efficient algorithms exist that avoid this constraint? Since the physical phenomena behind our results are believed to hold for $q \geq 3$ when $d \geq 3$, there is likely room for improvement.

2. On the torus, we obtained an FPRAS (as opposed to an FPTAS) for the partition function because of the estimate on Z_{tunnel} from Lemma 4.1: the contribution of Z_{tunnel} cannot be ignored when $\epsilon \leq \exp(-\Omega(n^{d-1}))$. Fortunately, it is exactly when ϵ is this small that the Glauber dynamics mix in time polynomial in $1/\epsilon$, but of course Markov chain Monte Carlo is a randomized algorithm. A method for systematically accounting for the interfaces that contribute to Z_{tunnel} would likely enable the development of an FPTAS. We leave this as an open problem.
3. Our algorithms have at least two other features that could be improved. The first is the running time: while our algorithms are polynomial time, the degree of the polynomial is not small. The second is that our algorithms rely on *a priori* knowledge of whether or not $\beta = \beta_c$.

   Both of these deficiencies have the potential to be addressed by Glauber-type dynamics as described in [13]; see also [25, Section 7.2]. Proving the efficiency of these proposed algorithms would be very interesting.
4. Our deterministic algorithms for $\beta > \beta_c$ (and $\beta < \beta_c$) have diverging running times as $\beta \downarrow \beta_c$ ($\beta \uparrow \beta_c$). Are there *deterministic* algorithms that do not suffer from this dependence?
5. The algorithmic adaptation of other sophisticated contour-based methods, for example, [36], would be also be quite interesting, particularly for applications to problems such as counting the number of proper q-colorings of a graph. For recent progress on approximation algorithms for q-colorings, see [4, 26, 33, 34].

ACKNOWLEDGMENTS

Part of this work was done while Will Perkins and Prasad Tetali were visiting Microsoft Research New England. Part of this work was done while Tyler Helmuth and Will Perkins were visiting the Simons Institute for the Theory of Computing. Tyler Helmuth was supported by EPSRC Grant EP/P003656/1. Will Perkins is supported in part by NSF Grants DMS-1847451 and CCF-1934915. Prasad Tetali is supported in part by the NSF Grant DMS-1811935. We thank Guus Regts and Ewan Davies for helpful comments on a draft of this article.

REFERENCES

1. K. S. Alexander, *Mixing properties and exponential decay for lattice systems in finite volumes*, Ann. Probab. **32** (2004), no. 1A, 441–487.
2. A. Barvinok, *Combinatorics and complexity of partition functions*, Algor. Comb. **30** (2017).
3. A. Barvinok and G. Regts, *Weighted counting of solutions to sparse systems of equations*, Comb. Probab. Comput. **28** (2019), no. 5, 696–719.
4. F. Bencs, E. Davies, V. Patel, and G. Regts, *On zero-free regions for the anti-ferromagnetic Potts model on bounded-degree graphs*, Ann. Inst. Henri Poincare (D) Comb. Phys. **8** (2021), no. 3, 459–489.
5. A. Blanca and A. Sinclair, *Random-cluster dynamics in Z^2*, Probab. Theory Relat. Fields **168** (2017), no. 3-4, 821–847.
6. C. Borgs, J. Chayes, T. Helmuth, W. Perkins, and P. Tetali, *Efficient sampling and counting algorithms for the Potts model on $\mathbb{Z}^d$ at all temperatures (extended abstract)*.,Proc. 52nd Annu. ACM SIGACT Symp. Theory Comput STOC 2020. New York, NY: Association for Computing Machinery, 2020, pp. 738–751.
7. C. Borgs, J. Chayes, J. Kahn, and L. Lovász, *Left and right convergence of graphs with bounded degree*, Random Struct. Algorithms **42** (2013), no. 1, 1–28.
8. C. Borgs, J. T. Chayes, and P. Tetali, *Tight bounds for mixing of the Swendsen–Wang algorithm at the Potts transition point*, Probab. Theory Relat. Fields **152** (2012), no. 3-4, 509–557.
9. C. Borgs and J. Z. Imbrie, *A unified approach to phase diagrams in field theory and statistical mechanics*, Commun. Math. Phys. **123** (1989), no. 2, 305–328.

10. C. Borgs, R. Kotecký, and S. Miracle-Solé, *Finite-size scaling for Potts models*, J. Stat. Phys. **62** (1991), no. 3-4, 529–551.
11. S. Cannon and W. Perkins, *Counting independent sets in unbalanced bipartite graphs*, Proc. 14th Annu. ACM-SIAM Sympos. Discr. Algor. (SODA), SIAM, 2020, pp. 1456–1466.
12. K. Casel, P. Fischbeck, T. Friedrich, A. Göbel, and J. Lagodzinski, *Zeros and approximations of Holant polynomials on the complex plane*. arXiv preprint arXiv:1905.03194, 2019.
13. Z. Chen, A. Galanis, L. A. Goldberg, W. Perkins, J. Stewart, and E. Vigoda, *Fast algorithms at low temperatures via Markov chains*, Random Struct. Algorithms **58** (2021), no. 2, 294–321.
14. H. Duminil-Copin, *Lectures on the Ising and Potts models on the hypercubic lattice*. In: Barlow, M., Slade, G. (eds) Random Graphs, Phase Transitions, and the Gaussian Free Field. SSPROB 2017, Springer, Cham, 2020, vol. 304, pp. 35–161.
15. H. Duminil-Copin, A. Raoufi, and V. Tassion, *Sharp phase transition for the random-cluster and Potts models via decision trees*, Ann. Math. **189** (2019), no. 1, 75–99.
16. M. Dyer, L. A. Goldberg, C. Greenhill, and M. Jerrum, *The relative complexity of approximate counting problems*, Algorithmica **38** (2004), no. 3, 471–500.
17. S. Friedli and Y. Velenik, Statistical mechanics of lattice systems: A concrete mathematical introduction, Cambridge University Press, Cambridge, 2017.
18. A. Galanis, D. Štefankovič, and E. Vigoda, *Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models*, Comb. Probab. Comput. **25** (2016), no. 4, 500–559.
19. A. Galanis, D. Stefankovic, E. Vigoda, and L. Yang, *Ferromagnetic Potts model: Refined #-BIS-hardness and related results*, SIAM J. Comput. **45** (2016), no. 6, 2004–2065.
20. R. Gheissari and E. Lubetzky, *Mixing times of critical two-dimensional Potts models*, Commun. Pure Appl. Math. **71** (2018), no. 5, 994–1046.
21. R. Gheissari and E. Lubetzky, *Quasi-polynomial mixing of critical two-dimensional random cluster models*, Random Struct. Algorithms **56** (2020), no. 2, 517–556.
22. C. Gruber and H. Kunz, *General properties of polymer systems*, Commun. Math. Phys. **22** (1971), no. 2, 133–161.
23. H. Guo and M. Jerrum, *Random cluster dynamics for the Ising model is rapidly mixing*, Ann. Appl. Probab. **28** (2018), no. 2, 1292–1313.
24. T. Helmuth, M. Jenssen, and W. Perkins, *Finite-size scaling, phase coexistence, and algorithms for the random cluster model on random graphs*. arXiv preprint arXiv:2006.11580, 2020.
25. T. Helmuth, W. Perkins, and G. Regts, *Algorithmic Pirogov-Sinai theory*, Probab. Theory Relat. Fields **176** (2020), 851–895.
26. M. Jenssen, P. Keevash, and W. Perkins, *Algorithms for #BIS-hard problems on expander graphs*, SIAM J. Comput. **49** (2020), no. 4, 681–710.
27. M. Jerrum and A. Sinclair, *Approximating the permanent*, SIAM J. Comput. **18** (1989), no. 6, 1149–1178.
28. M. Jerrum and A. Sinclair, *Polynomial-time approximation algorithms for the Ising model*, SIAM J. Comput. **22** (1993), no. 5, 1087–1116.
29. R. Koteckỳ, *Pirogov-sinai theory*, Ency. Math. Phys. **4** (2006), 60–65.
30. R. Kotecký and D. Preiss, *Cluster expansion for abstract polymer models*, Commun. Math. Phys. **103** (1986), no. 3, 491–498.
31. R. Koteckỳ and S. Shlosman, *First-order phase transitions in large entropy lattice models*, Commun. Math. Phys. **83** (1982), no. 4, 493–515.
32. L. Laanait, A. Messager, S. Miracle-Solé, J. Ruiz, and S. Shlosman, *Interfaces in the Potts model I: Pirogov-Sinai theory of the Fortuin-Kasteleyn representation*, Commun. Math. Phys. **140** (1991), no. 1, 81–91.
33. C. Liao, J. Lin, P. Lu, and Z. Mao, "*Counting independent sets and colorings on random regular bipartite graphs*," In: D. Achlioptas and L. A. Végh (eds.) Approximation, randomization, and combinatorial optimization. Algorithms and techniques (APPROX/RANDOM 2019), Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.
34. J. Liu, A. Sinclair, and P. Srivastava, *A deterministic algorithm for counting colorings with 2-Delta colors*, Proc. 2019 IEEE 60th Annu. Sympos. Found. Comput. Sci. (FOCS). IEEE, 2019, pp. 1380–1404.
35. F. Martinelli, E. Olivieri, and R. H. Schonmann, *For 2-d lattice spin systems weak mixing implies strong mixing*, Commun. Math. Phys. **165** (1994), no. 1, 33–47.
36. R. Peled and Y. Spinka, *Rigidity of proper colorings of* $\mathbb{Z}^d$. arXiv preprint arXiv:1808.03597, 2018.
37. S. A. Pirogov and Y. G. Sinai, *Phase diagrams of classical lattice systems*, Theor. Math. Phys. **25** (1975), no. 3, 1185–1192.
38. A. Sly, *Computational transition at the uniqueness threshold*, Proc. 51st Annu. IEEE Sympos. Found. Comput Science, FOCS 2010, pages 287–296. IEEE, 2010.
39. A. Sly and N. Sun, *Counting in two-spin models on d-regular graphs*, Ann. Probab. **42** (2014), no. 6, 2383–2416.
40. D. Štefankovič, S. Vempala, and E. Vigoda, *Adaptive simulated annealing: A near-optimal connection between sampling and counting*, J. ACM (JACM) **56** (2009), no. 3, 18.

41. M. Ullrich, *Comparison of Swendsen-Wang and heat-bath dynamics*, Random Struct. Algorithms **42** (2013), no. 4, 520–535.
42. D. Weitz, *Counting independent sets up to the tree threshold*, Proc. 38th Annu. ACM Sympos. Theory.Comput., STOC. ACM; 2006, pp. 140–149.

How to cite this article: C. Borgs, J. Chayes, T. Helmuth, W. Perkins, and P. Tetali, *Efficient sampling and counting algorithms for the Potts model on $\mathbb{Z}^d$ at all temperatures*, Random Struct. Alg. **63** (2023), 130–170. https://doi.org/10.1002/rsa.21131

APPENDIX A: COUPLING THE POTTS AND RANDOM CLUSTER MODELS

Here we review the standard Edwards–Sokal coupling between the Potts and random cluster models and indicate how one can obtain counting and sampling algorithms for the Potts model from counting and sampling algorithms for the random cluster model. For more details on the couplings between the Potts model and random-cluster measures, see [14, Section 1.2.2].

Let $G = (V, E(G))$ be a finite graph. Then the standard Edwards–Sokal coupling put the q-color Potts model at inverse temperature β on the same probability space as the random cluster model with parameters q and $p = 1-e^{-\beta}$. To obtain a Potts configuration we sample a random cluster configuration A, then assign one of the q colors uniformly at random to each of the connected components of the graph $G_A = (V, A)$; note that isolated vertices are connected components. Each vertex is then assigned the color of its connected component. This gives an efficient algorithm to sample from the Potts model given a sample from the random cluster model. Moreover,

$$Z_G^{\text{Potts}}(\beta) = e^{\beta|E(G)|} Z_G^{\text{RC}}(1 - e^{-\beta}, q)\,, \tag{A1}$$

which gives us an FPTAS (FPRAS) for Z^{Potts} given an FPTAS (FPRAS) for Z^{RC}.

We can also couple the Potts model with monochromatic boundary conditions to the random cluster model with wired boundary conditions. For this, let us specialize to finite induced subgraphs $(\Lambda, E(\Lambda))$ of $\mathbb{Z}^d$. Define the boundary of Λ to be $\partial\Lambda := \{i \in \Lambda : \exists j \in \Lambda^c, (i,j) \in E(\mathbb{Z}^d)\}$. Recall the definition of the random cluster model μ_Λ^f with wired boundary conditions from Section 1.3. Given a color $r \in [q]$, the allowed colorings for the Potts model with r-monochromatic boundary conditions on Λ are

$$\Omega_r(\Lambda) = \left\{\sigma \in [q]^\Lambda : \sigma_v = r \;\forall\; v \in \partial\Lambda\right\}\,. \tag{A2}$$

The corresponding Gibbs measure and partition function are:

$$\mu_\Lambda^{\text{Potts},r}(\sigma) = \frac{\prod_{(i,j)\in E(\Lambda)} e^{-\beta \mathbf{1}_{\sigma_i \neq \sigma_j}}}{Z_\Lambda^{\text{Potts},r}(\beta)}\,, \qquad \sigma \in \Omega_r(\Lambda)$$
$$Z_\Lambda^{\text{Potts},r}(\beta) = \sum_{\sigma\in\Omega_r(\Lambda)} e^{-\beta \mathbf{1}_{\sigma_i \neq \sigma_j}}\,.$$

A simple extension of the Edwards–Sokal coupling then gives the following facts. Given a sample A from μ_Λ^w one can obtain a sample from $\mu_\Lambda^{\text{Potts},r}$ by coloring all vertices in $\partial\Lambda$ or connected to $\partial\Lambda$ by

the edges in A with color r, and assigning one of the q colors uniformly at random to the remaining connected components of the graph (Λ, A). Moreover, we have the relation

$$qZ_{\Lambda}^{\text{Potts},r}(\beta) = e^{-\beta|E(\Lambda)|}Z_{\Theta}^{w}((1-e^{-\beta}, q)\ . \tag{A3}$$

Again this shows that efficient counting and sampling algorithms for the Potts model with monochromatic boundary conditions follow from efficient counting and sampling algorithms for the random cluster model with wired boundary conditions.

APPENDIX B: PROOFS FOR $\beta_h < \beta < \beta_c$

B.1 | LEMMA 5.1 (III)

The proof of Lemma 5.1 in the case $\beta_h < \beta < \beta_c$ is the same, *mutatis mutandis*, as for $\beta > \beta_c$. The necessary changes are that (i) the roles of the ordered and disordered contours are exchanged, and (ii) some of the ingredients from Sections 5 and 6 were stated only for $\beta > \beta_c$, and hence versions for $\beta_h < \beta < \beta_c$ are necessary. We outline how to obtain these versions here.

As explained in [8, Appendix A], [8, Lemma 6.3 (i) and (ii)] applies when [8, (A.1)] holds. In fact, the arguments apply if

$$\beta \geq \max\left\{C_1 \log(dC), \frac{3\log q}{4d}\right\}, \tag{B1}$$

where C is the constant from [8, Lemma 5.8] and C_1 is a sufficiently large constant depending only on d. To verify this it is enough to check that [8, (A.2)] holds (up to a change in the constant 8).[5]Thus for q_0 sufficiently large [8, Lemma 6.3 (i) and (ii)] apply when $\beta_h < \beta < \beta_c$. In particular, by following the proofs from $\beta > \beta_c$ we obtain that when $\beta_h < \beta < \beta_c$

1. The conclusions of Lemma 4.2 hold with the roles of ord and dis reversed. The fact that $a_{\text{ord}} > 0$ is contained in [8, Lemma A.3].
2. The conclusion of Lemma 4.5 holds with ord replaced by dis.
3. The conclusion of Lemma 4.6 holds with the roles of ord and dis reversed and

$$M \geq \frac{2}{a_{\text{ord}}}\log\frac{8q}{\epsilon} + \frac{2}{a_{\text{ord}}}(\kappa+4)\|\boldsymbol{\gamma}\|.$$

   The factor four (as opposed to three) in M arises in the computation of the lower bound on $Z_{\text{ord}}^{\text{flip}}(\boldsymbol{\Lambda}, M)$, as (in the notation of the proof of Lemma 4.6) Ext Γ may be of size $\|\boldsymbol{\gamma}\|$.

Lastly, the conclusion of Proposition 4.7 holds with dis changed to ord. The proof is very similar to the proof of Proposition 4.7, but using Lemmas 3.5 and 3.6 in place of Lemmas 3.7 and 3.8.

Proof of Lemma 5.1 (iii). Using the ingredients above, this follows exactly as in the proof of Lemma 5.1 (ii), that is, for $\beta > \beta_c$. ∎

[5]Our choice of 3/4 in (B1) is somewhat arbitrary; the same conclusion would hold for any number strictly larger than 2/3.

B.2 | THEOREMS 1.1 AND 1.2

These proofs are exactly as for $\beta > \beta_c$ provided the conclusions of Lemma 4.3 hold with dis replaced by ord. This is straightforward to obtain by imitating the proof of Lemma 4.3, using (as discussed in the previous section) that the conclusion of Lemma 4.2 hold with the roles of ord and dis reversed.

APPENDIX C: CONTOUR COMPUTATIONS USING SUBGRAPHS OF $(\frac{1}{2}\mathbb{T}_n^d)^\star$

The next lemma shows that computations relating to contours $\boldsymbol{\gamma}$ can be implemented using only γ, the connected subgraph of $(\frac{1}{2}\mathbb{T}_n^d)^\star$ that corresponds to $\boldsymbol{\gamma}$ by the construction in Section 3.1.

Lemma C.1. *Let $\boldsymbol{\gamma}$ and $\boldsymbol{\gamma}'$ be contours, and let γ and γ' be the corresponding subgraphs of $(\frac{1}{2}\mathbb{T}_n^d)^\star$. Then given γ, γ',*

1. *$d_\infty(\boldsymbol{\gamma}, \boldsymbol{\gamma}')$ can be computed in time $O(|\{V(\gamma)\}||\{V(\gamma')\}|)$,*
2. *The set $\operatorname{Int}\boldsymbol{\gamma} \cap \mathbb{T}_n^d$ can be computed in time $O(|\{V(\gamma)\}|^3)$,*
3. *$\|\boldsymbol{\gamma}\|$ can be computed in time $O(|\{V(\gamma)\}|)$.*

Proof. Each vertex in $(\frac{1}{2}\mathbb{T}_n^d)^\star$ corresponds to a $(d-1)$-dimensional hypercube in $\boldsymbol{T}_n^d$. For each pair of such hypercubes we can compute the distance between them in constant time, which implies the first claim. The third claim follows similarly, since the set of edges passing through a given $(d-1)$-dimensional hypercube can be determined in constant time.

For the second claim, we first determine the set of edges intersecting the $(d-1)$-dimensional hypercubes corresponding to $\boldsymbol{\gamma}$. We can then determine $\operatorname{Int}\boldsymbol{\gamma} \cap \mathbb{T}_n^d$ in time $O(\|\boldsymbol{\gamma}\|^3)$ by Lemma 3.10. ∎