Information and Computation 285 (2022) 104894

Contents lists available at ScienceDirect

Information and Computation

www.elsevier.com/locate/yinco

Fast mixing via polymers for random graphs with unbounded degree ☆

Andreas Galanis, Leslie Ann Goldberg, James Stewart *

Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford, OX1 3QD, UK

ARTICLE INFO

Article history:
Received 20 July 2021
Received in revised form 19 March 2022
Accepted 23 March 2022
Available online 28 March 2022

ABSTRACT

The polymer model framework is a classical tool from statistical mechanics that has recently been used to obtain approximation algorithms for spin systems on classes of bounded-degree graphs; examples include the ferromagnetic Potts model on expanders and on the grid. One of the key ingredients in the analysis of polymer models is controlling the growth rate of the number of polymers, which has been typically achieved so far by invoking the bounded-degree assumption. Nevertheless, this assumption is often restrictive and obstructs the applicability of the method to more general graphs. For example, sparse random graphs typically have bounded average degree and good expansion properties, but they include vertices with unbounded degree, and therefore are excluded from the current polymer-model framework.

We develop a less restrictive framework for polymer models that relaxes the standard bounded-degree assumption, by reworking the relevant polymer models from the edge perspective. The edge perspective allows us to bound the growth rate of the number of polymers in terms of the total degree of polymers, which in turn can be related more easily to the expansion properties of the underlying graph. To apply our methods, we consider random graphs with unbounded degrees from a fixed degree sequence (with minimum degree at least 3) and obtain approximation algorithms for the ferromagnetic Potts model, which is a standard benchmark for polymer models. Our techniques also extend to more general spin systems.

1. Introduction

The polymer model framework [21,14] is a classical tool from statistical mechanics which has recently been used to obtain efficient approximation algorithms for analysing spin systems (such as the Potts model) in parameter regimes where standard algorithmic approaches are provably inefficient/inaccurate on general graphs. These algorithms apply to certain classes of graphs that typically have sufficiently strong expansion properties relative to their local growth rates. Typically, the local growth rate is restricted by a bounded-degree assumption. Examples of such classes include bounded-degree expanders [20,22,7,5,6,2,11,15] and the d-dimensional grid [16,4,20,17]. The purpose of this work is to expand the current framework for applying polymer models by relaxing the bounded-degree assumption and using alternative methods to capture the growth of the graph.

☆ A preliminary short version of this paper (without proofs) will appear in the proceedings of RANDOM 2021.

* Corresponding author.
E-mail address: james.stewart@cs.ox.ac.uk (J. Stewart).

https://doi.org/10.1016/j.ic.2022.104894

To briefly review the current framework, we use as a running example the q-state ferromagnetic Potts model with parameter $\beta > 0$. For a graph $G = (V_G, E_G)$, the set $\Omega_{G,q}$ of configurations of the model is the set of all (not necessarily proper) q-colourings σ of V_G using the set of colours $[q] = \{1, \ldots, q\}$ where $q \geq 3$. The weight of a colouring σ is $w(\sigma) = e^{\beta m_G(\sigma)}$ where $m_G(\sigma)$ is the number of monochromatic edges under σ. The so-called partition function $Z = Z_{G,q,\beta}$ is the aggregate weight of all σ and the Gibbs distribution $\mu = \mu_{G,q,\beta}$ is the probability distribution on the set of all σ, in which each σ has mass proportional to its weight, i.e., $\mu(\sigma) = w(\sigma)/Z$. We will study the computational problems of approximating the partition function and approximately sampling from the Gibbs distribution. In general, these problems are computationally hard (#BIS-hard) when the parameter β is sufficiently large [13,12].

The recent works [16,20] introduced a framework based on polymer models that bypasses the worst-case hardness, on classes of bounded-degree graphs with expansion properties. The rough intuition for the Potts model is that, for large β, due to the expansion properties, the colourings with non-negligible weight are close to the so-called ground-states of the model, i.e., the q configurations in which all vertices get the same colour. Polymer models capture the deviation of configurations from these ground states. Given a ground state with colour r, a polymer is a connected set of vertices, none of which is coloured with r, and a polymer configuration (with respect to the ground state r) corresponds to the set of all polymers (see Example 4 for more details). The Potts model can then be decomposed into q polymer models, each of which can be studied using relatively streamlined algorithmic methods (based on interpolation [1] and Markov chains). This framework has already found multiple algorithmic applications in far more general settings [16,4,17,20,19,22,7,10,5,6,11,15].

Despite these advances, the current applications of polymer models rely crucially on the fact that the maximum degree of the underlying graph is bounded. This fact is used to control the number of polymers of a given size (which is crucially needed for the algorithmic analysis). As a result of this limitation, applications to several other interesting classes of graphs are ruled out, excluding for example sparse random graphs, which have bounded average degree and good expansion properties, but include vertices with unbounded degree.

1.1. Main results

In this paper, we propose a framework for polymer models that overcomes the bounded-degree limitations of previous algorithms, by revisiting the Markov chain approach of [7]. We introduce a new condition which requires that the weight of each polymer decays exponentially in its total degree (the sum of the degrees of the vertices in the polymer) instead of decaying exponentially in the polymer's size. This new condition allows us to prove rapid mixing for a Markov chain which is an adapted edge-version of the so-called polymer dynamics of [7]. Crucially, the fact that the new condition is formulated in terms of the total degree of a polymer allows us to relax the assumption that the instance has bounded degree.

As an application of our method, we consider the q-state ferromagnetic Potts model on sparse random graphs of unbounded degree with a given degree sequence, as detailed below.

Definition 1. Let d be a positive real number and n be a positive integer. We define $\mathcal{D}_{n,d}$ to be the set of all degree sequences $\{x_1, x_2, \ldots, x_n\}$ that satisfy

1. For all $i \in [n]$, $3 \leq x_i \leq n^{\rho}$ where $\rho = \frac{1}{50}$, and
2. $\sum_{i \in [n]} x_i^2 \leq dn$.

We write $G \sim \mathcal{G}(n, \vec{x})$ to indicate that G is a graph chosen uniformly at random from the set of all simple n-vertex graphs with degree sequence $\vec{x}$. We say that G satisfies a property with high probability (whp) if the probability that G satisfies the property is $1 - o(1)$, as a function of n, uniformly over $\vec{x}$.

Note that $\mathcal{D}_{n,d}$ is empty unless $d \geq 9$. The assumption that all degrees are greater than or equal to 3 (rather than 2) guarantees that the random graph G is connected and has good expansion properties. The degree lower bound also means that our results do not apply to Erdős-Rényi random graphs. The upper bound on the degrees is mild and can in fact be relaxed somewhat further (but in general cannot be made to be linear in n due to the sparsity assumption in Item 2).

We give an efficient algorithm for approximately sampling[1] from and approximating the partition function[2] of the ferromagnetic Potts model on random graphs with a given degree sequence for all sufficiently large β.

Theorem 2. *Let d be a real number and $q \geq 3$ be an integer. For the ferromagnetic Potts model, there is β_0 such that for all $\beta \geq \beta_0$ there is a poly-time approximate sampling algorithm for $\mu_{G,q,\beta}$ and an FPRAS for $Z_{G,q,\beta}$ that work with high probability on random graphs $G \sim \mathcal{G}(n, \vec{x})$ for any degree sequence $\vec{x} \in \mathcal{D}_{n,d}$.*

[1] A polynomial-time approximate sampling algorithm for $\mu_{G,q,\beta}$ is an algorithm that, given an accuracy parameter $\varepsilon > 0$ and a graph $G = (V_G, E_G)$ as input, outputs a sample from a probability distribution that is within total variation distance ε of $\mu_{G,q,\beta}$, in time poly$(|V_G|, 1/\varepsilon)$.

[2] Given an accuracy parameter $\varepsilon > 0$, we say that $\hat{Z}$ is an ε-approximation to the quantity Z if $e^{-\varepsilon} Z \leq \hat{Z} \leq e^{\varepsilon} Z$. A fully polynomial randomised approximation scheme (FPRAS) for $Z_{G,q,\beta}$ is a randomised algorithm that, given an accuracy parameter $\varepsilon > 0$ and a graph $G = (V_G, E_G)$ as input, outputs a random variable that is an ε-approximation to $Z_{G,q,\beta}$ with probability at least 3/4, in time poly$(|V_G|, 1/\varepsilon)$.

Remark 3. Note that β_0 depends on d and q, and our arguments later (see Remark 17) show that $\beta_0 = Cd\log d\log q$ for some $C > 0$ (independent of d or q). If the desired accuracy ε is at least e^{-n} then the running time of the sampling algorithm is $O\left(n\log\frac{n}{\varepsilon}\log\frac{1}{\varepsilon}\right)$ and the running time of the FPRAS is $O\left(n^2(\log\frac{n}{\varepsilon})^3\right)$.

We further remark that the bounded-degree assumption has also been relaxed in [17] for the ferromagnetic Potts model on lattice graphs; the approach therein however is tailored to a certain flow representation of the Potts model, which is used as a basis for the corresponding polymer models and therefore does not extend to general spin systems. Our approach applies to general polymer models as detailed in the next section and our focus on the ferromagnetic Potts model is mainly to illustrate the method without further technical overhead; the approach for example can be adapted to general spin systems on bipartite random graphs with a given degree sequence (analogously to [11]).

2. Polymers

The main technique that we use to prove Theorem 2 is to refine the polymer framework by introducing a new polymer sampling condition which requires that the weight of each polymer decays exponentially in its total degree. In order to state the new condition we first give some relevant definitions.

Let $G = (V_G, E_G)$ be a graph — we refer to G as the "host graph" of the polymer model. Let $[q] = \{1, \ldots, q\}$ be a set of spins and $g = \{g_v\}_{v\in V_G}$ specify a set of ground-state spins for the vertices, i.e., $g_v \subseteq [q]$ for each $v \in V_G$. A polymer is a pair $\gamma = (V_\gamma, \sigma_\gamma)$ consisting of a connected set of vertices V_γ and an assignment $\sigma_\gamma : V_\gamma \to [q]$ such that $\sigma_\gamma(v) \in [q]\backslash g_v$. Let $\mathcal{P}_G$ be the set of all polymers. A polymer model for the host graph G is specified by a subset of allowed polymers $\mathcal{C}_G \subseteq \mathcal{P}_G$, and a weight function $w_G : \mathcal{C}_G \to \mathbb{R}_{\geq 0}$. For polymers $\gamma, \gamma' \in \mathcal{P}_G$, we write $\gamma \sim \gamma'$ to denote that γ, γ' are compatible, i.e., that for every $u \in \gamma$ and $u' \in \gamma$ the graph distance in G between u and u' is at least 2. We define $\Omega_G = \{\Gamma \subseteq \mathcal{C}_G \mid \forall \gamma, \gamma' \in \Gamma, \gamma \sim \gamma'\}$ to be the set of all sets of mutually compatible polymers of $\mathcal{C}_G$; elements of Ω_G are called polymer configurations. We define the partition function as $Z_G = \sum_{\Gamma\in\Omega_G}\prod_{\gamma\in\Gamma} w_G(\gamma)$, and the Gibbs distribution on $\Gamma \in \Omega_G$ as $\mu_G(\Gamma) = \prod_{\gamma\in\Gamma} w_G(\gamma)/Z_G$. We use $(\mathcal{C}_G, w_G)$ to denote the polymer model.

Example 4 *(The polymer model $(\mathcal{C}^r_{G,q}, w_{G,\beta})$, [20]).* Consider the q-state ferromagnetic Potts model with parameter β, and let $r \in [q]$ be a colour. Let G be a graph and set $g_v = \{r\}$ for every $v \in V_G$. The weight of a polymer $\gamma = (V_\gamma, \sigma_\gamma)$ is defined as $w_{G,\beta}(\gamma) = \mathrm{e}^{-\beta B_\gamma}$, where B_γ denotes the number of edges from V_γ to $V_G\backslash V_\gamma$ plus the number of edges of G with both endpoints in V_γ that are bichromatic under σ_γ. We let $\mathcal{P}^r_{G,q}$ denote the set of all polymers and the set of allowed polymers $\mathcal{C}^r_{G,q}$ to be the set of polymers $\gamma \in \mathcal{P}^r_{G,q}$ such that $|V_\gamma| < |V_G|/2$. Note that a polymer configuration Γ consisting of the polymers $\gamma_1, \ldots, \gamma_k$ corresponds to a colouring σ, where a vertex v takes the colour $\sigma_{\gamma_i}(v)$ if $v \in V_{\gamma_i}$ for some $i \in [k]$, and the colour r otherwise; moreover, $e^{\beta|E_G|}\prod_{i\in[k]} w_G(\gamma_i) = w_G(\sigma)$.

Polymer models have been used to approximate the partition function of spin systems on bounded-degree host graphs. There are several existing algorithmic frameworks which can be used to sample from these resulting polymer models. One such deterministic algorithm uses the polynomial interpolation method of Barvinok [1] combined with the cluster expansion to approximate the partition function of the polymer model, see [16] for more details. Typical running times for these deterministic algorithms are of the form $n^{O(\log(\Delta))}$, where Δ is the maximum degree of the host graph, though for polymer models these have been improved to give a running time of $n^{1+o_\Delta(1)}$, see [20]. Another approach, which we will describe in detail in Section 4, uses a Markov chain called the polymer dynamics to sample from μ_G (see also [7] for more details). The running times of algorithms obtained using the Markov chain approach are usually faster and of the form $O(n\log n)$. Both of these approaches work for restricted ranges of parameters,[3] and the essential condition required is that the weight of each polymer decays exponentially in the number of vertices it contains. To obtain our results, we give a simple generic way to modify this condition, as detailed below.

For a vertex $v \in V_G$ we write $\deg_G(v)$ to denote the degree of v in G, and for a vertex subset $S \subseteq V_G$ we write $\deg_G(S)$ to denote $\sum_{v\in S}\deg_G(v)$.

Definition 5. Let $q \geq 2$ be an integer, $\mathcal{G}$ be a class of graphs, and $\mathcal{F}_\mathcal{G} = \{(\mathcal{C}_G, w_G) \mid G \in \mathcal{G}\}$ be a family of q-spin polymer models. We say that $\mathcal{F}_\mathcal{G}$ satisfies the *polymer sampling condition* with constant $\tau \geq 3\log(8\mathrm{e}^3(q-1))$ if $w_G(\gamma) \leq \mathrm{e}^{-\tau\deg_G(V_\gamma)}$ for all $G \in \mathcal{G}$ and all $\gamma \in \mathcal{C}_G$.[4]

Using Definition 5, we will show (Lemma 8, below) that if a "computationally feasible" family of polymer models on a class of graphs $\mathcal{G}$ satisfies this new condition, then there is an efficient algorithm which, given a graph $G \in \mathcal{G}$, approximately samples from the Gibbs distribution of the polymer model corresponding to G.

[3] For example, for the q-colour ferromagnetic Potts model with inverse temperature β on α-expander graphs of maximum degree Δ where $\alpha > 0$, $\Delta \geq 3$, $q \geq 2$, Theorem 3 of [20] applies when $\beta \geq (4 + 2\log(q\Delta))/\alpha$ whereas Theorem 9 of [7] requires $\beta \geq (5 + 3\log((q-1)\Delta)/\alpha$.

[4] Unless we specify otherwise, the base of all logarithms in this paper is assumed to be e.

The new polymer sampling condition in Definition 5 is analogous to the original one in [7] except that the original condition requires the weight of a polymer to decay exponentially in its size, and in particular that the constant τ is sufficiently big relative to the maximum degree of G. The new condition relaxes this, allowing us to choose the constant τ in Definition 5 so that it does not depend on the maximum degree of the host graph, which is how we overcome the limitations of previous work. Technically, the improvement comes from the fact that previous work relies on bounding the number of connected vertex subsets of a given size (with bounds that depend on the maximum degree of the graph), but here we are able to instead rely on the following lemma which bounds the number of connected vertex subsets with a given total degree and this enables us to avoid restricting the maximum degree of the graph. The new condition, which replaces the notion of "size" with total degree, fits well with the original abstract polymer model framework of [21], where the notion of the "size" of a polymer is an abstract function.

Lemma 6. *Let $G = (V_G, E_G)$ be a graph, $v \in V_G$, and $\ell \geq 1$ be an integer. The number of connected vertex subsets $S \subseteq V_G$ such that $v \in S$ and $\deg_G(S) = \ell$ is at most $(2\mathrm{e})^{2\ell-1}$.*

In addition to the bound on the number of connected vertex subsets in Lemma 6, we will use the fact that these connected vertex subsets can be enumerated in time exponential in the total degree ℓ (see Lemma 21). Although the bound in Lemma 6 is exponential in ℓ, this will be mitigated by the fact that the new polymer sampling condition ensures that the weight of each polymer is exponentially small in its total degree. The new polymer sampling condition therefore allows us to prove that the following condition holds — this condition is analogous to the polymer mixing condition of [7], except that we consider edges instead of vertices. For a polymer $\gamma \in \mathcal{P}_G$, let E_γ denote the set of edges of G with at least one endpoint in V_γ.

Definition 7. Let $q \geq 2$ be an integer, $\mathcal{G}$ be a class of graphs, and $\mathcal{F}_\mathcal{G} = \{(\mathcal{C}_G, w_G) \mid G \in \mathcal{G}\}$ be a family of q-spin polymer models. We say that $\mathcal{F}_\mathcal{G}$ satisfies the *polymer mixing condition* with constant $\theta \in (0,1)$ if $\sum_{\gamma' \nsim \gamma} |E_{\gamma'}| \cdot w_G(\gamma') \leq \theta |E_\gamma|$ for all $G \in \mathcal{G}$ and all $\gamma \in \mathcal{C}_G$.

In contrast to the conditions in [7], the two new conditions consider edges since we modify the polymer dynamics algorithm to sample edges instead of vertices. Subject to these new conditions, the techniques of [7] can be adapted to show that the modified polymer dynamics mixes rapidly, therefore giving the efficient algorithm for sampling from the Gibbs distribution of a polymer model. We give the relevant details in Section 4.

Finally, in order to use the modified polymer dynamics as an efficient algorithm for computing an approximate sample from μ_G, we will need a mild computational condition for polymers. More precisely, we say that a family of polymer models $\{(\mathcal{C}_G, w_G) \mid G \in \mathcal{G}\}$ is *computationally feasible* if for all $G \in \mathcal{G}$ and all $\gamma \in \mathcal{P}_G$, it is possible to decide whether $\gamma \in \mathcal{C}_G$ and to compute $w_G(\gamma)$, if it is, in $O(\mathrm{e}^{\deg_G(V_\gamma)})$ time. Computational feasibility serves exactly the same purpose as it did in Definition 3 of [7], which requires that the same operations are able to be carried out in time depending on $|V_\gamma|$ (instead of $\deg_G(V_\gamma)$ that we use here).

In Section 4, we prove the following lemma which gives an efficient algorithm for sampling[5] from the Gibbs distribution of a polymer model and for approximating its partition function. In order to prove the lemma, we extend the polymer dynamics algorithm of [7] to the unbounded degree setting. The proof of the lemma uses the fact (Lemma 18) that the polymer sampling condition implies the polymer mixing condition.

Lemma 8. *Let $q \geq 2$ be an integer, $\mathcal{G}$ be a class of graphs, and $\mathcal{F}_\mathcal{G}$ be a family of computationally feasible q-spin polymer models satisfying the polymer sampling condition.*

There are randomised algorithms which, given as input a graph $G \in \mathcal{G}$ with m edges and an accuracy parameter $\varepsilon > 0$, output an ε-sample from μ_G in $O\left(m \log \frac{m}{\varepsilon} \log \frac{1}{\varepsilon}\right)$ time, and an ε-approximation to Z_G, with probability at least $3/4$, in $O\left(m^2 \log(\frac{m}{\varepsilon})^3\right)$ time.

3. Application to unbounded-degree graphs

Let $\alpha > 0$ be a real number. We say that a graph G is an α-total-degree expander if, for all connected vertex subsets $S \subseteq V_G$ with $|S| \leq |V_G|/2$, we have $e_G(S, S^c) \geq \alpha \deg_G(S)$, where $e_G(S, S^c)$ denotes the number of edges with one endpoint in S and the other in $S^c := V_G \setminus S$. Let $\mathcal{G}_\alpha$ denote the set of all α-total-degree expanders. Note, every connected $G \in \mathcal{G}_\alpha$ is also an α-expander (i.e., $e_G(S, S^c) \geq \alpha |S|$).

When β is sufficiently large, the polymer model from Example 4 satisfies the polymer sampling condition (Definition 5) with constant $\tau = \alpha\beta$. To see this, consider $\gamma \in \mathcal{C}^r_{G,q}$ and observe that since $B_\gamma \geq e_G(V_\gamma, V_\gamma^c)$ and $|V_\gamma| < |V_G|/2$, it follows that

$$w_{G,\beta}(\gamma) \leq \exp\left\{-\alpha\beta \deg_G(V_\gamma)\right\} = \mathrm{e}^{-\tau \deg_G(V_\gamma)}, \tag{1}$$

[5] Given an accuracy parameter $\varepsilon > 0$, we say that a random variable X is an ε-sample from the probability distribution μ if the total variation distance between the distribution of X and μ is at most ε.

where $\tau \geq 3\log(8e^3(q-1))$ if $\beta \geq \frac{3}{\alpha}\log(8e^3(q-1))$.

We may therefore apply Lemma 8 in order to efficiently sample from the ferromagnetic Potts model and to estimate Z_G for $G \in \mathcal{G}_\alpha$, provided that β is sufficiently large. The proof of the following theorem is deferred until Section 5.

Theorem 9. *Let $\alpha > 0$ be a real number. Let $q \geq 3$ be an integer and $\beta \geq \frac{3}{\alpha}\log(8e^3(q-1))$ be a real. For the Potts model on $G \in \mathcal{G}_\alpha$, there is a poly-time approximate sampling algorithm for $\mu_{G,q,\beta}$ and an FPRAS for $Z_{G,q,\beta}$.*

In fact, for $n = |V_G|$ and $m = |E_G|$, if the desired accuracy ε satisfies $\varepsilon \geq e^{-n}$ then the running time of the sampler is $O(m\log\frac{m}{\varepsilon}\log\frac{1}{\varepsilon})$ and the running time of the FPRAS is $O\left(m^2(\log\frac{m}{\varepsilon})^3\right)$.

3.1. Expansion of random graphs with specified degree sequences

Let d be a real number. In this section, we will show that a graph $G \sim \mathcal{G}(n,\vec{x})$ for a degree sequence $\vec{x} \in \mathcal{D}_{n,d}$ is whp an α-total-degree-expander for some constant $\alpha > 0$, i.e., that $G \in \mathcal{G}_\alpha$.

To work with $G \sim \mathcal{G}(n,\vec{x})$, we consider the standard configuration model, where a random multigraph $H = (V_H, E_H)$ with the given degree sequence $\vec{x}$ is sampled by the following process. For each $i \in [n]$, we attach x_i half-edges to the vertex i. We then sample a uniformly random perfect matching on the half-edges to give E_H. This uniformly random perfect matching can be sampled by performing the following until no half-edges remain: choose any remaining half-edge, choose another remaining half-edge uniformly at random, then pair these two half-edges and remove them from the set of remaining half-edges. We write $H \sim \mathsf{CM}(n,\vec{x})$. Note, for two vertices $i, j \in V_H$ such that $i \neq j$, the probability that a half edge attached to i and a half edge attached to j are paired is

$$p_{\{i,j\}} = \frac{x_i x_j}{2m-1}, \text{where } m = \frac{1}{2}\sum_{k=1}^{n} x_k, \tag{2}$$

and similarly the probability that two half-edges of i are connected is $p_{\{i,i\}} = \frac{x_i(x_i-1)}{2m(2m-1)}$.

We first prove results about $\mathsf{CM}(n,\vec{x})$, since asymptotic properties of $\mathsf{CM}(n,\vec{x})$ can easily be transferred back to $\mathcal{G}(n,\vec{x})$ using the following straightforward consequence of [18, Theorem 1.1].

Lemma 10. *Let d be a positive real number. For every positive integer n, let $\mathcal{E}_n$ be a set of n-vertex multigraphs. If, for some $\vec{x} \in \mathcal{D}_{n,d}$, we have $G \sim \mathcal{G}(n,\vec{x})$ and $H \sim \mathsf{CM}(n,\vec{x})$ then the following is true. If $H \in \mathcal{E}_n$ with high probability, then $G \in \mathcal{E}_n$ with high probability.*

Proof. Suppose that H' is drawn from $\mathsf{CM}(n,\vec{x})$ conditioned on being simple. It is well-known (for example, see [24, Proposition 7.15]) that H' is a uniformly-random graph with degree sequence $\vec{x}$. Thus, it follows that

$$\Pr(G \notin \mathcal{E}_n) = \Pr(H' \notin \mathcal{E}_n) = \frac{\Pr(H \notin \mathcal{E}_n, H \text{ is simple})}{\Pr(H \text{ is simple})} \leq \frac{\Pr(H \notin \mathcal{E}_n)}{\Pr(H \text{ is simple})}. \tag{3}$$

By assumption, we have that $\Pr(H \notin \mathcal{E}_n) = o(1)$. Applying Theorem 1.1 of [18], whose conditions are satisfied by Item 2 of Definition 1, it follows that there is a positive p such that when n is sufficiently large $\Pr(H \text{ is simple}) > p$. By (3), we therefore have that $\Pr(G \notin \mathcal{E}_n) = o(1)$, and the result follows. □

For a (multi)graph $H = (V_H, E_H)$ we define the tree-excess to be $t_H = |E_H| - (|V_H| - 1)$; that is, the number of edges more than a tree that H has. First, we show that multigraphs drawn from the configuration model have locally bounded tree excess.

Lemma 11. *Let d be a positive real number. The following is true with high probability when $H = (V_H, E_H)$ is drawn from $\mathsf{CM}(n,\vec{x})$ uniformly over all degree sequences $\vec{x} \in \mathcal{D}_{n,d}$. For all connected vertex sets $S \subseteq V_H$ with $|S| \leq (\log n)^2$ and $\deg_H(S) \geq 36$, we have that $t_{H[S]} \leq \frac{1}{6}\deg_H(S)$.*

Proof. For positive integers k and ℓ, and a non-negative integer t, let the random variable $X_{k,\ell,t}$ denote the number of connected vertex subsets $S \subseteq V_H$ such that $|S| = k$, $\deg_H(S) = \ell$, and $t_{H[S]} = t$. To prove the lemma, we will show that whp

$$\sum_{k \leq \lfloor(\log n)^2\rfloor} \sum_{\ell \geq 36} \sum_{t \geq \lfloor \ell/6 \rfloor + 1} X_{k,\ell,t} = 0.$$

In fact, we can further restrict the range of summation. From the lower bound in Item 1 of Definition 1, we have that $x_i \geq 3$ for all i, and therefore $\ell \geq 3k$. Item 2 shows that $\sum_i x_i \leq dn$, and therefore $\ell \leq dn$ and $t \leq \ell/2 \leq dn/2$. So, consider any integer ℓ in the range $36 \leq \ell \leq dn$, any integer k in the range $1 \leq k \leq \min\{(\log n)^2, \ell/3\}$, and any integer $t > \ell/6$. There are at most $\binom{n}{k}$ vertex subsets $S \subseteq V_H$ with $|S| = k$ and $\deg_G(S) = \ell$. Let $j = k - 1 + t$ be the number of edges with both endpoints in S. Given such a set S, there are at most $\binom{\ell}{2j}$ possibilities for the set of half-edges in these j edges. On a given

set of $2j$ half-edges, there are $(2j-1)!! = \frac{(2j)!}{2^j j!}$ perfect matchings. Using the upper bound on the degrees from Item 1 of Definition 1, the probability that a set of j edges is present in H is at most

$$\frac{n^{2\rho}}{2m-1}\frac{n^{2\rho}}{2m-3}\cdots\frac{n^{2\rho}}{2m-2j+1} \le \left(\frac{n^{2\rho}}{2m-2j}\right)^j \le \left(\frac{n^{2\rho}}{n}\right)^j,$$

where the final inequality follows from the fact that $k \le (\log n)^2$ and therefore that $2m-2j \ge \deg_G(S^c) \ge 3|S^c| = 3(n-k) > n$ (as long as n is sufficiently big). We also have that

$$\binom{\ell}{2j}\cdot\frac{(2j)!}{2^j j!} < \frac{\ell!}{(\ell-2j)!\,j!} < \frac{\ell^{2j}\mathrm{e}^j}{j^j} \le \left(\frac{\mathrm{e}\ell^2}{t}\right)^j.$$

Putting everything together, it follows that

$$\mathbb{E}[X_{k,\ell,t}] \le \binom{n}{k}\left(\frac{\mathrm{e}\ell^2}{t}\right)^{k-1+t}\left(\frac{n^{2\rho}}{n}\right)^{k-1+t} < \left(\frac{\mathrm{e}^2\ell^2}{t}\right)^{k-1+t}\frac{n^{2\rho(k-1+t)}}{n^{t-1}}.$$

Furthermore, since $t > \ell/6$, $k < 2t$, and (by the upper bound in Item 1 of Definition 1) $\ell \le kn^\rho \le n^{2\rho}$, we have that

$$\mathbb{E}[X_{k,\ell,t}] < \frac{(6\mathrm{e}^2 n^{4\rho})^{3t-1}}{n^{t-1}}. \tag{4}$$

Let

$$X = \sum_{\ell=36}^{dn}\ \sum_{k=1}^{\lfloor \min\{(\log n)^2,\ell/3\}\rfloor}\ \sum_{t=\lfloor \ell/6\rfloor+1}^{\lfloor dn/2\rfloor} X_{k,\ell,t}.$$

Since $t > \ell/6 \ge 6$, it follows that $t \ge 7$. For big enough n, (4) shows that $\mathbb{E}[X_{k,\ell,t}] \le n^{13\rho t}/n^{t-1}$. Since $\rho \le 2/91$ and $t \ge 7$, $1-13\rho \ge 5/7 \ge 5/t$ so $13\rho t \le t-5$ and $\mathbb{E}[X_{k,\ell,t}]$ is at most n^{-4}. Taking a union bound over all permissible values for ℓ, k, and t, we find that $\mathbb{E}[X] = o(1)$. Applying Markov's inequality, we have that $\Pr(X > 0) = \Pr(X \ge 1) \le \mathbb{E}[X] = o(1)$, and the result follows. □

To obtain the expansion bounds in Lemmas 13 and 14, we will use the following result from Fountoulakis and Reed [9]. Although this result is stated in [9] in terms of the random graph model, it is first proved for the configuration model, so this is how we state it. Also, the result in [9] requires that the vector $\vec{x}$ be in $\mathcal{D}_{n,d}$ but this is only important for lifting their result to the random graph model, so it is not relevant for us.

Lemma 12 *([9, Proposition 4.5]). When $H = (V_H, E_H)$ is drawn from $\mathsf{CM}(n,\vec{x})$ for some length-n degree sequence $\vec{x}$ the following is true for any set $S \subseteq V_H$.*

$$\Pr(e_H(S,S^c) = 0) \le \binom{m}{\deg_H(S)/2}^{-1},$$

where $m = \frac{1}{2}\sum_i x_i$.

Note that Lemma 12 was stated in [9] for S such that $\deg_H(S)$ is even, but if $\deg_H(S)$ is odd, it is not possible to have $e_H(S,S^c) = 0$. Next, we show that in a multigraph H drawn from the configuration model, small vertex subsets satisfy certain expansion properties.

Lemma 13. *Let d be a positive real number. The following is true with high probability when $H = (V_H, E_H)$ is drawn from $\mathsf{CM}(n,\vec{x})$ uniformly over all degree sequences $\vec{x} \in \mathcal{D}_{n,d}$. For all connected vertex sets $S \subseteq V_H$ with $|S| \le (\log n)^2$, we have that $e_H(S,S^c) \ge |S|/4$.*

Proof. For positive integers k and ℓ, and a non-negative integer j, let the random variable $X_{k,j,\ell}$ denote the number of connected vertex subsets $S \subseteq V_H$ with $|S| = k$, $e_H(S,S^c) = j$, and $\deg_H(S) = \ell$. By Item 1 of Definition 1, we need only consider ℓ satisfying $3k \le \ell \le kn^\rho$. Let

$$X = \sum_{k=1}^{\lfloor (\log n)^2\rfloor}\ \sum_{j=0}^{\lfloor k/4\rfloor}\ \sum_{\ell=3k}^{\lfloor kn^\rho\rfloor} X_{k,j,\ell}.$$

To prove the lemma we will show that $X = 0$, whp. Consider any integer k in the range $1 \le k \le (\log n)^2$, any integer j in the range $0 \le j < k/4$, and any integer ℓ in the range $3k \le \ell \le kn^\rho$. There are at most $\binom{n}{k}$ candidates for vertex sets S with

$|S| = k$ and $\deg_H(S) = \ell$. There are then at most $\binom{\ell}{j}$ choices for the j half-edges emanating from vertices of S that will be matched with half-edges emanating from vertices of S^c, once H is drawn. Applying Lemma 12 to the degree sequence derived from $\vec{x}$ by removing the j half-edges (and their partners), the probability that the remaining $\ell - j$ half-edges are matched amongst themselves is at most

$$\binom{m'}{(\ell-j)/2}^{-1} \leq \left(\frac{(\ell-j)}{2m'}\right)^{\frac{(\ell-j)}{2}} \leq \left(\frac{kn^\rho}{n}\right)^{\frac{11k}{8}},$$

where $2m' = (\sum_{i=1}^n x_i) - 2j$ and the last inequality follows (for big enough n) since $11k/4 \leq \ell - j \leq kn^\rho$ and $2m' \geq 3n - 2j > n$. We therefore have that

$$\begin{aligned}
\mathbb{E}[X] &\leq \sum_{k=1}^{\lfloor (\log n)^2 \rfloor} \sum_{j=0}^{\lfloor k/4 \rfloor} \sum_{\ell=3k}^{\lfloor kn^\rho \rfloor} \binom{n}{k}\binom{\ell}{j}\left(\frac{kn^\rho}{n}\right)^{\frac{11k}{8}} \\
&\leq \sum_{k=1}^{\lfloor (\log n)^2 \rfloor} \sum_{j=0}^{\lfloor k/4 \rfloor} \sum_{\ell=3k}^{\lfloor kn^\rho \rfloor} \left(\frac{ne}{k}\right)^k \left(\frac{e\ell}{j}\right)^j \left(\frac{kn^\rho}{n}\right)^{\frac{11k}{8}} \\
&\leq \sum_{k=1}^{\lfloor (\log n)^2 \rfloor} \left(\frac{(\log n)^{O(1)} n^{\rho(2+11/8)}}{n^{3/8}}\right)^k.
\end{aligned}$$

This is $o(1)$ since $\rho < 1/9$. Applying Markov's inequality, we have that $\Pr(X > 0) = \Pr(X \geq 1) = o(1)$, and the result follows. □

We will also require the following lemma.

Lemma 14. *Let d be a positive real number. There is a positive real number α (depending on d) such that the following is true with high probability when $H = (V_H, E_H)$ is drawn from $\mathsf{CM}(n, \vec{x})$ uniformly over all degree sequences $\vec{x} \in \mathcal{D}_{n,d}$. For all connected vertex sets $S \subseteq V_H$ with $(\log n)^2 \leq |S| \leq n/2$, we have that $e_H(S, S^c) \geq \alpha \deg_H(S)$.*

Proof. We split the proof into three cases based on the degree of the vertex set $S \subseteq V_H$:

1. $\deg_H(S) \leq 100d|S|$,
2. $\deg_H(S) > 100d|S|$ and $\deg_H(S) \leq n/2$, and
3. $\deg_H(S) > 100d|S|$ and $\deg_H(S) > n/2$.

Throughout, we let $m = \frac{1}{2}\sum_{i=1}^n x_i$.

Case 1. The proof of this case is similar to that of Lemma 4.1 of [9], but we include it for completeness. We will show that there is a positive real number $\alpha' < 1/4$ such that whp, every vertex set $S \subseteq V_H$ with $(\log n)^2 \leq |S| \leq n/2$ and $\deg_H(S) \leq 100d|S|$ satisfies $e_H(S, S^c) \geq \alpha'|S| \geq \alpha' \deg_H(S)/(100d)$. We will eventually require α' to be sufficiently small depending on d.

For positive integers k and ℓ, and a non-negative integer j, let the random variable $X_{k,\ell,j}$ denote the number of connected vertex subsets $S \subseteq V_H$ with $|S| = k$, $e_H(S, S^c) = j$, and $\deg_H(S) = \ell$. Let

$$X = \sum_{k=\lceil (\log n)^2 \rceil}^{\lfloor n/2 \rfloor} \sum_{j=0}^{\lfloor \alpha' k \rfloor} \sum_{\ell=1}^{100dk} X_{k,\ell,j}.$$

Our aim will be to show that $\mathbb{E}[X] = o(1)$, and we begin by further restricting the range of summation in the above. Let $S \subseteq V_H$ be a connected vertex subset with $|S| = k$, $e_H(S, S^c) = j$, and $\deg_H(S) = \ell$. Since there are $(\ell - j)/2$ edges in $e_H(S, S)$, we need only consider ℓ such that $\ell - j$ is even. Furthermore, since S is connected, we can assume that $e_H(S, S) \geq k - 1$, so we need only consider values of ℓ satisfying $\ell \geq j + 2(k-1)$. Finally, since $x_i \geq 3$ for all i, and $\sum_i x_i = 2m$, we have that $2m \geq \ell + 3(n-k)$, so we need only consider values of ℓ satisfying $\ell \leq 2m - 3(n-k)$. Let $J_k = \{j \in \mathbb{Z} : 0 \leq j \leq \alpha' k\}$ and let $L_{k,j} = \{\ell \in \mathbb{Z} : j + 2(k-1) \leq \ell \leq \min\{2m - 3(n-k), 100dk\},\ \ell - j \text{ is even}\}$, we can therefore re-write X as follows

$$X = \sum_{k=\lceil (\log n)^2 \rceil}^{\lfloor n/2 \rfloor} \sum_{j \in J_k} \sum_{\ell \in L_{k,j}} X_{k,\ell,j}.$$

Given a set of k vertices, there are at most $\binom{\ell}{j}$ ways to choose the half-edges that will make up $E_H(S, S^c)$ once H is drawn. The remaining total degree of S after choosing these edges is $\ell - j$. By Lemma 12, the probability that every other edge of $E_H(S, V_H)$ is an edge of $E_H(S, S)$ is at most $\binom{m'}{(\ell-j)/2}^{-1}$, where $m' = m - j$. Since $\ell \in L_{k,j}$, $(\ell - j)/2 \geq k - 1$ and $(\ell - j)/2 \leq m - 3(n-k)/2 - j/2$. Since $j < \alpha' k < k/4 \leq n/8$ it follows that $n - k > j$ and therefore that $(\ell - j)/2 \leq m - 3(n-k)/2 - j/2 \leq m' - n + k$. Thus, we have that

$$\binom{m'}{(\ell-j)/2}^{-1} \leq \max\left\{\binom{m'}{k-1}^{-1}, \binom{m'}{n-k}^{-1}\right\}.$$

Using Stirling's approximation, there exists a constant C (see [9, Equation 7]) such that for all n, m', and k

$$\frac{\binom{n}{k}}{\binom{m'}{k-1}} \leq C \cdot \frac{m'}{k}\left(\frac{n}{m'}\right)^k \quad \text{and} \quad \frac{\binom{n}{k}}{\binom{m'}{n-k}} \leq C \cdot \frac{m'}{m'-n+k+1}\left(\frac{n}{m'}\right)^{n-k}.$$

Since $2m' = 2m - 2j \geq 3n - n/4$, it follows that $n/m' \leq 8/11$. Since $k \leq n/2$, it follows that $n - k \geq k$. Thus,

$$\frac{\binom{n}{k}}{\binom{m'}{k-1}} \leq C \cdot m'\left(\frac{8}{11}\right)^k \quad \text{and} \quad \frac{\binom{n}{k}}{\binom{m'}{n-k}} \leq C \cdot m'\left(\frac{8}{11}\right)^k,$$

therefore

$$\binom{n}{k}\binom{m'}{(\ell-j)/2}^{-1} \leq C \cdot m'\left(\frac{8}{11}\right)^k.$$

From the above, we have that

$$\begin{aligned}\mathbb{E}[X] &= O(1) \sum_{k=\lceil (\log n)^2 \rceil}^{\lfloor n/2 \rfloor} \sum_{j \in J_k} \sum_{\ell \in L_{k,j}} \binom{n}{k}\binom{\ell}{j}\binom{m'}{(\ell-j)/2}^{-1} \\ &= O(1) \sum_{k=\lceil (\log n)^2 \rceil}^{\lfloor n/2 \rfloor} \sum_{j \in J_k} \sum_{\ell \in L_{k,j}} m'\binom{\ell}{j}\left(\frac{8}{11}\right)^k.\end{aligned}$$

We wish to show that for each term $\binom{\ell}{j} \leq (9/8)^k$. Since $j \leq \alpha' k \leq \alpha' \ell < \ell/4$ we have $\binom{\ell}{j} \leq \binom{\ell}{\alpha' k} \leq (\frac{e\ell}{\alpha' k})^{\alpha' k}$ and since $\ell \leq 100dk$ this is at most $(100ed/\alpha')^{\alpha' k}$ which is at most $(9/8)^k$ if α' is sufficiently small, depending on d. (It suffices to take $\alpha' = 10^{-4}/\log d$, for example.)

We conclude that we can upper bound each term by $m'(9/11)^k \leq dn(9/11)^k$ and therefore

$$\mathbb{E}[X] = O(n) \sum_{k=\lceil (\log n)^2 \rceil}^{\lfloor n/2 \rfloor} \sum_{j \in J_k} \sum_{\ell \in L_{k,j}} \left(\frac{9}{11}\right)^k = O(n) \sum_{k=\lceil (\log n)^2 \rceil}^{\lfloor n/2 \rfloor} k^2 \left(\frac{9}{11}\right)^k.$$

It follows from the above that $\mathbb{E}[X] = o(1)$. Applying Markov's inequality, we have that $\Pr(X > 0) = \Pr(X \geq 1) = o(1)$, and the result follows.

Case 2. We will do more than is required, and show that every connected vertex set $S \subseteq V_H$ with $(\log n)^2 \leq \deg_H(S) \leq n/2$ satisfies $e_H(S, S^c) \geq \deg_H(S)/100$.

For every positive integer ℓ satisfying $(\log n)^2 \leq \ell \leq n/2$ and every non-negative integer j satisfying $j \leq \ell/100$, let $X_{\ell,j}$ denote the number of connected vertex subsets $S \subseteq V_H$ with $e_H(S, S^c) = j$ and $\deg_H(S) = \ell$. Let $J_\ell = \{j \in \mathbb{Z} : 0 \leq j \leq \ell/100, \ell - j \text{ is even}\}$ and let

$$X = \sum_{\ell=\lceil (\log n)^2 \rceil}^{\lfloor n/2 \rfloor} \sum_{j \in J_\ell} X_{\ell,j}.$$

Our aim will be to show that $\mathbb{E}[X] = o(1)$.

In order for a set $S \subseteq V_H$ to contribute to $X_{\ell,j}$ its size, s, must be at most $(\ell - j)/2 + 1$ (otherwise, the $(\ell - j)/2$ internal edges won't be able to connect S). The number of size-s sets S is at most $\binom{n}{s}$ which is at most $\binom{n}{(\ell-j)/2+1}$, as long as n is sufficiently large. The number of possible values of s is at most $(\ell - j)/2 \leq \ell$. Thus, the number of possibilities for S is at most $\ell \cdot \binom{n}{(\ell-j)/2+1}$

Given S, there are at most $\binom{\ell}{j}$ ways to choose the half-edges that will make up $E_H(S, S^c)$ once H is drawn. The remaining total degree of S after choosing these edges is $\ell - j$. By Lemma 12, the probability that every other edge of $E_H(S, V_H)$ is an edge of $E_H(S, S)$ is at most $\binom{m'}{(\ell-j)/2}^{-1}$, where $m' = m - j$. As noted in Case 1, it is shown using Stirling's approximation (see equation 7) in [9] that there exists a constant C such that for all n, ℓ, j and m',

$$\binom{n}{(\ell-j)/2+1}\binom{m'}{(\ell-j)/2}^{-1} \le Cm'\left(\frac{n}{m'}\right)^{\frac{\ell-j}{2}+1}.$$

Since $2m \ge 3n$ by Item 1 of Definition 1 and $2j \le 2\ell/100 \le n/100$, $2m' = 2m - 2j \ge 3n - n/100$ so $n/m' < 7/10$. Thus,

$$\binom{n}{(\ell-j)/2+1}\binom{m'}{(\ell-j)/2}^{-1} \le Cm'\left(\frac{7}{10}\right)^{\frac{\ell-j}{2}+1} \le Cm'\left(\frac{7}{10}\right)^{\frac{\ell-\ell/100}{2}} \le Cm'\left(\frac{17}{20}\right)^{\ell}.$$

Furthermore, we have that

$$\binom{\ell}{j} \le \binom{\ell}{\ell/100} \le (100\mathrm{e})^{\ell/100}.$$

Putting things together, it follows that

$$\begin{aligned}\mathbb{E}[X] &\le \sum_{\ell=\lceil(\log n)^2\rceil}^{\lfloor n/2\rfloor} \sum_{j\in J_\ell} \ell \cdot \binom{\ell}{j}\binom{n}{(\ell-j)/2+1}\binom{m'}{(\ell-j)/2}^{-1} \\ &\le C \sum_{\ell=\lceil(\log n)^2\rceil}^{\lfloor n/2\rfloor} \ell^2 m'\left(\frac{17}{20}\right)^{\ell}(100\mathrm{e})^{\ell/100}.\end{aligned}$$

Now $m' \le m$ and this is $O(n)$ by Item 2 of Definition 1, so

$$\mathbb{E}[X] = O(n) \sum_{\ell=\lceil(\log n)^2\rceil}^{\lfloor n/2\rfloor} \ell^2 \cdot \left(\frac{9}{10}\right)^{\ell} = o(1).$$

Applying Markov's inequality, we have that $\Pr(X > 0) = \Pr(X \ge 1) = o(1)$, and the result follows.

Case 3. Finally, we deal with vertex sets $S \subseteq V_H$ with $(\log n)^2 \le |S| \le n/2$ and $\deg_H(S) > \max\{100d|S|, n/2\}$. Let $C = 10^4 d$. By the Cauchy-Schwarz inequality, we have that $|S|\sum_{i\in S} x_i^2 \ge (\deg_H(S))^2 \ge 10^4 d^2|S|^2$, so using Item 2 of Definition 1 which ensures that $\sum_{i\in S} x_i^2 \le dn$, we find that $|S| \le n/C$.

Let $f = (\mathrm{e}C)^{1/C}$. The number of sets S satisfying $|S| \le n/C$ is at most $n\binom{n}{n/C} \le nf^n$ since there are at most n possibilities for $|S|$ to consider, and for each of them $\binom{n}{|S|} \le \binom{n}{n/C}$.

Fix any set $S \subseteq V_H$ with $|S| \le n/C$ and consider the random construction of H, starting from half-edges in S^c (and choosing their mates in the pairing). Let

$$j = \left\lfloor \frac{\deg_H(S^c)}{2} \right\rfloor \ge \left\lfloor \frac{3|S^c|}{2} \right\rfloor \ge \left\lfloor \frac{3n\left(1-\frac{1}{C}\right)}{2} \right\rfloor \ge \frac{3n\left(1-\frac{2}{C}\right)}{2},$$

where the first inequality uses the fact that each x_i is at least 3 (from Item 1 of Definition 1) and the final inequality uses the fact that n is sufficiently large.

Note that the process initiates a pairing from at least j half-edges in S^c. For each $i \in [j]$, let Y_i be the indicator random variable for the event that the i'th half-edge from which pairing is initiated connects to an endpoint in S (conditioned on the pairings of the first $i-1$ half-edges initiated from S^c).

Let $\varepsilon = 3(1-2/C)/(8\sqrt{d}) \le 1/2$. From Item 2 in Definition 1 and the Cauchy-Schwarz inequality, $(\sum x_i)^2 \le n\sum x_i^2 \le nd$ so $\sum_{i=1}^n x_i \le \sqrt{d}n$. For any $t \in [j]$ satisfying $\sum_{i=1}^{t-1} Y_t < \varepsilon n/2$ we have

$$\Pr(Y_t = 1) \ge \frac{\deg_H(S) - \varepsilon n/2}{\sqrt{d}n} > \frac{1-\varepsilon}{2\sqrt{d}} \ge \frac{1}{4\sqrt{d}}.$$

Now let $X_1, \ldots, X_j$ be i.i.d. Bernoulli random variables which are 1 with probability $1/(4\sqrt{d})$. We can couple the evolution of these variables so that, for any $t \in [j]$ satisfying $\sum_{i=1}^{t-1} Y_i < \varepsilon n/2$, we have $\sum_{i=1}^t Y_i \ge \sum_{i=1}^t X_i$. We conclude that $\Pr(\sum_{i=1}^j Y_i < \varepsilon n/2) \le \Pr(\sum_{i=1}^j X_i < \varepsilon n/2)$.

To conclude we will show that $nf^n \Pr(\sum_{i=1}^{j} X_i < \varepsilon n/2) = o(1)$, implying that we can take $\alpha'' = \varepsilon/(2\sqrt{d})$ since $\varepsilon n/2 = \alpha''\sqrt{d}n \geq \alpha'' \deg_H(S)$.

Let $X = \sum_{i=1}^{j} X_i$ and $\delta = 1/2$. Note that $\mathbb{E}[X] = j/(4\sqrt{d})$ and that

$$\frac{(1-\delta)j}{4\sqrt{d}} \geq \frac{(1-\delta)3n\left(1-\frac{2}{C}\right)}{8\sqrt{d}} = \frac{\varepsilon n}{2}.$$

By a Chernoff bound, $\Pr(X \leq \varepsilon n/2) \leq \Pr(X \leq (1-\delta)j/(4\sqrt{d})) \leq \exp(-j\delta^2/(8\sqrt{d}))$.

To conclude that $nf^n \exp(-j\delta^2/(8\sqrt{d})) = o(1)$ we observe that $f < \exp(3(1-2/C)\delta^2/(16\sqrt{d}))$. So with $\alpha'' = \varepsilon/(2\sqrt{d})$, we conclude Case 3.

To prove the lemma, combine the three cases by taking $\alpha = \min\{\alpha'/(100d), 1/100, \alpha''\}$. □

We can now prove the following result, which establishes the desired expansion properties of the multigraphs generated by the configuration model.

Lemma 15. *Let d be a positive real number. There is a positive real number α (depending on d) such that the following is true with high probability when $H = (V_H, E_H)$ is drawn from $\mathsf{CM}(n, \vec{x})$ uniformly over all degree sequences $\vec{x} \in \mathcal{D}_{n,d}$. For all connected vertex sets $S \subseteq V_H$ with $|S| \leq n/2$, we have that $e_H(S, S^c) \geq \alpha \deg_H(S)$.*

Proof. We consider three cases.

Case 1. Consider all connected subsets $S \subseteq V_H$ with $(\log n)^2 \leq |S| \leq n/2$. By Lemma 14 there is a positive real number α' such that, whp, every such subset S has $e_H(S, S^c) \geq \alpha' \deg_H(S)$.

Case 2. Consider all connected subsets $S \subseteq V_H$ with $|S| \leq (\log n)^2$ and $\deg_H(S) \geq 36$.

- Consider first those subsets S with $|S| \leq \frac{1}{6}\deg_H(S)$. We have that

$$e_H(S, S^c) = \deg_H(S) - 2(t_{H[S]} + |S| - 1) \geq \frac{2}{3}\deg_H(S) - 2|S| \geq \frac{1}{3}\deg_H(S),$$

  by Lemma 11 and our assumption on the size of S.
- Now consider those subsets S with $|S| > \frac{1}{6}\deg_H(S)$, then by Lemma 13, we have that $e_H(S, S^c) \geq |S|/4 \geq \deg_H(S)/24$.

Case 3. Finally, consider connected subsets $S \subseteq V_H$ with $|S| \leq (\log n)^2$ and $\deg_H(S) < 36$.

By Lemma 13, we have that $e_H(S, S^c) \geq |S|/4 \geq 1/4 = 36/144 > \deg_H(S)/144$.

The result follows from the three cases by taking $\alpha = \min\{1/144, \alpha'\} = \alpha'$. □

Using the definition of $\mathcal{G}_\alpha$ and Lemma 10, we have the following corollary of Lemma 15.

Corollary 16. *Let d be a real number. There is a positive real number α (depending on d) such that the following holds. With high probability, when $G \sim \mathcal{G}(n, \vec{x})$ for some $\vec{x} \in \mathcal{D}_{n,d}$, it holds that $G \in \mathcal{G}_\alpha$.*

Combining Corollary 16 with Theorem 9 implies our main theorem.

Theorem 2. *Let d be a real number and $q \geq 3$ be an integer. For the ferromagnetic Potts model, there is β_0 such that for all $\beta \geq \beta_0$ there is a poly-time approximate sampling algorithm for $\mu_{G,q,\beta}$ and an FPRAS for $Z_{G,q,\beta}$ that work with high probability on random graphs $G \sim \mathcal{G}(n, \vec{x})$ for any degree sequence $\vec{x} \in \mathcal{D}_{n,d}$.*

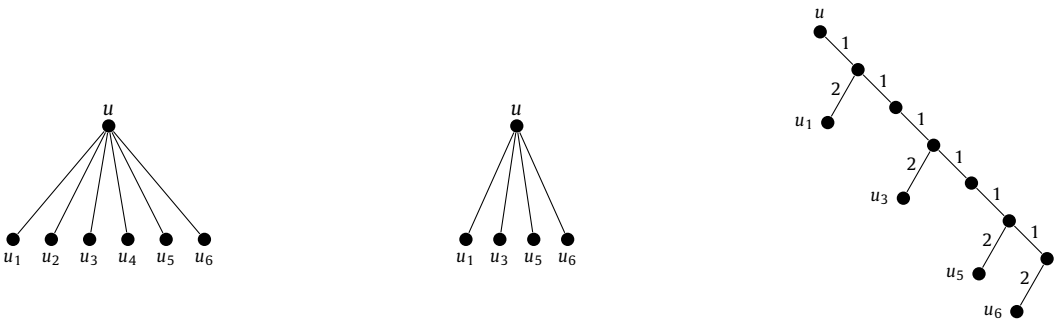
Proof. Let d be a real number and let $q \geq 2$ be an integer. Let α be the positive real number from Corollary 16. Let $\beta_0 = \frac{3}{\alpha}\log(8e^3(q-1))$.

Consider $\vec{x} \in \mathcal{D}_{n,d}$ and let G be drawn from $\mathcal{G}(n, \vec{x})$. By Corollary 16, $G \in \mathcal{G}_\alpha$ whp. The result then follows by using the algorithms from Theorem 9. □

Remark 17. The bounds on β in Remark 3 follow from the choice of β_0 in the proof of Theorem 2 and from the fact that $\alpha = \Omega(\frac{1}{d\log d})$ which follows from the proofs of Lemmas 14 and 15. The running time bounds in Remark 3 come from those in Theorem 9 using the fact that $|E_G| = O(n)$ which follows from Item 2 of Definition 1.

4. Details of the algorithms for the edge-based polymer model

We start with a lemma giving an upper bound on the number of vertex subsets of a given total degree. This lemma is used throughout this section.



(a) Neighbourhood of u in G (b) Neighbourhood of u in T (c) Neighbourhood of u in T'

Fig. 1. Neighbourhood of u.

Lemma 6. *Let $G = (V_G, E_G)$ be a graph, $v \in V_G$, and $\ell \geq 1$ be an integer. The number of connected vertex subsets $S \subseteq V_G$ such that $v \in S$ and $\deg_G(S) = \ell$ is at most $(2e)^{2\ell-1}$.*

Proof. Let $N(G, v, \ell)$ be the set of subtrees $T = (V_T, E_T)$ of G such that $v \in V_T$, $\deg_G(V_T) = \ell$. We will show that $|N(G, v, \ell)| \leq (2e)^{2\ell-1}$, which gives us the desired result for the following reason. Let $S \subseteq V_G$ be a connected vertex subset such that $v \in S$ and $\deg_G(S) = \ell$. Since S is connected, it has at least one spanning tree $T = (V_T = S, E_T)$ such that $v \in V_T$ and $\deg_G(V_T) = \ell$. Since S is the unique connected vertex subset that T spans, this gives us an injective map from the set of all connected vertex subsets containing v with total degree ℓ, to $N(G, v, \ell)$.

We now give an injective map from $N(G, v, \ell)$ to $T^*(2\ell, 3)$ – the set of subtrees of size 2ℓ that contain the root, of the infinite rooted 3-regular tree. By a result of Bollobás [3, p. 129], we know that $|T^*(2\ell, 3)|$ is at most $(2e)^{2\ell-1}$. Let $T = (V_T, E_T)$ be a subtree from $N(G, v, \ell)$. We will map T to a rooted subtree $T' = (V_{T'}, E_{T'})$ from $T^*(2\ell, 3)$. For each vertex of V_G, fix an ordering of its neighbours. In the infinite rooted 3-regular tree, label the edges incident to the root with $\{1, 2, 3\}$, and for each other vertex label the edges connecting it to its two children with $\{1, 2\}$. As we construct T', we will label its edges so that it is clear which subtree from $T^*(2\ell, 3)$ we are constructing, we will also label some of its vertices. We construct T' as follows (see Fig. 1 for an example of the following construction).

1. Add the root to $V_{T'}$ and label it v.
2. While there is a labelled vertex of T' (call its label u) such that u has a child w in T but no vertex of T' is labelled w, then we do the following. First, we create a path P of length $\deg_G(u)$ where each edge is labelled 1. We then connect the vertex of T' labelled u to P via an edge labelled 1. Finally, for $1 \leq i \leq \deg_G(u)$, we connect a vertex labelled w to the i^{th} vertex of P via an edge labelled 2, if w is the i^{th} neighbour of u in G and w is a child of u in T.

Each $T \in N(G, v, \ell)$ maps to a different $T' \in T^*(2\ell, 3)$. When constructing T', we used edge labels from $\{1, 2, 3\}$, therefore the maximum degree of T' is 3. For each $v \in V_T$, we added at most $2\deg_G(v)$ vertices to T', therefore the size of T' is at most $2\deg_G(V_T) = 2\ell$. □

The following lemma shows that if a family of polymer models satisfies the polymer sampling condition then it also satisfies the polymer mixing condition. This will be convenient when applying our algorithms, as it is in general easier to show that the polymer sampling condition holds for a given family of polymer models, than to show directly that the polymer mixing condition does. However, the polymer mixing condition is used to bound the mixing time of the polymer dynamics.

For a vertex subset $S \subseteq V_G$, we let $\partial_G S$ denote the vertices of S^c that are joined to S by an edge.

Lemma 18. *Let $q \geq 2$ be an integer, $\mathcal{G}$ be a class of graphs, and $\mathcal{F}_{\mathcal{G}} = \{(\mathcal{C}_G, w_G) \mid G \in \mathcal{G}\}$ be a family of q-spin polymer models. If $\mathcal{F}_{\mathcal{G}}$ satisfies the polymer sampling condition (Definition 5) then $\mathcal{F}_{\mathcal{G}}$ satisfies the polymer mixing condition (Definition 7) with constant $\theta = \frac{1}{e}$.*

Proof. Let $G \in \mathcal{G}$ and let $\gamma \in \mathcal{C}_G$ be an arbitrary polymer. We have that

$$\sum_{\gamma' \nsim \gamma} |E_{\gamma'}| \cdot w_G(\gamma') \leq \sum_{\{u,v\} \in E_\gamma} \left(\sum_{\gamma' : u \in V_{\gamma'}} |E_{\gamma'}| \cdot w_G(\gamma') + \sum_{\gamma' : v \in V_{\gamma'}} |E_{\gamma'}| \cdot w_G(\gamma') \right),$$

therefore the result will follow if we are able to show for all $v \in V_\gamma \cup \partial_G V_\gamma$ that

$$\sum_{\gamma' : v \in V_{\gamma'}} |E_{\gamma'}| \cdot w_G(\gamma') \leq \frac{1}{2\mathrm{e}}.$$

We can re-write the left-hand side of the above as follows

$$\sum_{\gamma' : v \in V_{\gamma'}} |E_{\gamma'}| \cdot w_G(\gamma') \leq \sum_{\ell \geq 1} \sum_{\substack{\gamma' :\ v \in V_{\gamma'}, \\ \deg_G(V_{\gamma'}) = \ell}} |E_{\gamma'}| \cdot w_G(\gamma').$$

By Lemma 6, we know that there are at most $(2\mathrm{e})^{2\ell-1}$ connected vertex subsets S such that $v \in S$, and $\deg_G(S) = \ell$. For each such S there are $(q-1)^{|S|} \leq (q-1)^{\ell}$ ways to assign spins to its vertices, therefore there are at most $(2\mathrm{e})^{2\ell-1}(q-1)^{\ell}$ polymers γ' such that $v \in V_{\gamma'}$ and $\deg_G(V_{\gamma'}) = \ell$. Furthermore, since $\mathcal{F}_{\mathcal{G}}$ satisfies the polymer sampling condition, we know that each of these polymers satisfies $w_G(\gamma') \leq \mathrm{e}^{-\tau\ell}$ where $\tau \geq 3\log(8\mathrm{e}^3(q-1))$. Combining these facts with the observation that $|E_{\gamma'}| \leq \deg_G(V_{\gamma'}) \leq \mathrm{e}^{\deg_G(V_{\gamma'})}$ for all $\gamma' \in \mathcal{C}_G$, we obtain that

$$\sum_{\gamma' : v \in V_{\gamma'}} |E_{\gamma'}| \cdot w_G(\gamma') \leq \sum_{\ell \geq 1} (2\mathrm{e})^{2\ell-1}(q-1)^{\ell}\mathrm{e}^{\ell}\mathrm{e}^{-\tau\ell} \leq \frac{1}{2\mathrm{e}} \sum_{\ell \geq 1} \left(\frac{4\mathrm{e}^3(q-1)}{\mathrm{e}^{\tau}} \right)^{\ell}.$$

Finally, since $\tau \geq \log(8\mathrm{e}^3(q-1))$, the result follows. □

4.1. Polymer dynamics

We will define a Markov chain (the polymer dynamics) to sample from μ_G, that mixes rapidly if the polymer mixing condition is met. This Markov chain is essentially the same as the polymer dynamics of [7], except that we sample edges instead of vertices. For $e \in E_G$ let $\mathcal{A}(e)$ be the set of polymers γ such that $e \in E_{\gamma}$. Let ν_e be the probability distribution on $\mathcal{A}(e) \cup \emptyset$ defined by

$$\nu_e(\gamma) = w_G(\gamma) \quad \text{and} \quad \nu_e(\emptyset) = 1 - \sum_{\gamma \in \mathcal{A}(e)} \nu_e(\gamma).$$

If the polymer mixing condition is satisfied, then ν_e gives a well-defined probability distribution. To see this, let $\gamma \in \mathcal{A}(e)$ be such that $|E_{\gamma}|$ is least. Applying the polymer mixing condition, we obtain that

$$\sum_{\substack{\gamma' \nsim \gamma \\ \gamma' \in \mathcal{A}(e)}} |E_{\gamma'}| \cdot w_G(\gamma') \leq \sum_{\gamma' \nsim \gamma} |E_{\gamma'}| \cdot w_G(\gamma') < |E_{\gamma}|.$$

Thus,

$$\sum_{\gamma' \in \mathcal{A}(e)} \nu_e(\gamma') = \sum_{\substack{\gamma' \nsim \gamma \\ \gamma' \in \mathcal{A}(e)}} w_G(\gamma') \leq \sum_{\substack{\gamma' \nsim \gamma \\ \gamma' \in \mathcal{A}(e)}} \frac{|E_{\gamma'}|}{|E_{\gamma}|} \cdot w_G(\gamma') \leq 1.$$

Let $t \geq 0$ be an integer. The polymer dynamics Markov chain transitions from $\Gamma_t \in \Omega_G$ to $\Gamma_{t+1} \in \Omega_G$ according to the following rules.

1. Choose $e \in E_G$ uniformly at random. Let γ_e be the (unique) polymer in $\Gamma_t \cap \mathcal{A}(e)$, if it exists; otherwise, let $\gamma_e = \emptyset$.
2. Mutually exclusively
   (a) with probability 1/2, let $\Gamma_{t+1} = \Gamma_t \setminus \{\gamma_e\}$, and
   (b) with probability 1/2, sample a polymer γ from ν_e, and let $\Gamma_{t+1} = \Gamma_t \cup \{\gamma\}$ if this is compatible; otherwise, let $\Gamma_{t+1} = \Gamma_t$.

The following lemma shows that the unique stationary distribution of the polymer dynamics is the Gibbs distribution of the polymer model.

Lemma 19. *Let $G = (V_G, E_G)$ be a graph and let $(\mathcal{C}_G, w_G)$ be a polymer model. The unique stationary distribution of the polymer dynamics is μ_G.*

Proof. One update of the polymer dynamics changes the polymer configuration by at most one polymer. Let $\Gamma' = \Gamma \cup \{\gamma\}$, then

$$\frac{\mu_G(\Gamma')}{\mu_G(\Gamma)} = w_G(\gamma) = \frac{w_G(\gamma)|E_\gamma|/2|E_G|}{|E_\gamma|/2|E_G|} = \frac{\Pr(\Gamma \to \Gamma')}{\Pr(\Gamma' \to \Gamma)}.$$

It therefore follows by detailed balance that μ_G is a stationary distribution of the polymer dynamics.

Finally, the polymer dynamics are irreducible since we can move from any $\Gamma \in \Omega_G$ to any $\Gamma' \in \Omega_G$ by adding and removing polymers (for example, via the empty set). The polymer dynamics are also aperiodic since there are self-loops. □

The following lemma shows that if the polymer mixing condition is satisfied, then the polymer dynamics Markov chain mixes rapidly.[6]

Lemma 20. *Let $\mathcal{G}$ be a class of graphs and let $\{(\mathcal{C}_G, w_G) \mid G \in \mathcal{G}\}$ be a family of polymer models that satisfies the polymer mixing condition (see Definition 7). For all $G \in \mathcal{G}$ and all $\varepsilon > 0$, the mixing time of the polymer dynamics is $\tau_{mix}(\varepsilon) = O(|E_G| \log(|E_G|/\varepsilon))$.*

Proof. Let $G \in \mathcal{G}$. We proceed by path coupling. Let $D(\cdot,\cdot)$ be a metric on Ω_G which we define by setting $D(\Gamma, \Gamma') = 1$ for polymer configurations $\Gamma, \Gamma' \in \Omega_G$ such that $\Gamma' = \Gamma \cup \{\gamma\}$ for some polymer $\gamma \in \mathcal{C}_G$. This can be extended to a shortest path metric on Ω_G (the symmetric difference).

Let (X_t, Y_t) be a coupling for the polymer dynamics where both chains make the same random choices. Suppose that $D(X_t, Y_t) = 1$ and that $X_t = Y_t \cup \{\gamma\}$ for some $\gamma \in \mathcal{C}_G$. We obtain (X_{t+1}, Y_{t+1}) from (X_t, Y_t) by performing an update according to the polymer dynamics, and consider the various possibilities for $D(X_{t+1}, Y_{t+1})$. With probability $|E_\gamma|/2|E_G|$ we choose γ and remove it from both X_t and Y_t – in this case $D(X_{t+1}, Y_{t+1}) = 0$. Alternatively, if we add a polymer γ' such that $\gamma' \nsim \gamma$, then the update will be rejected in one chain and performed in the other and $D(X_{t+1}, Y_{t+1}) \le 2$ – this occurs with probability $|E_{\gamma'}|/2|E_G| \cdot w_G(\gamma')$. Combining these two cases, we obtain that

$$\mathbb{E}[D(X_{t+1}, Y_{t+1}) \mid (X_t, Y_t)] \le 1 + \frac{1}{2|E_G|}\left(-|E_\gamma| + \sum_{\gamma' \nsim \gamma} |E_{\gamma'}| \cdot w_G(\gamma')\right),$$

and therefore by the polymer mixing condition that

$$\mathbb{E}[D(X_{t+1}, Y_{t+1}) \mid (X_t, Y_t)] \le 1 - \frac{|E_\gamma|(1-\theta)}{2|E_G|} \le 1 - \frac{(1-\theta)}{2|E_G|},$$

for some constant $\theta \in (0,1)$. Let W be the diameter of Ω_G with respect to the metric $D(\cdot,\cdot)$ and note that $W \le 2n$. By the path coupling lemma [8, Section 6], it follows that the mixing time is at most $\log(W/\varepsilon)2|E_G|/(1-\theta) = O(|E_G| \log(|E_G|/\varepsilon))$. □

4.2. Single polymer sampler

We now show how to use the polymer dynamics as part of an efficient algorithm for computing an approximate sample from μ_G. The input to this algorithm will be a graph and an accuracy parameter, and we will assume that we are given the graph both as an adjacency matrix and an array of linked adjacency lists, where each adjacency list contains all of the neighbours of a vertex. These data structures will allow us to decide whether two given vertices are adjacent in constant time (in the unit cost model), and will also allow us to access the next entry in any adjacency list in constant time. Other choices of data structures will increase the running time of our algorithms by at most a polynomial factor.

Fix $e \in E_G$. We now show how to sample a polymer from ν_e in expected constant time. Let $r = \tau - \log(12e^2(q-1)) > 0$, where τ is as in the polymer sampling condition (Definition 5). For $\ell \ge 0$ let $\mathcal{A}_\ell(e) = \{\gamma \in \mathcal{A}(e) : \deg_G(V_\gamma) \le \ell\}$. We use the following algorithm to sample a polymer from ν_e.

Algorithm 1: Sampling a polymer from ν_e.

1. Sample $\boldsymbol{\ell}$ from the following distribution: $\Pr(\boldsymbol{\ell} = \ell) = (1 - e^{-r})e^{-r\ell}$ for all integers $\ell \ge 1$. This is so that $\Pr(\boldsymbol{\ell} \ge \ell) = e^{-r\ell}$.
2. Enumerate $\mathcal{A}_{\boldsymbol{\ell}}(e)$ and compute $w_G(\gamma)$ for each $\gamma \in \mathcal{A}_{\boldsymbol{\ell}}(e)$.
3. Mutually exclusively, output each $\gamma \in \mathcal{A}_{\boldsymbol{\ell}}(e)$ with probability $w_G(\gamma) \cdot e^{r \deg_G(\gamma)}$. With all remaining probability, output $\emptyset$.

In order to perform the second step of the above algorithm, we must be able to enumerate all connected vertex subsets $S \subseteq V_G$ with $\deg_G(S) \le \boldsymbol{\ell}$, that contain an endpoint of e. In order to do this efficiently, we require the following result which is an adaptation of Lemma 3.4 of [23].

[6] Recall that for an ergodic Markov chain with a finite state space Ω and transition matrix P, the mixing time to its stationary distribution μ is defined as $\max_{\omega \in \Omega} \min\{t > 0 : \|P^t(\omega, \cdot) - \mu\|_{TV} \le \varepsilon\}$, where $\|\pi - \pi'\|_{TV}$ is the total variation distance between the probability distributions π and π'.

Lemma 21. *Let $G = (V_G, E_G)$ be a graph, let $v \in V_G$, and let $\ell \geq 1$. There is an algorithm which enumerates all connected vertex subsets $S \subseteq V_G$ such that $v \in S$ and $\deg_G(S) \leq \ell$. The running time of this algorithm is $O(\ell^7(2e)^{4\ell})$.*

Proof. Let $C(G, v, \ell)$ be the set of all connected vertex subsets $S \subseteq V_G$ such that $v \in S$ and $\deg_G(S) \leq \ell$. For $\ell < \deg_G(v)$, we simply output $\emptyset$. For $\ell \geq \deg_G(v)$, we construct $C(G, v, \ell)$ recursively. For the base case, observe that $C(G, v, \deg_G(v))$ is $\{v\}$. For $\ell' \geq \deg_G(v)$, in order to construct $C(G, v, \ell'+1)$ given $C(G, v, \ell')$, we first construct the multiset

$$C''(G, v, \ell'+1) = \{S \cup \{u\} : S \in C(G, v, \ell'),\ u \in \partial_G S,\ \deg_G(S \cup \{u\}) = \ell'+1\},$$

then remove the repeat elements from it to obtain $C'(G, v, \ell'+1)$, and finally set $C(G, v, \ell'+1) = C'(G, v, \ell'+1) \cup C(G, v, \ell')$.

To construct $C''(G, v, \ell'+1)$, we consider each element $S \in C(G, v, \ell')$ and each vertex $u \in \partial_G S$. Iterating through $C(G, v, \ell')$ requires $O(\ell'(2e)^{2\ell'-1})$ time, by Lemma 6. For each $S \in C(G, v, \ell')$, we can iterate through $\partial_G S$ in $O(\ell')$ time, given that we have constant-time access to the elements of the adjacency list of each vertex in S. The total time required is therefore $O(\ell'^2(2e)^{2\ell'-1})$. We then remove the repeat elements from $C''(G, v, \ell'+1)$ by pairwise comparison, which takes $O((\ell'+1)^2 \cdot |C''(G, v, \ell'+1)|^2) = O(\ell'^6(2e)^{4\ell'})$ time.

To construct $C(G, v, \ell)$, we begin with the base case and then perform the above $\ell - \deg_G(v)$ times. The total running time is therefore $O(\ell^7(2e)^{4\ell})$. We now show that the algorithm returns every element of $C(G, v, \ell)$. For the base case, it is clearly true. Furthermore, for all $\ell' > \deg_G(v)$ and all $S \in C(G, v, \ell')$, we know that S consists of $T \in C(G, v, \ell'')$ and $u \in \partial_G T$, for some $\ell'' < \ell'$. □

We can now show the correctness and expected constant running time of Algorithm 1.

Lemma 22. *Let $q \geq 2$ be an integer, $\mathcal{G}$ be a class of graphs, and $\{(\mathcal{C}_G, w_G) \mid G \in \mathcal{G}\}$ be a family of computationally feasible q-spin polymer models that satisfies the polymer sampling condition (see Definition 5). For all $G \in \mathcal{G}$ and all $e \in E_G$, Algorithm 1 samples a polymer from ν_e in expected constant time.*

Proof. We begin by showing that step 3 of the algorithm is well-defined, by showing that $\sum_{\gamma \in \mathcal{A}(e)} w_G(\gamma) \cdot e^{r \deg_G(V_\gamma)} \leq 1$ for all $e \in E_G$. For all $v \in V_G$, we have that

$$\sum_{\gamma : u \in V_\gamma} w_G(\gamma) \cdot e^{r \deg_G(V_\gamma)} \leq \sum_{\ell \geq 1} \sum_{\substack{\gamma : u \in V_\gamma, \\ \deg_G(V_\gamma) = \ell}} w_G(\gamma) \cdot e^{r\ell} \leq \sum_{\ell \geq 1} (4e^2)^\ell (q-1)^\ell e^{-\tau\ell} e^{r\ell} \leq \frac{1}{2},$$

where the second-to-last inequality follows from Lemma 6 and the final inequality follows from the fact that $4e^2(q-1)e^{-\tau+r} \leq 1/3$.

We now show that the expected running time of the algorithm is constant. For $\ell \geq 1$, the time taken to enumerate $\mathcal{A}_\ell(e)$ is $O((q-1)^\ell \ell^7 (2e)^{4\ell})$, by Lemma 21. The time taken to then iterate through $\mathcal{A}_\ell(e)$ and compute $w_G(\gamma)$ for each $\gamma \in \mathcal{A}_\ell(e)$ is $O((q-1)^\ell \ell (2e)^{2\ell-1} e^\ell)$, by Lemma 6 and the fact that the family of polymer models is computationally feasible. The expected running time of Algorithm 1 is therefore

$$= O\left(\sum_{\ell \geq 1} \Pr(\boldsymbol{\ell} = \ell) \left((q-1)^\ell \ell^7 (2e)^{4\ell} + (q-1)^\ell \ell (2e)^{2\ell-1} e^\ell \right) \right)$$

$$= O\left(\sum_{\ell \geq 1} \ell^7 \left(\frac{16e^4(q-1)}{e^r} \right)^\ell \right).$$

Since $r = \tau - \log(12e^2(q-1)) \geq 3\log(8e^3(q-1)) - \log(12e^2(q-1)) \geq 2\log(8e^3(q-1))$, it follows that $16e^4(q-1)/e^r < 1/2$, and therefore that the expected running time of Algorithm 1 is $O(1)$.

Finally, let $\gamma \in \mathcal{A}(e)$ be a polymer. In order for the algorithm to sample γ, it must first sample $\boldsymbol{\ell} \geq \deg_G(V_\gamma)$ in step 1 of the algorithm, then conditioned on this choice, it must output γ in step 3. This occurs with probability $e^{-r \deg_G(V_\gamma)} w_G(\gamma) \cdot e^{r \deg_G(V_\gamma)} = w_G(\gamma)$, and therefore the output distribution is ν_e. □

We now combine the polymer dynamics with Algorithm 1 to give an efficient algorithm for computing an approximate sample from the Gibbs distribution of a polymer model.

Lemma 8. *Let $q \geq 2$ be an integer, $\mathcal{G}$ be a class of graphs, and $\mathcal{F}_\mathcal{G}$ be a family of computationally feasible q-spin polymer models satisfying the polymer sampling condition.*

There are randomised algorithms which, given as input a graph $G \in \mathcal{G}$ with m edges and an accuracy parameter $\varepsilon > 0$, output an ε-sample from μ_G in $O\left(m \log \frac{m}{\varepsilon} \log \frac{1}{\varepsilon}\right)$ time, and an ε-approximation to Z_G, with probability at least $3/4$, in $O\left(m^2 \log(\frac{m}{\varepsilon})^3\right)$ time.

Proof. We focus on the sampling algorithm, since it is shown in [7, Section 3] how to convert this into the desired counting algorithm for the polymer model (although the approach in [7] was stated for bounded-degree graphs, it carries over to graphs of unbounded degree).

By Lemma 19, the unique stationary distribution of the polymer dynamics is μ_G, thus our sampling algorithm is based on the polymer dynamics. By Lemma 22, there is an integer $C_1 > 1$ (independent of G) such that the expected number of steps taken to execute a single update of the polymer dynamics is at most C_1. Lemma 18 shows that $\mathcal{F}_\mathcal{G}$ also satisfies the polymer mixing condition. This allows us to apply Lemma 20, which shows that there is an integer $C_2 > 1$ (also independent of G) such that the mixing time of the polymer dynamics satisfies $\tau_{\text{mix}}(\varepsilon/2) \leq C_2 m \log \frac{m}{\varepsilon}$.

We use the following algorithm to compute an ε-sample from μ_G. Repeat the following $\lceil \log(2/\varepsilon) \rceil$ times, and if no polymer configuration is returned, return $\emptyset$. Run the polymer dynamics for $3C_1C_2m\left\lceil \log \frac{m}{\varepsilon} \right\rceil$ steps, starting from $\emptyset$, and if at least $C_2 m \lceil \log \frac{m}{\varepsilon} \rceil$ updates of the polymer dynamics were executed, then return the configuration.

We claim that the distribution of the output configuration from the above algorithm is within total variation distance ε of μ_G. This will follow once we have shown that the probability that no run of the polymer dynamics returns a configuration, is at most $\varepsilon/2$. This is because the configuration we would output if one is returned by a run of the polymer dynamics, is within total variation distance $\varepsilon/2$ of μ_G.

Consider one of the $\lceil \log(2/\varepsilon) \rceil$ independent runs of the polymer dynamics that is made by the algorithm. Let the random variable X denote the total number of steps required to execute $C_2 m \left\lceil \log \frac{m}{\varepsilon} \right\rceil$ steps of the polymer dynamics. We have that $\mathbb{E}[X] \leq C_1 C_2 m \left\lceil \log \frac{m}{\varepsilon} \right\rceil$, and therefore by Markov's inequality that $\Pr(X \geq 3\mathbb{E}[X]) \leq 1/3 < 1/\mathrm{e}$. Since the runs are independent, the probability that no run performs at least $C_2 m \left\lceil \log \frac{m}{\varepsilon} \right\rceil$ updates of the polymer dynamics, is at most $(1/\mathrm{e})^{\log(2/\varepsilon)} = \varepsilon/2$. □

5. Details of polymer application for the ferromagnetic Potts model

It is known that if G is an α-expander then most of the weight of $Z_{G,q,\beta}$ is contributed by colourings that colour more than half of the vertices of G with a single colour. The following definitions apply to these colourings.

Definition 23. Let $G = (V_G, E_G)$ be a graph. Let q be a positive integer, let $r \in [q]$, and let $\beta > 0$. We define $\Omega^r_{G,q}$ to be set of q-colourings of G that colour more than half of the vertices of G with r. We also define $Z^r_{G,q,\beta} = \sum_{\sigma \in \Omega^r_{G,q}} \mathrm{e}^{\beta m_G(\sigma)}$.

The following result is due to Jenssen, Keevash, and Perkins [20]. We note that although this result is applied to bounded-degree graphs in [20], it remains true for arbitrary α-expanders.

Lemma 24. *[20, Lemma 12] Let $\alpha > 0$ and let $G = (V_G, E_G)$ be an α-expander. Let $q \geq 2$ be an integer, let $r \in [q]$ be any spin, and let $\beta > 2\log(eq)/\alpha$ be a real number. We have that $q \cdot Z^r_{G,q,\beta}$ is an $\mathrm{e}^{-|V_G|}$-approximation to $Z_{G,q,\beta}$.*

Consider the polymer model defined in Example 4. Let $\hat{\Omega}^r_{G,q}$ denote the set of all sets of mutually compatible allowed polymers, let $\hat{\mu}^r_{G,q,\beta}$ denote the Gibbs distribution of the polymer model, and let $\hat{Z}^r_{G,q,\beta}$ denote its partition function. Observe that there is a bijection between the polymer configurations of $\hat{\Omega}^r_{G,q}$ and the Potts configurations of $\Omega^r_{G,q}$, where a colouring $\sigma \in \Omega^r_{G,q}$ maps to the polymer configuration $\Gamma \in \hat{\Omega}^r_{G,q}$ consisting of the connected components of vertices that do not get colour r under σ. Moreover, the weight of each $\Gamma \in \hat{\Omega}^r_{G,q}$ is closely related to the weight of the Potts configuration $\sigma \in \Omega^r_{G,q}$ by

$$\mathrm{e}^{\beta|E_G|} \prod_{\gamma \in \Gamma} w_{G,\beta}(\gamma) = \mathrm{e}^{\beta m_G(\sigma)}.$$

Therefore sampling from this polymer model is equivalent to sampling from the Potts model restricted to colourings which colour more than half of the vertices with r.

We can now prove the following theorem, which gives an efficient algorithm for sampling from the low temperature Potts model on expanders.

Theorem 9. *Let $\alpha > 0$ be a real number. Let $q \geq 3$ be an integer and $\beta \geq \frac{3}{\alpha}\log(8\mathrm{e}^3(q-1))$ be a real. For the Potts model on $G \in \mathcal{G}_\alpha$, there is a poly-time approximate sampling algorithm for $\mu_{G,q,\beta}$ and an FPRAS for $Z_{G,q,\beta}$.*

In fact, for $n = |V_G|$ and $m = |E_G|$, if the desired accuracy ε satisfies $\varepsilon \geq \mathrm{e}^{-n}$ then the running time of the sampler is $O(m \log \frac{m}{\varepsilon} \log \frac{1}{\varepsilon})$ and the running time of the FPRAS is $O\left(m^2 (\log \frac{m}{\varepsilon})^3\right)$.

Proof. Assume without loss of generality that G is connected (otherwise, consider the connected components separately). Consider the family of polymer models defined in Example 4. It is computationally feasible since determining whether

$\gamma \in \mathcal{C}^r_{G,q}$ can be done in $O(|V_\gamma|) = O(\exp(\deg_G(V_\gamma)))$ time (one just needs to check whether $|V_\gamma| < n/2$). Computing $w_{G,\beta}(\gamma) = \mathrm{e}^{-\beta B_\gamma}$ can be done by examining all $O(\deg_G(V_\gamma))$ edges with endpoints in V_γ, by iterating through V_γ and counting the neighbours of each vertex that are not in V_γ. The total running time required to do this is therefore $O(|V_\gamma|^2 \deg_H(V_\gamma)) = O(\exp(\deg_H(V_\gamma)))$. As is shown in (1), since $\beta \geq \frac{3}{\alpha}\log(8\mathrm{e}^3(q-1))$, it also satisfies the polymer sampling condition (see Definition 5) with constant $\tau = \alpha\beta$.

The input to both algorithms is G and ε. If $\varepsilon < \mathrm{e}^{-n}$ then we can construct an ε-sample from $\mu_{G,q,\beta}$ in $\mathrm{poly}(n, \frac{1}{\varepsilon})$ time, by brute force. If $\varepsilon \geq \mathrm{e}^{-n}$ then we can construct an ε-sample from $\mu_{G,q,\beta}$ as follows. Choose a spin $r \in [q]$ uniformly at random. By Lemma 8 there is an algorithm which, given ε/q and G as input, outputs an (ε/q)-sample from $\hat{\mu}^r_{G,q,\beta}$ (and therefore also from $\mu^r_{G,q,\beta}$) in $O\left(m \log\frac{m}{\varepsilon}\log\frac{1}{\varepsilon}\right)$ time. Since G is also an α-expander, then by Lemma 24, this is an ε-sample from $\mu_{G,q,\beta}$. For the counting algorithm, if $\varepsilon < \mathrm{e}^{-n}$ then we can compute $Z_{G,q,\beta}$ exactly in $\mathrm{poly}(n, \frac{1}{\varepsilon})$ time, by brute force. If $\varepsilon \geq \mathrm{e}^{-n}$, then by Lemma 8, we can compute an ε-approximation to $Z^r_{G,q,\beta}$ in $O\left(m^2(\log\frac{m}{\varepsilon})^3\right)$ time. By Lemma 24, it follows that $q\mathrm{e}^{\beta m} \cdot Z^r_{G,q,\beta}$ is an ε-approximation to $Z_{G,q,\beta}$. □

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Alexander Barvinok, Combinatorics and Complexity of Partition Functions, vol. 9, Springer, 2016.
[2] Alexander Barvinok, Guus Regts, Weighted counting of solutions to sparse systems of equations, Comb. Probab. Comput. 28 (5) (2019) 696–719.
[3] Béla Bollobás, The Art of Mathematics: Coffee Time in Memphis, Cambridge University Press, 2006.
[4] Christian Borgs, Jennifer Chayes, Tyler Helmuth, Will Perkins, Prasad Tetali, Efficient sampling and counting algorithms for the Potts model on $\mathbb{Z}^d$ at all temperatures, in: Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, 2020, pp. 738–751.
[5] Sarah Cannon, Will Perkins, Counting independent sets in unbalanced bipartite graphs, in: Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 2020, pp. 1456–1466.
[6] Charles Carlson, Ewan Davies, Alexandra Kolla, Efficient algorithms for the Potts model on small-set expanders, arXiv preprint arXiv:2003.01154, 2020.
[7] Zongchen Chen, Andreas Galanis, Leslie Ann Goldberg, Will Perkins, James Stewart, Eric Vigoda, Fast algorithms at low temperatures via Markov chains, Random Struct. Algorithms 58 (2) (2021) 294–321, Theorems 5 and 6 from arXiv:1901.0665 [org/abs].
[8] Martin Dyer, Catherine Greenhill, Random Walks on Combinatorial Objects, London Mathematical Society Lecture Note Series, 1999, pp. 101–136.
[9] Nikolaos Fountoulakis, Bruce A. Reed, The evolution of the mixing rate of a simple random walk on the giant component of a random graph, Random Struct. Algorithms 33 (1) (2008) 68–86.
[10] Tobias Friedrich, Andreas Göbel, Martin S. Krejca, Marcus Pappik, A Spectral Independence View on Hard-Spheres via Block Dynamics, 2021.
[11] Andreas Galanis, Leslie Ann Goldberg, James Stewart, Fast algorithms for general spin systems on bipartite expanders, in: 45th International Symposium on Mathematical Foundations of Computer Science (MFCS 2020), 2020, 37.
[12] Andreas Galanis, Daniel Stefankovic, Eric Vigoda, Linji Yang, Ferromagnetic Potts model: refined #BIS-hardness and related results, SIAM J. Comput. 45 (6) (2016) 2004–2065.
[13] Leslie Ann Goldberg, Mark Jerrum, Approximating the partition function of the ferromagnetic Potts model, J. ACM 59 (5) (2012) 1–31.
[14] Christian Gruber, Hervé Kunz, General properties of polymer systems, Commun. Math. Phys. 22 (2) (1971) 133–161.
[15] Tyler Helmuth, Matthew Jenssen, Will Perkins, Finite-size scaling, phase coexistence, and algorithms for the random cluster model on random graphs, arXiv preprint arXiv:2006.11580, 2020.
[16] Tyler Helmuth, Will Perkins, Guus Regts, Algorithmic Pirogov–Sinai theory, Probab. Theory Relat. Fields 176 (3) (2020) 851–895.
[17] Jeroen Huijben, Viresh Patel, Guus Regts, Sampling from the low temperature Potts model through a Markov chain on flows, CoRR, arXiv:2103.07360 [abs], 2021.
[18] Svante Janson, The probability that a random multigraph is simple. II, J. Appl. Probab. 51 (A) (2014) 123–137.
[19] Matthew Jenssen, Peter Keevash, Homomorphisms from the Torus, 2020.
[20] Matthew Jenssen, Peter Keevash, Will Perkins, Algorithms for #BIS-hard problems on expander graphs, SIAM J. Comput. 49 (4) (2020) 681–710.
[21] Roman Kotecký, David Preiss, Cluster expansion for abstract polymer models, Commun. Math. Phys. 103 (3) (1986) 491–498.
[22] Chao Liao, Jiabao Lin, Pinyan Lu, Zhenyu Mao, Counting independent sets and colorings on random regular bipartite graphs, in: Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2019), 2019, 34.
[23] Viresh Patel, Guus Regts, Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials, SIAM J. Comput. 46 (6) (2017) 1893–1919.
[24] Remco Van Der Hofstad, Random Graphs and Complex Networks, vol. 1, Cambridge University Press, 2016.