CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Quasipolynomial-time algorithms for Gibbs point processes

Matthew Jenssen[1] (iD), Marcus Michelen[2] and Mohan Ravichandran[3]

[1]Department of Mathematics, King's College London, London, UK, [2]Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, Chicago, IL, USA, and [3]Department of Mathematics, Bogazici University, Istanbul, Turkey
Corresponding author: Matthew Jenssen; Email: matthew.jenssen@kcl.ac.uk

Abstract

We demonstrate a quasipolynomial-time deterministic approximation algorithm for the partition function of a Gibbs point process interacting via a stable potential. This result holds for all activities λ for which the partition function satisfies a zero-free assumption in a neighbourhood of the interval $[0, \lambda]$. As a corollary, for all finiterange stable potentials, we obtain a quasipolynomial-time deterministic algorithm for all $\lambda < 1/(e^{B+1}\hat{C}_\phi)$ where $\hat{C}_\phi$ is a temperedness parameter and B is the stability constant of ϕ. In the special case of a repulsive potential such as the hard-sphere gas we improve the range of activity by a factor of at least e^2 and obtain a quasipolynomial-time deterministic approximation algorithm for all $\lambda < e/\Delta_\phi$, where Δ_ϕ is the potential-weighted connective constant of the potential ϕ. Our algorithm approximates coefficients of the cluster expansion of the partition function and uses the interpolation method of Barvinok to extend this approximation throughout the zero-free region.

Keywords: Gibbs point process; deterministic approximation algorithm; cluster expansion

2020 MSC Codes: Primary: 60G55, 60K35, 68W25; Secondary: 82B30

1. Introduction

Gibbs point processes are a fundamental model of random spatial phenomena in the continuum. Most classically, such processes are used to model a gas under local interactions (see Ruelle's [31] text). Beyond that, Gibbs point processes are used to model various physical phenomena that often exhibit local repulsion, such as the locations of galaxies in the universe, the time and place of earthquakes, and the growth of trees in a forest; see [5, 22] for these applications and more. A simple and well-studied example of a Gibbs point process is the *hard-sphere model*, where one samples a Poisson point process in a set of finite volume in $\mathbb{R}^d$ and conditions on no two points having distance less than some parameter $r > 0$. In order to better understand these models one often wants to approximately compute the partition function of the model, which may be understood as a weighted count of allowable configurations of points. The partition function grows exponentially in the volume, making exact computation intractable even for basic examples. Additionally, the rate of exponential growth is equal to the infinite-volume pressure, a central quantity in statistical physics.

Approximating the partition function and sampling – either approximately or exactly – are the two main algorithmic problems associated with Gibbs point processes. Under very mild assumptions, polynomial-time approximate sampling of the point process can be used to provide a *randomised* approximation to the partition function. Many techniques have been applied

to Gibbs point processes for approximate and exact sampling in certain regimes; in fact, the seminal Metropolis-Hastings algorithm was developed to sample from the hard-sphere model in dimension 2 [21].

On the other hand, *deterministic* approximation algorithms for partition functions of Gibbs point processes are less well-understood. For Gibbs point processes, to our knowledge the only rigorous result giving a deterministic algorithm is that of Friedrich, Göbel, Katzmann, Krejca and Pappik [6], which shows that for the special case of hard spheres, one may approximate the partition function in quasipolynomial time for a certain range of parameters. The authors show that the hard-sphere model can be well approximated by its discrete analogue (the hard-core model) allowing them to apply known algorithmic results from the discrete setting (see Section 1.3 for more detail).

In this paper, we provide a quasipolynomial-time deterministic approximation algorithm for the partition function for a general class of stable Gibbs point processes. Our main result is stated only under the assumption of zero-freeness of the partition function; from there, we deduce two main corollaries using existing zero-freeness results from the literature, one which applies to all stable potentials and a stronger result that applies to the more restricted class of repulsive potentials. We defer formal statements to Section 1.2.

Our approach is via Barvinok's interpolation method [2] combined with use of the cluster expansion for Gibbs point processes. This allows us to work with the Gibbs process directly, rather than a discrete approximation of the process. By combining our result with the zero-free region for stable potentials guaranteed by the cluster expansion, we obtain what appears to be the first algorithmic result for stable, non-repulsive potentials. In the special case of a repulsive potential, combining our result with previous work of the second author and Perkins on zero-freeness [19] yields the first quasipolynomial-time deterministic approximation algorithm for a large class of repulsive potentials (which includes the hard-sphere model) and range of parameters. For the hard-sphere potential ϕ, we note that [19, Lemma 12] gives an explicit bound of $\Delta_\phi < C_\phi(1 - 8^{-d-1})$ where Δ_ϕ, C_ϕ denote the potential-weighted connective constant and temperedness constant of ϕ, respectively (defined in the next sections). This demonstrates that our algorithm works for a wider range of parameters than the previous deterministic algorithms [6]. Additionally, [7] argues that the connective constant of the discretization used in [6, 7] would not provide an improvement to their results; as such, working in the continuum and using the zero-freeness result of [19] gets around this issue.

1.1. Formal definition of the model

The point processes we consider are defined by three parameters:

- a measurable set $S \subset \mathbb{R}^d$ of finite volume,
- a parameter $\lambda \geqslant 0$ referred to as the *activity* or *fugacity*,
- a *pair potential* $\phi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ satisfying $\phi(x) = \phi(-x)$.

The *temperedness constant* of a potential is defined as

$$C_\phi = \int_{\mathbb{R}^d} |1 - e^{-\phi(x)}| \, dx. \tag{1}$$

The temperedness constant may be understood as a measure of the strength of the potential. We say that a potential ϕ is *tempered* if $C_\phi < \infty$, and always assume that ϕ is tempered. The *energy* of a configuration of points $\{x_1, \ldots, x_N\} \subset \mathbb{R}^d$ is defined by

$$H(x_1, \ldots, x_N) = \sum_{1 \leqslant i < j \leqslant N} \phi(x_i - x_j). \tag{2}$$

We will always assume that ϕ is *stable*, meaning that there is a constant $B \geqslant 0$ so that for all N and $x_1, \ldots, x_N$ we have

$$H(x_1, \ldots, x_N) \geqslant -BN. \tag{3}$$

The infimum over such B is called the *stability constant* of ϕ. The assumption of stability is used to show that the partition function – and thus the Gibbs point process – itself is well-defined. Under certain conditions on ϕ, the assumption of stability is a necessary condition for the point process to be well-defined (see [31, Section 3.2] for a detailed discussion and many examples). A potential ϕ is *repulsive* if $\phi(x) \geqslant 0$ for all x. In particular, repulsive potentials are stable with stability constant $B = 0$.

The *Gibbs point process* in S with potential ϕ at activity λ is the point process in S whose density against the Poisson point process of activity λ is proportional to $e^{-H(x_1, \ldots, x_N)}$. The *grand canonical partition function* at activity λ is defined by

$$Z_S(\lambda) = \sum_{k \geqslant 0} \frac{\lambda^k}{k!} \int_{S^k} e^{-H(x_1, \ldots, x_k)} \, dx_1 \, \ldots \, dx_k. \tag{4}$$

Throughout, we work with the case of $S = \Lambda_n := [-n, n]^d \subset \mathbb{R}^d$, i.e. the axis parallel box of side-length $2n$. One of the most studied examples of a Gibbs point process is the *hard-sphere model,* which is defined by the potential

$$\phi(x) = \begin{cases} +\infty & \text{if } \|x\|_2 < r \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

for fixed $r > 0$. The hard-sphere model is supported on configurations $\{x_1, \ldots, x_N\}$ such that $H(x_1, \ldots, x_N) = 0$ i.e. configurations consisting of the centres of a packing of spheres of radius $r/2$.

Another similar example is the *Strauss potential*, in which the $+\infty$ in the definition of the hard-sphere potential is replaced with a parameter $a > 0$.

Among the most common examples of a stable potential that is not repulsive is a *Lennard-Jones potential* (see [31, Section 3.2.10]). While there are many examples of potentials that are called Lennard-Jones potentials, they are characterised by being strongly repulsive at short distances and weakly attractive at far distances. A large family of widely used potentials is of the form

$$\phi(x) = A\|x\|_2^{-2\alpha} - B\|x\|_2^{-\alpha}$$

for parameters $A, B > 0$ and $\alpha > d$, where we recall d is the underlying dimension. One may also truncate this potential to be of finite range by simply setting it equal to 0 if $\|x\|_2 \geqslant T$ for some parameter T, and this still yields a stable potential.

1.2. Statement of results

Our results will require only two additional assumptions on the potential: first a basic assumption on its form so that we may approximately compute certain volumes; and second a zero-freeness assumption. The zero-freeness assumption we use for stable (non-repulsive) potentials is a classical result that follows from the cluster expansion [28, 31], while in the repulsive case, we will use the work of the second author and Perkins [19]. We begin with the assumption required for computational purposes.

Assumption 1.1. *There are compact centrally symmetric convex sets* $\{0\} = K_0 \subset K_1 \subset K_2 \subset \ldots \subset K_\ell$ *so that the function* $x \mapsto \exp(-\phi(x))$ *is L-Lipschitz on each of* $\Delta_j := K_j \setminus K_{j-1}$. *Additionally assume that there is an* $R > 0$ *so that* $\mathrm{supp}(\phi) \subset K_\ell \subset [-R, R]^d$ *and* $[-1/R, 1/R]^d \subset K_1$. *We work*

in a real-valued model of computation, and assume unit cost for elementary operations; we also assume that evaluation of $e^{-\phi}$ as well as checking if a point lies in a given K_j have unit cost.

Assumption 1.1 essentially has two elements: it assumes that the potential ϕ has a certain amount of (piecewise) regularity, and also assumes that ϕ is of finite range. We are not aware of any natural potential ϕ in the literature that fails to satisfy the piecewise regularity portion of the assumption. There are however many natural potentials that are not of finite range.

We remark that the hard-sphere potential (5), the Strauss potential, and the truncated Lennard-Jones potential are easily seen to satisfy Assumption 1.1.

1.2.1. A note on asymptotic notation

Throughout the paper, we think of the parameters d, ℓ, L, and R from Assumption 1.1 as fixed and allow the implicit constants in our asymptotic notation to depend on these parameters.

Our main theorem is that we obtain deterministic approximation algorithms for $Z_S(\lambda)$ for potentials satisfying Assumption 1.1 and the following zero-freeness assumption on $Z_S(\lambda)$.

Assumption 1.2. *We say a stable potential ϕ satisfies Assumption 1.2 at $\lambda_0 > 0$ if there exist constants $\delta, C > 0$ so that the following holds. For all bounded, measurable $S \subset \mathbb{R}^d$ we have*

$$Z_S(\lambda) \neq 0 \quad and \quad \frac{1}{|S|}|\log Z_S(\lambda)| \leqslant C \qquad for \ all \ \lambda \in \mathcal{N}_\delta([0, \lambda_0]), \tag{6}$$

where

$$\mathcal{N}_\delta([0, \lambda_0]) = \{z \in \mathbb{C} : d(z, [0, \lambda_0]) < \delta\}.$$

Assumption 1.2 may be understood as saying that the Gibbs point process exhibits no phase transition in the regime $[0, \lambda_0]$ in the Lee-Yang sense (see Section 1.3).

Given a number Z, an ε-approximation to Z is a value $\hat{Z}$ so that $e^{-\varepsilon}\hat{Z} \leqslant Z \leqslant e^{\varepsilon}\hat{Z}$. Our main theorem asserts a quasipolynomial-time approximation for the partition function under these two assumptions. Recall that $\Lambda_n := [-n, n]^d \subset \mathbb{R}^d$.

Theorem 1.3. *Let ϕ be a stable pair potential that satisfies the regularity Assumption 1.1. Suppose ϕ satisfies the zero-freeness Assumption 1.2 at $\lambda \geqslant 0$. Let $\varepsilon \in (0, 1)$, $n \in \mathbb{N}$. Then there is a deterministic algorithm to compute an ε-approximation to $Z_{\Lambda_n}(\lambda)$ with runtime at most $\exp(O(\log^3(|\Lambda_n|/\varepsilon)))$.*

We note that the implicit constant in the exponent depends on the potential ϕ as well as the activity λ.

We will deduce two main corollaries from Theorem 1.3. The first of which deduces an algorithm for all stable potentials in the regime of cluster expansion convergence. The fact that for $\lambda < (e^{1+2B}C_\phi)^{-1}$ a stable potential ϕ satisfies Assumption 1.2 is a direct consequence of the cluster expansion (see, e.g., (11) or [31, Thm 4.2.3]). This classical result was recently improved by Procacci and Yuhjtman [28] who showed that any stable tempered potential ϕ satisfies Assumption 1.2 at all $\lambda < (e^{1+B}\hat{C}_\phi)^{-1}$, where

$$\hat{C}_\phi = \int_{\mathbb{R}^d} 1 - e^{-|\phi(x)|}\, dx. \tag{7}$$

We note that $\hat{C}_\phi \leqslant C_\phi$ for all ϕ with equality if ϕ is repulsive.

Corollary 1.4. *Let ϕ be a stable tempered potential satisfying Assumption 1.1, let B denote its stability constant and let $\hat{C}_\phi$ be as in (7). Let $\varepsilon \in (0, 1)$, $n \in \mathbb{N}$. Then for all $\lambda < (e^{1+B}\hat{C}_\phi)^{-1}$, there is a deterministic algorithm to compute an ε-approximation to $Z_{\Lambda_n}(\lambda)$ with runtime at most $\exp(O(\log^3(|\Lambda_n|/\varepsilon)))$.*

Our second corollary applies for repulsive potentials – i.e. the case of $B = 0$ – for a much wider range of parameters. To apply Theorem 1.3 in the repulsive case, we will use pre-existing work on zero-freeness of Gibbs point process partition functions. In [19], the potential-weighted connective constant Δ_ϕ was introduced, which captures a relationship between the strength of the potential and the geometry of the underlying space.

First, define V_k via

$$V_k = \int_{(\mathbb{R}^d)^k} \prod_{j=1}^{k} \left[\exp\left(-\sum_{i=0}^{j-2} \mathbf{1}_{\|v_j - v_i\| < \|v_i - v_{i+1}\|} \phi(v_j - v_i) \right) \cdot (1 - e^{-\phi(v_j - v_{j-1})}) \right] d\mathbf{v} \qquad (8)$$

where we write $d\mathbf{v} = dv_1 \, dv_2 \, \ldots \, dv_k$ and interpret $v_0 = 0$, and in the case of $j = 1$ interpret the empty sum as equal to 0. Since the potential ϕ is repulsive, the sequence $\{V_k\}_{k \geqslant 1}$ is submultiplicative and so we may define the *potential-weighted connective constant* Δ_ϕ via

$$\Delta_\phi = \lim_{k \to \infty} V_k^{1/k} = \inf_{k \geqslant 1} V_k^{1/k}. \qquad (9)$$

For any repulsive potential ϕ we have $\Delta_\phi \leqslant C_\phi$ where C_ϕ is the temperedness constant defined at (1). So long as ϕ is non-trivial we in fact have $\Delta_\phi < C_\phi$. In [19] it was shown that Assumption 1.2 is satisfied for each tempered repulsive potential for all $\lambda_0 < e/\Delta_\phi$. We, therefore, obtain the following immediate corollary.

Corollary 1.5. *Let ϕ be a repulsive tempered potential satisfying Assumption 1.1 and let Δ_ϕ denote its potential-weighted connective constant. Let $\varepsilon \in (0, 1)$, $n \in \mathbb{N}$. Then for all $\lambda < e/\Delta_\phi$, there is a deterministic algorithm to compute an ε-approximation to $Z_{\Lambda_n}(\lambda)$ with runtime at most* $\exp(O(\log^3(|\Lambda_n|/\varepsilon)))$.

1.3. Context and related work

Much of the classical work on Gibbs point processes has consisted of showing the absence of a phase transition when λ is small. There are many definitions and notions of a phase transition. One of the most robust definitions is due to Lee and Yang [13, 33], which states that a phase transition is a point at which the pressure fails to be analytic. Often, various other definitions of phase transitions, e.g. in terms of infinite-volume Gibbs measures, can be shown to coincide with the Lee-Yang definition (see e.g. [4] for some rigorous equivalences in the discrete case).

It remains a major open problem to demonstrate – or rule out – a phase transition for even a single non-trivial pair potential ϕ.[1] Groenveld [9] used the cluster expansion to prove that no phase transition occurs for $\lambda < 1/(eC_\phi)$ for repulsive potentials; this was extended to the broader class of stable potentials by Penrose [24] and Ruelle [29]. For repulsive potentials, works of Meeron [14, 15] extend this regime to $1/C_\phi$. Taking inspiration from Weitz's groundbreaking work [32] on the hard-core model – a discrete repulsive spin system – the work of the second author and Perkins [18] proved that there is no phase transition for $\lambda < e/C_\phi$. This was extended further in [19], which showed that C_ϕ may be replaced with the potential-weighted connective constant Δ_ϕ.

In practice, various Markov chain Monte Carlo algorithms are used to sample from Gibbs point processes, both approximately and exactly. There is a large literature dedicated to this study, see for example [1, 3, 7, 8, 11, 12, 16, 17, 20, 21, 26, 27]. We refer to [20] and the references therein for more context and background on these results; we also note that [20] demonstrates a fast mixing approximate sampling algorithm and polynomial-time *random* approximation algorithm for a repulsive Gibbs point process in the regime $\lambda < e/\Delta_\phi$, the same regime in which the results of this work hold.

---

[1]The first example of a continuous system for which a phase transition was proven was the Widom-Rowlinson model, a result due to Ruelle [30]. However, this does not fit into the framework of indistinguishable particles that we consider in this paper.

When it comes to deterministic algorithms, there are several results for discrete systems. In this setting, three approaches for obtaining approximation algorithms have emerged in recent years. The first is due to Weitz [32] who pioneered an approach based on a notion of *correlation decay* (strong spatial mixing) related to the absence of phase transitions. The second is the *interpolation method* introduced by Barvinok [2] – a framework for showing that under a zero-freeness assumption, one can approximate the logarithm of a polynomial using a small number of Taylor coefficients (see also [25] for an important extension of this method). The third is a method based on the *cluster expansion* in statistical physics pioneered by Helmuth, Perkins, and Regts [10] that is closely related to Barvinok's method.

For Gibbs point processes, to our knowledge the only rigorous result giving a deterministic algorithm is that of Friedrich, Göbel, Katzmann, Krejca, and Pappik [6]. They show that for the hard-spheres model, one may approximately compute the partition function in quasipolynomial time for $\lambda < e/C_\phi$. The approach of [6] works by showing that one may approximate the partition function of the hard-sphere model with the partition function for the hard-core model on a graph given by discretizing Euclidean space with a small mesh. After this approximation is in place, an application of Weitz's method [32] provides an algorithm. We also note that [6] applies not only to the hard-sphere model but also to a class of multi-type Gibbs point processes with hard constraints, an example of this more general class being the Widom-Rowlinson model. Additionally, while both the algorithm of [6] and Theorem 1.3 have quasipolynomial runtime, the algorithm of [6] runs in time $\exp(O(\log^2(|\Lambda_n|/\varepsilon)))$ rather than our runtime of $\exp(O(\log^3(|\Lambda_n|/\varepsilon)))$.

To our knowledge, Theorem 1.3 marks the first approximation algorithm of any kind for the partition function for stable, non-repulsive potentials.

1.4. Our approach

Our approach is via Barvinok's interpolation method combined with use of the cluster expansion. The cluster expansion – also called the Meyer series – is a combinatorial description of the Taylor coefficients of $\log Z_S(\lambda)$. In particular for bounded, measurable $S \subset \mathbb{R}^d$ and $|\lambda| < (e^{1+B}\hat{C}_\phi)^{-1}$, we have

$$\log Z_S(\lambda) = \sum_{k \geqslant 1} \frac{\lambda^k}{k!} C_k(S) \tag{10}$$

$$C_k(S) = \sum_{G \in \mathcal{G}_k} \int_{S^k} \prod_{\{i,j\} \in E(G)} (e^{-\phi(x_i - x_j)} - 1) \, d\mathbf{x} \tag{11}$$

where $\mathcal{G}_k$ is the set of connected labelled graphs with k vertices and $E(G)$ is the set of edges in a graph G (see, e.g., [28] for this and similar expansions).

The algorithmic significance of the cluster expansion stems from two observations. The first is that in order to compute $\frac{1}{|S|} \log Z_S(\lambda)$ up to an additive error of ε, it suffices to compute the first $O(\log(|S|/\varepsilon))$ coefficients $C_k(S)$. This simple but crucial observation is at the heart of Barvinok's interpolation method and algorithmic applications of the cluster expansion. The second is that the description of the Taylor coefficients as a sum over *connected* graphs allows us to approximate them efficiently.

At first sight, it seems that the use of the cluster expansion restricts the range of λ for which we can obtain algorithms to the interval $(0, (e^{1+B}\hat{C}_\phi)^{-1})$ where the cluster expansion is known to converge. Moreover, the work of Groeneveld [9] and Penrose [24] implies that for a repulsive potential, the radius of convergence of the cluster expansion is at most $1/C_\phi$ (see Remark 3.7 in [23]). In order to use the cluster expansion to produce an algorithm throughout the zero-free region (which for repulsive potentials is known to include values of λ *outside* of the radius of

convergence), we use an idea of Barvinok and apply a well-chosen polynomial to map between the zero-free region and the disk. This is handled in Section 3.

The main technical contribution of this work is to approximate the coefficients of the cluster expansion (11) in quasipolynomial time. We note that the problem of approximating coefficients of the cluster expansion was pointed out in [6] as a central obstacle to obtaining an efficient deterministic algorithm that works entirely in the continuum. To approximate the coefficients $C_k(S)$, the main challenge is that for each graph $G \in \mathcal{G}_k$, the integrand in (11) need not be well-behaved; in particular, it need not be Lipschitz. To handle this, we break each such term into many subsequent terms, each of which will give us an integral of a function that is Lipschitz over its support. We then approximate each of these integrals by taking a sufficiently fine mesh and comparing the integral to a weighted sum over this mesh. Our assumption that the potential ϕ has finite range allows us to restrict our attention to approximating integrals over regions whose volume is independent of n. It would be interesting to extend the results of this paper to include stable potentials of infinite range.

2. Approximating coefficients in the cluster expansion

Throughout this section, we fix a stable potential ϕ with stability constant B that satisfies Assumption 1.1. Recall that $\Lambda_n := [-n, n]^d \subset \mathbb{R}^d$. Our goal in this section is to provide an approximation algorithm for $C_k(\Lambda_n)$ (as defined at (11)). In the next section, we show how to use these approximate coefficients to arrive at an approximation of the partition function $Z_{\Lambda_n}(\lambda)$.

Proposition 2.1. *Let ϕ be a stable potential satisfying Assumption 1.1. There are constants $C, c > 0$ depending only on ϕ and the dimension d so that the following holds. For each $\varepsilon \in (0, 1)$ and $k, n \in \mathbb{N}$, we may approximate $C_k(\Lambda_n)/|\Lambda_n|$ up to an additive error of ε in time at most $C|\Lambda_n|\varepsilon^{-dk}e^{ck^3}$.*

Note that Proposition 2.1 makes no use of a zero-freeness or correlation decay type assumption.

In order to prove Proposition 2.1, we will approximate each summand in (11). Further, we will break up the term for each graph into the terms that are Lipschitz on their support. The main work of this section is Lemma 2.3 below, which provides an approximation for each of these Lipschitz terms. Before stating this precisely, we make a few definitions.

Definition 2.2. Let $\Delta_1, \ldots, \Delta_\ell$ be defined as in Assumption 1.1.

1. Given a connected graph G on k vertices, an *edge-labelling* is a function $\sigma : E(G) \to = \{1, \ldots, l\}$.
2. For each labelling, σ and edge $\{i, j\} \in E(G)$ define the function $f_{i,j}^\sigma : (\mathbb{R}^d)^k \to \mathbb{R}$ via

$$f_{i,j}^\sigma(\mathbf{x}) := (e^{-\phi(x_i - x_j)} - 1)\mathbf{1}\{x_i - x_j \in \Delta_{\sigma(\{i,j\})}\}.$$

3. Define f^σ via

$$f^\sigma(\mathbf{x}) := \prod_{\{i,j\} \in E(G)} f_{i,j}^\sigma(\mathbf{x}).$$

4. Define $A_\sigma \subset (\mathbb{R}^d)^k$ via

$$A_\sigma := \{\mathbf{x} \in (\mathbb{R}^d)^k : x_i - x_j \in \Delta_{\sigma(\{i,j\})} \text{ for all } \{i, j\} \in E(G)\},$$

and note that $\text{supp}(f^\sigma) \subseteq A_\sigma$.

Recall that $\mathcal{G}_k$ is the set of connected labelled graphs with k vertices. Given a graph $G \in \mathcal{G}_k$ we have

$$\int_{\Lambda_n^k} \prod_{\{i,j\} \in E(G)} (e^{-\phi(x_i - x_j)} - 1) \, d\mathbf{x} = \sum_{\sigma} \int_{\Lambda_n^k} f^{\sigma}(\mathbf{x}) d\mathbf{x}, \tag{12}$$

where the sum ranges over all edge-labellings σ of G. We note that there are $\ell^{|E(G)|} \leqslant \ell^{k^2/2}$ such labellings, and there are at most $2^{k^2/2}$ connected graphs on k vertices.

With this setup in mind, in order to approximate $C_k(\Lambda_n)$, it is sufficient to approximate the integral of f^{σ} for each labelling σ of each graph $G \in \mathcal{G}_k$. In the next subsection, we will demonstrate such an approximation:

Lemma 2.3. *Let $n, k \in \mathbb{N}$. For any $G \in \mathcal{G}_k$, labelling σ, and $\delta < c_{d,R}$, there is a set of points $\mathcal{S}_{\sigma,\delta} \subset \Lambda_n^k$ with $|\mathcal{S}_{\sigma,\delta}| = O(|\Lambda_n|(2R)^{d(k-1)}\delta^{-dk})$ so that we have*

$$\left| \int_{\Lambda_n^k} f^{\sigma}(\mathbf{x}) \, d\mathbf{x} - \sum_{\mathbf{x} \in \mathcal{S}_{\sigma,\delta}} \delta^{dk} f^{\sigma}(\mathbf{x}) \right| = O(k^2 e^{Bk^2} \delta |\Lambda_n|(2R)^{d(k-1)}).$$

Further, the set $\mathcal{S}_{\sigma,\delta}$ may be computed in time $O(k^2 |\Lambda_n|(2R)^{d(k-1)}\delta^{-dk})$.

Proposition 2.1 follows quickly from here.

Proof of Proposition 2.1. Note that by (12)

$$C_k(\Lambda_n) = \sum_{G \in \mathcal{G}_k} \sum_{\sigma} \int_{\Lambda_n^k} f^{\sigma}(\mathbf{x}) d\mathbf{x}.$$

Taking $\delta = \varepsilon \exp(-Ck^2)$ in Lemma 2.3 with C sufficiently large as a function of R, d, ℓ and B we may approximate $C_k(\Lambda_n)/|\Lambda_n|$ up to additive error ε in time $O(k^2 e^{Bk^2} |\Lambda_n|(2R)^{d(k-1)}\delta^{-dk}) = O(|\Lambda_n|\varepsilon^{-dk} e^{ck^3})$ where c is a constant depending only on R, d and ℓ and B. □

2.1. Approximating the integral over a labelling

In this section, we prove Lemma 2.3. Fix $G \in \mathcal{G}_k$ and set $S = \Lambda_n$.

Recall from Assumption 1.1 that the sets Δ_j are defined via $\Delta_j = K_j \setminus K_{j-1}$. For a given $\gamma \geqslant 0$,

$$\Delta_j^{(\gamma)} := (1 - \gamma)K_j \setminus (1 + \gamma)K_{j-1}.$$

Define the set U_{γ} via

$$U_{\gamma} := \{\mathbf{x} \in S^k : x_i - x_j \in \Delta_{\sigma(\{i,j\})}^{(\gamma)}, \text{ for all } \{i,j\} \in E(G)\} \cap \text{supp}(f^{\sigma}).$$

Before continuing, let us motivate the definition of U_{γ}. We will show that if $\mathbf{x} \in U_{\gamma}$, then f^{σ} is Lipschitz on a small neighbourhood of $\mathbf{x}$ (see Lemmas 2.8 and 2.7 below). This will allow us to approximate the integral $\int_{U_{\gamma}} f^{\sigma}(\mathbf{x}) \, d\mathbf{x}$ by a sum over a mesh. Indeed, we define

$$\mathcal{S}_{\sigma,\delta} := U_{\gamma} \cap ((\delta\mathbb{Z})^d)^k,$$

and prove the following lemma.

Lemma 2.4.

$$\left| \int_{U_{\gamma}} f^{\sigma}(\mathbf{x}) \, d\mathbf{x} - \sum_{\mathbf{x} \in \mathcal{S}_{\sigma,\delta}} \delta^{dk} f^{\sigma}(\mathbf{x}) \right| = O(k^2 e^{Bk^2} \delta |S|(2R)^{d(k-1)}).$$

To complete the proof of Lemma 2.3 we show that $\int_{U_\gamma} f^\sigma(\mathbf{x})\,d\mathbf{x}$ closely approximates $\int_{\Lambda_n^k} f^\sigma(\mathbf{x})\,d\mathbf{x}$ and that the sum in Lemma 2.4 can be computed efficiently. We turn our attention first to the latter task.

Lemma 2.5. *The set $\mathcal{S}_{\sigma,\delta}$ may be enumerated in time $O(k^2|S|(2R)^{d(k-1)}\delta^{-dk})$ and satisfies $|\mathcal{S}_{\sigma,\delta}| = O(|S|(2R)^{d(k-1)}\delta^{-dk})$.*

Proof. Fix a spanning tree T of G and let

$$\mathcal{T} = \{\mathbf{x} \in S^k \cap ((\delta\mathbb{Z})^d)^k : \|x_i - x_j\|_\infty \leqslant 2R \text{ for all } \{i,j\} \in E(T)\}.$$

First observe that $\mathcal{S}_{\sigma,\delta} \subseteq \mathcal{T}$, since if $\mathbf{x} \in U_\gamma$ then $x_i - x_j \in \Delta_{\sigma(\{i,j\})}^{(\gamma)} \subseteq [-R, R]^d$ for all $\{i,j\} \in E(G)$. In particular, $\|x_i - x_j\|_\infty \leqslant 2R$ for all $\{i,j\} \in E(T)$.

Note that $|\mathcal{T}| = O(|S|(2R)^{d(k-1)}\delta^{-dk})$ and that $\mathcal{T}$ may be enumerated in time $O(|S|(2R)^{d(k-1)}\delta^{-dk})$ as well. Further, for each point in $\mathcal{T}$, we may check membership in $\mathcal{S}_{\sigma,\delta}$ in time $O(k^2)$, completing the lemma. $\square$

To show that $\int_{U_\gamma} f^\sigma(\mathbf{x})\,d\mathbf{x}$ closely approximates $\int_{S^k} f^\sigma(\mathbf{x})\,d\mathbf{x}$ it will be enough to show that the measure of

$$W_\gamma = S^k \cap \text{supp}(f^\sigma) \setminus U_\gamma,$$

is small.

Lemma 2.6. *For $\gamma \leqslant 1/d$, we have $|W_\gamma| \leqslant 4|S||E(G)|(2R)^{d(k-1)}d\gamma$.*

Proof. For each edge, $\{a, b\} \in E(G)$ define

$$W_{\gamma,a,b} = \left\{ \mathbf{x} \in S^k : x_i - x_j \in \Delta_{\sigma(\{i,j\})} \text{ for all } \{i,j\} \in E(G), \text{ and } x_a - x_b \notin \Delta_{\sigma(\{a,b\})}^{(\gamma)} \right\}.$$

Note then that $W_\gamma \subseteq \cup_{\{a,b\} \in E(G)} W_{\gamma,a,b}$. Thus, it is sufficient to prove

$$|W_{\gamma,a,b}| \leqslant 4|S|(2R)^{d(k-1)}d\gamma. \tag{13}$$

To see (13), fix a spanning tree T of G so that $\{a, b\} \in E(T)$. We then have that

$$W_{\gamma,a,b} \subseteq \left\{ \mathbf{x} \in (\mathbb{R}^d)^k : x_1 \in S, x_i - x_j \in [-R, R]^d \text{ for all } \{i,j\} \in E(T), x_a - x_b \in \Delta_{\sigma(\{a,b\})} \setminus \Delta_{\sigma(\{a,b\})}^{(\gamma)} \right\}.$$

This provides a bound of

$$|W_{\gamma,a,b}| \leqslant |S|(2R)^{d(k-2)}|\Delta_{\sigma(\{a,b\})} \setminus \Delta_{\sigma(\{a,b\})}^{(\gamma)}|.$$

Write $j = \sigma(\{a, b\})$ for notational simplicity. By convexity of the sets K_j and K_{j-1} we have $(1 - \gamma)K_j \subset K_j$ and $K_{j-1} \subset (1 + \gamma)K_{j-1}$. We may therefore bound

$$|\Delta_j \setminus \Delta_j^{(\gamma)}| \leqslant |K_j| - |(1 - \gamma)K_j| + |(1 + \gamma)K_{j-1}| - |K_{j-1}|$$

$$= |K_j|\left(1 - (1 - \gamma)^d\right) + |K_{j-1}|\left((1 + \gamma)^d - 1\right)$$

$$\leqslant (2R)^d \cdot 2d\gamma + (2R)^d \cdot 2d\gamma. \tag{14}$$

Combining the two previously displayed equations shows (13) and completes the lemma. $\square$

Assuming Lemma 2.4, the proof of Lemma 2.3 now follows quickly.

Proof of Lemma 2.3. Bound

$$\left| \int_{S^k} f^\sigma(\mathbf{x})\,d\mathbf{x} - \sum_{\mathbf{x} \in \mathcal{S}_{\sigma,\delta}} \delta^{dk} f^\sigma(\mathbf{x}) \right| \leqslant \left| \int_{S^k \setminus U_\gamma} f^\sigma(\mathbf{x})\,d\mathbf{x} \right| + \left| \int_{U_\gamma} f^\sigma(\mathbf{x})\,d\mathbf{x} - \sum_{\mathbf{x} \in \mathcal{S}_{\sigma,\delta}} \delta^{dk} f^\sigma(\mathbf{x}) \right|. \tag{15}$$

Since $|e^{-\phi(x)} - 1| \leqslant e^{2B}$, we have that $|f^\sigma| \leqslant e^{Bk^2}$. The first term on the RHS of (15) may thus be bounded by $e^{Bk^2}|W_\gamma|$, which we bound using Lemma 2.6, and the latter term may be bounded by Lemma 2.4. $\qquad\square$

It remains to prove Lemma 2.4. As discussed, a key step will be to show that if $\mathbf{x} \in U_\gamma$, then f^σ is Lipschitz on a small neighbourhood of $\mathbf{x}$. This is carried out in the following two lemmas.

Lemma 2.7. *The function f^σ is $(2e^{Bk^2}L|E(G)|)$-Lipschitz on A_σ.*

Proof. Let $\mathbf{x}, \mathbf{y} \in A_\sigma$. Then by definition

$$|f^\sigma(\mathbf{x}) - f^\sigma(\mathbf{y})| = \left| \prod_{\{i,j\} \in E(G)} (e^{-\phi(x_i - x_j)} - 1) - \prod_{\{i,j\} \in E(G)} (e^{-\phi(y_i - y_j)} - 1) \right|. \qquad (16)$$

To bound (16) we use the inequality

$$\left| \prod_{i=1}^m z_i - \prod_{i=1}^m w_i \right| \leqslant T^{m-1} \sum_{i=1}^m |z_i - w_i|,$$

which holds whenever $|z_i| \leqslant T$, $|w_i| \leqslant T$ for all i. This yields

$$|f^\sigma(\mathbf{x}) - f^\sigma(\mathbf{y})| \leqslant e^{Bk^2} \sum_{\{i,j\} \in E(G)} \left| e^{-\phi(x_i - x_j)} - e^{-\phi(y_i - y_j)} \right|$$

$$\leqslant e^{Bk^2}|E(G)|L \max_{\{i,j\} \in E(G)} \|(x_i - y_i) - (x_j - y_j)\|_2$$

$$\leqslant 2e^{Bk^2}|E(G)|L\|\mathbf{x} - \mathbf{y}\|_2,$$

where we used that $x \mapsto \exp(-\phi(x))$ is L-Lipschitz on each Δ_j by Assumption 1.1. $\qquad\square$

Set

$$\gamma := 2R\delta. \qquad (17)$$

Since we are working in the context of Lemma 2.3, we may also assume that $\gamma \leqslant 1/d$ by taking δ small enough as a function of d and R.

Lemma 2.8. *Let γ be as in (17). If $\mathbf{x} \in U_\gamma$, then*

$$B_\infty(\mathbf{x}; \delta) \subseteq A_\sigma,$$

where $B_\infty(\mathbf{x}; \delta)$ denotes the open ℓ_∞ ball of radius δ centred at $\mathbf{x}$.

Proof. Let $\mathbf{x} \in U_\gamma \subseteq \operatorname{supp}(f^\sigma)$. Suppose that $\mathbf{v} \in (\mathbb{R}^d)^k$ is such that $\|\mathbf{v}\|_\infty < \delta$. It suffices to show that $\mathbf{x} + \mathbf{v} \in A_\sigma$. Let $\{i, j\} \in E(G)$ and let $t = \sigma(\{i, j\})$. Our task is to verify that

$$x_i - x_j + v_i - v_j \in \Delta_t. \qquad (18)$$

Since $\mathbf{x} \in U_\gamma$ we have, by definition,

$$x_i - x_j \in \Delta_t^{(\gamma)}.$$

Moreover $\|v_i - v_j\|_\infty \leqslant 2\delta$. The statement (18) then follows from the following claim.

Claim 2.9. *If $\mathbf{y} \in \Delta_t^{(\gamma)}$, then $B_\infty(\mathbf{y}; 2\delta) \subseteq \Delta_t$.*

Proof. Recall that

$$\Delta_t = K_t \setminus K_{t-1} \quad \text{and} \quad \Delta_t^{(\gamma)} = (1 - \gamma)K_t \setminus (1 + \gamma)K_{t-1}.$$

We will show that $B_\infty(\mathbf{y}; 2\delta) \subseteq K_t$. The argument to show that $B_\infty(\mathbf{y}; 2\delta) \subseteq K_{t-1}^c$ is analogous. We let $\|\cdot\|$ denote the norm associated to the centrally symmetric convex set K_t, that is for $\mathbf{u} \in \mathbb{R}^d$,

$$\|\mathbf{u}\| := \inf\{\theta > 0 : \mathbf{u} \in \theta K_t\}.$$

We note that since $[-1/R, 1/R]^d \subseteq K_t \subseteq [R, R]^d$ by assumption, we have

$$\frac{1}{R}\|\mathbf{u}\|_\infty \leqslant \|\mathbf{u}\| \leqslant R\|\mathbf{u}\|_\infty. \tag{19}$$

With these observations in hand, we note that since $\mathbf{y} \in \Delta_t^{(\gamma)} \subseteq (1-\gamma)K_t$, we have $\|\mathbf{y}\| \leqslant (1-\gamma)$. Suppose now that $\|\mathbf{z}\|_\infty < 2\delta$, then by the triangle inequality, (17) and (19)

$$\|\mathbf{y} + \mathbf{z}\| \leqslant \|\mathbf{y}\| + \|\mathbf{z}\| < 1 - \gamma + 2R\delta = 1.$$

In other words $\mathbf{y} + \mathbf{z} \in K_t$ as desired. □

Applying Claim 2.9 verifies (18). □

With the previous two lemmas in hand, we are now in a position to prove Lemma 2.4.

Proof of Lemma 2.4. Given $\mathbf{x} \in (\mathbb{R}^d)^k$, let $r(\mathbf{x})$ denote the point in $((\delta\mathbb{Z})^d)^k$ closest to $\mathbf{x}$ in ℓ_∞ distance (breaking ties arbitrarily). Note that if $\mathbf{x} \in U_\gamma$ then $r(\mathbf{x}) \in A_\sigma$ by Lemma 2.8. It follows from Lemma 2.7 that

$$\left| \int_{U_\gamma} f^\sigma(\mathbf{x}) - f^\sigma(r(\mathbf{x})) \, d\mathbf{x} \right| \leqslant 2e^{Bk^2}LE(G)\sqrt{dk}\delta|U_\gamma| = O(e^{Bk^2}|E(G)|\delta|S|(2R)^{d(k-1)}). \tag{20}$$

Define

$$U_\gamma' := \bigcup_{\mathbf{x} \in \mathcal{S}_{\sigma,\delta}} B_\infty(\mathbf{x}, \delta/2)$$

and note that

$$\int_{U_\gamma'} f^\sigma(r(\mathbf{x})) \, d\mathbf{x} = \sum_{\mathbf{x} \in \mathcal{S}_{\sigma,\delta}} \delta^{dk} f^\sigma(\mathbf{x}). \tag{21}$$

Using the bound $|f^\sigma| \leqslant e^{Bk^2}$ gives

$$\left| \int_{U_\gamma} f^\sigma(r(\mathbf{x})) \, d\mathbf{x} - \int_{U_\gamma'} f^\sigma(r(\mathbf{x})) \, d\mathbf{x} \right| \leqslant e^{Bk^2}|U_\gamma' \triangle U_\gamma|. \tag{22}$$

Combining lines (20), (21) and (22), it suffices to show that $|U_\gamma' \triangle U_\gamma| = O\big(|E(G)|\delta|S|(2R)^{d(k-1)}\big)$ to complete the lemma. To this end note that

$$U_\gamma' \triangle U_\gamma \subseteq X_\gamma := \{\mathbf{x} \in (\mathbb{R}^d)^k : d_\infty(\partial U_\gamma, \mathbf{x}) \leqslant \delta/2\}.$$

If $\mathbf{y} \in \partial U_\gamma$, then $y_a - y_b \in \partial\Delta_{\sigma(\{a,b\})}^{(\gamma)}$ for some $\{a, b\} \in E(G)$. Thus, if $\mathbf{x} \in X_\gamma$, then

$$d_\infty(x_a - x_b, \partial\Delta_{\sigma(\{a,b\})}^{(\gamma)}) \leqslant \delta$$

for some $\{a, b\} \in E(G)$. Arguing as in Lemma 2.6, it suffices to show that for $t \in [\ell]$,

$$|\{z : d_\infty(z, \partial\Delta_t^{(\gamma)}) \leqslant \delta\}| = O((2R)^d\delta).$$

Suppose then that $d_\infty(z, \partial\Delta_t^{(\gamma)}) \leqslant \delta$. Since,

$$\partial\Delta_t^{(\gamma)} \subseteq \partial((1-\gamma)K_t) \cup \partial((1+\gamma)K_{t-1}),$$

let us suppose first that

$$d_\infty(z, \partial((1-\gamma)K_t)) \leqslant \delta. \tag{23}$$

As in the proof of Claim 2.9, let $\| \cdot \|$ denote the norm associated to K_t. Then by (23), there exists $p, q \in \mathbb{R}^d$ such that $z = p + q$, $\|p\| = (1 - \gamma)$, $\|q\|_\infty \leqslant \delta$. By the triangle inequality, (17) and (19)

$$\|z\| \leqslant \|p\| + \|q\| \leqslant 1 - \gamma + R\delta < 1.$$

Similarly $\|z\| > 1 - 2\gamma$. In other words,

$$z \in K_t \backslash (1 - 2\gamma)K_t.$$

If instead $d_\infty(z, \partial((1+\gamma)K_{t-1})) \leqslant \delta$, then by a similar argument we have that $z \in (1 + 2\gamma)K_{t-1} \backslash K_{t-1}$. By the same calculation as in (14), we conclude that

$$|\{z : d_\infty(z, \partial\Delta_t^{(\gamma)}) \leqslant \delta\}| \leqslant |K_t \backslash (1 - 2\gamma)K_t| + |(1 + 2\gamma)K_{t-1} \backslash K_{t-1}| = O((2R)^d \delta),$$

as desired.   □

3. Reducing cluster expansion coefficients

Here we show that one can approximate $\frac{1}{|\Lambda_n|} \log Z_{\Lambda_n}(\lambda)$ using approximations of the coefficients of the cluster expansion. A slight nuisance is that λ need not lie in the radius of convergence of the cluster expansion; in particular, it is known [9, 24] that for a repulsive potential, the radius of convergence of the cluster expansion is at most $1/C_\phi$ (see remark 3.7 in [23]). To get around this issue, we will use an idea of Barvinok and precompose our function $f = |\Lambda_n|^{-1} \log Z_{\Lambda_n}$ with a well-chosen polynomial map sending the unit disk into our zero-free region. This will result in a function that is analytic in the disk, and the appearance of the polynomial will end up providing only a polynomial fuzz to the efficiency of our algorithms. This approach is outlined in Barvinok's monograph [2, pg. 22] on partition functions; we isolate and prove an abstract version of this idea here.

Define $\mathcal{N}_\gamma := \{z \in \mathbb{C} : d(z, [0, 1]) < \gamma\}$ for all $\gamma \in (0, 1)$.

Theorem 3.1. *Let $\gamma \in (0, 1)$. Suppose f is analytic in $\mathcal{N}_\gamma$ and $|f(z)| \leqslant 1$ for all $z \in \mathcal{N}_\gamma$. Then there is a constant $C = C_\gamma > 0$ depending only on γ so that the following holds for all $\varepsilon \in (0, 1/2)$. If one can approximate each of $f^{(j)}(0)/j!$ up to an additive error of ε^C for all $j = 0, 1, \ldots, C \log(1/\varepsilon)$ in time at most T, then one can approximate $f(1)$ up to an additive error of ε in time $CT \log(1/\varepsilon) + \log^C(1/\varepsilon)$.*

We let $D = \{z \in \mathbb{C} : |z| < 1\}$ denote the open unit disk in $\mathbb{C}$. Our first step is finding a polynomial to map $\overline{D}$ into the region $\mathcal{N}_\gamma$ so that some point in D is mapped to 1. While there are many ways to find such a polynomial (e.g. [2, Lemma 2.2.3] gives an explicit construction) the form of the polynomial is not important for us here, and so we state the relevant properties:

Lemma 3.2. *For each $\gamma > 0$, we may find a polynomial $\Phi = \Phi_\gamma$ so that $\Phi(\overline{D}) \subset \mathcal{N}_\gamma$, $\Phi(0) = 0$ and there is a point $z_1 \in D$ so that $\Phi(z_1) = 1$.*

The point now will be to work with the function $g : \overline{D} \to \mathbb{C}$ defined by $g = f \circ \Phi$. Under the assumptions of Theorem 3.1, such a function g is analytic in D and $|g(z)| \leqslant 1$ for all $z \in \overline{D}$. The Cauchy integral formula will imply that we can approximate g by its Taylor polynomial provided we are not near the boundary:

Fact 3.3. *Let g be analytic in D with $|g(z)| \leqslant 1$ for all $z \in \overline{D}$. Then for all $z \in D$ and $k \in \mathbb{N}$, we have*

$$\left| g(z) - \sum_{j=0}^{k-1} \frac{g^{(j)}(0)}{j!} z^j \right| \leqslant \frac{|z|^k}{1 - |z|}.$$

Proof. Bound

$$\left| g(z) - \sum_{j=0}^{k-1} \frac{g^{(j)}(0)}{j!} z^j \right| \leqslant \sum_{j \geqslant k} \left| \frac{g^{(j)}(0)}{j!} \right| \cdot |z|^j$$

and note that by Cauchy's integral formula we have

$$\left| \frac{g^{(j)}(0)}{j!} \right| \leqslant 1.$$

Summing over $j \geqslant k$ completes the proof. □

Let z_1 be as in Lemma 3.2. It follows that in order to approximate $f(1) = g(z_1)$ up to an additive error of ε, it is sufficient to approximate the first $k = C_\gamma \log(1/\varepsilon)$ terms $g^{(j)}(0)/j!$ up to an error of $\varepsilon/(2k)$ each. To do so, we will relate the derivatives of g to those of f. Iterating the chain rule gives the classical Faà di Bruno formula:

Fact 3.4 (Faà di Bruno's formula). *Let Φ and f be analytic at 0 with $\Phi(0) = 0$ and set $g = f \circ \Phi$. Then for each $n \in \mathbb{N}$, we have*

$$g^{(n)}(0) = \sum_{k=1}^{n} f^{(k)}(0) B_{n,k}(\Phi'(0), \Phi''(0), \dots, \Phi^{(n-k+1)}(0)) \tag{24}$$

where $B_{n,k}(x_1, x_2, \dots, x_{n-k+1})$ are the Bell polynomials given by

$$B_{n,k}(x_1, x_2, \dots, x_{n-k+1}) = \sum \frac{n!}{j_1! j_2! \cdots j_{n-k+1}!} \left(\frac{x_1}{1!} \right)^{j_1} \left(\frac{x_2}{2!} \right)^{j_2} \cdots \left(\frac{x_{n-k+1}}{(n-k+1)!} \right)^{j_{n-k+1}} \tag{25}$$

where the sum is over all sequences of non-negative integers $j_1, j_2, \dots, j_{n-k+1}$ satisfying $\sum j_i = k$ and $\sum i j_i = n$.

To make use of this fact, we will need to evaluate the Bell polynomials $B_{n,k}$ at derivatives of our polynomial Φ. A simple term-by-term bound will be good enough for our purposes.

Lemma 3.5. *For each polynomial Φ, there is a constant $C = C_\Phi > 0$ so that the following holds. For all $n, k \in \mathbb{N}$, we may compute $B_{n,k}(\Phi'(0), \Phi''(0), \dots, \Phi^{(n-k+1)}(0))$ in time $O(k^C)$ and we have the bound*

$$\left| \frac{k!}{n!} B_{n,k}(\Phi'(0), \Phi''(0), \dots, \Phi^{(n-k+1)}(0)) \right| \leqslant e^{Ck}.$$

Proof. First compute all derivatives of Φ at 0, which takes $O_\Phi(1)$ time since Φ is a polynomial. Set $x_j = \Phi^{(j)}(0)$ and note that $x_j = 0$ for $j > d$ where d is the degree of Φ. We may thus reduce the sum in (25) to those for which $j_i = 0$ for all $i > d$ as all other summands are 0. Since the values $j_1, \dots, j_d$ are non-negative integers of size at most k, we have $(k+1)^d$ choices for the sequence $(j_1, \dots, j_d)$. Moreover, the constraints $\sum j_i = k$ and $\sum i j_i = n$ imply that $n \leqslant dk$ and so each summand in (25) can be computed in time $O_d(k^{C'})$ for some $C' = C'_d > 0$. We may, therefore, calculate the sum (25) in time $O_\Phi(k^C)$ for some $C = C_d > 0$.

To see the claimed bound, apply the triangle inequality to see

$$\left| \frac{k!}{n!} B_{n,k}(x_1, x_2, \ldots, x_{n-k+1}) \right| \leqslant \sum \binom{k}{j_1, \ldots, j_d} |x_1|^{j_1} |x_2|^{j_2} \cdots |x_d|^{j_d}.$$

Relaxing the sum to consist of sequences satisfying $\sum j_i = k$, the binomial theorem gives

$$\sum \binom{k}{j_1, \ldots, j_d} |x_1|^{j_1} |x_2|^{j_2} \cdots |x_d|^{j_d} = (|x_1| + |x_2| + \ldots + |x_d|)^k \leqslant e^{Ck}$$

for C large enough with respect to Φ. $\qquad\square$

The proof of Theorem 3.1 now follows from some bookkeeping.

Proof of Theorem 3.1. Apply Lemma 3.2 to find a polynomial $\Phi = \Phi_\gamma$ satisfying the hypotheses of the lemma and define $g = f \circ \Phi$. Then by Fact 3.3 we may take $k := C_\gamma \log(1/\varepsilon)$ large enough so that

$$\left| g(z_1) - \sum_{j=0}^{k-1} \frac{g^{(j)}(0)}{j!} z_1^j \right| \leqslant \varepsilon/2.$$

Since $g(z_1) = f(1)$, it is thus sufficient to approximate $g^{(j)}(0)/j!$ for $j = 0, 1, \ldots, k-1$ up to an additive error of at most $\delta := \varepsilon/(2k)$. By Fact 3.4 and Lemma 3.5 we may expand

$$\frac{g^{(j)}}{j!}(0) = \sum_{i=1}^{j} \frac{f^{(i)}(0)}{i!} B_{i,j}$$

where the coefficients $B_{i,j}$ satisfy $|B_{i,j}| \leqslant e^{Ci} \leqslant e^{Ck} = O(\varepsilon^{-C})$ and may be computed in time $O(k^C) = O(\log^C(1/\varepsilon))$. Thus, if we take C' large enough (as a function of γ) and approximate $f^{(i)}(0)/i!$ up to additive error $\varepsilon^{C'}$, we obtain an approximation to $g(z_1) = f(1)$ up to additive error ε. $\qquad\square$

4. Proof of Theorem 1.3

Proof of Theorem 1.3. Set $f(z) := \frac{1}{C|\Lambda_n|} \log Z_{\Lambda_n}(\lambda z)$, where C is as in Assumption 1.2. Then f satisfies the hypotheses of Theorem 3.1, and note that by (11) we have $f^{(j)}(0) = \lambda^j C_j(\Lambda_n)/(C|\Lambda_n|)$. If we set $\delta = \varepsilon/(C|\Lambda_n|)$, then we are seeking to approximate $f(1)$ up to an additive error of δ. By Propostion 2.1, for each $j = 0, 1, \ldots, O(\log(1/\delta))$, we may compute $f^{(j)}(0)$ up to additive error $\delta^{O(1)}$ in time at most $O(|\Lambda_n| \delta^{-O(j)} e^{O(j^3)}) = \exp(O(\log^3(1/\delta)))$. Applying Theorem 3.1 completes the proof. $\qquad\square$

Acknowledgements

References

[1] Alder, B. J. and Wainwright, T. E. (1957) Phase transition for a hard sphere system. *J. Chem. Phys.* **27**(5) 1208–1209.

[2] Barvinok, A. (2016) *Combinatorics and Complexity of Partition Functions*, Vol. 9. Springer.

[3] Bernard, E. P. and Krauth, W. (2011) Two-step melting in two dimensions: First-order liquid-hexatic transition. *Phys. Rev. Lett.* **107**(15) 155704.

[4] Dobrushin, R. L. and Shlosman, S. B. (1985) Completely analytical Gibbs fields. In *Statistical Physics and Dynamical Systems*. Springer, pp. 371–403.

[5] Daley, D. J. and Vere-Jones, D. (2007) *An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure*. Springer.

[6] Friedrich, T., Göbel, A., Katzmann, M., Krejca, M. S. and Pappik, M. (2022) Algorithms for hard-constraint point processes via discretization. In *International Computing and Combinatorics Conference (COCOON)*.

[7] Friedrich, T., Göbel, A., Krejca, M. and Pappik, M. (2021) A spectral independence view on hard spheres via block dynamics. *arXiv preprint arXiv:* 2102.07443.

[8] Garcia, N. L. (2000) Perfect simulation of spatial processes. *Resenhas do Instituto de Matemática e Estatística da Universidade de São Paulo* **4**(3) 283–325.

[9] Groeneveld, J. (1962) Two theorems on classical many-particle systems. *Phys. Lett.* **3** 50–51.

[10] Helmuth, T., Perkins, W. and Regts, G. (2020) Algorithmic pirogov–sinai theory. *Probab Theory Relat. Fields* **176**(3-4) 851–895.

[11] Huber, M. L. (2016) *Perfect Simulation*, Vol. 148. CRC Press.

[12] Häggström, O., Van Lieshout, M.-C. N. M. and Møller, J. (1999) Characterization results and Markov chain Monte Carlo algorithms including exact simulation for some spatial point processes. *Bernoulli* **5**(4) 641–658.

[13] Lee, T.-D. and Yang, C.-N. (1952) Statistical theory of equations of state and phase transitions. II. Lattice gas and ising model. *Phys. Rev.* **87**(3) 410.

[14] Meeron, E. (1962) Indirect exponential coupling in the classical many-body problems. *Phys. Rev.* **126**(3) 883.

[15] Meeron, E. (1970) Bounds, successive approximations, and thermodynamic limits for distribution functions, and the question of phase transitions for classical systems with non-negative interactions. *Phys. Rev. Lett.* **25**(3) 152.

[16] Møller, J. (1989) On the rate of convergence of spatial birth-and-death processes. *Ann. Inst. Stat. Math.* **41**(3) 565–581.

[17] Møller, J. (2001) A review of perfect simulation in stochastic geometry. *Lect. Notes Monogr. Ser.* **37** 333–355.

[18] Michelen, M. and Perkins, W. (2020) Analyticity for classical gasses via recursion. *arXiv preprint arXiv:* 2008.00972.

[19] Michelen, M. and Perkins, W. (2021) Potential-weighted connective constants and uniqueness of Gibbs measures. *arXiv preprint arXiv:* 2109.01094.

[20] Michelen, M. and Perkins, W. (2022) Strong spatial mixing for repulsive point processes. *J. Stat. Phys.* **189**(9).

[21] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E. (1953) Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**(6) 1087–1092.

[22] Møller, J. and Waagepetersen, R. P. (2007) Modern statistics for spatial point processes. *Scand. J. Stat.* **34**(4) 643–684.

[23] Nguyen, T. X. and Fernández, R. (2020) Convergence of cluster and virial expansions for repulsive classical gases. *J. Stat. Phys.* **179** 448–484.

[24] Penrose, O. (1963) Convergence of fugacity expansions for fluids and lattice gases. *J. Math. Phys.* **4**(10) 1312–1320.

[25] Patel, V. and Regts, G. (2017) Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials. *SIAM J. Comput.* **46**(6) 1893–1919.

[26] Preston, C. (1975) Spatial birth and death processes. *Adv. Appl. Probab.* **7**(3) 465–466.

[27] Propp, J. G. and Wilson, D. B. (1996) Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Struct. Algorithms* **9**(1-2) 223–252.

[28] Procacci, A. and Yuhjtman, S. A. (2017) Convergence of Mayer and virial expansions and the Penrose tree-graph identity. *Lett. Math. Phys.* **107**(1) 31–46.

[29] Ruelle, D. (1963) Correlation functions of classical gases. *Ann. Phys.* **25** 109–120.

[30] Ruelle, D. (1971) Existence of a phase transition in a continuous classical system. *Phys. Rev. Lett.* **27**(16) 1040.

[31] Ruelle, D. (1999) *Statistical Mechanics: Rigorous Results*. World Scientific.

[32] Weitz, D. (2006) Counting independent sets up to the tree threshold. In *Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing, STOC 2006*, ACM, pp. 140–149.

[33] Yang, C.-N. and Lee, T.-D. (1952) Statistical theory of equations of state and phase transitions. I. Theory of condensation. *Phys. Rev.* **87**(3) 404.