Combinatorics, Probability and Computing (2024), 1–27
doi:10.1017/S0963548324000282

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Sampling repulsive Gibbs point processes using random graphs

Tobias Friedrich[1], Andreas Göbel[1], Maximilian Katzmann[2], Martin S. Krejca[3] and Marcus Pappik[1]

[1]Hasso Plattner Institute, University of Potsdam, Potsdam, Germany, [2]Karlsruhe Institute of Technology, Karlsruhe, Germany, and [3]Laboratoire d'Informatique (LIX), CNRS, École Polytechnique, Institut Polytechnique de Paris, Palaiseau, France
Corresponding author: Andreas Göbel; Email: andreas.goebel@hpi.de

(Received 11 December 2023; revised 7 August 2024; accepted 13 August 2024)

Abstract
We study computational aspects of repulsive Gibbs point processes, which are probabilistic models of interacting particles in a finite-volume region of space. We introduce an approach for reducing a Gibbs point process to the hard-core model, a well-studied discrete spin system. Given an instance of such a point process, our reduction generates a random graph drawn from a natural geometric model. We show that the partition function of a hard-core model on graphs generated by the geometric model concentrates around the partition function of the Gibbs point process. Our reduction allows us to use a broad range of algorithms developed for the hard-core model to sample from the Gibbs point process and approximate its partition function. This is, to the extent of our knowledge, the first approach that deals with pair potentials of unbounded range.

Keywords: Partition functions; approximate sampling; Gibbs point processes; random graphs
2020 MSC Codes: Primary: 68W20; Secondary: 60G55, 82B21

1. Introduction

Gibbs point processes are a tool for modelling a variety of phenomena that can be described as distributions of random spatial events [4, 26]. Such phenomena include the location of stars in the universe, a sample of cells under the microscope, or the location of pores and cracks in the ground (see [25, 31] for more on applications of Gibbs point processes). In statistical physics, such point processes are frequently used as stochastic models for gases or liquids of interacting particles [28].

A Gibbs point process on a finite-volume region Λ is parameterized by a fugacity λ and a pair potential ϕ expressing the interactions between pairs of points. Every point configuration in the region is assigned a weight according to the pair interactions ϕ between all pairs of points in the configuration. One can then think of a Gibbs point process as a Poisson point process of intensity λ, where the density of each configuration is scaled proportionally to its weight. The density is normalized by the partition function, which is the integral of the weights over the configuration space [see Section 4.1 for a formal definition of the model]. The most famous example of such a process is the *hard-sphere model*, a model of a random packing of equal-sized spheres with radius r. The pair potential in the hard-sphere model defines hard-core interactions, that is configurations where two points closer than some distance $2r$ have weight zero, while all other configurations have weight one. In this article, we consider Gibbs point processes with *repulsive* potentials,

that is, pair potentials in which adding a point to a configuration does not increase its weight. The hard-sphere model, for example, does have a repulsive pair potential, however, we do not restrict ourselves to hard-core potentials and allow for soft-core interactions.

The two most fundamental algorithmic tasks considered on Gibbs point processes are to sample from the model and to compute its partition function, which are closely related. Understanding for which potentials and fugacities these two tasks are tractable is an ambitious endeavour. Towards this goal, there has been a plethora of algorithmic results on Gibbs point processes spanning several decades. Notably, the Markov chain Monte Carlo method was developed for sampling an instance of the hard-sphere model [22]. Since then, a variety of exact and approximate sampling algorithms for such point processes have been proposed in the literature, and their efficiency has been studied extensively both without [13, 16] and with rigorous running time guarantees [1, 15, 18, 23, 24]. The key objective of rigorous works is to identify a parameter regime for their respective model for which a randomized algorithm for sampling and approximating the partition function exists with running time polynomial in the volume of the region. In addition, deterministic algorithms for approximating the partition function have also appeared in the literature [11, 20] with running time quasi-polynomial in the volume of the region.

This recent flurry of algorithmic results on Gibbs point processes can be attributed to progress in understanding the computational properties of discrete spin systems, such as the hard-core model. Within these works, two main approaches can be identified for transferring insights from discrete spin systems to Gibbs point processes. The first one, which includes results such as [1, 17, 23], considers properties proven to hold in discrete spin systems and translates them to the continuous setting of Gibbs point processes. More precisely, these works consider the notion of strong spatial mixing, which has been strongly connected to algorithmic properties of discrete spin systems [9, 29, 33], and translate it to an analogous notion for Gibbs point processes to obtain algorithmic results. A common pattern in these works is that once the parameter regime for which strong spatial mixing holds is established, one needs to prove from scratch that this implies efficient algorithms for Gibbs point processes. In addition, the definition of strong spatial mixing for Gibbs points processes assumes that the pair interactions of two particles is always of bounded range, that is if two particles are placed at distance greater than some constant $r \in \mathbb{R}_{\geq 0}$, they do not interact with each other.

The second approach, used in [11, 12], is to discretize the model, that is reduce it to an instance of the hard-core model and then solve the respective algorithmic problem for the hard-core model. In this case, the algorithmic arsenal developed over the years for the hard-core model is now readily available for the instances resulting from this reduction. The main downside of these approaches is that they only apply to the hard-sphere model, a special case of bounded-range repulsive interactions.

1.1. Our contributions

We introduce a natural approach for reducing repulsive point processes to the hard-core model. Given an instance (Λ, λ, ϕ) of such a point process, we generate a random graph by sampling sufficiently many points independently and uniformly at random in Λ. Each point represents a vertex in the graph, and we connect each pair of vertices by an edge with a probability that depends on ϕ. We show that both computational problems considered can be reduced to studying the hard-core model on graphs generated by this model for an appropriately scaled vertex activity.

Our first step towards this result is to show that the partition function of the hard-core model on these graphs concentrates around the partition function of the Gibbs point process. Using existing algorithms for the hard-core model as a black box, our result immediately yields randomized approximation algorithms for the partition function of the point process in running time polynomial in the volume of Λ. Furthermore, we show that sampling an independent set from the

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

generated hard-core model and returning the positions of its vertices in Λ results in an approximate sampler from the distribution of the Gibbs point process. Our approach, in contrast to all previous algorithmic work in the literature, does not require the pair potential ϕ of the point process to be of bounded range. This includes various models of interest in statistical physics, such as the (hard-core) Yukawa model [8, 27], the Gaussian overlap model [5], the generalized exponential model [3], and the Yoshida–Kamakura model [34].

2. Notation

We write $\mathbb{N}$ for the non-negative integers (including 0) and $\mathbb{N}_{\geq 1}$ for the positive integers. Moreover, for $n \in \mathbb{N}$, we denote $[n] = \mathbb{N}_{\geq 1} \cap [0, n]$ (note that $[0] = \emptyset$). For a set S and $k \in \mathbb{N}$, we write $\binom{S}{k} := \{S' \subseteq S \mid 3 = k\}$ for the subsets of S of size k. Throughout the paper, we let $G = (V, E)$ denote undirected graphs with vertex set V and edge set $E \subseteq \binom{V}{2}$. Moreover, for $n \in \mathbb{N}$ we write $\mathcal{G}_n$ for the set of undirected graphs with vertex set $[n]$.

3. Concentration of the hard-core partition functions on random graphs

Before we discuss our results on Gibbs point processes, we first derive our main tool, which is a concentration result for the hard-core partition function on random graphs generated based on a graphon. We start by introducing the necessary notation and terminology.

3.1. Hard-core model

For a graph G we write $\mathcal{I}(G) \subseteq 2^V$ for its independent sets. That is, $S \in \mathcal{I}(G)$ if and only if no two vertices in S are adjacent in G. The *hard-core partition function* on G at *vertex activity* $z \in \mathbb{R}_{\geq 0}$ is defined as

$$Z_G(z) := \sum_{S \in \mathcal{I}(G)} z^{|S|}.$$

The *hard-core model* on G at activity z is a probability distribution on 2^V with support $\mathcal{I}(G)$ that is defined by

$$\mu_{G,z}(S) := \frac{z^{|S|}}{Z_G(z)}.$$

for every $S \in \mathcal{I}(G)$.

3.2. Graphons and random graphs

We will show that the hard-core partition $Z_G(z)$ of sufficiently large random graphs generated from a graphon concentrates around its expectation. Here, we refer to graphons in a very general sense, as defined in [21, Chapter 13]. That is, for a probability space $\mathcal{X} = (X, \mathcal{A}, \xi)$, a *graphon* is a symmetric function $W\colon X^2 \to [0, 1]$ that is measurable with respect to the product algebra $\mathcal{A}^2 = \mathcal{A} \otimes \mathcal{A}$. Note that, even though we call the function W the graphon, we mean implicitly that a graphon is a tuple of an underlying probability space and a suitable function W.

A graphon W naturally defines a family of random-graph models, sometimes called W-random graphs (see [21, Chapter 11]). For every $n \in \mathbb{N}_{\geq 1}$, we denote by $\mathbb{G}_{W,n}$ a distribution on $\mathcal{G}_n$ that is induced by the following procedure for generating a random graph:

1. Draw a tuple $(x_1, \ldots x_n) \in X^n$ according to the product distribution ξ^n.
2. For all $i, j \in [n]$, $i \neq j$, add the edge $\{i, j\}$ independently with probability $W(x_i, x_j)$.

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

Formally, this gives a probability distribution $\mathbb{G}_{W,n}$ on $\mathcal{G}_n$ with

$$\mathbb{G}_{W,n}(G) = \int_{X^n} \left(\prod_{\{i,j\} \in E} W(x_i, x_j) \right) \cdot \left(\prod_{\{i,j\} \in \binom{[n]}{2} \setminus E} \big(1 - W(x_i, x_j)\big) \right) \xi^n(\mathrm{d}\mathbf{x}).$$

for all $G \in \mathcal{G}_n$, where $\mathbf{x} = (x_i)_{i \in [n]}$ inside the integral. Observe that $\mathbb{G}_{W,n}$ encompasses classical random-graph models, such as Erdős–Rényi random graphs and geometric random graphs.

3.3 Efron–Stein inequality

For proving our concentration result, we use the Efron–Stein inequality. For $N \in \mathbb{N}_{\geq 1}$ let $\{(\Omega_i, \mathcal{F}_i, \mu_i)\}_{i \in [N]}$ be a collection of probability spaces and let $f : \Omega \to \mathbb{R}$ be a measurable function on the product space $(\Omega, \mathcal{F}, \mu) = \bigotimes_{i \in [N]} (\Omega_i, \mathcal{F}_i, \mu_i)$. For each $i \in [N]$ define a function $\Delta_i^{(f)} : \Omega \times \Omega_i \to \mathbb{R}_{\geq 0}$, where, for every $\mathbf{x} = (x_1, \ldots, x_N) \in \Omega$ and $y_i \in \Omega_i$, the value $\Delta_i^{(f)}(\mathbf{x}, y_i)$ is defined as the squared difference in f that is caused by replacing x_i in $\mathbf{x}$ with y_i. Formally, this is $\Delta_i^{(f)}(\mathbf{x}, y_i) = (f(\mathbf{x}) - f(\mathbf{y}))^2$ where $\mathbf{y} = (x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_N)$. The Efron–Stein inequality bounds the variance of f under μ based on the local squared deviations $\Delta_i^{(f)}(\mathbf{x}, y_i)$.

Theorem 3.1 (Efron – Stein inequalitye [7]). *Let $\{(\Omega_i, \mathcal{F}_i, \mu_i)\}_{i \in [N]}$ be probability spaces with product space $(\Omega, \mathcal{F}, \mu) = \bigotimes_{i \in [N]} (\Omega_i, \mathcal{F}_i, \mu_i])$. For every $\mathcal{F}$-measurable function $f : \Omega \to \mathbb{R}$ it holds that*

$$\mathrm{Var}_\mu[f] \leq \frac{1}{2} \sum\nolimits_{i \in [N]} \mathbb{E}_{\mu \times \mu_i}[\Delta_i^{(f)}].$$

Remark 3.1. The Efron–Stein inequality is usually stated for functions of independent real-valued random variables. However, it extends to functions on products of arbitrary probability spaces.

3.4. Proof of concentration

We start by deriving the following simple corollary of Theorem 3.1.

Corollary 3.1 (Corollary of the Efron – Stein inequality). *Let $\{(\Omega_i, \mathcal{F}_i, \mu_i])\}_{i \in [N]}$ be probability spaces with product space $(\Omega, \mathcal{F}, \mu) = \bigotimes_{i \in [N]} (\Omega_i, \mathcal{F}_i, \mu_i])$, and let $f : \Omega \to \mathbb{R}$ be an $\mathcal{F}$-measurable function. Assume that there are $c_i \in \mathbb{R}_{\geq 0}$ for $i \in [N]$ such that $C := \sum_{i \in [N]} c_i^2 < 2$ and, for all $\boldsymbol{x} = (x_j)_{j \in [N]} \in \Omega$ and $\boldsymbol{y} = (y_j)_{j \in [N]} \in \Omega$ that disagree only at position i, it holds that*

$$\big|f(\boldsymbol{x}) - f(\boldsymbol{y})\big| \leq c_i \cdot \min\{\big|f(\boldsymbol{x})\big|, \big|f(\boldsymbol{y})\big|\}.$$

Then, for all $\varepsilon \in \mathbb{R}_{>0}$, it holds that

$$\mathbb{P}\big[\big|f - \mathbb{E}_\mu[f]\big| \geq \varepsilon \mathbb{E}_\mu[f]\big] \leq \left(\frac{2}{2 - C} - 1\right) \frac{1}{\varepsilon^2}.$$

Proof. First, we observe that $\big|f(\mathbf{x}) - f(\mathbf{y})\big| \leq c_i \min\{\big|f(\mathbf{x})\big|, \big|f(\mathbf{y})\big|\} \leq c_i \big|f(\mathbf{x})\big|$ implies $\mathbb{E}_{\mu \times \mu_i}\left[\Delta_i^{(f)}\right] \leq c_i^2 \mathbb{E}_\mu\left[f^2\right]$ for all $i \in [N]$. Thus, by Theorem 3.1, we have $\mathrm{Var}_\mu[f] \leq \frac{C}{2} \mathbb{E}_\mu\left[f^2\right]$. Now, recall that by definition $\mathrm{Var}_\mu[f] = \mathbb{E}_\mu\left[f^2\right] - \mathbb{E}_\mu[f]^2$, which implies $\mathbb{E}_\mu\left[f^2\right] - \mathbb{E}_\mu[f]^2 \leq \frac{C}{2} \mathbb{E}_\mu\left[f^2\right]$. Rearranging for $\mathbb{E}_\mu[f^2]$ and using the fact that $\frac{C}{2} < 1$ yields $\mathbb{E}_\mu[f^2] \leq \frac{2}{2-C} \mathbb{E}_\mu[f]^2$. Substituting this back into the definition of the variance, we obtain

$$\mathrm{Var}_\mu[f] \leq \left(\frac{2}{2 - C} - 1\right) \mathbb{E}_\mu\left[f\right]^2.$$

The claim follows immediately by applying Chebyshev's inequality. □

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

Remark 3.2. Usually, we want to characterize concentration asymptotically in N. In this setting, Corollary 3.1 tells us that, if $c_i \in \mathrm{O}\left(N^{-\frac{1+\alpha}{2}}\right)$ for all $i \in [N]$ and some $\alpha > 0$, then, for all $\varepsilon \in \mathbb{R}_{>0}$ and $\delta \in (0, 1]$ such that $\varepsilon^2\delta < 1$, it is sufficient to choose $N \in \Theta\left(\delta^{-\frac{1}{\alpha}}\varepsilon^{-\frac{2}{\alpha}}\right)$ to ensure

$$\mathbb{P}\big[\big|f - \mathbb{E}_\mu[f]\big| \geq \varepsilon\mathbb{E}_\mu[f]\big] \leq \delta.$$

To apply Corollary 3.1 to the hard-core partition functions on random graphs from $\mathbb{G}_{W,n}$, we will need to bound how much the partition function changes when applying small modifications to the structure of a graph. More specifically, we want to get a bound on the relative change of the partition function, given that we

- add or remove a single edge, or
- add or remove a set of edges that are all incident to the same vertex.

The following two lemmas provide such bounds.

Lemma 3.1. *Let $G = (V, E)$ be a graph and, for some $e \in E$, let $G' = (V, E \setminus \{e\})$. For all $z \in \mathbb{R}_{\geq 0}$ it holds that*

$$Z_G(z) \leq Z_{G'}(z) \leq \left(1 + z^2\right) Z_G(z).$$

Proof. The lower bound follows immediately from observing that $\mathcal{I}(G) \subseteq \mathcal{I}\left(G'\right)$. For the upper bound, observe that

$$Z_{G'}(z) - Z_G(z) = \sum_{\substack{S \in \mathcal{I}(G'):\\ e \subseteq S}} z^{|S|} \leq z^2 \sum_{\substack{S' \in \mathcal{I}(G'):\\ S' \cap e = \emptyset}} z^{|S'|} \leq z^2 Z_G(z),$$

□

Lemma 3.2. *Let $G = (V, E)$, $G' = (V, E')$ be two graphs on a common vertex set V and assume there is some vertex $v \in V$ such that $v \in e$ for all edges e in the symmetric difference of E and E'. For all $z \in \mathbb{R}_{\geq 0}$ it holds that*

$$|Z_G(z) - Z_{G'}(z)| \leq z \min\{Z_G(z), Z_{G'}(z)\}.$$

Proof. Let $F = (E \cup E') \setminus \{\{v, w\} \mid w \neq v\}$ where v is as given by the claim. Let $F' = F \cup \{\{v, w\} \mid w \neq v\}$, and set $H = (V, F)$ and $H' = (V, F')$. Note that for every $S \in \mathcal{I}\left(H'\right)$ with $v \notin S$ we have $S \in \mathcal{I}(H)$ and $S \cup \{v\} \in \mathcal{I}(H)$. Thus, we get

$$Z_H(z) \leq (1 + z) \sum_{\substack{S \in \mathcal{I}(H'):\\ v \notin S}} z^{|S|} \leq (1 + z) Z_{H'}(z).$$

Further, we have $F \subseteq E$, $E' \subseteq F'$ and, by Lemma 3.1, we know that the hard-core partition function is non-decreasing under removing edges from the graph. Applying the triangle inequality we get

$$|Z_G(z) - Z_{G'}(z)| \leq Z_H(z) - Z_{H'}(z) \leq zZ_{H'}(z) \leq z \min\{Z_G(z), Z_{G'}(z)\}.$$

□

Based on Lemmas 3.1 and 3.2, we use Corollary 3.1 to prove the following statement.

Theorem 3.2. *Let W be a graphon on the probability space $\mathcal{X} = (X, \mathcal{A}, \xi)$. Let $(z_n)_{n\geq 1}$ be such that $z_n \leq z_0 n^{-\frac{1+\alpha}{2}}$ for some $z_0 \in \mathbb{R}_{\geq 0}$ and $\alpha \in \mathbb{R}_{>0}$. For all $\varepsilon \in (0, 1]$, $\delta \in (0, 1]$, $n \geq \left(2{z_0}^2\varepsilon^{-2}\delta^{-1}\right)^{\frac{1}{\alpha}}$, and $G \sim \mathbb{G}_{W,n}$, it holds that*

$$\mathbb{P}\left[|Z_G(z_n) - \mathbb{E}\left[Z_G(z_n)\right]| \geq \varepsilon\mathbb{E}\left[Z_G(z_n)\right]\right] \leq \delta.$$

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

Proof. We aim for applying Corollary 3.1 to prove our claim. To this end, for each $n \in \mathbb{N}_{\geq 1}$ we need to write the partition function $Z_G(z_n)$ for $G \sim \mathbb{G}_{W,n}$ as a function on a product of probability spaces (i.e., a function of independent random variables). At first, an obvious choice seems to be $\mathcal{X}^n$ together with $\binom{n}{2}$ additional binary random variables, one for each potential edge $\{i,j\} \in \binom{[n]}{2}$. However, note that the edges might not necessarily be independent, meaning that the resulting product distribution would not resemble $\mathbb{G}_{W,n}$. Instead, let $\mathcal{Y} = ([0,1], \mathcal{B}([0,1]), u)$, where $\mathcal{B}([0,1])$ is the Borel algebra restricted to $[0,1]$ and u is the uniform distribution on that interval. We consider the probability space $\mathcal{X}^n \otimes \mathcal{Y}^{\binom{n}{2}}$.

For $\mathbf{x} \in X^n$ and $\mathbf{y} \in [0,1]^{\binom{n}{2}}$ let $\mathbf{x} \circ \mathbf{y} \in X^n \times [0,1]^{\binom{n}{2}}$ denote the concatenation of $\mathbf{x}$ and $\mathbf{y}$. We construct a measurable function $g{:}X^n \times [0,1]^{\binom{n}{2}} \to \mathcal{G}_n$ by mapping every $\mathbf{z} = \mathbf{x} \circ \mathbf{y} \in X^n \times [0,1]^{\binom{n}{2}}$ with $\mathbf{x} = (x_i)_{i \in [n]} \in X^n$ and $\mathbf{y} = (y_{i,j})_{1 \leq i < j \leq n} \in [0,1]^{\binom{n}{2}}$ to $g(\mathbf{z}) = ([n], E)$ such that, for all $i < j$, it holds that $\{i,j\} \in E$ if and only if $W(x_i, x_j) \geq y_{i,j}$. Simple calculations show that, for $\mathbf{z} \sim \xi^n \times u^{\binom{n}{2}}$, it holds that $g(z) \sim \mathbb{G}_{W,n}$. Now, let $f : X^n \times [0,1]^{\binom{n}{2}} \to \mathbb{R}$ with $\mathbf{z} \mapsto Z_{g(z)}(z_n)$. In order to apply Corollary 3.1, we need to bound the relative change of $f(\mathbf{z})$ if we change one component of $\mathbf{z}$. Let $\mathbf{x}' = (x_1, \cdots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n) \in X^n$ for any $i \in [n]$. Then $g(\mathbf{x}' \circ \mathbf{y})$ can only differ from $g(\mathbf{z})$ on edges that are incident to vertex i. Thus, by Lemma 3.2, it holds that

$$\left| f(\mathbf{z}) - f(\mathbf{x}' \circ \mathbf{y}) \right| \leq z_n \min\{f(\mathbf{z}), f(\mathbf{x}' \circ \mathbf{y})\}.$$

Now, let $\mathbf{y}' = (y'_{i,j})_{1 \leq i < j \leq n} \in [0,1]^{\binom{n}{2}}$ such that $y'_{i,j} = y_{i,j}$ except for one pair $1 \leq i < j \leq n$. Note that $g(\mathbf{z})$ and $g(\mathbf{x} \circ \mathbf{y}')$ differ by at most one edge. By Lemma 3.1, we have

$$\left| f(\mathbf{z}) - f(\mathbf{x} \circ \mathbf{y}') \right| \leq z_n^2 \min\{f(\mathbf{z}), f(\mathbf{x} \circ \mathbf{y}')\}.$$

For $\varepsilon \leq 1$ and $\delta \leq 1$ it holds that $n \geq \left(2 {z_0}^2 \varepsilon^{-2} \delta^{-1}\right)^{\frac{1}{\alpha}} > {z_0}^{\frac{2}{\alpha}}$. Hence, for $z_n \leq z_0 n^{-\frac{1+\alpha}{2}}$, it holds that

$$C := n z_n^2 + \binom{n}{2} z_n^4 \leq {z_0}^2 n^{-\alpha} + {z_0}^4 n^{-2\alpha} \leq 2 {z_0}^2 n^{-\alpha} < 2.$$

By Corollary 3.1 we obtain

$$\mathbb{P}\left[\left| Z_G(z_n) - \mathbb{E}\left[Z_G(z_n) \right] \right| \geq \varepsilon \mathbb{E}\left[Z_G(z_n) \right] \right] \leq \left(\frac{1}{1 - {z_0}^2 n^{-\alpha}} - 1 \right) \frac{1}{\varepsilon^2} = \frac{{z_0}^2}{\left(n^\alpha - {z_0}^2 \right) \varepsilon^2}.$$

Using again that $n \geq \left(2 {z_0}^2 \varepsilon^{-2} \delta^{-1}\right)^{\frac{1}{\alpha}}$ we get

$$\mathbb{P}\left[\left| Z_G(z_n) - \mathbb{E}\left[Z_G(z_n) \right] \right| \geq \varepsilon \mathbb{E}\left[Z_G(z_n) \right] \right] \leq \delta,$$

which concludes the proof. □

4. Approximating the partition function of Gibbs point processes

We now turn towards the main goal of this paper, which is applying Theorem 3.2 to approximate repulsive Gibbs point processes using the hard-core model on random graphs. To this end, we start by formally defining the class of point processes that we are interested in.

4.1. Gibbs point processes

Let $(\mathbb{X}, d)$ be a complete, separable metric space and let $\mathcal{B} = \mathcal{B}(\mathbb{X})$ be the Borel algebra of that space. Let ν be a locally finite reference measure on $(\mathbb{X}, \mathcal{B})$ such that all bounded measurable sets have finite measure. Denote by $\mathcal{N}$ the set of all locally finite counting measures on $(\mathbb{X}, \mathcal{B})$. Formally, this is the set of all measures η on $(\mathbb{X}, \mathcal{B})$ with values in $\mathbb{N} \cup \{\infty\}$ such that $\nu(A) < \infty$

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

implies $\eta[A] < \infty$ for all $A \in \mathcal{B}$. For each $A \in \mathcal{B}$, define a map $N_A : \mathcal{N} \to \mathbb{N} \cup \{\infty\}$ with $\eta \mapsto \eta[A]$ and let $\mathcal{R}$ be the sigma algebra on $\mathcal{N}$ that is generated by the set of those maps $\{N_A \mid A \in \mathcal{B}\}$. A *point process* on $\mathbb{X}$ is now a measurable map from some probability space to the measurable space $(\mathcal{N}, \mathcal{R})$. With some abuse of terminology, we call any probability distribution on $(\mathcal{N}, \mathcal{R})$ a point process. Moreover, a point process is called *simple* if $N_x(\eta) \leq 1$ with probability 1, where we write N_x for $N_{\{x\}}$.

Note that every counting measure $\eta \in \mathcal{N}$ is associated with a multiset of points in $\mathbb{X}$. To see this, define $X_(\eta) = \{x \in \mathbb{X} \mid N_x(\eta) > 0\}$. Then η can be expressed as a weighted sum of Dirac measures

$$\eta = \sum_{x \in X_(\eta)} N_x(\eta)\delta_x.$$

In this sense, η is associated with a multiset of points $x \in X_(\eta)$, each occurring with finite multiplicity $N_x(\eta)$. We may use such a *point configuration* interchangeably with its corresponding counting measure.

An important example for point processes are Poisson point processes. A *Poisson point process* with intensity $\kappa \in \mathbb{R}_{\geq 0}$ on $(\mathbb{X}, d)$ is uniquely defined by the following properties

- for all bounded measurable $A \subseteq \mathbb{X}$ it holds that N_A is Poisson distributed with intensity $\kappa\, \nu(A)$ and
- for all $m \in \mathbb{N}_{\geq 2}$ and disjoint measurable $A_1, \ldots, A_m \subseteq \mathbb{X}$ it holds that $N_{A_1}, \ldots, N_{A_m}$ are independent.

Generally speaking, a *Gibbs point process* is a point process that is absolutely continuous with respect to a Poisson point process. For a bounded measurable $\Lambda \subseteq \mathbb{X}$ let $\mathcal{N}(\Lambda)$ denote the set of locally finite counting measures $\eta \in \mathcal{N}$ that satisfy $N_A(\eta) = 0$ for all measurable $A \subseteq \mathbb{X} \setminus \Lambda$. In this work we are interested in Gibbs point processes $P_\Lambda^{(\lambda,\phi)}$ on bounded measurable regions $\Lambda \subseteq \mathbb{X}$ that are parameterized by a *fugacity* parameter $\lambda \in \mathbb{R}_{\geq 0}$ and non-negative, symmetric, measurable *potential function* $\phi : \mathbb{X}^2 \to \mathbb{R}_{\geq 0} \cup \{\infty\}$. Formally, such a process $P_\Lambda^{(\lambda,\phi)}$ is defined by having a density with respect to a Poisson point process with intensity λ of the form

$$\frac{\mathrm{d}P_\Lambda^{(\lambda,\phi)}}{\mathrm{d}\mathrm{d}Q_\lambda}(\eta) = \frac{\mathbb{1}_{\eta \in \mathcal{N}_\Lambda} \mathrm{e}^{-H(\eta)} \mathrm{e}^{\lambda\, \nu(\Lambda)}}{\Xi_\Lambda(\lambda, \phi)}$$

where $H : \mathcal{N} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ is the *Hamiltonian* defined by

$$H(\eta) = \sum_{\{x,y\} \in \binom{X_\eta}{2}} N_x(\eta) N_y(\eta) \phi(x, y) + \sum_{x \in X_\eta} \frac{N_x(\eta)\,(N_x(\eta) - 1)}{2} \phi(x, x).$$

The normalizing constant $\Xi_\Lambda(\lambda, \phi)$ is usually called the *(grand-canonical) partition function* and can be written explicitly as

$$\begin{aligned}\Xi_\Lambda(\lambda, \phi) &= 1 + \sum_{k \in \mathbb{N}_{\geq 1}} \frac{\lambda^k}{k!} \int_{\Lambda^k} \mathrm{e}^{-H} \left(\delta_{x_1} + \cdots + \delta_{x_k}\right) \nu^k(\mathrm{d}\mathbf{x}) \\ &= 1 + \sum_{k \in \mathbb{N}_{\geq 1}} \frac{\lambda^k}{k!} \int_{\Lambda^k} \prod_{\{i,j\} \in \binom{[k]}{2}} \mathrm{e}^{-\phi(x_i, x_j)} \nu^k(\mathrm{d}\mathbf{x}).\end{aligned}$$

For a more detailed overview on the theory of Gibbs point processes, see [19].

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

4.2. Concentration of the partition function

The rest of this paper will use the following assumptions: $(\mathbb{X}, d)$ is a complete, separable metric space with Borel algebra $\mathcal{B} = \mathcal{B}(\mathbb{X})$, and ν is a locally-finite reference measure on $(\mathbb{X}, \mathcal{B})$. Moreover, $\Lambda \subseteq \mathbb{X}$ is bounded and measurable, $\lambda \in \mathbb{R}_{\geq 0}$, and $\phi : \mathbb{X}^2 \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ is a symmetric repulsive potential.

Under the assumptions above, we introduce the following way of obtaining a graphon from a Gibbs point process.

Definition 4.1. *The graphon associated with a repulsive Gibbs point process, denoted by W_ϕ, is defined as follows: We consider the probability space $(\Lambda, \mathcal{B}_\Lambda, u_\Lambda)$, where $\mathcal{B}_\Lambda$ is the trace of Λ in $\mathcal{B}$ and u_Λ is the probability measure on $(\Lambda, \mathcal{B}_\Lambda)$ that is defined via the constant density $\frac{1}{\nu(\Lambda)}$ with respect to ν restricted to Λ. We define $W_\phi : \Lambda^2 \to [0, 1]$ via $W_\phi(x, y) := 1 - \mathrm{e}^{-\phi(x,y)}$ for $x, y \in \Lambda$.*

We are particularly interested in the random graph model $\mathbb{G}_{W_\phi,n}$. Note that, given $n \in \mathbb{N}$, a random graph from $n \in \mathbb{N}$ can be generated by the following procedure:

1. Draw $x_1, \ldots, x_n$ independently from u_Λ.
2. Output a random graph on vertex set $[n]$, where $i \neq j$ share an edge independently with probability $W_\phi(x_i, x_j) = 1 - \mathrm{e}^{-\phi(x_i,x_j)}$.

Our goal is to reduce the problem of approximate sampling from $P_\Lambda^{(\lambda,\phi)}$ to approximate sampling from $\mu_{G,z}$ and, similarly, to reduce the problem of approximating $\Xi_\Lambda(\lambda, \phi)$ to approximating $Z_{G(z)}$ for a random graph $G \sim \mathbb{G}_{W_\phi,n}$ at an appropriate vertex activity z. To this end, our first step it to derive the following concentration result.

Theorem 4.1. *Let W_ϕ be as in Definition 4.1 and let $z_n = \lambda \frac{\nu(\Lambda)}{n}$ for $n \in \mathbb{N}_{\geq 1}$. For all $\varepsilon \in (0, 1]$, $\delta \in (0, 1]$ and $n \geq 4\varepsilon^{-2}\delta^{-1} \max\left\{\mathrm{e}^6\lambda^2\nu(\Lambda)^2, \ln\left(4\varepsilon^{-1}\right)^2\right\}$, it holds that, for $G \sim \mathbb{G}_{W_\phi,n}$,*

$$\mathbb{P}\left[|Z_G(z_n) - \Xi_\Lambda(\lambda, \phi)| \geq \varepsilon\, \Xi_\Lambda(\lambda, \phi)\right] \leq \delta.$$

To prove Theorem 4.1, we start by relating the expected hard-core partition function with the partition function of the Gibbs point process.

Lemma 4.1. *Let W_ϕ be as in Definition 4.1 and let $z_n = \lambda \frac{\nu(\Lambda)}{n}$ for $n \in \mathbb{N}_{\geq 1}$. For all $\varepsilon \in \mathbb{R}_{>0}$ and $n \geq 2\varepsilon^{-1} \max\left\{\mathrm{e}^6\lambda^2\nu(\Lambda)^2, \ln\left(2\varepsilon^{-1}\right)^2\right\}$ it holds that, for $G \sim \mathbb{G}_{W_\phi,n}$,*

$$(1 - \varepsilon)\, \Xi_\Lambda(\lambda, \phi) \leq \mathbb{E}\left[Z_G(z_n)\right] \leq \Xi_\Lambda(\lambda, \phi).$$

Proof. We start by rewriting the hard-core partition function as

$$Z_G(z_n) = 1 + \sum_{k=1}^{n} \lambda^k \frac{\nu(\Lambda)^k}{n^k} \sum_{S \in \binom{[n]}{k}} \mathbb{1}_{S \in \mathcal{I}(G)}.$$

Thus, by linearity of expectation we have

$$\mathbb{E}[Z_G(z_n)] = 1 + \sum_{k=1}^{n} \lambda^k \frac{\nu(\Lambda)^k}{n^k} \sum_{S \in \binom{[n]}{k}} \mathbb{P}[S \in \mathcal{I}(G)].$$

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

By the definition of W_ϕ we observe that for all $S \in \binom{[n]}{k}$ with $|S| = k$

$$\begin{aligned}\mathbb{P}[S \in \mathcal{I}(G)] &= \int_{\Lambda^n} \prod_{\{i,j\} \in \binom{S}{2}} \mathrm{e}^{-\phi(x_i,x_j)} u_\Lambda^n(\mathrm{d}\mathbf{x}) \\ &= \frac{\nu(\Lambda)^{n-k}}{\nu(\Lambda)^n} \int_{\Lambda^k} \prod_{\{i,j\} \in \binom{[k]}{2}} \mathrm{e}^{-\phi(x_i,x_j)} \nu^k(\mathrm{d}\mathbf{x}) \\ &= \frac{1}{\nu(\Lambda)^k} \int_{\Lambda^k} \prod_{\{i,j\} \in \binom{[k]}{2}} \mathrm{e}^{-\phi(x_i,x_j)} \nu^k(\mathrm{d}\mathbf{x}),\end{aligned}$$

where the second equality uses that, by symmetry, the value of the integral does not depend on the particular choice of S. We obtain

$$\mathbb{E}[Z_G(z_n)] = 1 + \sum_{k=1}^{n} \lambda^k \frac{\binom{n}{k}}{n^k} \int_{\Lambda^k} \prod_{\{i,j\} \in \binom{[k]}{2}} \mathrm{e}^{-\phi(x_i,x_j)} \nu^k(\mathrm{d}\mathbf{x}).$$

Writing $\binom{n}{k}/n^k = \frac{1}{k!} \prod_{i=0}^{k-1} (1 - i/n)$ we immediately obtain the upper bound

$$\mathbb{E}[Z_G(z_n)] \le \Xi_\Lambda(\lambda, \phi).$$

For the lower bound, we set

$$S_m = 1 + \sum_{k=1}^{m} \frac{\lambda^k}{k!} \int_{\Lambda^k} \prod_{\{i,j\} \in \binom{[k]}{2}} \mathrm{e}^{-\phi(x_i,x_j)} \nu^k(\mathrm{d}\mathbf{x}).$$

for any $1 \le m \le n$ and observe that

$$\mathbb{E}[Z_G(z_n)] \ge \left(1 - \frac{m}{n}\right)^m S_m.$$

In particular, for $n \ge 2\varepsilon^{-1}m^2$, Bernoulli's inequality yields $\mathbb{E}[Z_G(z_n)] \ge \left(1 - \frac{\varepsilon}{2}\right) S_m$. Moreover, using that ϕ is non-negative, we note that

$$\Xi_\Lambda(\lambda, \phi) - S_m = \sum_{k=m+1}^{\infty} \frac{\lambda^k}{k!} \int_{\Lambda^k} \prod_{\{i,j\} \in \binom{[k]}{2}} \mathrm{e}^{-\phi(x_i,x_j)} \nu^k(\mathrm{d}\mathbf{x}) \le \sum_{k=m+1}^{\infty} \frac{\lambda^k \nu(\Lambda)^k}{k!}.$$

Using Lagrange's remainder formula, we obtain

$$\Xi_\Lambda(\lambda, \phi) - S_m \le \frac{\mathrm{e}^{\lambda\nu(\Lambda)}}{(m+1)!} \left(\lambda\nu(\Lambda)\right)^{m+1}.$$

Choosing $m \ge \max\left\{\mathrm{e}^3 \lambda \nu(\Lambda), \ln\left(2\varepsilon^{-1}\right)\right\}$ and using the fact that $(m+1)! > \left(\frac{m+1}{\mathrm{e}}\right)^{m+1}$ yields

$$\Xi_\Lambda(\lambda, \phi) - S_m \le \left(\frac{\mathrm{e}^2 \lambda \nu(\Lambda)}{m+1}\right)^{m+1} \le \mathrm{e}^{-(m+1)} \le \frac{\varepsilon}{2},$$

and, as $\Xi_\Lambda(\lambda, \phi) \ge 1$, we get $S_m \ge \left(1 - \frac{\varepsilon}{2}\right) \Xi_\Lambda(\lambda, \phi)$. Hence, for $n \ge 2\varepsilon^{-1}m^2 = 2\varepsilon^{-1} \max\left\{\mathrm{e}^6 \lambda^2 \nu(\Lambda)^2, \ln\left(2\varepsilon^{-1}\right)^2\right\}$, we obtain

$$\mathbb{E}[Z_G(z_n)] \ge \left(1 - \frac{\varepsilon}{2}\right) S_m \ge \left(1 - \frac{\varepsilon}{2}\right)^2 \Xi_\Lambda(\lambda, \phi) \ge (1 - \varepsilon)\, \Xi_\Lambda(\lambda, \phi),$$

which proves the claim. □

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

Proof of Theorem 4.1. By setting $\alpha = 1$ and $z_0 = \lambda\nu(\Lambda)$ and using the fact that

$$n \geq 4\varepsilon^{-2}\delta^{-1} \max\left\{e^6\lambda^2\nu(\Lambda)^2, \ln\left(4\varepsilon^{-1}\right)^2\right\} \geq \left(2{z_0}^2\left(\frac{\varepsilon}{2}\right)^{-2}\delta^{-1}\right)^{\frac{1}{\alpha}},$$

Theorem 3.2 yields

$$\mathbb{P}\left[|Z_G(z_n) - \mathbb{E}[Z_G(z_n)]| \geq \frac{\varepsilon}{2}\mathbb{E}[Z_G(z_n)]\right] \leq \delta.$$

Furthermore, by Lemma 4.1 we know that for

$$n \geq 4\varepsilon^{-2}\delta^{-1} \max\left\{e^6\lambda^2\nu(\Lambda)^2, \ln\left(4\varepsilon^{-1}\right)^2\right\} \geq 2\left(\frac{\varepsilon}{2}\right)^{-1} \max\left\{e^6\lambda^2\nu(\Lambda)^2, \ln\left(2\left(\frac{\varepsilon}{2}\right)^{-1}\right)^2\right\}.$$

it holds that

$$\left(1 - \frac{\varepsilon}{2}\right)\Xi_\Lambda(\lambda,\phi) \leq \mathbb{E}[Z_G(z_n)] \leq \Xi_\Lambda(\lambda,\phi).$$

Combining both, we obtain

$$\mathbb{P}\left[|Z_G(z_n) - \Xi_\Lambda(\lambda,\phi)| \geq \varepsilon\,\Xi_\Lambda(\lambda,\phi)\right] \leq \delta,$$

which proves the claim. □

4.3. Approximation algorithm for the partition function

One of the main applications of Theorem 4.1 is that it yields a rather simple randomized procedure for approximating $\Xi_\Lambda(\lambda,\phi)$. That is, we will use Theorem 4.1 to *ε-approximate* $\Xi_\Lambda(\lambda,\phi)$: compute some value $x \in \mathbb{R}$ such that $(1-\varepsilon)\Xi_\Lambda(\lambda,\phi) \leq x \leq (1+\varepsilon)\Xi_\Lambda(\lambda,\phi)$. More specifically, we will obtain a *randomized ε-approximation* algorithm, i.e. an algorithm that outputs an ε-approximation of $\Xi_\Lambda(\lambda,\phi)$ with probability at least $\frac{2}{3}$.[1]

The rough idea of the approximation algorithm is as follows:

1. For $n \in \mathbb{N}$ sufficiently large, sample $G \sim \mathbb{G}_{W_\phi,n}$ where W_ϕ is as in Definition 4.1.
2. Approximate $Z_G(z_n)$ for $z_n = \lambda\frac{\nu(\Lambda)}{n}$ and use the result as an approximation for $\Xi_\Lambda(\lambda,\phi)$.

We are especially interested in obtaining an algorithm that is asymptotically efficient in the volume $\nu(\Lambda)$, as this gives a natural way to parameterize the algorithmic problem. More specifically, we want to characterize the regime of the fugacity λ in terms of the potential ϕ for which we can get a randomized ε-approximation of $\Xi_\Lambda(\lambda,\phi)$ in time polynomial in $\nu(\Lambda)$ and $\frac{1}{\varepsilon}$. We characterize this fugacity regime in terms of the temperedness constant

$$C_\phi = \operatorname{ess\,sup}_{x_1\in\mathbb{X}} \int_{\mathbb{X}} \left|1 - e^{-\phi(x_1,x_2)}\right| \nu(dx_2),$$

where ess sup denotes the essential supremum (i.e., an upper bound that holds almost everywhere). We obtain the following result.

Theorem 4.2. *Assume for $n \in \mathbb{N}$ there is a sampler for $\mathbb{G}_{W_\phi,n}$ with running time $t_{\Lambda,\phi}(n)$.*

If $\lambda < \frac{e}{C_\phi}$, then, for all $\varepsilon \in (0,1]$, there is a randomized ε-approximation algorithm for $\Xi_\Lambda(\lambda,\phi)$ with running time in $\widetilde{O}\left(\nu(\Lambda)^4\varepsilon^{-6}\right) + t_{\Lambda,\phi}\left(\widetilde{O}\left(\nu(\Lambda)^2\varepsilon^{-2}\right)\right)$.

[1] The choice of the constant $\frac{2}{3}$ is rather arbitrary here, as the error probability can be made smaller than every $\delta \in \mathbb{R}_{>0}$ by taking the median of $O(\log(\delta^{-1}))$ independent runs, as long as the error probability of each run is some constant smaller than $\frac{1}{2}$.

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

To prove Theorem 4.2 we need two ingredients. First, we need to bound how large n needs to be chosen to ensure that $Z_G(z_n)$ is close to $\Xi_\Lambda(\lambda, \phi)$ with high probability. Second, we need to ensure that $Z_G(z_n)$ can be approximated in time polynomial in $\nu(\Lambda)$. To tackle the first part, we use Theorem 4.1. For the second part, we will use some well known results on approximating the hard-core partition function.

Theorem 4.3 [(30, Corollary 8.4] and [2, Theorem 1)]. *Let $G = (V, E)$ be an undirected graph with maximum vertex degree bounded by $d_G \in \mathbb{N}_{\geq 2}$ and let $z \in \mathbb{R}_{\geq 0}$ with*

$$z < z_c\left(d_G\right) = \frac{\left(d_G - 1\right)^{d_G - 1}}{\left(d_G - 2\right)^{d_G}}.$$

Then, for all $\varepsilon \in (0, 1]$, there is a randomized ε-approximation algorithm for the hard-core partition function $Z_G(z)$ with running time $\widetilde{\mathrm{O}}\left(|V|^2\, \varepsilon^{-2}\right)$.

Remark 4.1. In [30] the result above is only stated for $z < \frac{2}{d_G}$ as an older mixing time result for Glauber dynamics from [32] is used. Combining their approach with the more recent mixing time bound in [2] gives the desired bound of $z < z_c\left(d_G\right)$.

Thus, arguing that Z_G for $G \sim \mathbb{G}_{W_\phi, n}$ can be approximated in time poly (n) boils down to obtaining a probabilistic upper bound on d_G. We use the following simple lemma.

Lemma 4.2. *Assume $C_\phi > 0$. For $\alpha \in \mathbb{R}_{>0}$, $q \in (0, 1]$, $n \geq 3 \max\left\{\alpha^{-1}, \alpha^{-2}\right\} \ln\left(q^{-1}\right) C_\phi^{-1} \nu(\Lambda) + 1$ and $G \sim \mathbb{G}_{W_\phi, n}$ it holds that*

$$\mathbb{P}\left[d_G \geq (1+\alpha)\frac{n-1}{\nu(\Lambda)} C_\phi\right] \leq qn.$$

Proof. By union bound, it is sufficient to argue that, for each $i \in [n]$ it holds that

$$\mathbb{P}\left[d_G(i) \geq (1+\alpha)\frac{n-1}{\nu(\Lambda)} C_\phi\right] \leq q,$$

where $d_G[i]$ denotes the degree of vertex $i \in [n]$ in G. Now, observe that the random variables $d_G[i]$ for $i \in [n]$ are identically distributed. Thus, we can focus on $d_G[n]$ for ease of notation. By definition, it holds for $k \in [n-1] \cup \{0\}$ that

$$\begin{aligned}
\mathbb{P}\left[d_G(n) = k\right] &= \sum_{S \in \binom{[n-1]}{k}} \int_{\Lambda^n} \left(\prod_{i \in S} W_\phi(x_n, x_i)\right) \cdot \left(\prod_{i \in [n-1] \setminus S} \left(1 - W_\phi(x_n, x_i)\right)\right) u_\Lambda^n(\mathrm{d}\mathbf{x}) \\
&= \int_\Lambda \binom{n-1}{k} \left(\int_\Lambda W_\phi(x_1, x_2), u_\Lambda(\mathrm{d}x_2)\right)^k \\
&\quad \left(1 - \int_\Lambda W_\phi(x_1, x_2)\, u_\Lambda(\mathrm{d}x_2)\right)^{n-1-k} u_\Lambda(\mathrm{d}x_1).
\end{aligned}$$

For every $x_1 \in \Lambda$, let B_{x_1} be a binomial random variable with $n-1$ trials and with success probability $\int_\Lambda W_\phi(x_1, x_2)\, u_\Lambda(\mathrm{d}x_2)$. We obtain

$$\mathbb{P}[d_G(n) = k] = \int_\Lambda \mathbb{P}[B_{x_1} = k]\, u_\Lambda(\mathrm{d}x_1).$$

Next, let B be a binomial random variable with $n-1$ trials and success probability $\frac{C_\phi}{\nu(\Lambda)}$. Observe that, by the definition of C_ϕ, it holds for ν-almost all $x_1 \in \Lambda$ that $\int_\Lambda W_\phi(x_1, x_2)\, u_\Lambda(\mathrm{d}x_2) \leq \frac{C_\phi}{\nu(\Lambda)}$. Thus, we have that B stochastically dominates B_{x_1} for u_Λ almost all $x_1 \in \Lambda$. Consequently,

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

we obtain

$$\mathbb{P}[d_G(n) \geq a] \leq \int_\Lambda \mathbb{P}[B \geq a]\, u_\Lambda(\mathrm{d}x_1) = \mathbb{P}[B \geq a].$$

Observing that $\mathbb{E}(B) = \frac{n-1}{\nu(\Lambda)} C_\phi$ and applying Chernoff bound yields

$$\mathbb{P}\left[d_G(n) \geq (1+\alpha)\frac{n-1}{\nu(\Lambda)} C_\phi\right] \leq \mathrm{e}^{-\frac{\min\{\alpha,\alpha^2\} C_\phi (n-1)}{3\nu(\Lambda)}}.$$

Setting $n \geq 3\max\{\alpha^{-1}, \alpha^{-2}\} \ln\left(q^{-1}\right) C_\phi^{-1} \nu(\Lambda) + 1$ we have $\mathbb{P}\left[d_G(n) \geq (1+\alpha)\frac{n-1}{\nu(\Lambda)} C_\phi\right] \leq q$, which proves the claim. □

Combining Theorem 4.1, Lemma 4.2, and Theorem 4.3, we prove our main algorithmic result for approximating the partition function of repulsive Gibbs point processes.

Proof of Theorem 4.2. We start by giving a more precise outline of the algorithmic idea. To this end, we define

$$N = \max\left\{ 324\varepsilon^{-2} \max\left\{\mathrm{e}^6 \lambda^2 \nu(\Lambda)^2, \ln\left(4\varepsilon^{-1}\right)^2\right\}, \right.$$
$$\left. 24 \max\left\{\frac{1}{\mathrm{e}^{-\lambda C_\phi}}, \frac{\lambda C_\phi}{\left(\mathrm{e}^{-\lambda C_\phi}\right)^2}\right\} \lambda \nu(\Lambda) \ln\left(24 \max\left\{\frac{1}{\mathrm{e}^{-\lambda C_\phi}}, \frac{\lambda C_\phi}{\left(\mathrm{e}^{-\lambda C_\phi}\right)^2}\right\} \lambda \nu(\Lambda)\right)^2 \right\}.$$

We now use the following procedure to approximate $\Xi_\Lambda(\lambda, \phi)$:

1. Choose some integer $n \geq N$ and draw a graph $G \sim \mathbb{G}_{W_\phi, n}$ for W_ϕ as in Definition 4.1.
2. If $d_G \geq \frac{\mathrm{e}n}{\lambda \nu(\Lambda)}$, return an arbitrary value.
3. Else, use the algorithm from Theorem 4.3 to $\frac{\varepsilon}{3}$-approximate $Z_G(z_n)$ for $z_n = \lambda \frac{\nu(\Lambda)}{n}$ with an error probability of at most $\frac{1}{9}$ and return the result.

We proceed by arguing that this procedure yields an ε-approximation of $\Xi_\Lambda(\lambda, \phi)$ in time poly $\left(\nu(\Lambda)\varepsilon^{-1}\right)$. We start by bounding the probability that the computed value is not an ε-approximation.

First, we assume that, whenever $d_G \geq \frac{\mathrm{e}n}{\lambda \nu(\Lambda)}$, the algorithm returns no ε-approximation in step 2. Let A be the event that this happens. Second, let B denote the event that the hard-core partition function $Z_G(z_n)$ of G as in step 1 is not an $\frac{\varepsilon}{3}$-approximation of $\Xi_\Lambda(\lambda, \phi)$. Finally, let C denote the event we do not manage to compute an $\frac{\varepsilon}{3}$-approximation of $Z_G(z_n)$ in step 3. Note that the probability that the above procedure does not output an ε-approximation for $\Xi_\Lambda(\lambda, \phi)$ is upper bounded by

$$\mathbb{P}\left[A \cup (B \cap \overline{A}) \cup (C \cap \overline{B} \cap \overline{A})\right] \leq \mathbb{P}[A] + \mathbb{P}(B) + \mathbb{P}[C\overline{A}].$$

We proceed with bounding each of these probabilities separately.

To bound $\mathbb{P}[A]$, let $x = 24 \max\left\{\frac{1}{\mathrm{e}^{-\lambda C_\phi}}, \frac{\lambda C_\phi}{\left(\mathrm{e}^{-\lambda C_\phi}\right)^2}\right\} \lambda \nu(\Lambda)$. As we are interested in asymptotic behavior in terms of $\nu(\Lambda)$, we may assume that $\nu(\Lambda)$ is sufficiently large to ensure $x \geq 5$. Note that for this, we have to exclude the case $\lambda = 0$, which trivially yields $\Xi_\Lambda(\lambda, \phi) = 1$. Now, observe that for $x \geq 5$ it holds that $x \ln(x)^2 \geq x \ln\left(x \ln(x)^2\right)$. Next, observe that $n \geq x \ln(x)^2$. Thus, we have

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

$n \geq x \ln(n)$. Furthermore, by $n \geq 5 \ln(5)^2 \geq \mathrm{e} \geq 2$, we have

$$\begin{aligned} n - 1 &\geq 12 \max\left\{\frac{1}{\mathrm{e} - \lambda C_\phi}, \frac{\lambda C_\phi}{\left(\mathrm{e} - \lambda C_\phi\right)^2}\right\} \lambda \nu(\Lambda) \ln(n) \\ &= 3\left(\ln(9)\ln(n) + \ln(n)\right) \max\left\{\frac{1}{\mathrm{e} - \lambda C_\phi}, \frac{\lambda C_\phi}{\left(\mathrm{e} - \lambda C_\phi\right)^2}\right\} \lambda \nu(\Lambda) \\ &\geq 3 \ln(9n) \max\left\{\frac{1}{\left(\mathrm{e} - \lambda C_\phi\right)}, \frac{\lambda C_\phi}{\left(\mathrm{e} - \lambda C_\phi\right)^2}\right\} \lambda \nu(\Lambda). \end{aligned}$$

Thus, we obtain

$$n \geq 3 \max\left\{\frac{\lambda C_\phi}{\mathrm{e} - \lambda C_\phi}, \left(\frac{\lambda C_\phi}{\mathrm{e} - \lambda C_\phi}\right)^2\right\} \ln(9n)\, C_\phi^{-1} \nu(\Lambda) + 1$$

and by Lemma 4.2

$$\mathbb{P}\left[d_G \geq \frac{\mathrm{en}}{\lambda \nu(\Lambda)}\right] \leq \mathbb{P}\left[d_G \geq \left(1 + \frac{\mathrm{e} - \lambda C_\phi}{\lambda C_\phi}\right) \frac{n-1}{\nu(\Lambda)} C_\phi\right] \leq \frac{1}{9}.$$

To bound $\mathbb{P}(B)$, note that for $n \geq 324\varepsilon^{-2} \max\left\{\mathrm{e}^6 \lambda^2 \nu(\Lambda)^2, \ln\left(4\varepsilon^{-1}\right)^2\right\}$ Theorem 4.1 yields

$$\mathbb{P}(B) = \mathbb{P}\left[|Z_G(z_n) - \Xi_\Lambda(\lambda, \phi)| \geq \frac{\varepsilon}{3} \Xi_\Lambda(\lambda, \phi)\right] \leq \frac{1}{9}.$$

Finally, note that, by Theorem 4.3, we can obtain an $\frac{\varepsilon}{3}$-approximation of $Z_G(z_n)$ with error probability at most $\frac{1}{9}$ in time $\widetilde{\mathrm{O}}\left(n^2\varepsilon^{-2}\right)$ as long as $z_n < z_{\mathrm{c}}\left(d_G\right)$, meaning that $\mathbb{P}[C|\overline{A}] \leq \frac{1}{9}$.

We obtain that the error probability is bounded by $\frac{1}{3}$. To finish the proof, we need to argue that our algorithm has the desired running time. To this end, note that $N \in \mathrm{O}\left(\nu(\Lambda)^2\varepsilon^{-2}\right)$. Thus, we can also choose $n \in \mathrm{O}(\nu(\Lambda)^2\varepsilon^{-2})$. By assumption, step 1 can be computed in time $t_{\Lambda,\phi}(n) = t_{\Lambda,\phi}\left(\mathrm{O}\left(\nu(\Lambda)^2\varepsilon^{-2}\right)\right)$. Furthermore, step 2 can be computed in time $\widetilde{\mathrm{O}}\left(n^2\nu(\Lambda)^{-1}\right) = \widetilde{\mathrm{O}}\left(\nu(\Lambda)^3\varepsilon^{-4}\right)$ and, by Theorem 4.3, step 3 runs in time $\widetilde{\mathrm{O}}\left(n^2\varepsilon^{-2}\right) = \widetilde{\mathrm{O}}\left(\nu(\Lambda)^4\varepsilon^{-6}\right)$ for $z_n < z_{\mathrm{c}}\left(d_G\right)$. Consequently, the overall running time is in $\widetilde{\mathrm{O}}\left(\nu(\Lambda)^4\varepsilon^{-6}\right) + t_{\Lambda,\phi}\left(\mathrm{O}\left(\nu(\Lambda)^2\varepsilon^{-2}\right)\right)$. □

5. Sampling from Gibbs point processes

In this section, we propose an approximate sampling algorithm for the Gibbs measure of a repulsive Gibbs point process. To this end, consider the setting of Section 4.2. Formally, the problem of *ε-approximate sampling* from $P_\Lambda^{(\lambda,\phi)}$ is defined as producing a random point configuration with a distribution that has a total variation distance of at most ε to $P_\Lambda^{(\lambda,\phi)}$. As for approximating the partition function, we consider an ε-approximate sampler efficient if the running time is polynomial in the volume $\nu(\Lambda)$ and in ε^{-1}.

With respect to sampling from $P_\Lambda^{(\lambda,\phi)}$, it is less obvious how Theorem 4.1 can be utilized. However, under mild assumptions, we obtain an approximate sampler, given in Algorithm 1.

Our main theorem in this section is as follows.

Theorem 5.1. *Assume we can sample from the uniform distribution u_Λ in time t_Λ and, for every $x, y \in \Lambda$, evaluate $\phi(x, y)$ in time t_ϕ. If the Gibbs point process $P_\Lambda^{(\lambda,\phi)}$ is simple and $\lambda < \frac{\mathrm{e}}{C_\phi}$ then, for every $\varepsilon \in \mathbb{R}_{>0}$, Algorithm 1 samples ε-approximately from $P_\Lambda^{(\lambda,\phi)}$ and has running time $\widetilde{\mathrm{O}}\left(\nu(\Lambda)^2\varepsilon^{-4} + \nu(\Lambda)^2\varepsilon^{-3}t_\Lambda + \nu(\Lambda)^4\varepsilon^{-6}t_\phi\right)$.*

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

Algorithm 1: Approximate sampling algorithm for a repulsive point process (Λ, λ, ϕ).

Data: Instance of a repulsive Gibbs point process (Λ, λ, ϕ), error bound $\varepsilon \in (0, 1]$
Result: multiset of points in Λ

set

$$n = \left\lceil \max\left\{ \begin{array}{c} 8\frac{18^2 \cdot 12}{\varepsilon^3} \max\left\{e^6 \lambda^2 \nu(\Lambda)^2, \ln\left(\frac{4 \cdot 18}{\varepsilon}\right)\right\}, \\ 6\ln\left(\frac{4e}{\varepsilon}\right) \max\left\{\frac{1}{e-\lambda C_\phi}, \frac{\lambda C_\phi}{(e-\lambda C_\phi)^2}\right\} \lambda \nu(\Lambda) \ln\left(3\ln\left(\frac{4e}{\varepsilon}\right) \max\left\{\frac{1}{e-\lambda C_\phi}, \frac{\lambda C_\phi}{(e-\lambda C_\phi)^2}\right\} \lambda \nu(\Lambda)\right)^2 \end{array} \right\} \right\rceil;$$

for each $i \in [n]$ draw $X_i \sim u_\Lambda$ independently;
draw $E \subseteq \binom{[n]}{2}$ s.t. $\{i, j\} \in E$ with probability $W_\phi(X_i, X_j) = 1 - e^{-\phi(X_i, X_j)}$ independently;
set $G = ([n], E)$;
if *maximum degree* $d_G \geq \frac{en}{\lambda \nu(\Lambda)}$ **then**
    set $X = \emptyset$;
else
    sample $S \in 2^{[n]}$ $\frac{\varepsilon}{4}$-approximately from the hard-core distribution μ_{G, z_n} where $z_n = \frac{\lambda \nu(\Lambda)}{n}$;
    set $X = \{X_i \mid i \in S\}$ (possibly multiset);
end
return X;

Algorithm 2: Modified sampling process

Data: Instance of a repulsive Gibbs point process (Λ, λ, ϕ), error bound $\varepsilon \in (0, 1]$

Result: multiset of points in Λ

set

$$n = \left\lceil \max\left\{ \begin{array}{c} 8\frac{18^2 \cdot 12}{\varepsilon^3} \max\left\{e^6 \lambda^2 \nu(\Lambda)^2, \ln\left(\frac{4 \cdot 18}{\varepsilon}\right)\right\}, \\ 6\ln\left(\frac{4e}{\varepsilon}\right) \max\left\{\frac{1}{e-\lambda C_\phi}, \frac{\lambda C_\phi}{(e-\lambda C_\phi)^2}\right\} \lambda \nu(\Lambda) \ln\left(3\ln\left(\frac{4e}{\varepsilon}\right) \max\left\{\frac{1}{e-\lambda C_\phi}, \frac{\lambda C_\phi}{(e-\lambda C_\phi)^2}\right\} \lambda \nu(\Lambda)\right)^2 \end{array} \right\} \right\rceil$$

for each $i \in [n]$ draw $X_i \sim u_\Lambda$ independently;
draw $E \subseteq \binom{[n]}{2}$ s.t. $\{i, j\} \in E$ with probability $W_\phi(X_i, X_j) = 1 - e^{-\phi(X_i, X_j)}$ independently;
set $G = ([n], E)$;
sample $S' \sim \mu_{G, z_n}$ where $z_n = \frac{\lambda \nu(\Lambda)}{n}$;
set $Y = \{X_i \mid i \in S'\}$ (possibly multiset);
Return Y;

To prove Theorem 5.1, we start by analyzing a simplified algorithm, given in Algorithm 2. The main difference between Algorithm 1 and Algorithm2 is that the latter one does not check if the maximum degree of the sampled graph G is bounded and that is assumes access to a perfect sampler for μ_{G, z_n}. It is not clear if such a perfect sampler for the hard-core Gibbs distribution can be realized in polynomial time, especially for arbitrary vertex degrees. Therefore, Algorithm 2 is not

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

suitable for algorithmic applications. However, the main purpose of Algorithm 2 is that the distribution of point multisets that it outputs are much easier to analyse. We use this, together with a coupling argument, to bound the total variation distance between the output of Algorithm 1 and $P_\Lambda^{(\lambda,\phi)}$. Once this is done, it remains to show that Algorithm 1 satisfies the running time requirements, given in Theorem 5.1.

To analyse the output distribution of Algorithm 2, we start by considering the resulting distribution of multisets of points (or counting measures respectively) when conditioning on the event that the hard-core partition function $Z_G(z_n)$ of the drawn graph G is close to the partition function of the continuous process $\Xi_\Lambda(\lambda,\phi)$. More specifically, for any given n and $\alpha \in \mathbb{R}_{\geq 0}$, let $A_\alpha^{(n)} = \{H \in \mathcal{G}_n \mid |Z_H(z_n) - \Xi_\Lambda(\lambda,\phi)| \leq \alpha\, \Xi_\Lambda(\lambda,\phi)\}$. We derive an explicit density for the output of Algorithm 2 with respect to a Poisson point process under the condition that $G \in A_\alpha^{(n)}$ for some sufficiently small α. To this end, we use the following characterization of simple point processes via so called *void probabilities*.

Theorem 5.2 (Rényi–Mönch, see [6, Theorem 9.2.XII]). *Let P and Q be simple point process on $(\mathbb{X}, d)$. If, for $\eta_P \sim P$ and $\eta_Q \sim Q$ and for all bounded $B \in \mathcal{B}$, it holds that*

$$\mathbb{P}\left[\eta_P(B) = 0\right] = \mathbb{P}\left[\eta_Q(B) = 0\right],$$

then $P = Q$.

Theorem 5.2 greatly simplifies proving that a given candidate function actually is a valid density for the point process in question, as it implies that it is sufficient to check if it yields the correct void probabilities.

Lemma 5.1. *For $\varepsilon \in (0,1]$ let $\widehat{P}_\varepsilon$ be the point process produced by Algorithm 2 conditioned on $G \in A_{\frac{\varepsilon}{12}}^{(n)}$, and let Q_λ denote a Poisson point process with intensity λ. If the Gibbs point process $P_\Lambda^{(\lambda,\phi)}$ is simple, then $\widehat{P}_\varepsilon$ has a density with respect to Q_λ of the form*

$$g_\varepsilon(\eta) = \mathbb{1}_\eta(\Lambda) \leq n \cdot \mathbb{1}_\eta \in \mathcal{N}(\Lambda) \cdot \mathbb{P}\left[G \in A_{\frac{\varepsilon}{12}}^{(n)}\right]^{-1} \left(\prod_{i=0}^{\eta(\Lambda)-1} 1 - \frac{i}{n}\right) \mathrm{e}^{-H(\eta)} \Psi_n^{(\eta(\Lambda))}\left(\varphi(\eta)\right) \mathrm{e}^{\lambda \nu(\Lambda)},$$

where φ maps every finite counting measure η to an arbitrary but fixed tuple $(x_1, \ldots, x_{\eta(\mathbb{X})})$ such that $\eta = \sum_{i=1}^{\eta(\mathbb{X})} \delta_{x_i}$ and

$$\Psi_n^{(k)}(\boldsymbol{x}) = \sum_{\substack{H \in A_{\frac{\varepsilon}{12}}^{(n)}: \\ [k] \in \mathcal{I}(H)}} \frac{1}{Z_H(z_n)} \int_{\Lambda^{n-k}} \left(\prod_{\substack{(i,j) \in [k] \times [n-k]: \\ \{i,j+k\} \in E_H}} 1 - \mathrm{e}^{-\phi(x_i,y_j)}\right) \left(\prod_{\substack{(i,j) \in [k] \times [n-k]: \\ \{i,j+k\} \notin E_H}} \mathrm{e}^{-\phi(x_i,y_j)}\right)$$
$$\left(\prod_{\substack{\{i,j\} \in \binom{[n-k]}{2}: \\ \{i+k,j+k\} \in E_H}} 1 - \mathrm{e}^{-\phi(y_i,y_j)}\right) \left(\prod_{\substack{\{i,j\} \in \binom{[n-k]}{2}: \\ \{i+k,j+k\} \notin E_H}} \mathrm{e}^{-\phi(y_i,y_j)}\right) u_\Lambda^{n-k}(d\boldsymbol{y})$$

for all $\boldsymbol{x} = (x_1, \ldots, x_k) \in \Lambda^k$.

Proof. First, observe that in Algorithm 2 it holds that $G \sim \mathbb{G}_{W_\phi,n}$ for W_ϕ as in Definition 4.1. As $n \geq 4\frac{12^3}{\varepsilon^3} \max\left\{\mathrm{e}^6 \lambda^2 \nu(\Lambda)^2, \ln\left(4\frac{12}{\varepsilon}\right)^2\right\}$, Theorem 4.1 implies that $\mathbb{P}\left[G \in A_{\frac{\varepsilon}{12}}^{(n)}\right] \geq 1 - \frac{\varepsilon}{12} > 0$. Therefore, conditioning on the event $G \in A_{\frac{\varepsilon}{12}}^{(n)}$ is well defined.

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

Next, note that, for all $x \in \Lambda$ it holds that

$$\mathbb{P}\left[\eta(\{x\}) \geq 2\right] \geq \frac{\mathrm{e}^{-\phi(x,x)}\lambda^2 \nu\left(\{x\}\right)^2}{\Xi_\Lambda(\lambda,\phi)} \geq \frac{\mathrm{e}^{-\phi(x,x)}\lambda^2 \nu(\{x\})^2}{\mathrm{e}^{\lambda\nu(\Lambda)}}$$

for $\eta \sim P_\Lambda^{(\lambda,\phi)}$. Thus, if $P_\Lambda^{(\lambda,\phi)}$ is simple (i.e., $\mathbb{P}\left[\eta(\{x\}) \geq 2\right] = 0$ for all $x \in \Lambda$), it holds that $\lambda = 0$ or, for all $x \in \Lambda$, $\nu(\{x\}) = 0$ or $\phi(x,x) = \infty$. This implies that the output of Algorithm 2 is simple as well, and consequently $\widehat{P_\varepsilon}$ is a simple point process.

Knowing that $\widehat{P_\varepsilon}$ is simple, Theorem 5.2 implies that, in order to verify that g_ε is indeed a density for $\widehat{P_\varepsilon}$, it suffices to prove that it yields the correct void probabilities. Formally, this means showing that for all bounded $B \in \mathcal{B}$ it holds that

$$\mathbb{P}\left[Y \cap B = \emptyset | G \in A_{\frac{\varepsilon}{\overline{12}}}^{(n)}\right] = \int_{\mathcal{N}} \mathbb{1}_\eta(B) = 0 \cdot g_\varepsilon(\eta) Q_\lambda(\mathrm{d}\eta)$$

for Y and G as in Algorithm 2.

To prove this, we first write

$$\begin{aligned}
&\int_{\mathcal{N}} \mathbb{1}_\eta(B) = 0 \cdot g_\varepsilon(\eta) Q_\lambda(\mathrm{d}\eta) \\
&\quad = \int_{\mathcal{N}(\Lambda)} \mathbb{1}_\eta(B) = 0 \cdot g_\varepsilon(\eta) Q_\lambda(\mathrm{d}\eta) \\
&\quad = \mathrm{e}^{-\lambda\nu(\Lambda)} \cdot \left(g_\varepsilon(\mathbf{0}) + \sum_{k \in \mathbb{N}_{\geq 1}} \frac{\lambda^k}{k!} \int_{\Lambda^k} \mathbb{1}_{\forall i \in [k]: x_i \notin B} \cdot g_\varepsilon\left(\sum_{i \in [k]} \delta_{x_i}\right) \nu^k(\mathrm{d}\mathbf{x}) \right),
\end{aligned}$$

where $\mathbf{0}$ denotes the constant 0 measure on $\mathbb{X}$. Note that

$$\begin{aligned}
&\mathrm{e}^{-\lambda\nu(\Lambda)} g_\varepsilon(\mathbf{0}) \\
&= \mathbb{P}\left[G \in A_{\frac{\varepsilon}{\overline{12}}}^{(n)}\right]^{-1} \cdot \sum_{H \in A_{\frac{\varepsilon}{\overline{12}}}^{(n)}} \frac{1}{Z_H(z_n)} \int_{\Lambda^n} \left(\prod_{\substack{\{i,j\} \in \binom{[n]}{2}: \\ \{i,j\} \in E_H}} 1 - \mathrm{e}^{-\phi(y_i,y_j)} \right) \left(\prod_{\substack{\{i,j\} \in \binom{[n]}{2}: \\ \{i,j\} \notin E_H}} \mathrm{e}^{-\phi(y_i,y_j)} \right) u_\Lambda^n(\mathrm{d}\mathbf{y}) \\
&= \mathbb{P}\left[G \in A_{\frac{\varepsilon}{\overline{12}}}^{(n)}\right]^{-1} \cdot \sum_{H \in A_{\frac{\varepsilon}{\overline{12}}}^{(n)}} \mathbb{P}\left[S' = \emptyset | G = H\right] \mathbb{P}\left[G = H\right] \\
&= \frac{\mathbb{P}\left[S' = \emptyset \wedge G \in A_{\frac{\varepsilon}{\overline{12}}}^{(n)}\right]}{\mathbb{P}\left[G \in A_{\frac{\varepsilon}{\overline{12}}}^{(n)}\right]} \\
&= \mathbb{P}\left[S' = \emptyset | G \in A_{\frac{\varepsilon}{\overline{12}}}^{(n)}\right]
\end{aligned}$$

for S' as in Algorithm 2. We proceed with a case distinction based on k. For every $k > n$ and $(x_1, \ldots, x_k) \in \Lambda^k$ we have $g_\varepsilon\left(\sum_{i \in [k]} \delta_{x_k}\right) = 0$. Therefore, we get

$$\int_{\Lambda^k} \mathbb{1}_{\forall i \in [k]: x_i \notin B} \cdot g_\varepsilon\left(\sum_{i \in [k]} \delta_{x_i}\right) \nu^k(\mathrm{d}\mathbf{x}) = 0$$

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

for all $k > n$. Now, consider $k \in [n]$ and observe that for all $\mathbf{x} = (x_1, \dots, x_k) \in \Lambda^k$ we have

$$\Psi_n^{(k)}\left(\varphi\left(\sum_{i\in[k]}\delta_{x_i}\right)\right) = \Psi_n^{(k)}(\mathbf{x})$$

by symmetry. Moreover, it holds that

$$\frac{\lambda^k}{k!}\left(\prod_{i=0}^{k-1} 1 - \frac{i}{n}\right) = \binom{n}{k}\frac{z_n^k}{\nu(\Lambda)^k}.$$

Therefore, we have

$$\begin{aligned}
&\mathrm{e}^{-\lambda\nu(\Lambda)}\frac{\lambda^k}{k!}\int_{\Lambda^k} \mathbb{1}_{\forall i\in[k]:x_i\notin B}\cdot g_\varepsilon\left(\sum_{i\in[k]}\delta_{x_i}\right)\nu^k(\mathrm{d}\mathbf{x}) \\
&\qquad = \mathbb{P}\left[G \in A^{(n)}_{\frac{\varepsilon}{12}}\right]^{-1}\binom{n}{k}z_n^k\int_{\Lambda^k}\mathbb{1}_{\forall i\in[k]:x_i\notin B}\cdot\left(\prod_{i,j\in\binom{[k]}{2}}\mathrm{e}^{-\phi(x_i,x_j)}\right)\Psi_n^{(k)}(\mathbf{x})u_\Lambda^k(\mathrm{d}\mathbf{x}).
\end{aligned}$$

Next, note that

$$\begin{aligned}
&z_n^k\int_{\Lambda^k}\mathbb{1}_{\forall i\in[k]:x_i\notin B}\cdot\left(\prod_{i,j\in\binom{[k]}{2}}\mathrm{e}^{-\phi(x_i,x_j)}\right)\Psi_n^{(k)}(\mathbf{x})u_\Lambda^k(\mathrm{d}\mathbf{x}) \\
&= \sum_{H\in A^{(n)}_{\frac{\varepsilon}{12}}}\mathbb{1}_{[}k]\in\mathcal{I}\,(H)\frac{z_n^k}{Z_H(z_n)}\int_{\Lambda^n}\mathbb{1}_{\forall i\in[k]:x_i\notin B}\cdot\left(\prod_{\substack{\{i,j\}\in\binom{[n]}{2}:\\ \{i,j\}\in E_H}}1-\mathrm{e}^{-\phi(x_i,x_j)}\right) \\
&\left(\prod_{\substack{\{i,j\}\in\binom{[n]}{2}:\\ \{i,j\}\notin E_H}}\mathrm{e}^{-\phi(x_i,x_j)}\right)u_\Lambda^n(\mathrm{d}\mathbf{x}) \\
&= \sum_{H\in A^{(n)}_{\frac{\varepsilon}{12}}}\mathbb{P}\left[S'=[k]|G=H\right]\mathbb{P}\left[G=H\wedge\forall i\in[k]:X_i\notin B\right]
\end{aligned}$$

for $X_1, \dots, X_n$ as in Algorithm 2. Furthermore, because the event $S' = [k]$ is independent of $X_1, \dots, X_n$ given G, it holds that

$$\begin{aligned}
&\sum_{H\in A^{(n)}_{\frac{\varepsilon}{12}}}\mathbb{P}\left[S'=[k]|G=H\right]\mathbb{P}\left[G=H\wedge\forall i\in[k]:X_i\notin B\right] \\
&= \sum_{H\in A^{(n)}_{\frac{\varepsilon}{12}}}\mathbb{P}\left[S'=[k]\wedge G=H\wedge\forall i\in[k]:X_i\notin B\right] \\
&= \mathbb{P}\left[S'=[k]\wedge G\in A^{(n)}_{\frac{\varepsilon}{12}}\wedge\forall i\in[k]:X_i\notin B\right]
\end{aligned}$$

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

and

$$\begin{aligned}
&\mathrm{e}^{-\lambda\nu(\Lambda)}\frac{\lambda^k}{k!}\int_{\Lambda^k}\mathbb{1}_{\forall i\in[k]:x_i\notin B}\cdot g_\varepsilon\left(\sum_{i\in[k]}\delta_{x_i}\right)\nu^k(\mathrm{d}\mathbf{x})\\
&\quad=\binom{n}{k}\frac{\mathbb{P}\left[S'=[k]\wedge G\in A^{(n)}_{\frac{\varepsilon}{12}}\wedge\forall i\in[k]:X_i\notin B\right]}{\mathbb{P}\left[G\in A^{(n)}_{\frac{\varepsilon}{12}}\right]}\\
&\quad=\binom{n}{k}\mathbb{P}\left[S'=[k]\wedge\forall i\in[k]:X_i\notin B|G\in A^{(n)}_{\frac{\varepsilon}{12}}\right]\\
&\quad=\sum_{V'\in\binom{[n]}{k}}\mathbb{P}\left[S'=V'\wedge\forall i\in V':X_i\notin B|G\in A^{(n)}_{\frac{\varepsilon}{12}}\right],
\end{aligned}$$

where the last equality is due to symmetry. Combining everything yields

$$\begin{aligned}
\int_{\mathcal{N}}\mathbb{1}_{\eta(B)=0}\cdot g_\varepsilon(\eta)Q_\lambda(\mathrm{d}\eta)&=\mathbb{P}\left[S'=\emptyset|G\in A^{(n)}_{\frac{\varepsilon}{12}}\right]\\
&\quad+\sum_{k=1}^{n}\sum_{V'\in\binom{[n]}{k}}\mathbb{P}\left[S'=V'\wedge\forall i\in V':X_i\notin B|G\in A^{(n)}_{\frac{\varepsilon}{12}}\right]\\
&=\sum_{V'\in 2^{[n]}}\mathbb{P}\left[S'=V'\wedge\forall i\in V':X_i\notin B|G\in A^{(n)}_{\frac{\varepsilon}{12}}\right]\\
&=\mathbb{P}\left[\forall i\in S':X_i\notin B|G\in A^{(n)}_{\frac{\varepsilon}{12}}\right]\\
&=\mathbb{P}\left[Y\cap B=\emptyset|G\in A^{(n)}_{\frac{\varepsilon}{12}}\right],
\end{aligned}$$

which concludes the proof. □

We proceed by upper and lower bounding the density $g_\varepsilon(\eta)$ in terms of the density of $P^{(\lambda,\phi)}_\Lambda$. To this end, we use the following basic facts about the partition function of the hard-core model.

Observation 5.1 (see [11]). *For every undirected graph $G=(V,E)$ the following holds:*

1. *For all $z_1,z_2\in\mathbb{R}_{\geq 0}$*
$$Z_G[z_1]\leq Z_G[z_1+z_2]\leq \mathrm{e}^{z_2|V|}Z_G[z_1].$$
2. *For all $z\in\mathbb{R}_{\geq 0}$ and $U\subseteq V$*
$$Z_{G-U}(z)\leq Z_G(z)\leq \mathrm{e}^{z|U|}Z_{G-U}(z),$$
where $G-U$ denotes the subgraph of G that is induced by $V\setminus U$.

Using Observation 5.1 we derive the following bounds.

Lemma 5.2. *Consider the setting of Lemma 5.1 and let f denote the density of $P^{(\lambda,\phi)}_\Lambda$ with respect to Q_λ. For n as in Algorithm 2 and all $\eta\in\mathcal{N}$ with $\eta(\Lambda)\leq\min\left\{\sqrt{\frac{\varepsilon n}{12}},\frac{\varepsilon}{40\lambda\nu(\Lambda)+\varepsilon}n\right\}$ it holds that*

$$\left(1-\frac{\varepsilon}{4}\right)f(\eta)\leq g_\varepsilon(\eta)\leq\left(1+\frac{\varepsilon}{4}\right)f(\eta).$$

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

Proof. First, recall that, when $P_\Lambda^{(\lambda,\phi)}$ is simple, its density with respect to Q_λ can be expressed as

$$f(\eta) = \mathbb{1}_{\eta\in\mathcal{N}(\Lambda)} \cdot \frac{\mathrm{e}^{-H(\eta)}}{\Xi_\Lambda(\lambda,\phi)} \mathrm{e}^{\lambda\nu(\Lambda)}$$

for every $\eta \in \mathcal{N}$. Therefore, we have

$$g_\varepsilon(\eta) = \mathbb{1}_{\eta(\Lambda)\le n} \cdot \mathbb{P}\left[G \in A^{(n)}_{\frac{\varepsilon}{12}}\right]^{-} 1 \left(\prod_{i=0}^{\eta(\Lambda)-1} 1 - \frac{i}{n}\right) \Psi_n^{(\eta(\Lambda))}\left(\varphi(\eta)\right) \Xi_\Lambda(\lambda,\phi) f(\eta).$$

As we focus on η with $\eta(\Lambda) \le \sqrt{\frac{\varepsilon n}{12}} \le n$, we omit the indicator $\mathbb{1}_\eta(\Lambda) \le n$ from now on.

We proceed by deriving an upper bound on $g_\varepsilon(\eta)$ for $\eta \in \mathcal{N}$ with $\eta(\Lambda) \le \min\left\{\sqrt{\frac{\varepsilon n}{12}}, \frac{\varepsilon}{40\lambda\nu(\Lambda)+\varepsilon} n\right\}$. To this end, note that

$$\left(\prod_{i=0}^{\eta(\Lambda)-1} 1 - \frac{i}{n}\right) \le 1.$$

Moreover, for $G \sim \mathbb{G}_{W_\phi,n}$ for W_ϕ as in Definition 4.1 and $n \ge 4\frac{12^3}{\varepsilon^3} \max\left\{\mathrm{e}^6\lambda^2\nu(\Lambda)^2, \ln\left(4\frac{12}{\varepsilon}\right)^2\right\}$ Theorem 4.1 yields $\mathbb{P}\left[G \in A^{(n)}_{\frac{\varepsilon}{12}}\right] \ge 1 - \frac{\varepsilon}{12}$. Finally, observe that for all $\mathbf{x} \in \Lambda^k$ for $k \le n$ we have

$$\begin{aligned}
\Psi_n^{(k)}(\mathbf{x}) &\le \frac{1}{\left(1-\frac{\varepsilon}{12}\right)\Xi_\Lambda(\lambda,\phi)} \sum_{\substack{H\in A^{(n)}_{\frac{\varepsilon}{12}}:\\ [k]\in\mathcal{I}(H)}} \int_{\Lambda^{n-k}} \left(\prod_{\substack{(i,j)\in[k]\times[n-k]:\\ \{i,j+k\}\in E_H}} 1-\mathrm{e}^{-\phi(x_i,y_j)}\right) \\
&\quad \cdot \left(\prod_{\substack{(i,j)\in[k]\times[n-k]:\\ \{i,j+k\}\notin E_H}} \mathrm{e}^{-\phi(x_i,y_j)}\right)\left(\prod_{\substack{\{i,j\}\in\binom{[n-k]}{2}:\\ \{i+k,j+k\}\in E_H}} 1-\mathrm{e}^{-\phi(y_i,y_j)}\right)\left(\prod_{\substack{\{i,j\}\in\binom{[n-k]}{2}:\\ \{i+k,j+k\}\notin E_H}} \mathrm{e}^{-\phi(y_i,y_j)}\right) u_\Lambda^{n-k}(\mathrm{d}\mathbf{y}) \\
&= \frac{1}{\left(1-\frac{\varepsilon}{12}\right)\Xi_\Lambda(\lambda,\phi)} \int_{\Lambda^{n-k}} \sum_{\substack{H\in A^{(n)}_{\frac{\varepsilon}{12}}:\\ [k]\in\mathcal{I}(H)}} \left(\prod_{\substack{(i,j)\in[k]\times[n-k]:\\ \{i,j+k\}\in E_H}} 1-\mathrm{e}^{-\phi(x_i,y_j)}\right) \\
&\quad \cdot \left(\prod_{\substack{(i,j)\in[k]\times[n-k]:\\ \{i,j+k\}\notin E_H}} \mathrm{e}^{-\phi(x_i,y_j)}\right)\left(\prod_{\substack{\{i,j\}\in\binom{[n-k]}{2}:\\ \{i+k,j+k\}\in E_H}} 1-\mathrm{e}^{-\phi(y_i,y_j)}\right)\left(\prod_{\substack{\{i,j\}\in\binom{[n-k]}{2}:\\ \{i+k,j+k\}\notin E_H}} \mathrm{e}^{-\phi(y_i,y_j)}\right) u_\Lambda^{n-k}(\mathrm{d}\mathbf{y}) \\
&\le \frac{1}{\left(1-\frac{\varepsilon}{12}\right)\Xi_\Lambda(\lambda,\phi)} \int_{\Lambda^{n-k}} 1\, u_\Lambda^{n-k}(\mathrm{d}\mathbf{y}) \\
&\le \frac{1}{\left(1-\frac{\varepsilon}{12}\right)\Xi_\Lambda(\lambda,\phi)}.
\end{aligned}$$

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

Given that $\varepsilon \leq 1$ we get

$$g_\varepsilon(\eta) \leq \left(1 - \frac{\varepsilon}{12}\right)^{-2} f(\eta) \leq \left(1 + \frac{\varepsilon}{11}\right)^2 f(\eta) \leq \left(1 + \frac{\varepsilon}{4}\right) f(\eta),$$

which proves the upper bound.

For the lower bound, note that

$$\mathbb{P}\left[G \in A^{(n)}_{\frac{\varepsilon}{12}}\right]^{-1} \geq 1$$

and for $\eta(\Lambda) \leq \sqrt{\frac{\varepsilon n}{12}}$

$$\left(\prod_{i=0}^{\eta(\Lambda)-1} 1 - \frac{i}{n}\right) \geq \left(1 - \frac{\eta(\Lambda)}{n}\right)^{\eta(\Lambda)} \geq 1 - \frac{\eta(\Lambda)^2}{n} \geq 1 - \frac{\varepsilon}{12}.$$

We proceed by lower bounding $\Psi_n^{(k)}(\mathbf{x})$. First, observe that

$$\begin{aligned}
\Psi_n^{(k)}(\mathbf{x}) &\geq \frac{1}{\left(1 + \frac{\varepsilon}{12}\right) \Xi_\Lambda(\lambda, \phi)} \sum_{\substack{H \in A^{(n)}_{\frac{\varepsilon}{12}}: \\ [k] \in \mathcal{I}(H)}} \int_{\Lambda^{n-k}} \left(\prod_{\substack{(i,j) \in [k] \times [n-k]: \\ \{i,j+k\} \in E_H}} 1 - \mathrm{e}^{-\phi(x_i, y_j)} \right) \\
&\quad \cdot \left(\prod_{\substack{(i,j) \in [k] \times [n-k]: \\ \{i,j+k\} \notin E_H}} \mathrm{e}^{-\phi(x_i, y_j)} \right) \left(\prod_{\substack{\{i,j\} \in \binom{[n-k]}{2}: \\ \{i+k,j+k\} \in E_H}} 1 - \mathrm{e}^{-\phi(y_i, y_j)} \right) \left(\prod_{\substack{\{i,j\} \in \binom{[n-k]}{2}: \\ \{i+k,j+k\} \notin E_H}} \mathrm{e}^{-\phi(y_i, y_j)} \right) u_\Lambda^{n-k}(\mathrm{d}\mathbf{y}) \\
&= \frac{1}{\left(1 + \frac{\varepsilon}{12}\right) \Xi_\Lambda(\lambda, \phi)} \sum_{\substack{H \in \mathcal{G}_n: \\ [k] \in \mathcal{I}(H)}} \mathbb{1}_H \in A^{(n)}_{\frac{\varepsilon}{12}} \int_{\Lambda^{n-k}} \left(\prod_{\substack{(i,j) \in [k] \times [n-k]: \\ \{i,j+k\} \in E_H}} 1 - \mathrm{e}^{-\phi(x_i, y_j)} \right) \\
&\quad \cdot \left(\prod_{\substack{(i,j) \in [k] \times [n-k]: \\ \{i,j+k\} \notin E_H}} \mathrm{e}^{-\phi(x_i, y_j)} \right) \left(\prod_{\substack{\{i,j\} \in \binom{[n-k]}{2}: \\ \{i+k,j+k\} \in E_H}} 1 - \mathrm{e}^{-\phi(y_i, y_j)} \right) \left(\prod_{\substack{\{i,j\} \in \binom{[n-k]}{2}: \\ \{i+k,j+k\} \notin E_H}} \mathrm{e}^{-\phi(y_i, y_j)} \right) u_\Lambda^{n-k}(\mathrm{d}\mathbf{y}).
\end{aligned}$$

Next, for each graph $H \in \mathcal{G}_n$, let $H' = ([n-k], E')$ denote the subgraph that results from $H - [k]$ after relabelling each vertex in $i \in [n] \setminus [k]$ to $i - k \in [n-k]$ (note that this relabelling is formally required for $H' \in \mathcal{G}_{n-k}$). By Observation 5.1 and the fact that $z_n \leq z_{n-k}$ and $Z_{H'}(z) = Z_{H-[k]}(z)$ for all $z \in \mathbb{R}_{\geq 0}$ we have

$$Z_H(z_n) \leq \mathrm{e}^{z_n k} Z_{H'}(z_n) \leq \mathrm{e}^{\frac{k}{n} \lambda \nu(\Lambda)} Z_{H'}(z_{n-k}).$$

On the other hand, note that

$$\begin{aligned}
z_{n-k} = \frac{\lambda \nu(\Lambda)}{n-k} &= \frac{n}{n(n-k)} \lambda \nu(\Lambda) = \left(\frac{n-k}{n(n-k)} + \frac{k}{n(n-k)} \right) \lambda \nu(\Lambda) \\
&= \left(\frac{1}{n} + \frac{k}{n(n-k)} \right) \lambda \nu(\Lambda) = z_n + \frac{k}{n(n-k)} \lambda \nu(\Lambda).
\end{aligned}$$

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

Therefore, Observation 5.1 yields

$$Z_H(z_{n-k}) \le \mathrm{e}^{\frac{k}{n-k}\lambda\nu(\Lambda)} Z_H(z_n)$$

and

$$Z_H(z_n) \ge \mathrm{e}^{-\frac{k}{n-k}\lambda\nu(\Lambda)} Z_H(z_{n-k}) \ge \mathrm{e}^{-\frac{k}{n-k}\lambda\nu(\Lambda)} Z_{H'}(z_{n-k}).$$

Thus, for $k \le \frac{\varepsilon}{40\lambda\nu(\Lambda)+\varepsilon} n$ we have

$$\mathrm{e}^{-\frac{\varepsilon}{40}} Z_{H'}(z_{n-k}) \le Z_H(z_n) \le \mathrm{e}^{\frac{\varepsilon}{40}} Z_{H'}(z_{n-k}).$$

As $\mathrm{e}^{-\frac{\varepsilon}{40}}\left(1-\frac{\varepsilon}{18}\right) \ge \left(1-\frac{\varepsilon}{12}\right)$ and $\mathrm{e}^{\frac{\varepsilon}{40}}\left(1+\frac{\varepsilon}{18}\right) \le \left(1+\frac{\varepsilon}{12}\right)$ for all $\varepsilon \in [0,1]$, this means that $H' \in A^{(n-k)}_{\frac{\varepsilon}{18}}$ is a sufficient condition for $H \in A^{(n)}_{\frac{\varepsilon}{12}}$ and

$$\begin{aligned}
&\sum_{\substack{H\in\mathcal{G}_n:\\ [k]\in\mathcal{I}(H)}} \mathbb{1}_{H\in A^{(n)}_{\frac{\varepsilon}{12}}} \int_{\Lambda^{n-k}} \left(\prod_{\substack{(i,j)\in[k]\times[n-k]:\\ \{i,j+k\}\in E_H}} 1-\mathrm{e}^{-\phi(x_i,y_j)}\right)\left(\prod_{\substack{(i,j)\in[k]\times[n-k]:\\ \{i,j+k\}\notin E_H}} \mathrm{e}^{-\phi(x_i,y_j)}\right)\\
&\left(\prod_{\substack{\{i,j\}\in\binom{[n-k]}{2}:\\ \{i+k,j+k\}\in E_H}} 1-\mathrm{e}^{-\phi(y_i,y_j)}\right)\left(\prod_{\substack{\{i,j\}\in\binom{[n-k]}{2}:\\ \{i+k,j+k\}\notin E_H}} \mathrm{e}^{-\phi(y_i,y_j)}\right) u_\Lambda^{n-k}(\mathrm{d}\mathbf{y})\\
&\ge \sum_{H'\in\mathcal{G}_{n-k}} \mathbb{1}_{H'\in A^{(n-k)}_{\frac{\varepsilon}{18}}} \int_{\Lambda^{n-k}} \left(\prod_{\substack{\{i,j\}\in\binom{[n-k]}{2}:\\ \{i,j\}\in E_{H'}}} 1-\mathrm{e}^{-\phi(y_i,y_j)}\right)\left(\prod_{\substack{\{i,j\}\in\binom{[n-k]}{2}:\\ \{i,j\}\notin E_{H'}}} \mathrm{e}^{-\phi(y_i,y_j)}\right)\\
&\sum_{F\subseteq[k]\times[n-k]} \left(\prod_{\substack{(i,j)\in[k]\times[n-k]:\\ (i,j)\in F}} 1-\mathrm{e}^{-\phi(x_i,y_j)}\right)\left(\prod_{\substack{(i,j)\in[k]\times[n-k]:\\ (i,j)\notin F}} \mathrm{e}^{-\phi(x_i,y_j)}\right) u_\Lambda^{n-k}(\mathrm{d}\mathbf{y})\\
&= \sum_{H'\in\mathcal{G}_{n-k}} \mathbb{1}_{H'\in A^{(n-k)}_{\frac{\varepsilon}{18}}} \int_{\Lambda^{n-k}} \left(\prod_{\substack{\{i,j\}\in\binom{[n-k]}{2}:\\ \{i,j\}\in E_{H'}}} 1-\mathrm{e}^{-\phi(y_i,y_j)}\right)\left(\prod_{\substack{\{i,j\}\in\binom{[n-k]}{2}:\\ \{i,j\}\notin E_{H'}}} \mathrm{e}^{-\phi(y_i,y_j)}\right) u_\Lambda^{n-k} = (\mathrm{d}\mathbf{y})\\
&= \mathbb{P}\left[G' \in A^{(n-k)}_{\frac{\varepsilon}{18}}\right]
\end{aligned}$$

for $G' \sim \mathbb{G}_{W_\phi,n-k}$. Next, observe that $n \ge 1$ we have $k \le \min\left\{\sqrt{\frac{\varepsilon n}{12}}, \frac{\varepsilon}{40\lambda\nu(\Lambda)+\varepsilon} n\right\} \le \frac{n}{2}$ and $n-k \ge \frac{n}{2}$. Therefore, for $n \ge 8\frac{18^2\cdot 12}{\varepsilon^3}\max\left\{\mathrm{e}^6\lambda^2\nu(\Lambda)^2, \ln\left(\frac{4\cdot 18}{\varepsilon}\right)^2\right\}$ Theorem 4.1 yields $\mathbb{P}\left[G' \in A^{(n-k)}_{\frac{\varepsilon}{18}}\right] \ge 1-\frac{\varepsilon}{12}$ for $G' \sim \mathbb{G}_{W_\phi,n-k}$. Consequently, we have

$$\Psi_n^{(k)}(\mathbf{x}) \ge \frac{1-\frac{\varepsilon}{12}}{1+\frac{\varepsilon}{12}} \cdot \frac{1}{\Xi_\Lambda(\lambda,\phi)}$$

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

and

$$g_\varepsilon(\eta) \geq \left(1-\frac{\varepsilon}{12}\right)^2 \left(1+\frac{\varepsilon}{12}\right)^{-1} f(\eta) \geq \left(1-\frac{\varepsilon}{12}\right)^3 f(\eta) \geq \left(1-\frac{\varepsilon}{4}\right) f(\eta),$$

which concludes the proof. □

We proceed by using Lemmas 5.1 and 5.2 to bound the total variation distance between $P_\Lambda^{(\lambda,\phi)}$ and the output distribution of Algorithm 2. However, as Lemma 5.2 only provides information for point sets that are sufficiently small compared to n, we need a different way to deal with large point configurations. To this end, the following two results are useful. The first lemma is a domination result for the size of independent sets, drawn from a hard-core model.

Lemma 5.3. *Let $G \in \mathcal{G}_n$ for some $n \in \mathbb{N}$ and let $z \in \mathbb{R}_{\geq 0}$. For $S \sim \mu_{G,z}$ it holds that $|S|$ is stochastically dominated by a binomial random variable with n trials and success probability $\frac{z}{1+z}$.*

Proof. We use a coupling argument to prove this statement. Consider the following procedure for sampling a set $S_n \subseteq [n]$:

1. Start with $S_0 = \emptyset$.
2. For each $i \in [n]$, set $S_i = S_{i-1} \cup \{i\}$ with probability $\mathbb{P}[i \in S | S_{i-1} \subseteq S]$ and $S_i = S_{i-1}$ otherwise.

Note that S_n has law $\sigma \sim \mu_{G,z}$. Due to the definition of this process, it suffices to consider sequences $(S_i)_{i\in[n]\cup\{0\}}$ such that the event $S_i \subseteq S$ has non-zero probability for all i. Further, note that for all $i \in [n]$

$$\mathbb{P}[i \in S | S_{i-1} \subseteq S] \leq \frac{z}{1+z}$$

Now, we consider a modified process $(S'_i)_{i\in[n]\cup\{0\}}$ with $S'_0 = \emptyset$ and $S'_i = S'_{i-1} \cup \{i\}$ with probability $\frac{z}{1+z}$. Observe that $(S_i)_{i\in[n]\cup\{0\}}$ and $(S'_i)_{i\in[n]\cup\{0\}}$ can be coupled in such a way that $S_i \subseteq S'_i$ whenever $S_{i-1} \subseteq S'_{i-1}$ for all $i \in [n]$. As initially $S_0 = S'_0$, the same coupling yields $S_n \subseteq S'_n$. Finally, observing that $|S'_n|$ follows a binomial distribution with n trials and success probability $\frac{z}{1+z}$ concludes the proof. □

The second lemma is the analog of Lemma 5.3 for repulsive point processes. However, proving it is slightly more technically involved. We start by introducing some additional notation and terminology. For two counting measures $\eta_1, \eta_2 \in \mathcal{N}$, we write $\eta_1 \leq \eta_2$ if $\eta_1(B) \leq \eta_2(B)$ for every $B \in \mathcal{B}$. A measurable function $h{:}\mathcal{N} \to \mathbb{R}$ is called *increasing* if $h(\eta_1) \leq h(\eta_2)$ for all $\eta_1 \leq \eta_2$. Moreover, for some $\kappa \in \mathbb{R}_{\geq 0}$, let Q_κ denote the Poisson point process with intensity κ and let P be a point process that has a density f_P with respect to Q_κ. A function $\zeta : \mathcal{N} \times \mathbb{X} \to \mathbb{R}_{\geq 0}$ is called a *Papangelou intensity* for P (w.r.t. Q_κ) if, for all $\eta \in \mathcal{N}$ and $x \in \mathbb{X}$, it holds that

$$f_P(\eta + \delta_x) = \zeta(\eta, x) f_P(\eta).$$

The domination lemma we are aiming for is implied by the following result.

Theorem 5.3 [(14, Theorem 1.1)]. *Let Q_κ be a Poisson point process of intensity $\kappa \in \mathbb{R}_{\geq 0}$ and let P_1, P_2 be point processes that are absolutely continuous with respect to Q_κ. Assume P_1 and P_2 have Papangelou intensities ζ_1 and ζ_2. If, for all $x \in \mathbb{X}$ and $\eta_1, \eta_2 \in \mathcal{N}$ with $\eta_1 \leq \eta_2$, $\zeta_1(\eta_1, x) \leq \zeta_2(\eta_2, x)$, then, for all increasing $h : \mathcal{N} \to \mathbb{R}$, it holds that*

$$\int_{\mathcal{N}} h(\eta) P_1(d\eta) \leq \int_{\mathcal{N}} h(\eta) P_2(d\eta).$$

With that, we show the following simple domination result.

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

Lemma 5.4. *For $\eta \sim P_\Lambda^{(\lambda,\phi)}$ it holds that $\eta(\Lambda)$ is dominated by a Poisson random variable with parameter $\lambda\nu(\Lambda)$.*

Proof. Let Q_λ denote a Poisson point process with intensity λ. Note that

$$f_1(\eta) = \mathbb{1}_\eta \in \mathcal{N}(\Lambda) \cdot \frac{1}{\Xi_\Lambda(\lambda,\phi)} \left(\prod_{\{x,y\} \in \binom{X(\eta)}{2}} \mathrm{e}^{-N_x(\eta)N_y(\eta)\phi(x,y)} \right) \left(\prod_{x \in X(\eta)} \mathrm{e}^{-\frac{N_x(\eta)(N_x(\eta)-1)}{2}\phi(x,x)} \right) \mathrm{e}^{\lambda\nu(\Lambda)}$$

is a density for $P_\Lambda^{(\lambda,\phi)}$ with respect to Q_λ. Therefore,

$$\zeta_1, (\eta, x) = \mathbb{1}_{x\in\Lambda} \prod_{y \in X(\eta)} \mathrm{e}^{-N_y(\eta)\phi(x,y)}$$

is a Papangelou intensity for $P_\Lambda^{(\lambda,\phi)}$. Moreover, let P denote the point process defined by the density $f_2(\eta) = \mathbb{1}_\eta \in \mathcal{N}(\Lambda)$ and observe that $\zeta_2(\eta, x) = \mathbb{1}_x \in \Lambda$ is a Papangelou intensity for P.

For all $k \in \mathbb{N}$, let $h_k(\eta) = \mathbb{1}_{\eta(\Lambda)\geq k}$ and observe that h_k is increasing. Further, note that, for all $x \in \mathbb{X}$ and $\eta_1, \eta_2 \in \mathcal{N}$, it holds that

$$\zeta_1, (\eta, x) = \mathbb{1}_{x\in\Lambda} \prod_{y \in X_{\eta_1}} \mathrm{e}^{-N_y(\eta_1)\phi(x,y) \leq \mathbb{1}_{x\in\Lambda}} = \zeta_1, (\eta_2, x).$$

By Theorem 5.3, this implies that for all $k \in \mathbb{N}$

$$\int_{\mathcal{N}} h_k(\eta) P_\Lambda^{(\lambda,\phi)}(\mathrm{d}\eta) \leq \int_{\mathcal{N}} h_k(\eta) P(\mathrm{d}\eta).$$

Consequently, for $\eta \sim P_\Lambda^{(\lambda,\phi)}$ and $\xi \sim P$ and for all $k \in \mathbb{N}$, it holds that

$$\mathbb{P}\left[\eta(\Lambda) \geq k\right] \leq \mathbb{P}\left[\xi(\Lambda) \geq k\right]$$

and observing that $\xi(\Lambda)$ follows a Poisson distribution with parameter $\lambda\nu(\Lambda)$ concludes the proof. □

We now bound the total variation distance between the output of Algorithm 2 and $P_\Lambda^{(\lambda,\phi)}$.

Lemma 5.5. *For every given $\varepsilon \in (0, 1]$, Algorithm 2 is an $\frac{\varepsilon}{2}$-approximate sampler from $P_\Lambda^{(\lambda,\phi)}$.*

Proof. We start by bounding the total variation distance between $d_{\mathrm{tv}}\left(P_\Lambda^{(\lambda,\phi)}, \widehat{P_\varepsilon}\right)$ for $\widehat{P_\varepsilon}$ as in Lemma 5.1. The statement then follows from a coupling argument. Let Q_λ denote a Poisson point process of intensity λ. Let g_ε be the density of $\widehat{P_\varepsilon}$ with respect to Q_λ as given in Lemma 5.1 and let f denote the density of $P_\Lambda^{(\lambda,\phi)}$ with respect to Q_λ. Moreover, set $m = \min\left\{\sqrt{\frac{\varepsilon n}{12}}, \frac{\varepsilon}{40\lambda\nu(\Lambda)+\varepsilon} n\right\}$. Note that the total variation distance can be expressed as

$$\begin{aligned} d_{\mathrm{tv}}\left(P_\Lambda^{(\lambda,\phi)}, \widehat{P_\varepsilon}\right) &= \int_{\mathcal{N}} \left|f(\eta) - g_\varepsilon(\eta)\right| Q_\lambda(\mathrm{d}\eta) \\ &= \int_{\mathcal{N}} \mathbb{1}_{\eta(\Lambda)\leq m} \left|f(\eta) - g_\varepsilon(\eta)\right| Q_\lambda(\mathrm{d}\eta) + \int_{\mathcal{N}} \mathbb{1}_{\eta(\Lambda)>m} \left|f(\eta) - g_\varepsilon(\eta)\right| Q_\lambda(\mathrm{d}\eta). \end{aligned}$$

By Lemma 5.2, we get

$$\int_{\mathcal{N}} \mathbb{1}_{\eta(\Lambda)\leq m} \left|f(\eta) - g_\varepsilon(\eta)\right| Q_\lambda(\mathrm{d}\eta) \leq \frac{\varepsilon}{4} \int_{\mathcal{N}} \mathbb{1}_{\eta(\Lambda)\leq m} f(\eta) Q_\lambda(\mathrm{d}\eta) \leq \frac{\varepsilon}{4}.$$

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

Further, it holds that

$$\int_{\mathcal{N}} \mathbb{1}_{\eta(\Lambda)>m} \left|f(\eta) - g_\varepsilon(\eta)\right| Q_\lambda(\mathrm{d}\eta) \leq \int_{\mathcal{N}} \mathbb{1}_{\eta(\Lambda)>m} f(\eta) Q_\lambda(\mathrm{d}\eta) + \int_{\mathcal{N}} \mathbb{1}_{\eta(\Lambda)>m} g_\varepsilon(\eta) Q_\lambda(\mathrm{d}\,\eta)$$
$$= \mathbb{P}\left[\xi(\Lambda) > m\right] + \mathbb{P}\left[|Y| > m | G \in A^{(n)}_{\frac{\varepsilon}{12}}\right]$$

for $\xi \sim P^{(\lambda,\phi)}_\Lambda$, and G and Y as in Algorithm 2.

We proceed by bounding each of these probability separately. Note that, by our choice of n it holds that $m \geq \frac{12}{\varepsilon}\lambda\nu(\Lambda)$. By Lemma 5.4, we have $\mathbb{E}[\xi(\Lambda)] \leq \lambda\nu(\Lambda)$. Thus, Markov's inequality yields $\mathbb{P}\left[\xi(\Lambda) > m\right] \leq \frac{\varepsilon}{12}$. Moreover, note that $|Y| = |S'|$ for S' as in Algorithm 2. As $S' \sim \mu_{H,zn}$ for some $H \in A^{(n)}_{\frac{\varepsilon}{12}} \subseteq \mathcal{G}_n$ and Lemma 5.3 applies to all such graphs, we get

$$\mathbb{E}\left[|Y| | G \in A^{(n)}_{\frac{\varepsilon}{12}}\right] \leq \frac{z_n}{1+z_n} n \leq z_n n = \lambda\nu(\Lambda).$$

Again, applying Markov's inequality gives $\mathbb{P}\left[|Y| > m | G \in A^{(n)}_{\frac{\varepsilon}{12}}\right] \leq \frac{\varepsilon}{12}$. Consequently, we have

$$d_{\mathrm{tv}}\left(P^{(\lambda,\phi)}_\Lambda, \widehat{P}_\varepsilon\right) \leq \frac{\varepsilon}{4} + \frac{\varepsilon}{6} = \frac{5}{12}\varepsilon.$$

To finish the proof, we now relate the output of Algorithm 2 with $\widehat{P}_\varepsilon$ by using a coupling argument. To this end, note that Algorithm 2 can be used to sample from $\widehat{P}_\varepsilon$ by simply restarting the sampler whenever $G \notin A^{(n)}_{\frac{\varepsilon}{12}}$. For our choice of n we know that with a probability of $\mathbb{P}\left[G \in A^{(n)}_{\frac{\varepsilon}{12}}\right] \geq 1 - \frac{\varepsilon}{12}$ only a single run of Algorithm 2 is required. By this coupling, the total variation distance between the output of Algorithm 2 and $\widehat{P}_\varepsilon$ is at most $\frac{\varepsilon}{12}$. Finally, applying triangle inequality shows that the total variation distance between the output of Algorithm 2 and $P^{(\lambda,\phi)}_\Lambda$ is bounded by $\frac{5}{12}\varepsilon + \frac{\varepsilon}{12} = \frac{\varepsilon}{2}$, which concludes the proof. □

Using Lemma 5.5, we are able to prove that Algorithm 1 is an ε-approximate sampler for $P^{(\lambda,\phi)}_\Lambda$. In order to argue that Algorithm 1 also satisfies the running time requirements, given in Theorem 5.1, we require an efficient approximate sampler from the hard-core distribution μ_{G,z_n}. To this end, we use the following known result.

Theorem 5.4 [(2, Theorem 5)]. *Let $G = (V, E)$ be an undirected graph with maximum vertex degree bounded by $d_G \in \mathbb{N}_{\geq 2}$ and let $z \in \mathbb{R}_{\geq 0}$ with*

$$z < z_c\left(d_G\right) = \frac{\left(d_G - 1\right)^{d_G - 1}}{\left(d_G - 2\right)^{d_G}}.$$

Then, for all $\varepsilon \in (0, 1]$, there is an ε-approximate sampler for the hard-core Gibbs distribution $\mu_{G,z}$ with an expected running time of $\mathrm{O}\left(|V| \ln\left(\frac{|V|}{\varepsilon}\right)\right)$.

Proof of Theorem 5.1. We start by arguing that Algorithm 1 is an ε-approximate sampler for $P^{(\lambda,\phi)}_\Lambda$. To this end, we show that the total variation distance between the output distributions of Algorithm 1 and Algorithm 2 is bounded by $\frac{\varepsilon}{2}$. Using the triangle inequality and Lemma 5.5 then yields the desired result. To bound the total variation distance between the Algorithm 1 and Algorithm 2 by $\frac{\varepsilon}{2}$, it suffices to construct a coupling of both algorithms such that their output coincides with a probability of at least $1 - \frac{\varepsilon}{2}$. This is, we want to find a coupling of both algorithms such that $X \neq Y$ with probability at most $\frac{\varepsilon}{2}$, where X and Y are as in Algorithm 1 and Algorithm 2.

To construct such a coupling, we start by letting both algorithms draw the same points $X_1, \ldots, X_n$ and construct the same graph G. If $d_G \geq \frac{\mathrm{en}}{\lambda\nu(\Lambda)}$, then we may just assume $X \neq Y$.

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

Otherwise, if $d_G < \frac{\mathrm{en}}{\lambda\nu(\Lambda)}$, then $S = S'$ is a sufficient condition for $X = Y$. As S' is drawn from μ_{G,z_n} and S is drawn from an $\frac{\varepsilon}{4}$ approximation of that distribution, they can be coupled in such a way that $\mathbb{P}\left[S' \neq S\right] \leq \frac{\varepsilon}{4}$. Using this coupling of Algorithm 1 and Algorithm 2, we have

$$\mathbb{P}\left[X \neq Y\right] \leq \mathbb{P}\left[d_G \geq \frac{\mathrm{en}}{\lambda\nu(\Lambda)}\right] + \frac{\varepsilon}{4} \cdot \mathbb{P}\left[d_G < \frac{\mathrm{en}}{\lambda\nu(\Lambda)}\right] \leq \mathbb{P}\left[d_G \geq \frac{\mathrm{en}}{\lambda\nu(\Lambda)}\right] + \frac{\varepsilon}{4}.$$

Therefore, it remains to prove that $d_G \geq \frac{\mathrm{en}}{\lambda\nu(\Lambda)}$ with probability at most $\frac{\varepsilon}{4}$, where $G \sim \mathbb{G}_{W_\phi,n}$ for W_ϕ as in Definition 4.1. We follow a similar arguments as in the proof of Theorem 4.2. Note that, for our choice of n, there exists $x \geq 3\ln\left(\frac{4\mathrm{e}}{\varepsilon}\right)\max\left\{\frac{1}{\mathrm{e}-\lambda C_\phi}, \frac{\lambda C_\phi}{(\mathrm{e}-\lambda C_\phi)^2}\right\}\lambda\nu(\Lambda)$ such that $n = 2x\ln(x)^2$. Moreover, we have $n \geq \mathrm{e} \geq 2$ and, for $\nu(\Lambda)$ (consequently x) sufficiently large, it holds that $2x\ln(x)^2 \geq 2x\ln\left(2x\ln(x)^2\right) = 2x\ln(n)$. Therefore, we have

$$\begin{aligned} n - 1 &\geq \frac{n}{2} \\ &\geq 3\ln\left(\frac{4\mathrm{e}}{\varepsilon}\right)\max\left\{\frac{1}{\mathrm{e}-\lambda C_\phi}, \frac{\lambda C_\phi}{\left(\mathrm{e}-\lambda C_\phi\right)}^2\right\}\lambda\nu(\Lambda)\ln(n) \\ &\geq 3\ln\left(\frac{4n}{\varepsilon}\right)\max\left\{\frac{1}{\mathrm{e}-\lambda C_\phi}, \frac{\lambda C_\phi}{\left(\mathrm{e}-\lambda C_\phi\right)}^2\right\}\lambda\nu(\Lambda) \end{aligned}$$

and by Lemma 4.2

$$\mathbb{P}\left[d_G \geq \frac{\mathrm{en}}{\lambda\nu(\Lambda)}\right] \leq \mathbb{P}\left[d_G \geq \left(1 + \frac{\mathrm{e}-\lambda C_\phi}{\lambda C_\phi}\right)\frac{n-1}{\nu(\Lambda)}C_\phi\right] \leq \frac{\varepsilon}{4}.$$

To prove Theorem 5.1, it remains to show that Algorithm 1 satisfies the given running time requirements. To this end, note that, for all $\lambda < \frac{\mathrm{e}}{C_\phi}$, it holds that $n \in \widetilde{\mathrm{O}}\left(\nu(\Lambda)^2\varepsilon^{-3}\right)$. Therefore, sampling $X_1, \ldots, X_n$ requires a running time of $\widetilde{\mathrm{O}}\left(nt_\Lambda\right) = \widetilde{\mathrm{O}}\left(\nu(\Lambda)^2\varepsilon^{-3}t_\Lambda\right)$. Moreover, the graph can be constructed in time $\widetilde{\mathrm{O}}\left(n^2t_\phi\right) = \widetilde{\mathrm{O}}\left(\nu(\Lambda)^4\varepsilon^{-6}t_\phi\right)$ and $d_G \geq \frac{\mathrm{en}}{\lambda\nu(\Lambda)}$ can be checked in $\mathrm{O}(1)$ if we keep track of d_G while constructing the graph. Finally, for $d_G < \frac{\mathrm{en}}{\lambda\nu(\Lambda)}$ it holds that

$$z_n = \frac{\lambda\nu(\Lambda)}{n} < \frac{\mathrm{e}}{d_G} < z_{\mathrm{c}}\left(d_G\right).$$

Thus, Theorem 5.4 guarantees the existence of an $\frac{\varepsilon}{8}$-approximate sampler from μ_{G,z_n} with an expected running time in $\mathrm{O}\left(n\ln\left(\frac{n}{\varepsilon}\right)\right) = \mathrm{O}\left(\nu(\Lambda)^2\varepsilon^{-3}\ln\left(\frac{\nu(\Lambda)}{\varepsilon}\right)\right)$. Note that, by Markov's inequality, the probability that this sampler takes more than $\frac{8}{\varepsilon}$ times its expected running time is bounded by $\frac{\varepsilon}{8}$. Therefore, if we run the sampler from Theorem 5.4 with an error bound of $\frac{\varepsilon}{8}$ and, whenever the algorithm takes more than $\frac{8}{\varepsilon}$ times its expected running time, stop it and return an arbitrary spin configuration, this results in an $\frac{\varepsilon}{4}$-approximate sampler with a guaranteed running time in $\widetilde{\mathrm{O}}\left(\nu(\Lambda)^2\varepsilon^{-4}\right)$. Consequently, Algorithm 1 runs in time $\widetilde{\mathrm{O}}\left(\nu(\Lambda)^2\varepsilon^{-4} + \nu(\Lambda)^2\varepsilon^{-3}t_\Lambda + \nu(\Lambda)^4\varepsilon^{-6}t_\phi\right)$, which concludes the proof. □

6. Discussion and outlook

The most commonly studied setting for algorithms for Gibbs point processes is when $\Lambda = [0, \ell]^d$ is a box of side length ℓ in d-dimensional Euclidean space. Usually, the following computational assumptions are made:

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

- We can perform arithmetic operations and comparison with arbitrary floating-point precision.
- We have oracle access to a uniform sampler for the continuous interval $[0, 1]$.
- We have oracle access to $e^{-\phi(x,y)}$ for any $x, y \in \mathbb{R}^d$.

In this setting, Theorem 4.2 and Theorem 5.1 provide an approximation algorithm for the partition function and an approximate sampling algorithm with running time $\widetilde{O}\left(\nu(\Lambda)^4\right)$ up to a fugacity of $\lambda < e/C_\phi$ and for arbitrary repulsive potentials ϕ.

Recently, Michelen and Perkins [23] provide a $\widetilde{O}\left(\nu(\Lambda)\right)$ approximate sampler and a $\widetilde{O}\left(\nu(\Lambda)^2\right)$ approximation for the partition function in the setting of bounded-range potentials up to a slightly better fugacity threshold of e/Δ_ϕ, where Δ_ϕ is called the potential-weighted connective constant. This result is obtained by directly applying a Markov chain Monte-Carlo algorithm on the continuous problem, which is efficient as long as the Gibbs point process satisfies a property called strong spatial mixing. The strong spatial mixing assumption restricts their approach to bounded-range potentials. This raises the question if there are efficient approximation algorithms for repulsive Gibbs point processes without bounded-range assumption for every $\lambda < e/\Delta_\phi$. As a first step towards this, we argue in an extended version of this paper [10] that a bound for the connective constant of our discretization can be obtained in terms of Δ_ϕ. However, while this is known to imply some notion of spatial mixing for the discrete hard-core model [29], it is not clear if it suffices for a polynomial-time approximation.

Moreover, Jenssen, Michelen and Ravichandran [20] recently obtained a deterministic quasi-polynomial time approximation algorithm for the partition function of repulsive Gibbs point processes with sufficiently smooth bounded-range potentials up to $\lambda < e/\Delta_\phi$. Their result is based on using a discretization to approximate the coefficients of a suitable expansion of the logarithm of the partition function. To perform the involved enumeration task efficiently, they require the potential to have bounded range. It would be interesting to see if such a quasi-polynomial deterministic approximation can be obtained without bounded-range assumptions. In particular, we wonder if a derandomized version of our discretization under similar smoothness assumptions on the potential as in [20] could give such a result.

Acknowledgements

Andreas Göbel was funded by the project PAGES (project No. 467516565) of the German Research Foundation (DFG). Marcus Pappik was funded by the HPI Research School on Data Science and Engineering.

Competing interests

The authors declare no competing interests.

References

[1] Anand, K., Göbel, A., Pappik, M. and Perkins, W. (2023) Perfect sampling for hard spheres from strong spatial mixing. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques* (N. Megow and A. D. Smith, eds), APPROX/RANDOM, September 11–13, 2023, Atlanta, Georgia, pp. 38:1–38:18.

[2] Anari, N., Jain, V., Koehler, F., Pham, H. T. and Vuong, T. (2021) Entropic independence II: optimal sampling and concentration via restricted modified log-sobolev inequalities. *CoRR* abs/2111.03247 1–53.

[3] Bacher, A. K., Schrøder, T. B. and Dyre, J. C. (2014) Explaining why simple liquids are quasi-universal. *Nat. Commun.* **5**(1) 1–7.

[4] Baddeley, A., Gregori, P., Mateu, J., Stoica, R. and Stoyan, D. (2006) *Case Studies in Spatial Point Process Modeling*, **185**, Springer.

https://doi.org/10.1017/S0963548324000282 Published online by Cambridge University Press

[5] Berne, B. J. and Pechukas, P. (1972) Gaussian model potentials for molecular interactions. *J. Chem. Phy.* **56**(8) 4213–4216.
[6] Daley, D. J.andVere-Jones, D. (2008) *An Introduction to the Theory of Point Processes. Volume II: General Theory and Structure*, Springer
[7] Efron, B. and Stein, C. (1981) The jackknife estimate of variance. *Ann. Stats.* 586–596.
[8] El Azhar, F., Baus, M., Ryckaert, J.-P. and Meijer, E. J. (2000) Line of triple points for the hard-core Yukawa model: a computer simulation study. *J. Chem. Phy.* **112**(11) 5121–5126.
[9] Feng, W., Guo, H. and Yin, Y. (2022) Perfect sampling from spatial mixing. *Random Struct. Algor.* **61**(4) 678–709.
[10] Friedrich, T., Göbel, A., Katzmann, M., Krejca, M. and Pappik, M. (2022) Using random graphs to sample repulsive Gibbs point processes with arbitrary-range potentials. *CoRR* abs/2204.01793 1–47
[11] Friedrich, T., Göbel, A., Katzmann, M., Krejca, M. S. and Pappik, M. (2023) Algorithms for hard-constraint point processes via discretization. In Computing and Combinatorics: 28th International Conference, COCOON 2022, Shenzhen, China, October 22-24, 2022. Springer, pp. 242–254.
[12] Friedrich, T., Göbel, A., Krejca, M. and Pappik, M. (2021) A spectral independence view on hard spheres via block dynamics, *48th International Colloquium on Automata, Languages, and Programming (ICALP)*, **198**, p. 15.
[13] Garcia, N. L. (2000) Perfect simulation of spatial processes, *Resenhas do Instituto de Matemática e Estatística da Universidade de São Paulo*, **4**, pp. 283–325.
[14] Georgii, H.-O. and Küneth, T. (1997) Stochastic comparison of point random fields. *J. Appl. Probab.* **34**(4) 868–881.
[15] Guo, H. and Jerrum, M. (2021) Perfect simulation of the hard disks model by partial rejection sampling. *Annales de l'Institut Henri Poincaré D.* **8** 159–177.
[16] Häggström, O., Van Lieshout, M.-C. N. and Møller, J. (1999) Characterization results and markov chain monte carlo algorithms including exact simulation for some spatial point processes, *Bernoulli*, 641–658.
[17] Helmuth, T., Perkins, W. and Petti, S. (2022) Correlation decay for hard spheres via Markov chains. *Ann. Appl. Probab.* **32**(3) 2063–2082.
[18] Huber, M. (2012) Spatial birth–death swap chains, Bernoulli, 1031–1041.
[19] Jansen, S. (2018) Gibbsian point processes, pp. 1–106
[20] Jenssen, M., Michelen, M. and Ravichandran, M. (2023) Quasipolynomial-time algorithms for Gibbs point processes. *Comb. Probab. Comp.* **33**(1) 1–15.
[21] Lovász, L. (2012) *Large networks and graph limits*. American Mathematical Society, **60**.
[22] 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. Phy.* **21**(6) 1087–1092.
[23] Michelen, M. and Perkins, W. (2022) Strong spatial mixing for repulsive point processes. *J. Stat. Phys.* **189**(1) 189–189.
[24] Møller, J. (1989) On the rate of convergence of spatial birth-and-death processes. *Ann. I. Stat. Math.* **41**(3) 565–581.
[25] Møller, J. and Waagepetersen, R. P. (2003) *Statistical inference and simulation for spatial point processes*. CRC press.
[26] Møller, J. and Waagepetersen, R. P. (2007) Modern statistics for spatial point processes. *Scand. J. Stat.* **34**(4) 643–684.
[27] Rowlinson, J. S. (1989) The Yukawa potential. *Physica A: Stat. Mech. Appl.* **156**(1) 15–34.
[28] Ruelle, D. (1999) *Statistical Mechanics: Rigorous Results*. World Scientific.
[29] Sinclair, A., Srivastava, P., Štefankovič, D. and Yin, Y. (2017) Spatial mixing and the connective constant: optimal bounds. *Probab. Theory. Rel.* **168**(1-2) 153–197.
[30] Štefankovič, D., Vempala, S. and Vigoda, E. (2009) Adaptive simulated annealing: A near-optimal connection between sampling and counting. *J. ACM* **56**(3) 1–36.
[31] Van Lieshout, M.-C. N. (2000) *Markov point processes and their applications*. World Scientific.
[32] Vigoda, E. (2001) A note on the glauber dynamics for sampling independent sets. *Electron. J Combina.* **8**(1) 1–8.
[33] Weitz, D. (2006) Counting independent sets up to the tree threshold. In Proceedings of the thirty-eighth annual ACM symposium on Theory of computing. 140–149,
[34] Yoshida, T. and Kamakura, S. (1974) Liquid-solid transitions in systems of soft repulsive forces: Softness of potentials and a maximum in melting curves. *Prog. Theor. Phys.* **52**(3) 822–839.

Cite this article: Friedrich T, Göbel A, Katzmann M, Krejca MS, and Pappik M (2024). Sampling repulsive Gibbs point processes using random graphs. *Combinatorics, Probability and Computing*. https://doi.org/10.1017/S0963548324000282