A SPECTRAL INDEPENDENCE VIEW ON HARD SPHERES VIA BLOCK DYNAMICS[*]

TOBIAS FRIEDRICH[†], ANDREAS GÖBEL[†], MARTIN S. KREJCA[‡], AND MARCUS PAPPIK[†]

Abstract. The hard-sphere model is one of the most extensively studied models in statistical physics. It describes the continuous distribution of spherical particles, governed by hard-core interactions. An important quantity of this model is the normalizing factor of this distribution, called the *partition function*. We propose a Markov chain Monte Carlo algorithm for approximating the grand canonical partition function of the hard-sphere model in d dimensions. Up to a fugacity of $\lambda < e/2^d$, the runtime of our algorithm is polynomial in the volume of the system. Key to our approach is to define a discretization that closely approximates the partition function of the continuous model. This results in a discrete hard-core instance that is exponential in the size of the initial hard-sphere model. Our approximation bound follows directly from the correlation decay threshold of an infinite regular tree with degree equal to the maximum degree of our discretization. To cope with the exponential blow-up of the discrete instance, we use block dynamics, a Markov chain that generalizes the more frequently studied Glauber dynamics by grouping the vertices of the graph into blocks and updating an entire block instead of a single vertex in each step. We prove rapid mixing of block dynamics, based on disjoint cliques as blocks, up to the tree threshold of the univariate hard-core model. This is achieved by adapting the spectral expansion method, which was recently used for bounding the mixing time of Glauber dynamics within the same parameter regime.

Key words. hard-sphere model, Markov chain, partition function, Gibbs distribution, approximate counting, spectral independence

MSC codes. 60J10, 60G55, 68Q87, 68W20, 68W25, 82-08

DOI. 10.1137/21M143697X

1. Introduction. Statistical physics models particle systems as probability distributions. One of the most fundamental and mathematically challenging models in this area is the hard-sphere model, which plays a central role in understanding the thermodynamic properties of monoatomic gases and liquids [7, 27]. It is a continuous model that studies the distribution and macroscopic behavior of indistinguishable spherical particles, assuming only hard-core interactions, i.e., no two particles can occupy the same space.

We focus on computational properties of the *grand canonical ensemble* of the hard-sphere model in a finite d-dimensional cubic region $\mathbb{V} = [0, \ell)^d$ in the Euclidean space. In the grand canonical ensemble, the system can exchange particles with its surrounding based on a fugacity parameter λ, which is inverse to the temperature of the system. For the rest of the paper, we make the common assumption that the

[†]Hasso Plattner Institute, University of Potsdam, Potsdam, Germany (tobias.friedrich@hpi.de, https://hpi.de/friedrich/, andreas.goebel@hpi.de, https://hpi.de/friedrich/people/andreas-goebel.html, marcus.pappik@hpi.de, https://hpi.de/friedrich/people/marcus-pappik.html).
[‡]Sorbonne University, CNRS, LIP6, Paris, France (martin.krejca@lip6.fr, https://webia.lip6.fr/~krejca/).

system is normalized such that the particles have unit volume. That means we fix their radii to $r = (1/v_d)^{1/d}$, where v_d is the volume of a unit sphere in d dimensions.

A simple probabilistic interpretation of the distribution of particles in the grand canonical ensemble is that centers of points that are drawn from a Poisson point process on $\mathbb{V}$ with intensity λ, conditioned on the event that no two particles overlap (i.e., every pair of centers has distance at least $2r$). The resulting distribution over particle configurations in $\mathbb{V}$ is called the *Gibbs distribution* of the model. An important quantity in such models is the so-called partition function $Z(\mathbb{V}, \lambda)$, which can be seen as the normalizing constant of the Gibbs distribution. Formally, it is defined as

$$Z(\mathbb{V}, \lambda) = 1 + \sum_{k \in \mathbb{N}_{>0}} \frac{\lambda^k}{k!} \int_{\mathbb{V}^k} D\left(x^{(1)}, \ldots, x^{(k)}\right) \mathrm{d}\nu^{d \times k},$$

where

$$D\left(x^{(1)}, \ldots, x^{(k)}\right) = \begin{cases} 1 & \text{if } d\left(x^{(i)}, x^{(j)}\right) \geq 2r \text{ for all } i, j \in [k] \text{ with } i \neq j, \\ 0 & \text{otherwise} \end{cases}$$

and $\nu^{d \times k}$ is the Lebesgue measure on $\mathbb{R}^{d \times k}$. Commonly, two computational tasks are associated with the grand canonical hard-sphere model: (1) approximating its partition function $Z(\mathbb{V}, \lambda)$, and (2) approximately sampling from the Gibbs distribution.

Studying computational aspects of the hard-sphere model carries a historical weight, as in the seminal work of Metropolis [38], the Monte Carlo method was introduced to investigate a two-dimensional hard-sphere model. Approximate-sampling Markov chain approaches have been mainly focused on the canonical ensemble of the model, that is, the system does not exchange particles with its surrounding and thus the distribution is defined over a fixed number of spheres [29, 32, 31]. Considering the grand canonical ensemble, exact sampling algorithms have appeared in the literature for the two-dimensional model without asymptotic runtime guarantees [33, 34, 42]. A result that is more aligned with theoretical computer science was given in [26], where the authors introduced an exact sampling algorithm for the grand canonical hard-sphere model in d-dimensions. Their algorithm is based on partial rejection sampling with a runtime linear in the volume of the system $|\mathbb{V}|$ when assuming a continuous computational model and access to a sampler from a continuous Poisson point process. Their approach is guaranteed to apply for $\lambda < 2^{-(d+1/2)}$.

Besides such sampling results, there is an ongoing effort to improve the known fugacity regime where the Gibbs measure is unique and correlations decay exponentially fast [18, 12, 30, 39]. Note that for many discrete spin systems, such as the hard-core model, correlation decay is closely related to the applicability of different methods for efficient approximation of the partition function [49, 23, 52]. Recently, the correlation decay bounds for the hard-sphere model were improved in [30] to $\lambda < 2/2^{-d}$, using probabilistic arguments, and in [39] to $\lambda < \mathrm{e}/2^d$, based on an analytic approach. A common feature of [30] and [39] is that they translated tools originally developed in theoretical computer science for investigating the discrete hard-core model to the continuous domain.

Our work is in line with the computational view on the hard-sphere model but more algorithmic in nature. We investigate the range of the fugacity λ for which an approximation of $Z(\mathbb{V}, \lambda)$ can be obtained efficiently in terms of the volume of the system $|\mathbb{V}|$, assuming a discrete computational model. To this end, we state all running time asymptotics in terms of $|\mathbb{V}|$ and treat the fugacity λ as a constant. Our main

result is that for all $\lambda < \mathrm{e}/2^d$ there is a randomized algorithm for ε-approximating the partition function in time polynomial in $|\mathbb{V}|$ and $1/\varepsilon$.

THEOREM 1.1. *Let $(\mathbb{V}, \lambda)$ be an instance of the continuous hard-sphere model with $\mathbb{V} = [0, \ell)^d$. If there is a $\delta \in (0, 1]$ such that*

$$\lambda \leq (1 - \delta)\frac{\mathrm{e}}{2^d},$$

then for each $\varepsilon \in (0, 1]$ there is a randomized ε-approximation of $Z(\mathbb{V}, \lambda)$ computable in time polynomial in $|\mathbb{V}|^{1/\delta^2}$ and $\frac{1}{\varepsilon}$.

Note that this bound on λ precisely coincides with the bound for uniqueness of the Gibbs measure in the thermodynamic limit, established in [39]. For many discrete spin systems, such as the hard-core model or general antiferromagnetic 2-state spin systems, the region of efficient approximation of the partition function is closely related to uniqueness of the Gibbs measure. More precisely, it can be shown that the partition function of every graph of maximum degree Δ can be approximated efficiently if the corresponding Gibbs distribution on an infinite Δ regular tree is unique [36, 51]. A detailed discussion for the discrete hard-core model can be found in the next subsection. In a sense, Theorem 1.1 can be seen as the algorithmic counterpart of the uniqueness result for the continuous hard-sphere model in [39]. This answers an open question asked in [39].

The way we prove our result is quite contrary to [30] and [39]. Instead of translating discrete tools from computer science into the continuous domain, we rather discretize the hard-sphere model. By this, existing algorithmic and probabilistic techniques for discrete models become available, and we avoid continuous analysis.

Our applied discretization scheme is fairly intuitive and results in an instance of the discrete hard-core model on a suitably constructed graph. The hard-core model has been extensively studied in the computer science community. However, as the resulting graph is exponential in the size of the continuous system $|\mathbb{V}|$, existing approaches for approximating its partition function, such as Markov chain Monte Carlo methods based on Glauber dynamics, are not feasible. We overcome this problem by grouping the vertices of the graph into blocks and applying a Markov chain Monte Carlo approach based on *block dynamics*, which update an entire block in each step rather than a single vertex. By adapting a recently introduced technique for bounding the mixing time of Markov chains based on local spectral expansion [3], we then prove that if the blocks consist of disjoint cliques, the mixing time of the block dynamics does only depend on the number of blocks that are used. Using the fact that the graph that results from our discretization scheme can always be partitioned into $O(|\mathbb{V}|)$ disjoint cliques, this yields a mixing time that is polynomial in the size of the continuous system. Together with a known self-reducibility scheme based on cliques, this results in the desired approximation algorithm.

Note that we aim for a rigorous algorithmic result for approximating the partition function of the continuous hard-sphere model. To be in line with commonly used discrete computational models, our Markov chain Monte Carlo algorithm does not assume access to a continuous sampler but instead samples approximately from a discretized version of the Gibbs distribution. Note that sampling from the continuous hard-sphere partition function cannot be done using a discrete computation model as this would involve infinite floating-point precision. For practical matters, our discretization of the Gibbs distribution can be seen as an approximation of the original continuous Gibbs measure. However, a rigorous comparison between both

distributions based on total variation distance is not applicable, due to the fact that one is discrete whereas the other is continuous in nature.

In subsections 1.1 to 1.3, we discuss our technical contributions in more detail and explain how they relate to the existing literature. Finally, in subsection 1.4, we discuss subsequent work and open problems.

1.1. Discretization and hard-core model. Our discretization scheme expresses the hard-sphere partition function as the partition function of an instance of the (univariate) hard-core model. An instance of the hard-core model is a tuple (G, λ) where $G = (V, E)$ is an undirected graph and $\lambda \in \mathbb{R}_{>0}$. Its partition function is defined as

$$Z(G, \lambda) := \sum_{I \in \mathcal{I}(G)} \lambda^{|I|},$$

where $\mathcal{I}(G)$ denotes the independent sets of G. A common way to obtain an approximation for the partition function is by applying a Markov chain Monte Carlo algorithm. This involves sampling from the Gibbs distribution $\mu^{(G,\lambda)}$ of (G, λ), which is a probability distribution on $\mathcal{I}(G)$ that assigns each independent set $I \in \mathcal{I}(G)$ the probability

$$\mu^{(G,\lambda)}(I) = \frac{\lambda^{|I|}}{Z(G, \lambda)}.$$

Conditions for efficient approximation of the hard-core partition function have been studied extensively in the theoretical computer science community. Due to hardness results in [49] and [23], it is known that for general graphs of maximum degree $\Delta \geq 3$ there is a critical parameter value $\lambda_c(\Delta) = (\Delta - 1)^{\Delta-1}/(\Delta - 2)^{\Delta}$ such that there is no FPRAS for the partition function of (G, λ) for $\lambda > \lambda_c(\Delta)$, unless RP = NP. On the other hand, in [52] it was proven that there is a deterministic algorithm for approximating the partition function of (G, λ) for $\lambda < \lambda_c(\Delta)$ that runs in time $|V|^{O(\log(\Delta))}$. The critical value $\lambda_c(\Delta)$ is especially interesting, as it precisely coincides with the upper bound on λ for which the hard-core model on an infinite Δ-regular tree exhibits strong spatial mixing and a unique Gibbs distribution [52]. For this reason, it is also referred to as the *tree threshold*. This relation between computational hardness and phase transition in statistical physics is one of the most celebrated results in the area. Both the hardness results [24, 4] and the approximation algorithms [45, 28] were later partially generalized for complex λ.

Note that the computational hardness above the tree threshold $\lambda_c(\Delta)$ for general graphs of maximum degree Δ applies to both randomized and deterministic algorithms. However, in the randomized setting, Markov chain Monte Carlo methods are known to improve the runtime of the algorithm introduced in [52]. Those approaches use the vertexwise self-reducibility of the hard-core model to construct a randomized approximation of the partition function based on an approximate sampler for the Gibbs distribution. Commonly, a Markov chain on the state space $\mathcal{I}(G)$, called *Glauber dynamics*, is used to construct the sampling scheme. At each step, a vertex $v \in V$ is chosen uniformly at random. With probability $\lambda/(1 + \lambda)$ the chain tries to add v to the current independent set and otherwise it tries to remove it. The resulting Markov chain is ergodic and reversible with respect to the Gibbs distribution, meaning that it eventually converges to $\mu^{(G,\lambda)}$. A sequence of results has shown that for all $\Delta \geq 3$ there is a family of graphs with maximum degree Δ such that the Glauber dynamics are torpidly mixing for $\lambda > \lambda_c(\Delta)$ [14, 25, 44]. Whether the Glauber dynamics are rapidly mixing for the entire regime $\lambda < \lambda_c(\Delta)$ remained a

long-standing open problem, until recently the picture was completed [3]. By relating spectral expansion properties of certain random walks on simplicial complexes to the Glauber dynamics, it was shown that the mixing time is polynomial in $|V|$ below the tree threshold. In a series of subsequent papers [11, 8, 2], the mixing time was further improved for a broader class of spin systems by combining simplicial complexes with entropy factorization and using the modified log-Sobolev inequality.

By mapping the hard-sphere model to an instance of the hard-core model we can make use of the existing results about approximation and sampling below the tree threshold. Roughly, our discretization scheme restricts the positions of sphere centers to an integer grid, while scaling the radii of spheres and the fugacity appropriately. For a hard-sphere instance $(\mathbb{V}, \lambda)$ with $\mathbb{V} = [0, \ell)^d$ the hard-core representation for resolution $\rho \in \mathbb{R}_{\geq 1}$ is a hard-core instance (G_ρ, λ_ρ) with $G_\rho = (V_\rho, E_\rho)$. Each vertex $v \in V_\rho$ represents a grid point in the finite integer lattice of side length $\rho\ell$. Two distinct vertices in V_ρ are connected by an edge in E_ρ if the Euclidean distance of the corresponding grid points is less than $2\rho r$. Furthermore, we set $\lambda_\rho = \lambda/\rho^d$. We provide the following result on the rate of convergence of $Z(G_\rho, \lambda_\rho)$ to the hard-sphere partition function $Z(\mathbb{V}, \lambda)$ in terms of ρ.

LEMMA 1.2. *Let $(\mathbb{V}, \lambda)$ be an instance of the continuous hard-sphere model in d dimensions. For each resolution $\rho \geq 2\sqrt{d}$ it holds that*

$$1 - \rho^{-1}\mathrm{e}^{\Theta(|\mathbb{V}| \ln |\mathbb{V}|)} \leq \frac{Z(\mathbb{V}, \lambda)}{Z(G_\rho, \lambda_\rho)} \leq 1 + \rho^{-1}\mathrm{e}^{\Theta(|\mathbb{V}| \ln |\mathbb{V}|)}.$$

Although proving this rate of convergence involves some detailed geometric arguments, there is an intuitive reason why the partition functions converge eventually as $\rho \to \infty$. Increasing the resolution ρ also linearly increases the side length of the grid and the minimum distance that sphere centers can have. This is equivalent to putting a grid into $\mathbb{V}$ with increasing granularity but fixing the radii of spheres instead. However, only scaling the granularity of this grid increases the number of possible configurations by roughly ρ^d, which would cause the partition function of the hard-core model to diverge as $\rho \to \infty$. To compensate for this, we scale the weight of each vertex in the hard-core model by the inverse of this factor.

Using this discretization approach, the fugacity bound from Theorem 1.1 results from simply considering Δ_ρ, the maximum degree of G_ρ, and comparing λ_ρ with the corresponding tree threshold $\lambda_c(\Delta_\rho)$. Recall that we assume $r = (1/v_d)^{1/d}$. A simple geometric argument shows that Δ_ρ is roughly upper bounded by $2^d \rho^d$ for sufficiently large ρ. Now, observe that

$$\lambda_\rho = \frac{\lambda}{\rho^d} < \lambda_c\big(2^d \rho^d\big)$$

for $\lambda < \rho^d \lambda_c\big(2^d \rho^d\big)$. This follows from the fact that $\rho^d \lambda_c\big(2^d \rho^d\big)$ converges to $\mathrm{e}/2^d$ from above for $\rho \to \infty$. Thus, the approximation bound from Theorem 1.1 and the uniqueness bound in [39] coincide with the regime of λ, for which λ_ρ is below the tree threshold $\lambda_c(\Delta_\rho)$ in the limit $\rho \to \infty$.

The arguments above show that for a sufficiently high resolution ρ the partition function of the hard-sphere model $Z(\mathbb{V}, \lambda)$ is well approximated by the partition function of our discretization (G_ρ, λ_ρ) and that (G_ρ, λ_ρ) is below the tree threshold for $\lambda < \mathrm{e}/2^d$. However, this does not immediately imply an approximation algorithm within the desired runtime bounds. Based on Lemma 1.2, we still need to choose ρ exponentially large in the volume $|\mathbb{V}|$. Note that the number of vertices in G_ρ is roughly $|V_\rho| \in \Theta\big(\rho^d |\mathbb{V}|\big)$. Even without explicitly constructing the graph, this causes

problems, as the best bound for the mixing time of the Glauber dynamics is polynomial in $|V_\rho|$ and thus exponential in $|\mathbb{V}|$. Intuitively, the reason for this mixing time is that the Glauber dynamics only change one vertex at each step. Assuming that each vertex should be updated at least once to remove correlations with the initial state, any mixing time that is sublinear in the number of vertices is unlikely. We circumvent this problem by applying dynamics that update multiple vertices at each step but still allow each step to be computed efficiently without constructing the graph explicitly.

1.2. Block dynamics. Most of the results that we discuss from now on apply to the multivariate version of the hard-core model, that is, each vertex $v \in V$ has its own weight λ_v. For a given graph $G = (V, E)$ we denote the set of such vertex weights by $\lambda = \{\lambda_v\}_{v \in V}$ and write (G, λ) for the resulting multivariate hard-core instance. In the multivariate setting, the contribution of an independent set $I \in \mathcal{I}(G)$ to the partition function is defined as the product of its vertex weights (i.e., $\prod_{v \in I} \lambda_v$), where the contribution of the empty set is fixed to 1. Similar to the univariate hard-core model, the Gibbs distribution assigns a probability to each independent set proportionally to its contribution to the partition function. For a formal definition, see subsection 2.2.

As we discussed before, the main problem with approximating the partition function of our discretization (G_ρ, λ_ρ) is that the required graph G_ρ is exponential in the volume of the original continuous system $|\mathbb{V}|$. As the Glauber dynamics Markov chain only updates a single vertex at each step, the resulting mixing time is usually polynomial in the size of the graph, which is not feasible in our case.

We circumvent this problem by studying block dynamics, which are a natural generalization of Glauber dynamics that allows for updating arbitrary sets of vertices in each step. For a given graph $G = (V, E)$, we call a set $\Lambda = \{\Lambda_i\}_{i \in [m]} \subseteq 2^V$ a block cover of size m if and only if its union covers all vertices V. We refer to the elements of Λ as blocks. At each step, the block dynamics Markov chain $\mathcal{B}(G, \lambda, \Lambda)$ chooses a block $\Lambda_i \in \Lambda$ uniformly at random. Then, the current independent set is updated on Λ_i based on the projection of the Gibbs distribution onto Λ_i and conditioned on the current independent set outside Λ_i. For a formal definition, see subsection 2.3.

In fact, block dynamics are defined for a much more general class of spin systems than the hard-core model. However, due to the fact that each step of the Markov chain involves sampling from a conditional Gibbs distribution, block dynamics are rarely used as an algorithmic tool on its own. Instead, they are usually used to deduce rapid mixing of other dynamics.

For spin systems on lattice graphs, close connections between the mixing time of block dynamics and Glauber dynamics are known [37]. Such connections were, for example, applied to improve the mixing time of Glauber dynamics of the monomer dimer model on torus graphs [50]. Moreover, block dynamics were used to improve conditions for rapid mixing of Glauber dynamics on specific graph classes, such as proper colorings [13, 15, 16, 43] or the hard-core model [15, 43] in sparse random graphs. A very general result for the mixing time of block dynamics was achieved in [5], which proved that for all spin systems on a finite subgraph of the d-dimensional integer lattice the mixing time of block dynamics is polynomial in the number of blocks if the spin system exhibits strong spatial mixing. This result was later generalized in [6] for the Ising model on arbitrary graphs. Very recently, block dynamics based random equally sized blocks where used in [11] to prove entropy factorization and improve the mixing time of Glauber dynamics for a variety of discrete spin systems up to the tree threshold.

In this paper, we will mainly focus on the setting of block dynamics based on a so-called clique cover. For a given graph $G = (V, E)$, we call a set $\Lambda = \{\Lambda_i\}_{i \in [m]} \subseteq 2^V$ a clique cover of size m if and only if its union covers all vertices V and each $\Lambda_i \in \Lambda$ induces a clique in G. More specifically, for our application it will be sufficient to consider clique covers consisting of pairwise disjoint cliques Λ_i. In this case, we call Λ a *disjoint clique cover*.

Our main goal in this regard is to show sufficient conditions such that the mixing time of block dynamics based on a disjoint clique cover is polynomial in the size of that cover. To see why this is helpful, consider our discretization (G_ρ, λ_ρ) for a given resolution $\rho \in \mathbb{R}_{>0}$. Set $a = \frac{2\rho}{\sqrt{d}} r$ and divide the d-dimensional integer lattice of side length $\rho\ell$ into cubic regions of side length a. Every pair of integer points within such a cubic region has Euclidean distance less than $2\rho r$, meaning that the corresponding vertices in G_ρ are adjacent. Thus, each such cubic region forms a clique, resulting in a clique cover of size $(\rho\ell/a)^d \in \mathrm{O}(|\mathbb{V}|)$. This means there is always a clique cover with size linear in $|\mathbb{V}|$ and independent of the resolution ρ. By showing that, for the univariate hard-core model, the mixing time of block dynamics based on a disjoint clique cover is polynomial in the size of that cover for all $\lambda_\rho < \lambda_c(\Delta_\rho)$, we obtain a Markov chain with mixing time polynomial in $|\mathbb{V}|$ independent of the resolution ρ.

Mixing of block dynamics based on a disjoint clique cover. We analyze the mixing time of block dynamics, based on a disjoint clique cover. This is done by investigating a notion of pairwise influence between vertices that has also been used to establish rapid mixing of Glauber dynamics up to the tree threshold [3]. Let $\mathbb{P}_G[w]$ denote the probability of the event that a vertex $w \in V$ is in an independent set drawn from $\mu^{(G,\lambda)}$. Further, let $\mathbb{P}_G[\overline{w}]$ denote the probability that w is not in an independent set. We extend this abuse of notation to conditional probabilities, so $\mathbb{P}_G[\cdot \mid \overline{w}]$, for example, denotes the probability of some event conditioned on w not being in an independent set. For a pair of vertices $v, w \in V$ the influence $\Psi_G(v, w)$ of v on w is defined as

$$\Psi_G(v, w) = \begin{cases} 0 & \text{if } v = w, \\ \mathbb{P}_G[w \mid v] - \mathbb{P}_G[w \mid \overline{v}] & \text{otherwise.} \end{cases}$$

The following condition in terms of pairwise influence is central to our analysis.

Condition 1.3. Let (G, λ) be an instance of the multivariate hard-core model. There is a constant $C \in \mathbb{R}_{>0}$ and a function $q \colon V \to \mathbb{R}_{>0}$ such that for all $S \subseteq V$ and $r \in S$ it holds that

$$\sum_{v \in S} |\Psi_G(r, v)| q(v) \leq Cq(r).$$

Note that this condition appeared before in [10], where it was used for bounding the mixing time of Glauber dynamics for antiferromagnetic spin systems. Given Condition 1.3, we obtain the following result for the mixing time of block dynamics based on a disjoint clique cover.

THEOREM 1.4. *Let (G, λ) be an instance of the multivariate hard-core model that satisfies Condition 1.3. Let Λ be a disjoint clique cover for G of size m, and let $Z_{\max} = \max_{i \in [m]} \{Z(G[\Lambda_i], \lambda[\Lambda_i])\}$. The mixing time of the block dynamics $\mathcal{B}(G, \lambda, \Lambda)$, starting from $\emptyset \in \mathcal{I}(G)$, is bounded by*

$$\tau_{\mathcal{B}}^{(\emptyset)}(\varepsilon) \leq m^{\mathrm{O}((2+C)C)} Z_{\max}^{\mathrm{O}((2+C)C)} \ln\left(\frac{1}{\varepsilon}\right).$$

Using a bound for the sum of absolute pairwise influences that was recently established in [10], it follows that the univariate hard-core model satisfies Condition 1.3 up to the tree threshold. As a result, we know that the mixing time of block dynamics is polynomial in m and $Z_{\max}$ for any clique cover of size m.

A side journey: Comparison to multivariate conditions. In fact, Theorem 1.4 is sufficient for our application to the hard-sphere model. However, to set Condition 1.3 into context with other conditions for rapid mixing of similar dynamics, we compare it to a strict version of the clique dynamics condition, originally introduced in [22]. The clique dynamics condition is a sufficient condition for rapid mixing of clique dynamics, which is a Markov chain that was initially introduced in the setting of abstract polymer model and which is closely related to block dynamics based on a clique cover. It turns out that this strict version of the clique dynamics condition directly implies Condition 1.3. This is especially interesting, as the clique dynamics condition was initially introduced as a local condition (only considering the neighborhood of each vertex) and is based on a coupling argument. However, we show that it can as well be understood as a sufficient condition for the global decay of pairwise influence with increasing distance between vertices.

Formally, we say that the strict clique dynamics condition is satisfied for an instance of the multivariate hard-core model (G, λ) if there is a function $f \colon V \to \mathbb{R}_{>0}$ and a constant $\alpha \in (0, 1)$ such that, for all $v \in V$, it holds that

$$\sum_{w \in N(v)} \frac{\lambda_w}{1 + \lambda_w} f(w) \le (1 - \alpha) f(v),$$

where $N(v)$ is the neighborhood of v in G. This is a strict version of the clique dynamics condition in that the original clique dynamics condition would correspond to the case $\alpha = 0$ (i.e., the strict clique dynamics condition requires some strictly positive slack α).

The result of our comparison is summarized in the following statement.

LEMMA 1.5. *Let (G, λ) be an instance of the multivariate hard-core model. If (G, λ) satisfies the strict clique dynamics condition for a function f and a constant α, then it also satisfies Condition 1.3 for $q = f$ and $C = \frac{1}{\alpha}$.*

Lemma 1.5 is proven by translating the calculation of pairwise influences to the self-avoiding walk tree of the graph, based on a result in [10], and applying a recursive argument on this tree. The technical details are given in section 5.

Despite being an interesting relationship between local coupling arguments and global pairwise influence, Lemma 1.5 also implies that, from an algorithmic perspective, Theorem 1.4 can be used to produce similar results as those obtained in [22] for abstract polymer models. Further, note that for the univariate model, using pairwise influence yields strictly better results than any coupling approach in the literature. This raises the question of whether a refined argument based on pairwise influences can be used in the multivariate setting to improve on the clique dynamics condition, leading to better approximation results on abstract polymer models.

1.3. Analyzing spectral expansion. As a core technique for obtaining Theorem 1.4, we adapt an approach for bounding the mixing time that was recently used to prove rapid mixing of Glauber dynamics for the entire regime below the tree threshold for several applications, such as the hard-core model [3], general two-state spin systems [10], and proper colorings [9, 17]. The idea is to map the desired distribution to a weighted simplicial complex.

A simplicial complex X over a groundset U is a set family $X \subseteq 2^U$ such that for each $\tau \in X$ every subset of τ is also in X. We call the elements $\tau \in X$ the faces of X and refer to its cardinality $|\tau|$ as dimensionality.

For a hard-core instance (G, λ), the authors of [3] construct a simplicial complex over a ground set U that contains two elements $x_v, x_{\overline{v}} \in U$ for each vertex $v \in V$. For every independent set $I \in \mathcal{I}(G)$, a face $\tau_I \in X$ is introduced such that $x_v \in \tau_I$ if $v \in I$ and $x_{\overline{v}} \in \tau_I$ otherwise. The simplicial complex is completed by taking the downward closure of these faces. Note that by construction all maximum faces of the resulting complex are $|V|$-dimensional and there is a one-to-one correspondence between the maximum faces and the independent sets in $\mathcal{I}(G)$. By assigning each maximum face $\tau_I \in X$ an appropriate weight, the Glauber dynamics can be represented as a random walk on those maximum faces, which is sometimes referred to as the two-step walk or down-up walk. Using a local-to-global theorem [1], the mixing time of this two-step walk can then be bounded based on certain local expansion properties of the simplicial complex X (see section 2 for the technical details). It is then proved that such local expansion properties are well captured by the largest eigenvalue of the pairwise influence matrix Ψ_G, which is a $|V| \times |V|$ matrix that contains the pairwise influence $\Psi_G(v, w)$ for all $v, w \in V$. Finally, by bounding those influences a bound on this largest eigenvalue of Ψ_G is obtained. This analysis was later refined and generalized in [10] to general two-state spin systems.

This method was independently extended in [9] and [17] to the non-Boolean domain by applying it to the Glauber dynamics for proper colorings. The main differences to the Boolean domain are that elements of the simplicial complex now represent combinations of a vertex and a color. Furthermore, the bound on the local spectral expansion was obtained by using a different influence matrix, which captures the effect of selecting a certain color for one vertex on the distribution of colors for another vertex.

Although we are dealing with the hard-core model, which is Boolean in nature, the way that we model block dynamics is mainly inspired by the existing work on proper colorings [9]. Assume we have an instance of the multivariate hard-core model (G, λ) and let Λ be a clique cover for G of size m such that every pair of distinct cliques is vertex-disjoint (i.e., Λ is a partition of G into cliques). We construct a simplicial complex X based on a ground set U that contains one element $x_v \in U$ for each vertex $v \in V$ and one additional element $\emptyset_i$ for each clique $\Lambda_i \in \Lambda$. We introduce a face $\tau_I \in X$ for each independent set $I \in \mathcal{I}(G)$ such that for every $\Lambda_i \in \Lambda$ we have $\emptyset_i \in \tau_I$ if $\Lambda_i \cap I = \emptyset$ and $x_v \in \tau_I$ if $\Lambda_i \cap I = \{v\}$ for some $v \in \Lambda_i$. The simplicial complex is completed by taking the downward closure of these faces. As we discuss in subsection 3.1, all maximum faces of the resulting complex are m-dimensional and there is a bijection between the maximum faces and the independent sets of G. Furthermore, there is a natural partitioning $\{U_i\}_{i \in [m]}$ of the ground set U, each partition U_i corresponding to a clique Λ_i, such that every maximum face in X contains exactly one element from each partition U_i.

By weighting each maximum face of X by the contribution of the corresponding independent set to the partition function, the block dynamics based on Λ are equivalent to the two-step walk on X. Thus, in order to bound the mixing time of the block dynamics, it is sufficient to study the local expansion properties of X. To this end, we adapt the influence matrix used for proper colorings in [9]. For $x \in U$, let $\mathbb{P}_G[x]$ denote the probability that $x \in \tau_I$ for an independent set $I \in \mathcal{I}(G)$ drawn from $\mu^{(G, \lambda)}$ and corresponding maximum face $\tau_I \in X$. Similarly as for defining pairwise influences, we extend this notation to conditional probabilities. The clique influence

matrix $\Phi_{G,\Lambda}$ contains an entry $\Phi_{G,\Lambda}(x,y)$ for each $x, y \in U$ with

$$\Phi_{G,\Lambda}(x,y) = \begin{cases} 0 & \text{if } x, y \in U_i \text{ for some } i \in [m], \\ \mathbb{P}_G[y \mid x] - \mathbb{P}_G[y] & \text{otherwise.} \end{cases}$$

By using similar linear-algebraic arguments as in [9] we prove that the maximum eigenvalue of $\Phi_{G,\Lambda}$ can be used to upper bound the local spectral expansion of X. To obtain Theorem 1.4 it is then sufficient to relate Condition 1.3 to the maximum eigenvalue of $\Phi_{G,\Lambda}$. The following lemma establishes this connection.

LEMMA 1.6. *Let (G, λ) be an instance of the multivariate hard-core model that satisfies Condition* 1.3 *for a function q and a constant C. For every $S \subseteq V$ and every disjoint clique cover Λ of $G[S]$ it holds that the largest eigenvalue of $\Phi_{G[S],\Lambda}$ is at most $(2 + C)C$.*

Note that our simplicial-complex representation is only given under the assumption that the cliques in the clique cover Λ are pairwise disjoint. Indeed, this is a necessary requirement to map the block dynamics to the two-step walk such that the local-global theorem from [1] can be applied. Since our discretization of the hard-sphere model will yield a graph with a small disjoint clique cover, this is sufficient for our application. However, in fact the restriction to disjoint clique covers can be overcome without significantly increasing the mixing time by using a Markov chain comparison, which can be seen in the conference version of this paper [21].

We note that in subsequent papers, the method of using simplicial complexes to bound the mixing time of Glauber dynamics was further refined to obtain improved bounds on the mixing time [11, 8, 2]. We leave it as an open question for future work if those refinements could also be applied in our setting to improve the running time of our approximation algorithm for the hard-sphere partition function.

1.4. Subsequent work and open questions. Subsequent to the conference version of this paper [21], various new results appeared that partially improve the results in this work.

One of the most notable recent improvements is that Michelen and Perkins [40] achieved a better bound for the regime of uniqueness of the Gibbs measure for Gibbs point processes with repulsive pair potentials. Since the hard-sphere model is a special case of such point processes, this leads to a broader uniqueness regime for this model. Specifically, the improved bound in d-dimensional Euclidean space is $\lambda < e/((1 - 1/8^{d+1})2^d)$. Based on this result, it was then shown in [41] that an efficient approximate sampler for Gibbs point processes with repulsive finite-range pair potentials is obtained by using a connection between rapid mixing of a local Markov chain and strong spatial mixing of the point process. This leads to an efficient approximate sampler for the hard-sphere model and a randomized approximation of the hard-sphere partition function in the same fugacity regime. In contrast to our result, their algorithms require us to assume a continuous model of computation. Further, note that the gap between our fugacity regime and the results in [41] is largest if the number of dimensions d is small, and both bounds converge as d increases. Both the improved bound for uniqueness of the Gibbs measure and the algorithmic results are based on the notion of the potential-weighted connective constant. It is inspired by the connective constant as studied in graph theory, which was previously used to obtain better algorithmic bounds for the discrete hard-core model [48]. This raises the question of whether studying the connective constant of our discretization G_ρ could yield a similar improvement of the fugacity regime for the approach presented in this

paper. For this, we would require the connective constant of G_ρ to be at least by a constant factor small than its maximum degree Δ_ρ. Unfortunately, due to a result in [46], this is not the case. This means that a similar improvement would require studying a different structural property of G_ρ than the connective constant.

Moreover, there are several new results when it comes to discretization-based approaches of approximation. In [20], we improve on Lemma 1.2 by providing stronger bounds on the rate of convergence of the hard-core partition function of our discretization. Specifically, we show that $\Theta(|\mathbb{V}|^2\varepsilon^{-1})$ points in the grid are in fact sufficient to obtain a relative ε-approximation of the hard-sphere partition function, which is a tight bound. This improved rate of convergence not only makes Glauber dynamics applicable to the discretization, which simplifies the randomized algorithm presented in this paper and improves its running time, but it also allows for the first deterministc approximation algorithm for the hard-sphere partition function with running time quasi-polynomial in $|\mathbb{V}|$. The existence of a fully polynomial deterministic approximation algorithm remains an open question (see [20] for a more detailed discussion). Moreover, we generalize the discretization scheme presented in this paper to a broader class of point processes that are characterized by hard-constraint interactions, such as the continuous Widom–Rowlinson model [53], and we introduce an approach for approximate sampling that is based on first drawing an independent set from the hard-core distribution of the discretization G_ρ and then adding small random perturbations to the corresponding grid points.

In another follow-up paper [19], we study the possibility of using hard-core models on geometric random graphs based on random point sets for discretization. This has several advantages compared to the fixed grid-like structure presented in this paper. First, it allows for more general regions $\mathbb{V}$ in arbitrary Polish spaces whereas the results in this paper are restricted to cubic regions in Euclidean space. Moreover, it allows us to cover Gibbs point processes with arbitrary repulsive pair potentials. However, in general, this approach requires a continuous model of computation and, due to the random nature of the resulting graphs, it only yields a randomized approximation algorithm for the partition function of such a point process. Obtaining efficient deterministic approximations for the partition functions of repulsive Gibbs point processes remains an open problem.

2. Preliminaries. We denote the set of all natural numbers (including 0) by $\mathbb{N}$ and the set of all real numbers by $\mathbb{R}$. For each $n \in \mathbb{N}$, let $[n]$ denote the interval $[1, n] \cap \mathbb{N}$. Further, for a graph $G = (V, E)$, we write $N_G(v)$ for the open neighborhood of a vertex $v \in V$ (i.e., all $w \in V$ with $(v, w) \in E$) and $N[v] = N_G(v) \cup \{v\}$ for the closed neighborhood. We might omit the graph if it is clear from the context.

2.1. Markov chains and spectral properties. For any (time-homogeneous) Markov chain $\mathcal{M}$, we denote its state space by $\Omega_\mathcal{M}$ and its transition probabilities by $P_\mathcal{M}$. If $\mathcal{M}$ has a unique stationary distribution, we denote it by $\pi_\mathcal{M}$. Assume $|\Omega_\mathcal{M}| = N \in \mathbb{N}_{>0}$. It is well known that if $\mathcal{M}$ is reversible with respect to $\pi_\mathcal{M}$, this implies that $P_\mathcal{M}$ has N real eigenvalues

$$1 = \beta_1(P_\mathcal{M}) \geq \beta_2(P_\mathcal{M}) \geq \cdots \geq \beta_N(P_\mathcal{M}) \geq -1.$$

We write $\beta^*(P_\mathcal{M}) = \max\{\beta_1(P), |\beta_N(P)|\}$ for the largest absolute eigenvalue and call $1 - \beta^*(P_\mathcal{M})$ the *spectral gap* of $P_\mathcal{M}$. We extend these notations to general matrices A with real eigenvalues, e.g., we denote the largest eigenvalue by $\beta_1(A)$.

If $\mathcal{M}$ is ergodic, we define its mixing time starting from some state $x \in \Omega_{\mathcal{M}}$ as

$$\tau_{\mathcal{M}}^{(x)}(\varepsilon) = \inf\{t \in \mathbb{N} \mid d_{\mathrm{TV}}\big(P_{\mathcal{M}}^t(x, \cdot), \pi_{\mathcal{M}}\big) \leq \varepsilon\},$$

where $P_{\mathcal{M}}^t(x, \cdot)$ is the distribution of $\mathcal{M}$ on $\Omega_{\mathcal{M}}$ after t steps, starting from x, and where $d_{\mathrm{TV}}(\cdot, \cdot)$ denotes the total variation distance. Recall that for any ergodic, reversible Markov chain $\mathcal{M}$ and every state $x \in \Omega_{\mathcal{M}}$, it holds that

$$(2.1) \qquad \tau_{\mathcal{M}}^{(x)}(\varepsilon) \leq \frac{1}{1 - \beta^*(P_{\mathcal{M}})} \ln\left(\frac{1}{\pi_{\mathcal{M}}(x) \cdot \varepsilon}\right).$$

For further details on Markov chains please refer to [35].

2.2. The multivariate hard-core model. Let $G = (V, E)$ be an undirected graph, and let $\mathcal{I}(G)$ denote the set of independent sets in G; if the graph is clear from the context, we only write $\mathcal{I}$. The *multivariate hard-core model* is a tuple (G, λ), where $\lambda = \{\lambda_v\}_{v \in V}$ is a set of weights, containing one weight $\lambda_v \in \mathbb{R}_{>0}$ for each vertex $v \in V$. The *partition function* of (G, λ) is defined as

$$Z(G, \lambda) \coloneqq \sum_{I \in \mathcal{I}} \prod_{v \in I} \lambda_v.$$

The *Gibbs distribution* $\mu^{(G, \lambda)}$ is a probability distribution on $\mathcal{I}$, assigning each independent set $I \in \mathcal{I}$ the probability

$$\mu^{(G, \lambda)}(I) = \frac{\prod_{v \in I} \lambda_v}{Z(G, \lambda)}.$$

If the model (G, λ) is clear, we only write Z and μ.

Large parts of our analysis consider the Gibbs distributions and the partition functions of induced subgraphs $G[S]$ for $S \subseteq V$ while keeping the weights of the respective vertices in S. In this case, we might omit the set of weights and write $Z(G[S])$ for $Z(G[S], \lambda[S])$ or $\mu^{(G[S])}$ for $\mu^{(G[S], \lambda[S])}$. Further, for any nonempty set of vertices $S \subseteq V$, we define $\mu_{|S}^{(G)}$ to be the distribution of the independent sets in $\mathcal{I}(G[S])$ induced by $\mu^{(G)}$. Formally, this means $\mu_{|S}^{(G)}$ assigns every independent set $I \in \mathcal{I}(G[S])$ the probability

$$\mu_{|S}^{(G)}(I) = \sum_{I' \in \mathcal{I}(G)} \mathbb{1}_{I \subseteq I'} \mu^{(G)}(I').$$

We associate every independent set $I \in \mathcal{I}$ with a spin assignment $\sigma^{(I)} \colon V \to \{0, 1\}$ such that $\big(\sigma^{(I)}\big)^{-1}(1) = I$. We extend this notation to restrictions on subsets $S \subseteq V$. For any independent set $I \in \mathcal{I}$, the *partial configuration* on S corresponding to I is a spin assignment $\sigma_{|S}^{(I)} \colon S \to \{0, 1\}$ such that $\big(\sigma_{|S}^{(I)}\big)^{-1}(1) = I \cap S$. By abuse of notation, we use these spin assignments as events (e.g., for conditioning on partial configurations). Further, for all $S \subseteq V$ let $\mathbf{0}_{|S} \colon S \to \{0\}$ be the partial configuration that fixes all vertices in S not to be in the independent set (i.e., $\mathbf{0}_{|S} = \sigma_{|S}^{(\emptyset)}$).

Finally, for each for $v \in V$, we write $\mathbb{P}_G[v]$ to denote the probability of the event that $v \in I$ for $I \sim \mu$, and $\mathbb{P}_G[\overline{v}]$ to denote the probability of the event $v \notin I$ for $I \sim \mu$. Formally,

$$\mathbb{P}_G[v] = \mu_{|\{v\}}^{(G)}(\{v\}) \text{ and } \mathbb{P}_G[\overline{v}] = \mu_{|\{v\}}^{(G)}(\emptyset).$$

2.3. Block dynamics and block covers. For any graph $G = (V, E)$, we call a set $\Lambda = \{\Lambda_i\}_{i \in [m]} \subseteq 2^V$ a *block cover* of G with size $m \in \mathbb{N}_{>0}$ if $\bigcup_{i \in [m]} \Lambda_i = V$. We refer to the elements of Λ as blocks.

DEFINITION 2.1 (block dynamics). *Let (G, λ) be a multivariate hard-core model, and let Λ be a block cover of G with size m. We define the* block dynamics $\mathcal{B}(G, \lambda, \Lambda)$ *to be the following Markov chain with state space $\mathcal{I}(G)$. Let $(X_t)_{t \in \mathbb{N}}$ denote a (random) sequence of states of $\mathcal{B}(G, \lambda, \Lambda)$, where X_0 is arbitrary. Then, for all $t \in \mathbb{N}$ and any $X_t = I$ with $I \in \mathcal{I}(G)$, the transitions of $\mathcal{B}(G, \lambda, \Lambda)$ are as follows:*

1. *choose $i \in [m]$ uniformly at random;*
2. *choose $I_+ \in \mathcal{I}(G[\Lambda_i])$ according to $\mu_{|\Lambda_i}(\ \cdot \ |\sigma^{(I)}_{|V \setminus \Lambda_i})$;*
3. *$X_{t+1} = (I \setminus \Lambda_i) \cup I_+$;*

The block dynamics Markov chain $\mathcal{B}(G, \lambda, \Lambda)$ is ergodic with stationary distribution μ, independent of the chosen block cover Λ. If $\Lambda = \{\{v\} \mid v \in V\}$, then the block dynamics correspond to the Glauber dynamics. In this paper, we are mostly interested in block dynamics based on a clique cover. Recall that, for a graph $G = (V, E)$, a set $\Lambda = \{\Lambda_i\}_{i \in [m]} \subseteq 2^V$ is called a clique cover of G with size $m \in \mathbb{N}_{>0}$ if $\bigcup_{i \in [m]} \Lambda_i = V$ and each $\Lambda_i \in \Lambda$ induces a clique in G. Further, we call Λ a disjoint clique cover if every pair of distinct cliques in Λ is vertex-disjoint. Obviously, any clique cover is also a block cover by definition.

2.4. Pairwise influence. Let $v, w \in V$ and let $S \subset V$ such that $v, w \notin S$. Further, let $\sigma_{|S} = \sigma^{(I)}_{|S}$ be a partial configuration on S corresponding to any independent set $I \in \mathcal{I}$. The *pairwise influence* of v on w in G under condition $\sigma_{|S}$ is defined as

$$\Psi_G^{\sigma_{|S}}(v, w) = \begin{cases} 0 & \text{if } v = w, \\ \mathbb{P}_G[w \mid v, \sigma_{|S}] - \mathbb{P}_G[w \mid \overline{v}, \sigma_{|S}] & \text{otherwise.} \end{cases}$$

For the case $S = \emptyset$, we also write $\Psi_G(v, w)$. Furthermore, we denote by $\Psi_G^{\sigma_{|S}}$ and Ψ_G the corresponding $(n - |S|) \times (n - |S|)$ matrices.

2.5. Simplicial complexes and local spectral expansion. Let U denote a set. A *simplicial complex* (over U) is a family of subsets $X \subseteq 2^U$ such that, for all $\tau \in X$ and all $\tau' \subseteq \tau$, it holds that $\tau' \in X$. We call the elements $\tau \in X$ *faces,* and we call $|\tau|$ the *dimension* of a face τ. We denote the set of all k-dimensional faces in X by $X(k)$. A simplicial complex is *pure d-dimensional* if and only if the set of all inclusion-maximal faces is exactly $X(d)$. Last, we say that a pure d-dimensional simplicial complex is *d-partite* if and only if there is a partition $\{U_i\}_{i \in [d]}$ such that, for all $i \in [d]$ and all $\tau \in X(d)$, it holds that $|U_i \cap \tau| = 1$.

We extend the definition of a pure d-dimensional simplicial complex X to a *weighted* simplicial complex (X, w) with a weight function $w \colon X \to \mathbb{R}_{>0}$ in the following inductive manner. Each face $\tau \in X(d)$ is assigned a weight $w(\tau) \in \mathbb{R}_{>0}$. Each nonmaximal face $\tau' \in X$ has the weight

$$w(\tau') = \sum\nolimits_{\tau \in X(d) : \tau' \subset \tau} w(\tau).$$

We are interested in two types of Markov chains on a weighted pure d-dimensional simplicial complex (X, w).

(1) The *two-step random walk* $\mathcal{V}(X, w)$, which is a Markov chain on the state space $X(d)$. Let $\tau_t \in X(d)$ be the state of $\mathcal{V}(X, w)$ at time $t \in \mathbb{N}$; then τ_{t+1} is chosen according the following transition rule:

1. choose $x \in \tau_t$ uniformly at random, let $\tau' = \tau_t \setminus \{x\}$, and
2. choose $\tau_{t+1} \in \{\tau \in X(d) \mid \tau' \subset \tau\}$ proportionally to their weights $w(\tau)$.

(2) The *1-skeleton* of (X, w) is an edge-weighted graph with vertices $V_X = \{x \in U \mid \{x\} \in X\}$, edges $E_X = \{(x, y) \in V_X^2 \mid \{x, y\} \in X\}$, and weights $w(\{x, y\})$. The *skeleton walk* on (X, w), denoted by $\mathcal{S}(X, w)$, is the nonlazy random walk on its 1-skeleton.

For a face $\tau \in X$, the *link* of τ is a weighted pure $(d - |\tau|)$-dimensional simplicial complex (X_τ, w_τ), where $X_\tau = \{\tau' \setminus \tau \mid \tau' \in X, \tau \subseteq \tau'\}$ and, for all $\tau' \in X$, we have $w_\tau(\tau') = w(\tau' \cup \tau)$.

DEFINITION 2.2 (local expander). *Let (X, w) be a weighted pure d-dimensional simplicial complex, and let $\alpha \in \mathbb{R}_{>0}$. We say that a face $\tau \in X$ is a* local α-expander *if and only if the second largest eigenvalue of its skeleton walk $\mathcal{S}_\tau = \mathcal{S}(X_\tau, w_\tau)$ is at most α (i.e., $\beta_2(P_{\mathcal{S}_\tau}) \leq \alpha$). Further, we say (X, w) is a* local $(\alpha_0, \ldots, \alpha_{d-2})$-expander *if and only if, for all $k \in \{0\} \cup [d - 2]$, each face $\tau \in X(k)$ is a local α_k-expander.*

In [1] the authors relate local expansion and two-step walks. We use the following formulation of their result.

THEOREM 2.3 ([3, Theorem 1.3]). *Let (X, w) be a weighted pure d-dimensional simplicial complex. If (X, w) is a local $(\alpha_0, \ldots, \alpha_{d-2})$-expander, then for the second-largest eigenvalue of the two-step walk $\mathcal{V} = \mathcal{V}(X, w)$, it holds that*

$$\beta_2(P_{\mathcal{V}}) \leq 1 - \frac{1}{d} \prod_{k \in \{0\} \cup [d-2]} (1 - \alpha_k).$$

3. Mixing time of block dynamics for clique covers.

3.1. Simplicial-complex representation. Let (G, λ) be an instance of the multivariate hard-core model and let Λ be a disjoint clique cover of size m. We construct the *simplicial-complex representation* as follows. For each clique $\Lambda_i \in \Lambda$, we have a set U_i that consists of an element $\emptyset_i \in U_i$ and one element $x_v \in U_i$ for each vertex $v \in \Lambda_i$. The ground set of the simplicial complex is $U = \bigcup_{i \in [m]} U_i$. Further, the complex X contains a face $\tau_I \in X$ for each independent set $I \in \mathcal{I}$ where

- for each $i \in [m]$ and $x_v \in U_i$, we have $x_v \in \tau_I$ if and only if $v \in I$, and
- for each $i \in [m]$, we have $\emptyset_i \in \tau_I$ if and only if $I \cap \Lambda_i = \emptyset$.

Note that each independent set contains at most one vertex $v \in \Lambda_i$ for each clique in the clique cover $\Lambda_i \in \Lambda$. As Λ is a disjoint cover, each of the faces $\tau_I \in X$ contains exactly one element from each U_i for $i \in [m]$. We complete X by taking the downward closure of these faces. We make the following observation.

Observation 3.1. The simplicial-complex representation X for an instance of the multivariate hard-core model (G, λ) with disjoint clique cover Λ of size m is pure m-dimensional and m-partite with partition $\{U_i\}_{i \in [m]}$ as constructed above. Further, there is a one-to-one correspondence between the independent sets of G and the maximum faces $X(m)$.

We continue by equipping $X(m)$ with weights, which induces weights for all other faces in X as well. For a face $\tau_I \in X(m)$, corresponding to the independent set $I \in \mathcal{I}$, we set $w(\tau_I) = \mu(I)$. We now observe the following relation to block dynamics.

Observation 3.2. Let (G, λ) be an instance of the multivariate hard-core model, and let Λ be a disjoint clique cover of G. Further, let (X, w) be the corresponding simplicial-complex representation. It holds that the two-step walk $\mathcal{V}(X, \lambda)$ is equivalent to the block dynamics $\mathcal{B}(G, \lambda, \Lambda)$ in the sense that there is a bijection between

both state spaces that preserves transition probabilities. Consequently, $\mathcal{V}(X, \lambda)$ is ergodic and reversible and has a unique stationary distribution $\pi_{\mathcal{V}}(\tau_I) = \mu(I)$ for every maximum face $\tau_I \in X(m)$, corresponding to the independent set $I \in \mathcal{I}(G)$.

Based on Observation 3.1, applying Theorem 2.3, we obtain a lower bound on the spectral gap of $\mathcal{V} = \mathcal{V}(X, w)$ in terms of local expansion. Moreover, for an independent set $I \sim \mu$, it holds that

$$\Pr[x_v \in \tau_I] = \mathbb{P}_G[v] \text{ for all } v \in V \text{ and } \Pr[\emptyset_i \in \tau_I] = \mathbb{P}_G\left[\bigcap_{v \in \Lambda_i} \overline{v}\right] \text{ for all } i \in [m].$$

For simplicity, we also write $\mathbb{P}_G[x_v]$ and $\mathbb{P}_G[\emptyset_i]$ for these probabilities.

3.2. Bounding local expansion by clique influence. Let (G, λ) be an instance of the multivariate hard-core model with disjoint clique cover Λ of size m. Further, let (X, w) be the resulting simplicial-complex representation with ground set U and partition $\{U_i\}_{i \in [m]}$. The clique influence matrix $\Phi_{G,\Lambda}$ contains an entry $\Phi_{G,\Lambda}(x, y)$ for each $x, y \in U$ with

$$(3.1) \qquad \Phi_{G,\Lambda}(x, y) = \begin{cases} 0 & \text{if } x, y \in U_i \text{ for some } i \in [m], \\ \mathbb{P}_G[y \mid x] - \mathbb{P}_G[y] & \text{otherwise.} \end{cases}$$

Note that this definition includes the cases where $x \in \{\emptyset_i \mid i \in [m]\}$ or $y \in \{\emptyset_i \mid i \in [m]\}$. The following lemma and its proof are an adapted version of [9, Theorem 8].

LEMMA 3.3. *Let (G, λ) be an instance of the multivariate hard-core model with a disjoint clique cover Λ of size m. Further, let (X, w) be the resulting simplicial-complex representation. Denote by $\mathcal{S} = \mathcal{S}(X, w)$ be the skeleton walk on (X, w), and let $\Phi_{G,\Lambda}$ be the clique influence matrix as defined in* (3.1). *Then*

$$(3.2) \qquad \beta_2(P_{\mathcal{S}}) \leq \frac{1}{m-1}\beta_1(\Phi_{G,\Lambda}).$$

Proof. We first take a detailed look at the entries of the transitions $P_{\mathcal{S}}$. Note that by definition $\Omega_{\mathcal{S}} = U$ and $P_{\mathcal{S}}(x, x) = 0$ for all $x \in U$. Further, it holds for each $x \in U$ that

$$\sum_{\substack{z \in U: \\ z \neq x}} w(\{x, z\}) = \sum_{\substack{z \in U: \\ z \neq x}} \mathbb{P}_G[x, z] = \sum_{i \in [m]} \sum_{\substack{z \in U_i: \\ z \neq x}} \mathbb{P}_G[x, z]$$

$$= \sum_{\substack{i \in [m]: \\ x \notin U_i}} \sum_{z \in U_i} \mathbb{P}_G[x, z] = (m-1)\mathbb{P}_G[x],$$

where the third equality comes from the fact that for all $i \in [m]$ and every pair $x, z \in U_i$ it holds that $\mathbb{P}_G[x, z] = 0$. Thus, we get for each $y \in U$ with $y \neq x$ the transition probability

$$P_{\mathcal{S}}(x, y) = \frac{w(\{x, y\})}{\sum_{\substack{z \in U: \\ z \neq x}} w(\{x, z\})} = \frac{\mathbb{P}_G[x, y]}{(m-1)\mathbb{P}_G[x]} = \frac{1}{m-1}\mathbb{P}_G[y \mid x].$$

Note that this especially implies $P_{\mathcal{S}}(x, y) = 0$ if $x, y \in U_i$ for some $i \in [m]$.

Let D be the matrix with $D(x,x) = \frac{1}{m}\mathbb{P}_G[x]$ for each $x \in U$ and 0 everywhere else, and let d be the vector of its diagonal, that is, for all $x \in U$, we have $d(x) = D(x,x)$. Note that $\mathcal{S}$ satisfies the detailed-balance equations with d, that is, it is reversible with respect to d. Thus, $A = D^{1/2}P_{\mathcal{S}}D^{-1/2}$ is symmetric. This implies that for each eigenvector z of A with eigenvalue β there is a left eigenvector $z^{\mathrm{T}}D^{1/2} = (D^{1/2}z)^{\mathrm{T}}$ and a right eigenvector $D^{-1/2}z$ of $P_{\mathcal{S}}$ for the same eigenvalue β. Thus, if $z' = D^{-1/2}z$ is such a right eigenvector of $P_{\mathcal{S}}$, then $Dz' = DD^{-1/2}z = D^{1/2}z$ yields a (transposed) left eigenvector for the same eigenvalue.

We investigate the eigenvalues of $P_{\mathcal{S}}$ more carefully. Consider the column vector $\mathbf{1}$ with $\mathbf{1}(x) = 1$ for all $x \in U$. Note that $\mathbf{1}$ is a right eigenvector of $P_{\mathcal{S}}$ for eigenvalue 1, which is the maximum eigenvalue, since $P_{\mathcal{S}}$ is a transition matrix. We denote by $\pi = (D\mathbf{1})^{\mathrm{T}}$ the corresponding left eigenvector with $\pi(x) = \frac{1}{m}\mathbb{P}_G[x]$ for all $x \in U$. Further, we define a set of column vectors $\{\mathbf{1}_i\}_{i \in [m]}$ and a set of row vectors $\{\pi_i\}_{i \in [m]}$ such that for each $i \in [m]$ and each $x \in U$ it holds that

$$\mathbf{1}_i(x) = \begin{cases} 1 & \text{if } x \in U_i, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad \pi_i(x) = \begin{cases} \frac{1}{m}\mathbb{P}_G[x] & \text{if } x \in U_i, \\ 0 & \text{otherwise.} \end{cases}$$

Note that for each $i,j \in [m]$, $i \neq j$, and each $x \in U_i$ it holds that $\sum_{y \in U_j} P_{\mathcal{S}}(x,y) = \frac{1}{m-1}$ and $\sum_{y \in U_i} P_{\mathcal{S}}(x,y) = 0$. Thus, for all $i \in [m]$, we have $P_{\mathcal{S}}\mathbf{1}_i = \frac{1}{m-1}(\mathbf{1} - \mathbf{1}_i)$. It follows that

$$P_{\mathcal{S}}\left(\frac{1}{m}\mathbf{1} - \mathbf{1}_i\right) = \frac{1}{m}\mathbf{1} - \frac{1}{m-1}(\mathbf{1} - \mathbf{1}_i) = -\frac{1}{m(m-1)}\mathbf{1} + \frac{1}{m-1}\mathbf{1}_i$$
$$= -\frac{1}{m-1}\left(\frac{1}{m}\mathbf{1} - \mathbf{1}_i\right),$$

which shows that for each $i \in [m]$ the vector $\frac{1}{m}\mathbf{1} - \mathbf{1}_i$ is a right eigenvector with eigenvalue $-\frac{1}{m-1}$. Similarly, the vector $\left(D\left(\frac{1}{m}\mathbf{1} - \mathbf{1}_i\right)\right)^{\mathrm{T}} = \frac{1}{m}\pi - \pi_i$ is a left eigenvector for this eigenvalue.

We use these vectors to construct an eigenbasis of $P_{\mathcal{S}}$. Let $i \in [m]$, and consider the set

$$S = \{\mathbf{1}\} \cup \bigcup_{\substack{j \in [m]: \\ j \neq i}} \left\{\frac{1}{m}\mathbf{1} - 1_j\right\}.$$

Note that S is a set of m linearly independent right eigenvectors of $P_{\mathcal{S}}$. By the relation between eigenvectors of $P_{\mathcal{S}}$ and A, we construct a set $S_A = \{D^{1/2}|z z \in S\}$ of independent eigenvectors of A. As A is symmetric, such a set can always be extended to an eigenbasis $\overline{S_A}$ of A, such that the vectors in $\overline{S_A}\backslash S_A$ are orthogonal to the vectors in S_A. This gives us an eigenbasis $\overline{S} = \{D^{-1/2}z|z \in \overline{S_A}\}$ of right eigenvectors of $P_{\mathcal{S}}$.

We proceed by relating the eigenvalues of $P_{\mathcal{S}}$ to the eigenvalues of $\Phi_{G,\Lambda}$, using $\overline{S}$. Note that both $P_{\mathcal{S}}$ and $\Phi_{G,\Lambda}$ are $(n+m) \times (n+m)$ matrices. We first show (Claim 1) that all vectors of S are in the kernel of $\Phi_{G,\Lambda}$. Since $|S| = m+1$ and since the vectors of S are linearly independent, the kernel of $\Phi_{G,\Lambda}$ has a dimension of at least $m+1$. Thus, $\Phi_{G,\Lambda}$ has at least $m+1$ eigenvectors associated with the eigenvalue 0. Then (Claim 2) we show that all vectors of $\overline{S} \setminus S$, which are the remaining eigenvectors of $P_{\mathcal{S}}$ in our consideration, are also right eigenvectors of $\Phi_{G,\Lambda}$ but with eigenvalues scaled by $(m-1)$. Last (Claim 3), we conclude that (3.2) holds.

Claim 1. Let $z \in S$, and let 0 denote the vector with $0(x) = 0$ for all $x \in U$. If $z = \mathbf{1}$, then $\Phi_{G,\Lambda} z = 0$ because for all $j \in [m]$ and every $x \in U_j$ it holds that

$$\Phi_{G,\Lambda} \mathbf{1}(x) = \sum_{\substack{k \in [m]:\\ k \neq j}} \sum_{y \in U_k} (\mathbb{P}_G[y \mid x] - \mathbb{P}_G[y])$$

$$= \sum_{\substack{k \in [m]:\\ k \neq j}} \sum_{y \in U_k} \mathbb{P}_G[y \mid x] - \sum_{\substack{k \in [m]:\\ k \neq j}} \sum_{y \in U_k} \mathbb{P}_G[y] = 0.$$

If $z = 1_j$ for some $j \in [m]$, again, we have $\Phi_{G,\Lambda} z = 0$ because for all $k \in [m]$ and every $x \in U_k$ it holds that

$$\Phi_{G,\Lambda} 1_j(x) = \sum_{\substack{l \in [m]:\\ l \neq k}} \sum_{y \in U_l} \Phi_{G,\Lambda}(x,y) \mathbb{1}\{k = j, l = j\} = 0.$$

Claim 2. We first show that all vectors $z \in \overline{S} \setminus S$ are orthogonal to $\frac{1}{m}\pi - \pi_j$ for all $j \in [m]$. Let $z \in \overline{S} \setminus S$. First, note that

$$\pi z = (D\mathbf{1})^{\mathrm{T}} z = \mathbf{1}^{\mathrm{T}} Dz = \left(D^{1/2}\mathbf{1}\right)^{\mathrm{T}} D^{1/2} z = 0,$$

where the last equality is due to $D^{1/2}\mathbf{1} \in S_A$ and $D^{1/2}z \in \overline{S_A} \setminus S_A$. Similarly, we obtain for each $j \in [m]$ with $j \neq i$ that

$$\left(\frac{1}{m}\pi - \pi_j\right) z = \left(D^{1/2}\left(\frac{1}{m}\mathbf{1} - 1_j\right)\right)^{\mathrm{T}} D^{1/2} z = 0.$$

Finally, note that $\frac{1}{m}\pi - \pi_i$ can be obtain as a linear combination from π and the vectors $\frac{1}{m}\pi - \pi_j$ for $j \neq i$, implying that it is orthogonal to z as well.

Assume that z has eigenvalue β. We define the matrix $\Pi = \mathbf{1}\pi$ and the matrices $\Pi_j = 1_j\pi_j$ for $j \in [m]$ and note that

$$\Phi_{G,\Lambda} = (m-1)P_{\mathcal{S}} - m\left(\Pi - \sum_{j \in [m]} \Pi_j\right).$$

Since z is orthogonal to all vectors $\frac{1}{m}\pi - \pi_j$ for $j \in [m]$, we have for every $k \in [m]$ and $x \in U_k$ that

$$\left(\left(\Pi - \sum_{j \in [m]} \Pi_j\right) z\right)(x) = \sum_{y \in U} \sum_{j \in [m]} \left(\left(\frac{1}{m}\Pi(x,y) - \Pi_j(x,y)\right) z\right)(y)$$

$$= \sum_{j \in [m]} \left(\frac{1}{m}\pi - \pi_j\right) z = 0.$$

This implies that $(\Pi - \sum_{j \in [m]} \Pi_j)z = 0$, and thus it holds that

$$\Phi_{G,\Lambda} z = (m-1)P_{\mathcal{S}} z = (m-1)\beta z.$$

Claim 3. Recall that $\beta_1(P_\mathcal{S}) = 1$ and that all eigenvectors of $P_\mathcal{S}$ from $S \setminus \{\mathbf{1}\}$ have a negative eigenvalue. We make a case distinction with respect to the sign of $\beta_2(P_\mathcal{S}) = \beta$. If $\beta > 0$, then there is an eigenvector $z \in \overline{S} \setminus S$ corresponding to β, as the eigenvalues of vectors from $S \setminus \{\mathbf{1}\}$ are negative. By Claim 2, there is an eigenvalue β' of $\Phi_{G,\Lambda}$ such that $\beta = \beta'/(m-1)$. Since $\beta' \leq \beta_1(\Phi_{G,\Lambda})$, (3.2) holds.

If $\beta \leq 0$, then (3.2) follows immediately, as the kernel of $\Phi_{G,\Lambda}$ is nontrivial and, thus, $\beta_1(\Phi_{G,\Lambda}) \geq 0$. This concludes the proof. $\square$

3.3. Bounding clique influence. We prove an upper bound for $\beta_1(\Phi_{G',\Lambda})$ for all induced subgraphs G' of G and every disjoint clique cover Λ of G', given that (G, λ) satisfy Condition 1.3.

LEMMA 3.4. *Let (G, λ) be an instance of the multivariate hard-core model that satisfies Condition 1.3 for a function q and a constant C. For every $S \subseteq V$ and every disjoint clique cover Λ of $G[S]$ it holds that the largest eigenvalue of $\Phi_{G[S],\Lambda}$ is at most $(2 + C)C$.*

In the proof of Lemma 1.6, we apply the following lemma that was already used in [10].

LEMMA 3.5. *Let $n \in \mathbb{N}$, let $A \in \mathbb{C}^{n \times n}$, and let $\rho(A)$ denote the spectral radius of A. Assume that there is a $\xi \in \mathbb{R}$ and a $p \colon [n] \to \mathbb{R}_{>0}$ such that for all $i \in [n]$ it holds that $\sum_{j \in [n]} |A(i,j)|p(j) \leq \xi p(i)$. Then $\rho(A) \leq \xi$.*

Since in [10] the above lemma is only stated and proven for a specific matrix $A \in \mathbb{C}^{n \times n}$, we restate the proof for completion.

Proof of Lemma 3.5. Let $P \in \mathbb{R}^{n \times n}$ with $P(i,i) = p(i)$ for all $i \in [n]$ and $P(i,j) = 0$ for all $i \neq j$. Observe that

$$\sum_{j \in [n]} |A(i,j)|p(j) \leq \xi p(i)$$

implies $\left\| P^{-1}AP \right\|_\infty \leq \xi$. Consequently, we have $\rho(A) = \rho\left(P^{-1}AP\right) \leq \xi$. $\square$

Note that, by Lemma 3.5, Condition 1.3 implies $\beta_1\left(\Psi_{G[S]}\right) \leq C$ for all $S \subseteq V$.

We show that Condition 1.3 implies the existence of a ξ from Lemma 3.5 such that for all induced subgraphs G' of G and every disjoint clique cover Λ of G' there is a function p that satisfies the conditions of Lemma 3.5 for $\Phi_{G',\Lambda}$. To this end, we use the following lemmas.

LEMMA 3.6. *Let (G, λ) be an instance of the multivariate hard-core model with disjoint clique cover Λ of size m. Further, let (X, w) be the corresponding simplicial-complex representation with ground set U and partition $\{U_i\}_{i \in [m]}$. For all $i, j \in [m]$ and $x \in U_i$ it holds that*

$$\Phi_{G,\Lambda}(x, \emptyset_j) = -\sum_{v \in \Lambda_j} \Phi_{G,\Lambda}(x, x_v).$$

Proof. By definition,

$$\begin{aligned}
\Phi_{G,\Lambda}(x, \emptyset_j) &= \mathbb{P}_G\left[\bigcap_{w \in \Lambda_j} \overline{w} \;\middle|\; x\right] - \mathbb{P}_G\left[\bigcap_{w \in \Lambda_j} \overline{w}\right] \\
&= -\left(\mathbb{P}_G\left[\bigcup_{w \in \Lambda_j} w \;\middle|\; x\right] - \mathbb{P}_G\left[\bigcup_{w \in \Lambda_j} w\right]\right).
\end{aligned}$$

Note that for any pair of vertices from the same clique $w_1, w_2 \in \Lambda_j$ with $w_1 \neq w_2$ the events that w_1 is in an independent set and that w_2 is in an independent set are disjoint. Thus, we obtain

$$-\left(\mathbb{P}_G\left[\bigcup_{w \in \Lambda_j} w \mid x\right] - \mathbb{P}_G\left[\bigcup_{w \in \Lambda_j} w\right]\right) = -\sum_{w \in \Lambda_j}\left(\mathbb{P}_G[w \mid x] - \mathbb{P}_G[w]\right)$$
$$= -\sum_{w \in \Lambda_j} \Phi_{G,\Lambda}(x,w). \qquad \square$$

LEMMA 3.7. *Let (G, λ) be an instance of the multivariate hard-core model with disjoint clique cover Λ of size m. Further, let (X, w) be the corresponding simplicial-complex representation with ground set U and partition $\{U_i\}_{i \in [m]}$. For all $i, j \in [m]$ with $i \neq j$ and all $v \in \Lambda_i, w \in \Lambda_j$ it holds that*

$$\Phi_{G,\Lambda}(x_v, x_w) = \mathbb{P}_G[\overline{v}]\Psi_G(v,w).$$

Proof. By the law of total probability,

$$\begin{aligned}
\Phi_{G,\Lambda}(x_v, x_w) &= \mathbb{P}_G[w \mid v] - \mathbb{P}_G[w] \\
&= \mathbb{P}_G[w \mid v] - \mathbb{P}_G[w \mid v]\mathbb{P}_G[v] - \mathbb{P}_G[w \mid \overline{v}]\mathbb{P}_G[\overline{v}] \\
&= \mathbb{P}_G[\overline{v}](\mathbb{P}_G[w \mid v] - \mathbb{P}_G[w \mid \overline{v}]) \\
&= \mathbb{P}_G[\overline{v}]\Psi_G(v,w). \qquad \square
\end{aligned}$$

LEMMA 3.8. *Let (G, λ) be an instance of the multivariate hard-core model with disjoint clique cover Λ of size m. Further, let (X, w) be the corresponding simplicial-complex representation with ground set U and partition $\{U_i\}_{i \in [m]}$. For all $i, j \in [m]$ with $i \neq j$ and every $w \in \Lambda_j$ it holds that*

$$\Phi_{G,\Lambda}(\emptyset_i, x_w) = \sum_{v \in \Lambda_i} \mathbb{P}_G[v]\Psi_G^{\mathbf{0}_{|\Lambda_i \setminus \{v\}}}(v,w) = \sum_{v \in \Lambda_i} \mathbb{P}_G[v]\Psi_{G_v}(v,w),$$

where $G_v = G[V \setminus (\Lambda_i \setminus \{v\})]$.

Proof. Let $\overline{\emptyset_i}$ denote the complementary event to $\emptyset_i$, meaning that some vertex $u \in \Lambda_i$ is an independent set drawn from the Gibbs distribution. By the law of total probability,

$$\begin{aligned}
\Phi_{G,\Lambda}(\emptyset_i, x_w) &= \mathbb{P}_G[w \mid \emptyset_i] - \mathbb{P}_G[w] \\
&= \mathbb{P}_G[w \mid \emptyset_i] - \mathbb{P}_G[w \mid \emptyset_i]\mathbb{P}_G[\emptyset_i] - \mathbb{P}_G\left[w \mid \overline{\emptyset_i}\right]\mathbb{P}_G\left[\overline{\emptyset_i}\right] \\
&= \mathbb{P}_G\left[\overline{\emptyset_i}\right]\left(\mathbb{P}_G[w \mid \emptyset_i] - \mathbb{P}_G\left[w \mid \overline{\emptyset_i}\right]\right) \\
&= \mathbb{P}_G\left[\bigcup_{u \in \Lambda_i} u\right]\left(\mathbb{P}_G\left[w \mid \bigcap_{u \in \Lambda_i} \overline{u}\right] - \mathbb{P}_G\left[w \mid \bigcup_{u \in \Lambda_i} u\right]\right).
\end{aligned}$$

Because the events that two distinct vertices from the same clique are in an independent set are disjoint, we get

$$\mathbb{P}_G\left[w \mid \bigcup_{u \in \Lambda_i} u\right] = \sum_{v \in \Lambda_i} \mathbb{P}_G[w \mid v]\frac{\mathbb{P}_G[v]}{\mathbb{P}_G\left[\bigcup_{u \in \Lambda_i} u\right]} \text{ and}$$

$$\mathbb{P}_G\Big[w \mid \bigcap_{u \in \Lambda_i} \overline{u}\Big] = \mathbb{P}_G\Big[w \mid \bigcap_{u \in \Lambda_i} \overline{u}\Big] \frac{\mathbb{P}_G\big[\bigcup_{v \in \Lambda_i} v\big]}{\mathbb{P}_G\big[\bigcup_{u \in \Lambda_i} u\big]}$$

$$= \sum_{v \in \Lambda_i} \mathbb{P}_G\Big[w \mid \bigcap_{u \in \Lambda_i} \overline{u}\Big] \frac{\mathbb{P}_G[v]}{\mathbb{P}_G\big[\bigcup_{u \in \Lambda_i} u\big]}.$$

Thus, we obtain

$$\mathbb{P}_G\Big[\bigcup_{u \in \Lambda_i} u\Big]\Big(\mathbb{P}_G\Big[w \mid \bigcap_{u \in \Lambda_i} \overline{u}\Big] - \mathbb{P}_G\Big[w \mid \bigcup_{u \in \Lambda_i} u\Big]\Big)$$
$$= \sum_{v \in \Lambda_i} \mathbb{P}_G[v]\Big(\mathbb{P}_G\Big[w \mid \bigcap_{u \in \Lambda_i} \overline{u}\Big] - \mathbb{P}_G[w \mid v]\Big).$$

Note that for each $v \in \Lambda_i$ it holds that

$$\mathbb{P}_G\Big[w \mid \bigcap_{u \in \Lambda_i} \overline{u}\Big] = \mathbb{P}_G\Big[w \mid \overline{v}, \bigcap_{u \in \Lambda_i \setminus \{v\}} \overline{u}\Big].$$

Further, because v being in the independent set implies that no other vertex $u \in \Lambda_i$ can be in the independent set too, it also holds that

$$\mathbb{P}_G[w \mid v] = \mathbb{P}_G\Big[w \mid v, \bigcap_{u \in \Lambda_i \setminus \{v\}} \overline{u}\Big].$$

Consequently, we conclude that

$$\Phi_{G,\Lambda}(\emptyset_i, x_w) = \sum_{v \in \Lambda_i} \mathbb{P}_G[v] \Psi_G^{\mathbf{0}_{|\Lambda_i \setminus \{v\}}}(v, w) = \sum_{v \in \Lambda_i} \mathbb{P}_G[v] \Psi_{G_v}(v, w). \qquad \Box$$

LEMMA 3.9. *Let (G, λ) be an instance of the multivariate hard-core model, and let $v \in V$ and $w \in N_G(v)$. Then $\mathbb{P}_G[w] \leq -\Psi_G(v, w) = |\Psi_G(v, w)|$.*

Proof. Since $w \in N_G(v)$, it holds that $\Psi_G(v, w) = -\mathbb{P}_G[w \mid \overline{v}]$. By $Z(G) \geq Z(G[V \setminus \{v\}])$ we conclude that

$$\mathbb{P}_G[w \mid \overline{v}] = \lambda_w \frac{Z(G[V \setminus N_G[w]])}{Z(G[V \setminus \{v\}])} \geq \lambda_w \frac{Z(G[V \setminus N_G[w]])}{Z(G)} = \mathbb{P}_G[w],$$

which proves the claim. $\qquad \Box$

We now prove the main lemma of this subsection.

Proof of Lemma 1.6. To simplify notation, set $G' = G[S]$ and $m = |\Lambda|$. Let (X, w) be the simplicial-complex representation of $(G', \lambda[S])$ with clique cover Λ and let U be the corresponding ground set of (X, w) with partition $\{U_i\}_{i \in [m]}$.

As we aim to prove our claim using Lemma 3.5, we need to construct a function $p \colon U \to \mathbb{R}_{>0}$ such that for all $x \in U$ it holds that

$$\sum_{y \in U} |\Phi_{G',\Lambda}(x, y)| p(y) \leq (2 + C)C p(x).$$

To this end, we set $p(x_v) = q(v)$ for all $v \in S$ and $p(\emptyset_i) = \sum_{v \in \Lambda_i} \mathbb{P}_{G'}[v] q(v)$ for all $i \in [m]$. By Lemma 3.9 we have for all $i \in [m]$ and $w \in \Lambda_i$ that

$$\sum_{v \in \Lambda_i} \mathbb{P}_{G'}[v] q(v) \leq \mathbb{P}_{G'}[w] q(w) + \sum_{v \in N_{G'}(w)} |\Psi_{G'}(w, v)| q(v),$$

which, by Condition 1.3, implies

$$(3.3) \qquad\qquad p(\emptyset_i) < (1+C)q(w).$$

Without loss of generality, assume $x \in U_i$ for some $i \in [m]$. Recall that by definition $\Phi_{G',\Lambda}(x,y) = 0$ for all $y \in U_i$. By Lemma 3.6, we obtain

$$\sum_{y \in U} |\Phi_{G',\Lambda}(x,y)| p(y) = \sum_{\substack{j \in [m]: \\ j \neq i}} \left(|\Phi_{G',\Lambda}(x,\emptyset_j)| p(\emptyset_j) + \sum_{w \in \Lambda_j} |\Phi_{G',\Lambda}(x,x_w)| p(x_w) \right)$$

$$= \sum_{\substack{j \in [m]: \\ j \neq i}} \left(\left| \sum_{w \in \Lambda_j} \Phi_{G',\Lambda}(x,x_w) \right| p(\emptyset_j) + \sum_{w \in \Lambda_j} |\Phi_{G',\Lambda}(x,x_w)| p(x_w) \right)$$

$$\leq \sum_{\substack{j \in [m]: \\ j \neq i}} \sum_{w \in \Lambda_j} |\Phi_{G',\Lambda}(x,x_w)| (p(\emptyset_j) + p(x_w)),$$

where the last step follows from the triangle inequality. Further, by our choice of p and by (3.3), we obtain

$$\sum_{\substack{j \in [m]: \\ j \neq i}} \sum_{w \in \Lambda_j} |\Phi_{G',\Lambda}(x,x_w)| (p(\emptyset_j) + p(x_w)) < (2+C) \sum_{\substack{j \in [m]: \\ j \neq i}} \sum_{w \in \Lambda_j} |\Phi_{G',\Lambda}(x,x_w)| q(w).$$

We proceed with a case distinction based on x. Assume that $x = x_v$ for some $v \in \Lambda_i$. By Lemma 3.7, we have

$$(2+C) \sum_{\substack{j \in [m]: \\ j \neq i}} \sum_{w \in \Lambda_j} |\Phi_{G',\Lambda}(x_v,x_w)| q(w) = (2+C)\mathbb{P}_{G'}[\overline{v}] \sum_{\substack{j \in [m]: \\ j \neq i}} \sum_{w \in \Lambda_j} |\Psi_{G'}(v,w)| q(w)$$

$$\leq (2+C) \sum_{\substack{j \in [m]: \\ j \neq i}} \sum_{w \in \Lambda_j} |\Psi_{G'}(v,w)| q(w).$$

Using that the cliques are disjoint and applying Condition 1.3, we get

$$(2+C) \sum_{\substack{j \in [m]: \\ j \neq i}} \sum_{w \in \Lambda_j} |\Psi_{G'}(v,w)| q(w) \leq (2+C) \sum_{w \in S} |\Psi_{G'}(v,w)| q(w) \leq (2+C)Cq(v)$$

$$= (2+C)Cp(x_v).$$

Now, assume that $x = \emptyset_i$. By Lemma 3.8, we have

$$(2+C) \sum_{\substack{j \in [m]: \\ j \neq i}} \sum_{w \in \Lambda_j} |\Phi_{G',\Lambda}(\emptyset_i,x_w)| q(w) = (2+C) \sum_{\substack{j \in [m]: \\ j \neq i}} \sum_{w \in \Lambda_j} \left| \sum_{v \in \Lambda_i} \mathbb{P}_{G'}[v] \Psi_{G'_v}(v,w) \right| q(w)$$

$$\leq (2+C) \sum_{v \in \Lambda_i} \mathbb{P}_{G'}[v] \sum_{\substack{j \in [m]: \\ j \neq i}} \sum_{w \in \Lambda_j} \left| \Psi_{G'_v}(v,w) \right| q(w),$$

where the last step follows from the triangle inequality and $G'_v = G'[S \setminus (\Lambda_i \setminus \{v\})] = G[S \setminus (\Lambda_i \setminus \{v\})]$. As the cliques are disjoint and G'_v is a subgraph of G, we apply Condition 1.3 to obtain

$$
\begin{aligned}
(2+C) \sum_{v \in \Lambda_i} \mathbb{P}_{G'}[v] &\sum_{\substack{j \in [m]: \\ j \neq i}} \sum_{w \in \Lambda_j} \big|\Psi_{G'_v}(v,w)\big| q(w) \\
&= (2+C) \sum_{v \in \Lambda_i} \mathbb{P}_{G'}[v] \sum_{w \in S \setminus (\Lambda_i \setminus \{v\})} \big|\Psi_{G'_v}(v,w)\big| q(w) \\
&\leq (2+C) \sum_{v \in \Lambda_i} \mathbb{P}_{G'}[v] C q(v) \\
&= (2+C)C \sum_{v \in \Lambda_i} \mathbb{P}_{G'}[v] q(v) \\
&= (2+C)C p(\emptyset_i),
\end{aligned}
$$

which concludes the proof. □

3.4. Canonical paths in skeleton walks. The previous section shows that we can bound local expansion of the simplicial-complex representation of a disjoint clique cover based on pairwise influence between vertices. However, note that Theorem 2.3 only yields a nontrivial bound if all α_k are sufficiently small. Since we only obtain such bounds on α_k if k is small, we introduce a more crude bound on the second largest eigenvalue of the skeleton walk by applying the canonical-path method. Although the resulting bound is worse if k is small compared to m, it is guaranteed to be less than 1, which is sufficient to cover the cases where using pairwise influence fails.

We start by giving a short overview on the canonical-path method. Let $\mathcal{M}$ be a Markov chain that is reversible with respect to its stationary distribution $\pi_\mathcal{M}$. Let $E(\mathcal{M}) = \{(x,y) \in \Omega_\mathcal{M}^2 \mid x \neq y, P_\mathcal{M}(x,y) > 0\}$ denote the edges of the Markov chain excluding self-loops, and let $E^*(\mathcal{M}) = \{(x,y) \in \Omega_\mathcal{M}^2 \mid P_\mathcal{M}(x,y) > 0\}$ be the set of edges including self-loops. For each $(x,y) \in E^*(\mathcal{M})$, we set $Q_\mathcal{M}(x,y) = \pi_\mathcal{M}(x) P_\mathcal{M}(x,y)$. The idea of the canonical-path method is to construct a path $\gamma = (x_0 = x, x_1, \ldots, x_l = y)$ for every $x,y \in \Omega_\mathcal{M}$ with $x \neq y$ using the edges in $E(\mathcal{M})$ (i.e., $(x_{i-1}, x_i) \in E(\mathcal{M})$ for all $i \in [l]$). We denote by $E(\gamma_{xy})$ the set of edges that are used by the path γ_{xy} and by $|\gamma_{xy}|$ the length of a path. Further, we call a set of paths $\Gamma = \{\gamma_{xy} \mid x,y \in \Omega_\mathcal{M}, x \neq y\}$ canonical if and only if it contains exactly one path for each $x,y \in \Omega_\mathcal{M}$ with $x \neq y$, and its *congestion* is defined to be

$$
\rho(\Gamma) = \max_{(w,z) \in E(\mathcal{M})} \frac{1}{Q_\mathcal{M}(w,z)} \sum_{\substack{x,y \in \Omega_\mathcal{M}: \\ (w,z) \in E(\gamma_{xy})}} |\gamma_{xy}| \pi_\mathcal{M}(x) \pi_\mathcal{M}(y).
$$

THEOREM 3.10 ([47, Theorem 5]). *For any reversible Markov chain $\mathcal{M}$ and every set of canonical paths Γ for $\mathcal{M}$ it holds that*

$$
\beta_2(P_\mathcal{M}) \leq 1 - \frac{1}{\rho(\Gamma)}.
$$

By applying the canonical-path method to the skeleton walk, we obtain the following lemma.

LEMMA 3.11. *Let (G, λ) be an instance of the multivariate hard-core model with disjoint clique cover Λ of size m, and let $Z_{\max} = \max_{i \in [m]}\{Z(G[\Lambda_i])\}$. Further, let (X, w) be the resulting simplicial-complex representation, and let $\mathcal{S} = \mathcal{S}(X, w)$ be the skeleton walk on (X, w). Then $\beta_2(P_\mathcal{S}) \leq 1 - \frac{1}{12Z_{\max}^2}$.*

Proof. As discussed in the proof of Lemma 3.3, $P_\mathcal{S}$ is reversible with respect to its stationary distribution $\pi_\mathcal{S}$, where $\pi_\mathcal{S}(x) = \frac{1}{m}\mathbb{P}_G[x]$. Thus, $Q_\mathcal{S}(x, y) = \frac{1}{m(m-1)}\mathbb{P}_G[x, y]$.

We start by constructing the paths $\Gamma = \{\gamma_{xy} \mid x, y \in U, x \neq y\}$. To this end, let p be a fixed-point-free permutation of $[m]$. Our construction goes as follows:

- $\gamma_{\emptyset_i \emptyset_j} = (\emptyset_i, \emptyset_j)$ for $i \neq j$,
- $\gamma_{x_v \emptyset_i} = (x_v, \emptyset_i)$, $\gamma_{\emptyset_i x_v} = (\emptyset_i, x_v)$ for all $v \in V$ with $v \notin \Lambda_i$,
- $\gamma_{x_v x_w} = (x_v, \emptyset_i, \emptyset_j, x_w)$ for $v \in \Lambda_j$, $w \in \Lambda_i$ with $i \neq j$,
- $\gamma_{xy} = (x, \emptyset_{p(i)}, y)$ for $x, y \in U_i$.

Let $E(\Gamma) = \bigcup_{\gamma \in \Gamma} E(\gamma)$ and note that for all $x \neq y$ we have $|\gamma_{xy}| \leq 3$. It suffices to upper bound

$$\rho(\Gamma) \leq 3 \max_{e \in E(\Gamma)} \frac{1}{Q_\mathcal{S}(e)} \sum_{\substack{x,y \in U: \\ e \in E(\gamma_{xy})}} \pi_\mathcal{S}(x)\pi_\mathcal{S}(y).$$

We derive such an upper bound by partitioning $E(\Gamma)$ into the three following types of edges:

- $A = \{(x_v, \emptyset_i) \in E(\Gamma) \mid v \notin \Lambda_i\}$,
- $B = \{(\emptyset_i, x_v) \in E(\Gamma) \mid v \notin \Lambda_i\}$,
- $C = \{(\emptyset_i, \emptyset_j) \in E(\Gamma) \mid i \neq j\}$.

We make a case distinction with respect to these types.

Case A. Let $(x_v, \emptyset_i) \in A$ and without loss of generality assume $v \in \Lambda_j$ for $j \neq i$. If $p(j) \neq i$, then $(x_v, \emptyset_i)$ is only used by paths that start at x_v and go to any element in U_i (including $\emptyset_i$). Further, if $p(j) = i$, then it is also used by paths from x_v to any $y \in U_j$ with $y \neq x_v$. Thus, we obtain

$$\frac{1}{Q_\mathcal{S}(x_v, \emptyset_i)} \sum_{\substack{x,y \in U: \\ (x_v, \emptyset_i) \in E(\gamma_{xy})}} \pi_\mathcal{S}(x)\pi_\mathcal{S}(y) \leq \frac{m(m-1)}{m^2} \frac{1}{\mathbb{P}_G[\emptyset_i \mid v]} \left(\sum_{y \in U_i} \mathbb{P}_G[y] + \sum_{\substack{y \in U_j: \\ y \neq x_v}} \mathbb{P}_G[y] \right)$$

$$\leq \frac{2}{\mathbb{P}_G[\emptyset_i \mid v]},$$

where the second inequality comes from the fact that we have $\sum_{y \in U_k} \mathbb{P}_G[y] = \mathbb{P}_G[\bigcup_{y \in U_k} y] = 1$ for all $k \in [m]$. By the submultiplicativity of Z,

$$\mathbb{P}_G[\emptyset_i \mid v] = \frac{Z(G[V \setminus (N[v] \cup \Lambda_i)])}{Z(G[V \setminus N[v]])} \geq \frac{Z(G[V \setminus (N[v] \cup \Lambda_i)])}{Z(G[V \setminus (N[v] \cup \Lambda_i)])Z(G[\Lambda_i \setminus N[v]])}$$

$$\geq \frac{1}{Z(G[\Lambda_i])}.$$

We obtain the bound

$$(3.4) \qquad \frac{1}{Q_\mathcal{S}(x_v, \emptyset_i)} \sum_{\substack{x,y \in U: \\ (x_v, \emptyset_i) \in E(\gamma_{xy})}} \pi_\mathcal{S}(x)\pi_\mathcal{S}(y) \leq 2Z(G[\Lambda_i]).$$

Case B. For $(\emptyset_i, x_v) \in B$, by symmetry, this case is analogous to Case *A*. Thus, we get

$$(3.5) \qquad \frac{1}{Q_\mathcal{S}(\emptyset_i, x_v)} \sum_{\substack{x,y \in U: \\ (\emptyset_i, x_v) \in E(\gamma_{xy})}} \pi_\mathcal{S}(x)\pi_\mathcal{S}(y) \leq 2Z(G[\Lambda_i]).$$

Case C. Finally, consider $(\emptyset_i, \emptyset_j) \in C$. If $p(j) \neq i$ and $p(i) \neq j$, then this edge is only used by paths from x_v to x_w for any pair $v \in \Lambda_j, w \in \Lambda_i$ and for the direct transition from $\emptyset_i$ to $\emptyset_j$. If $p(j) = i$, then it is also used by paths from any x_v for $v \in \Lambda_j$ to $\emptyset_j$. Symmetrically, if $p(i) = j$, then it is also used by paths from $\emptyset_i$ to x_v for $v \in \Lambda_i$. Thus, we obtain

$$\frac{1}{Q_\mathcal{S}(\emptyset_i, \emptyset_j)} \sum_{\substack{x,y \in U: \\ (\emptyset_i, \emptyset_j) \in E(\gamma_{xy})}} \pi_\mathcal{S}(x)\pi_\mathcal{S}(y)$$

$$\leq \frac{m(m-1)}{m^2} \frac{1}{\mathbb{P}_G[\emptyset_i, \emptyset_j]} \left(\sum_{\substack{v \in \Lambda_j: \\ w \in \Lambda_i}} \mathbb{P}_G[v]\mathbb{P}_G[w] + \mathbb{P}_G[\emptyset_i]\mathbb{P}_G[\emptyset_j] \right.$$

$$\left. + \mathbb{P}_G[\emptyset_j] \sum_{v \in \Lambda_j} \mathbb{P}_G[v] + \mathbb{P}_G[\emptyset_i] \sum_{v \in \Lambda_i} \mathbb{P}_G[v] \right)$$

$$\leq \frac{4}{\mathbb{P}_G[\emptyset_i, \emptyset_j]}.$$

Now, observe that

$$\mathbb{P}_G[\emptyset_i, \emptyset_j] = \frac{Z(G[V \setminus (\Lambda_i \cup \Lambda_j)])}{Z(G)} \geq \frac{Z(G[V \setminus (\Lambda_i \cup \Lambda_j)])}{Z(G[V \setminus (\Lambda_i \cup \Lambda_j)])Z(G[\Lambda_i])Z(G[\Lambda_j])}$$

$$= \frac{1}{Z(G[\Lambda_i])Z(G[\Lambda_j])}.$$

Thus,

$$(3.6) \qquad \frac{1}{Q_\mathcal{S}(\emptyset_i, \emptyset_j)} \sum_{\substack{x,y \in U \\ \text{s.t. } (\emptyset_i, \emptyset_j) \in E(\gamma_{xy})}} \pi_\mathcal{S}(x)\pi_\mathcal{S}(y) \leq 3Z(G[\Lambda_i])Z(G[\Lambda_j]).$$

Combining (3.4)–(3.6) we get $\rho(\Gamma) \leq 12Z_{\max}^2$. By Theorem 3.10 this implies $\beta_2(P_\mathcal{S}) \leq 1 - \frac{1}{12Z_{\max}^2}$, which concludes the proof. $\qquad\square$

Note that Lemmas 1.6, 3.3, and 3.11 only consider the skeleton walk on the complex (X, w). However, in order to bound the local expansion, we need to investigate the skeleton walk on all links (X_τ, w_τ) for every face $\tau \in X(k)$ with $0 \leq k \leq m - 2$. To achieve this, we map the link for any such face to the simplicial complex representation of a smaller instance, such that we can apply Theorem 2.3. To this end, we introduce the following lemma.

LEMMA 3.12. *Let X be a pure d-dimensional simplicial complex for $d \geq 2$, let w and w' be two weight functions for X, and let $\mathcal{S} = \mathcal{S}(X, w)$ and $\mathcal{S}' = \mathcal{S}(X, w')$. Further, if there is an $r \in \mathbb{R}_{>0}$ such that for all maximum faces $\tau \in X(d)$ we have $w'(\tau) = rw(\tau)$, then $\mathcal{S} = \mathcal{S}'$ and, in particular, $\beta_2(P_\mathcal{S}) = \beta_2(P_{\mathcal{S}'})$.*

Proof. We prove this statement by showing equality of the state spaces and transition probabilites. The fact that $\Omega_{\mathcal{S}} = \Omega_{\mathcal{S}'}$ follows directly from the fact that both walks are on the 1-skeleton of the same complex. Now, let $\tau' \in X$ be any face of X. Note that

$$w'(\tau') = \sum_{\substack{\tau \in X(d): \\ \tau' \subseteq \tau}} w'(\tau) = r \sum_{\substack{\tau \in X(d): \\ \tau' \subseteq \tau}} w(\tau) = rw(\tau').$$

Thus, the weights of all faces differ by the same factor r. Let $\{x\}, \{y\} \in X$ with $x \neq y$. If $\{x, y\} \notin X$, then $P_{\mathcal{S}}(x, y) = 0 = P_{\mathcal{S}'}(x, y)$. Otherwise, if $\{x, y\} \in X$, then

$$P_{\mathcal{S}'}(x, y) = \frac{w'(\{x, y\})}{\displaystyle\sum_{\substack{\{z\} \in X: \\ \{x, z\} \in X}} w'(\{x, z\})} = \frac{rw(\{x, y\})}{r \displaystyle\sum_{\substack{\{z\} \in X: \\ \{x, z\} \in X}} w(\{x, z\})} = P_{\mathcal{S}}(x, y).$$

As both $\mathcal{S}$ and $\mathcal{S}'$ have self-loop probabilities of 0, it follows $P_{\mathcal{S}} = P_{\mathcal{S}'}$. This implies $\beta_2(P_{\mathcal{S}}) = \beta_2(P_{\mathcal{S}'})$. $\qquad\square$

3.5. Bounding the mixing time. We now have everything to state and prove our main theorem on the mixing time of block dynamics.

THEOREM 1.4. *Let (G, λ) be an instance of the multivariate hard-core model that satisfies Condition* 1.3. *Let Λ be a disjoint clique cover for G of size m, and let $Z_{\max} = \max_{i \in [m]} \{Z(G[\Lambda_i], \lambda[\Lambda_i])\}$. The mixing time of the block dynamics $\mathcal{B}(G, \lambda, \Lambda)$, starting from $\emptyset \in \mathcal{I}(G)$, is bounded by*

$$\tau_{\mathcal{B}}^{(\emptyset)}(\varepsilon) \leq m^{\mathrm{O}((2+C)C)} Z_{\max}^{\mathrm{O}((2+C)C)} \ln\left(\frac{1}{\varepsilon}\right).$$

Proof. By (2.1) it is sufficient to lower bound the spectral gap of $P_{\mathcal{B}}$ by $\frac{1}{\mathrm{poly}(Z_{\max})}$ and $\frac{1}{\mathrm{poly}(m)}$ to prove our claim. Further, transforming the chain into a lazy version only results in constant overhead in the mixing time. Thus, we focus on lower-bounding $1 - \beta_2(P_{\mathcal{B}})$, which is equivalent to upper-bounding $\beta_2(P_{\mathcal{B}})$.

Next, let (X, w) be the simplicial-complex representation based on Λ with ground-set U and partitions $\{U_i\}_{i \in [m]}$, and let $\mathcal{V} = \mathcal{V}(X, w)$ denote the two-step walk on (X, w). By Observation 3.2, we have $\beta_2(P_{\mathcal{B}}) = \beta_2(P_{\mathcal{V}})$ and it suffices to obtain an upper bound on $\beta_2(P_{\mathcal{V}})$.

To this end, we aim to apply Theorem 2.3, which involves upper-bounding local expansion of the simplicial-complex representation. Let C be the constant for which (G, λ) satisfies Condition 1.3. We proceed by proving that (X, w) is a local $(\alpha_0, \ldots, \alpha_{m-2})$-expander, where

(3.7) $$\alpha_k \leq \min\left\{1 - \frac{1}{12Z_{\max}^2}, \frac{(2+C)C}{m-k-1}\right\} \text{ for } 0 \leq k \leq m-2.$$

We start by arguing both bounds for the case $k = 0$. Then we generalize our arguments for the cases $k \in [m-2]$.

Case $k = 0$. Let $\mathcal{S} = \mathcal{S}(X, w)$ be the local skeleton walk on (X, w). By definition, we have $\alpha_0 = \beta_2(P_{\mathcal{S}})$. Note that the first bound $\alpha_0 \leq 1 - \frac{1}{12Z_{\max}^2}$ follows directly from Lemma 3.11. To prove the second bound, we apply Lemma 3.3, which gives us

$$\alpha_0 \leq \frac{\beta_1(\Phi_{G,\Lambda})}{m-1}.$$

By Lemma 1.6, we conclude that $\alpha_0 \leq \frac{(2+C)C}{m-1}$.

Case $k \in [m-2]$. By definition, we have to show for all $\tau \in X(k)$ that the skeleton walk $\mathcal{S}_\tau = \mathcal{S}(X_\tau, w_\tau)$ on the link (X_τ, w_τ) satisfies

$$\beta_2(P_{\mathcal{S}_\tau}) \leq \min\left\{1 - \frac{1}{12Z_{\max}^2}, \frac{(2+C)C}{m-k-1}\right\}.$$

We construct a subset of vertices $S \subseteq V$ such that

$$S = \bigcup_{x \in \tau} \begin{cases} N[v] & \text{if } x = x_v \text{ for some } v \in V, \\ \Lambda_i & \text{if } x = \emptyset_i \text{ for some } i \in [m]. \end{cases}$$

Let $G' = G[V \setminus S]$ be the subgraph induced by $V \setminus S$ and let $\lambda' = \lambda[V \setminus S]$ be the corresponding vertex weights. Further, let $\Lambda' = \{\Lambda_i \setminus S \mid i \in [m] \wedge \tau \cap U_i = \emptyset\}$ and note that Λ' is a disjoint clique cover of the multivariate hard-core instance (G', λ'). Let (X', w') be the corresponding simplicial-complex representation and let $\mathcal{S}' = \mathcal{S}(X', w')$ be the skeleton walk on (X', w'). Note that $X' = X_\tau$ and that for every maximum face $\tau' \in X_\tau(m - |\tau|)$ it holds that

$$\begin{aligned}
w_\tau(\tau') &= w(\tau' \cup \tau) \\
&= \frac{1}{Z(G, \lambda)}\left(\prod_{w \in V \setminus S} \lambda_w \mathbb{1}_{x_w \in \tau'}\right)\left(\prod_{v \in S} \lambda_v \mathbb{1}_{x_v \in \tau}\right) \\
&= \frac{1}{Z(G', \lambda')}\left(\prod_{w \in V \setminus S} \lambda'_w \mathbb{1}_{x_w \in \tau'}\right)\frac{Z(G', \lambda')}{Z(G, \lambda)}\left(\prod_{v \in S} \lambda_v \mathbb{1}_{x_v \in \tau}\right) \\
&= w'(\tau')Z(G', \lambda')w(\tau),
\end{aligned}$$

where $Z(G', \lambda')w(\tau) > 0$. Thus, by Lemma 3.12, we obtain $\beta_2(P_{\mathcal{S}_\tau}) = \beta_2(P_{\mathcal{S}'})$.

To upper bound $\beta_2(P_{\mathcal{S}'})$, observe that

- $|\Lambda'| = m - |\tau| = m - k$,
- $\max_{\Lambda'_i \in \Lambda'}\{Z(G[\Lambda'_i])\} \leq Z_{\max}$, and
- G' is and induced subgraph of G.

Thus, analogous to the case $k = 0$, applying Lemma 3.11 yields

$$\alpha_k \leq 1 - \frac{1}{12\max_{\Lambda'_i \in \Lambda'}\{Z(G[\Lambda'_i])\}^2} \leq 1 - \frac{1}{12Z_{\max}^2}.$$

Together, Lemmas 1.6 and 3.3 result in

$$\alpha_k \leq \frac{(2+C)C}{|\Lambda'| - 1} = \frac{(2+C)C}{m-k-1}.$$

From (3.7) and Theorem 2.3, we conclude

$$\begin{aligned}
\beta_2(P_\mathcal{V}) &\leq 1 - \frac{1}{m} \prod_{0 \leq k \leq m-2}\left(1 - \min\left\{1 - \frac{1}{12Z_{\max}^2}, \frac{(2+C)C}{m-k-1}\right\}\right) \\
&= 1 - \frac{1}{m} \prod_{0 \leq k \leq m-2} \max\left\{\frac{1}{12Z_{\max}^2}, 1 - \frac{(2+C)C}{m-k-1}\right\}.
\end{aligned}$$

Let $k_0 = m - 2(2 + C)C - 1$ and observe that for $k \leq k_0$ it holds that

$$1 - \frac{(2 + C)C}{m - k - 1} \geq \frac{1}{2} > \frac{1}{12} \geq \frac{1}{12Z_{\max}^2}.$$

Thus, we have

$$\frac{1}{m} \prod_{0 \leq k \leq m-2} \max\left\{ \frac{1}{12Z_{\max}^2}, 1 - \frac{(2 + C)C}{m - k - 1} \right\}$$

$$\geq \frac{1}{m} \left(\prod_{0 \leq k \leq k_0} \left(1 - \frac{(2 + C)C}{m - k - 1} \right) \right) \left(\prod_{k_0 < k \leq m-2} \frac{1}{12Z_{\max}^2} \right)$$

$$= \frac{1}{m} \left(\frac{1}{12Z_{\max}^2} \right)^{m-2-k_0} \prod_{0 \leq k \leq k_0} \left(1 - \frac{(2 + C)C}{m - k - 1} \right)$$

$$= \frac{1}{m} \left(\frac{1}{\sqrt{12}Z_{\max}} \right)^{4(2+C)C-2} \prod_{0 \leq k \leq k_0} \left(1 - \frac{(2 + C)C}{m - k - 1} \right).$$

Further, because $\ln(1 - x) \geq \frac{-x}{1-x}$ for $x < 1$, we have

$$\ln\left(\prod_{0 \leq k \leq k_0} \left(1 - \frac{(2 + C)C}{m - k - 1} \right) \right) = \sum_{0 \leq k \leq k_0} \ln\left(1 - \frac{(2 + C)C}{m - k - 1} \right)$$

$$\geq \sum_{0 \leq k \leq k_0} \frac{-(2 + C)C/(m - k - 1)}{1 - (2 + C)C/(m - k - 1)}$$

$$= -(2 + C)C \sum_{0 \leq k \leq k_0} \frac{1}{m - (2 + C)C - k - 1}$$

$$= -(2 + C)C \sum_{(2+C)C \leq j \leq m-(2+C)C-1} \frac{1}{j}$$

$$\geq -(2 + C)C \ln(m).$$

We obtain

$$(3.8) \qquad \beta_2(P_{\mathcal{V}}) \leq 1 - \frac{1}{m} \left(\frac{1}{\sqrt{12}Z_{\max}} \right)^{4(2+C)C-2} e^{-(2+C)C \ln(m)}$$

$$= 1 - \left(\frac{1}{m} \right)^{(2+C)C+1} \left(\frac{1}{\sqrt{12}Z_{\max}} \right)^{4(2+C)C-2}$$

and thus

$$1 - \beta_2(P_{\mathcal{B}}) \geq \frac{1}{O\left(Z_{\max}^{4(2+C)C-2} m^{(2+C)C+1} \right)}.$$

As C is assumed to be a constant, this implies the desired mixing time. □

4. Univariate model: Mixing up to uniqueness. We consider the univariate hard-core model, often just referred to as the hard-core model, in which all vertices $v \in V$ have the same weight $\lambda_v = \lambda$ for some $\lambda \in \mathbb{R}_{>0}$. We denote an instance of this model by (G, λ).

We define $\lambda_c(\Delta) = \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^\Delta}$ to be the *critical weight* of the hard-core model. As we discussed in the introduction, $\lambda_c(\Delta)$ is the threshold for correlation decay on general graphs and a tight upper bound for rapid mixing of Glauber dynamics.

We show that the univariate model (G, λ) satisfies Condition 1.3 for all $\lambda \leq \lambda_c(\Delta)$. To do so, we use the following recently established result.

LEMMA 4.1 ([10, Theorem 5][1]). *Let (G, λ) be an instance of the univariate hard-core model and assume that the maximum degree of G is bounded by Δ. If there is a constant $\delta > 0$ such that $\lambda \leq (1 - \delta)\lambda_c(\Delta)$, then there is a constant $C \in \mathrm{O}\big(\frac{1}{\delta}\big)$ such that for all $S \subseteq V$ it holds that $\big\|\Psi_{G[S]}\big\|_\infty \leq C$.*

This implies the following result immediately.

LEMMA 4.2. *Let (G, λ) be an instance of the univariate hard-core model and assume that the maximum degree of G is bounded by Δ. If there is a constant $\delta > 0$ such that $\lambda \leq (1 - \delta)\lambda_c(\Delta)$, then (G, λ) satisfies Condition 1.3 for a constant $C \in \mathrm{O}\big(\frac{1}{\delta}\big)$.*

Proof. By Lemma 4.1, there is a $C \in \mathrm{O}\big(\frac{1}{\delta}\big)$ such that, for all $S \in V$ and $r \in S$, it holds that

$$\sum_{v \in S}\big|\Psi_{G[S]}(r, v)\big| \leq \big\|\Psi_{G[S]}\big\|_\infty \leq C.$$

Thus, Condition 1.3 is satisfied for the same constant C and $q(v) = 1$ for all $v \in V$.□

The following claim is a direct consequence of Theorem 1.4 and Lemma 4.2.

COROLLARY 4.3. *Let (G, λ) be an instance of the univariate hard-core model and assume that the maximum degree of G is bounded by Δ. Let Λ be a disjoint clique cover for G of size m, and let $Z_{\max} = \max_{i \in [m]}\{Z(G[\Lambda_i])\}$. If there is a constant $\delta > 0$ such that $\lambda \leq (1 - \delta)\lambda_c(\Delta)$, then the mixing time of the block dynamics $\mathcal{B}(G, \lambda, \Lambda)$, starting from $\emptyset \in \mathcal{I}(G)$, is bounded by*

$$\tau_\mathcal{B}^{(\emptyset)}(\varepsilon) \leq m^{\mathrm{O}\big(1/\delta^2\big)} Z_{\max}^{\mathrm{O}\big(1/\delta^2\big)} \ln\bigg(\frac{1}{\varepsilon}\bigg).$$

5. Multivariate model: Comparison to clique dynamics condition. In this section, we relate Condition 1.3 to a strict version of the clique dynamics condition, first introduced in [22].

DEFINITION 5.1 (strict clique dynamics condition). *An instance of the multivariate hard-core model (G, λ) satisfies the strict clique dynamics condition for a function $f\colon V \to \mathbb{R}_{>0}$ and a constant $\alpha \in (0, 1)$ if and only if for all $v \in V$ it holds that*

$$\sum_{w \in N(v)} \frac{\lambda_w}{1 + \lambda_w} f(w) \leq (1 - \alpha)f(v).$$

We show that the strict clique dynamics condition is sufficient to imply Condition 1.3. This yields a mixing-time bound for block dynamics based on the strict clique dynamics condition and also bounds the eigenvalues of the pairwise influence matrix, which might be of independent interest. To obtain our result, we translate the original instance (G, λ) to the self-avoiding-walk tree and apply a recursive proof on this tree.

---

[1]In fact, the statement is part of the proof of [10, Theorem 5].

Influence in self-avoiding-walk trees. For any instance of the multivariate hard-core model (G, λ) and any vertex $r \in V$, let $\mathbb{T}(G, r)$ denote the tree of self-avoiding walks as defined in [52], which is constructed as follows. Assume there is a total order of vertices in $v_1, \ldots, v_n$, where $n = |V|$. A *self-avoiding walk* of length $l \geq 2$ is a simple path $v_{i_1}, v_{i_2}, \ldots, v_{i_l}$ in G. Further, a *closed self-avoiding walk* $v_{i_1}, \ldots, v_{i_{l-1}}, v_{i_j}$ of length $l \geq 3$ consists of a self-avoiding walk $v_{i_1}, \ldots, v_{i_{l-1}}$ of length $l - 1$ and an appended vertex v_{i_j} such that $j \in [l - 2]$ and $(v_{i_{l-1}}, v_{i_j}) \in E$. That is, the edge $(v_{i_{l-1}}, v_{i_j})$ closes a cycle $v_{i_j}, v_{i_{j+1}}, \ldots, v_{i_{l-1}}$. The graph $\mathbb{T}(G, r)$ consists of all closed self-avoiding walks with $v_{i_1} = r$, and all self-avoiding walks with $v_{i_1} = r$ and v_{i_l} having degree 1 in G. Note that any vertex $v \in V$ with $v \neq r$ might have multiple copies in $\mathbb{T}(G, r)$.

For any root $r \in V$, the multivariate hard-core model (G, λ) is translated to a multivariate hard-core model on $\mathbb{T}(G, r)$ as follows. Let $v_{i_1}, \ldots, v_{i_{l-1}}, v_{i_j}$ be a closed self-avoiding walk. We fix v_{i_j} always to be in the independent set (fix spin to 1) if $i_{j+1} > i_{l-1}$, and we fix it always to be excluded from the independent (fix spin to 0) otherwise. We call such vertices *fixed copies*. For each $v \in V$, let $C_{\mathbb{T}(G,r)}(v)$ denote the set of all *unfixed copies of v* in $\mathbb{T}(G, r)$. We write $C(v)$ if the tree $\mathbb{T}(G, r)$ is clear from the context. In the multivariate hard-core model on $\mathbb{T}(G, r)$, each such copy $\widehat{v} \in C_{\mathbb{T}(G,r)}(v)$ has weight $\lambda_{\widehat{v}} = \lambda_v$.

This way of translating (G, λ) to the tree of self-avoiding walks for some root $r \in V$ was shown to have a variety of useful properties. One of them is that pairwise influences are preserved in the following sense.

LEMMA 5.2 ([10, Lemma 8]). *Let (G, λ) be an instance of the multivariate hard-core model. For all $r, v \in V$ and $T = \mathbb{T}(G, r)$ it holds that*

$$\Psi_G(r, v) = \sum_{\widehat{v} \in C_T(v)} \Psi_T(r, \widehat{v}).$$

Lemma 5.2 states that it suffices to discuss the pairwise influence on the self-avoiding walk tree instead of the original graph. This allows us to use the following multiplicative property for pairwise influence along paths in tree graphs.

LEMMA 5.3 ([3, Lemma B.2]). *Let $T = (V, E)$ be a tree and (T, λ) be a multivariate hard-core model on T. Further, let $v, w \in V$ be a pair of distinct, nonadjacent vertices, and let $u \in V$ with $u \neq v$ and $u \neq w$ be any vertex on the unique path between v and w. Then*

$$\Psi_T(v, w) = \Psi_T(v, u)\Psi_T(u, w).$$

Bounding pairwise influence via the strict clique dynamics condition. We start by proving that the influence of the root on a certain layer in the self-avoiding-walk tree exhibits the following exponential decay in terms of depth. For a tree T and integer k let $L_T(k) \subseteq \widehat{V}$ denote the set of vertices in T at layer $k \in \mathbb{N}$.

LEMMA 5.4. *Let (G, λ) be a multivariate hard-core model, and let $r \in V$. Furthermore, let $T = \mathbb{T}(r, G)$ and let $\widehat{V} = \bigcup_{v \in V} C_T(v)$. Assume that (G, λ) satisfies the strict clique dynamics condition for a function f and a constant α, and define the function $\widehat{f} : \widehat{V} \to \mathbb{R}_{>0}$ with $\widehat{f}(\widehat{v}) = f(v)$ for all $\widehat{v} \in C_T(v)$ and $v \in V$. Then for all $k \in \mathbb{N}_{>0}$ it holds that*

$$\sum_{w \in L_T(k)} |\Psi_T(r, w)| \widehat{f}(w) \leq (1 - \alpha)^k f(r).$$

Proof. Note that if (G, λ) satisfies the strict clique dynamics condition for a function f and a constant α, then the corresponding multivariate hard-core instance on T satisfies the strict clique dynamics condition for $\widehat{f}$ and the same constant α. Based on that, we prove our claim by induction on k.

Base case: $k = 1$. Note that $L_T(1) = N_T(r)$. Further, we have for each $w \in N_T(r)$ by definition

$$|\Psi_T(r, w)| = |\mathbb{P}_T[w \mid r] - \mathbb{P}_T[w \mid \overline{r}]| \leq \frac{\lambda_w}{1 + \lambda_w}.$$

Thus, by the strict clique dynamics condition, we obtain

$$\sum_{w \in L_T(1)} |\Psi_T(r, w)| \widehat{f}(w) \leq \sum_{w \in L_T(1)} \frac{\lambda_w}{1 + \lambda_w} \widehat{f}(w) \leq (1 - \alpha)\widehat{f}(r) = (1 - \alpha)f(r),$$

which proves the case $k = 1$.

Induction step: $k > 1$. Assume that the statement holds for $k - 1$. For every $u \in L_T(k-1)$, let T_u denote the subtree rooted at u, and let $L_{T_u}(l)$ denote the vertices at layer $l \in \mathbb{N}$ in T_u. Note that the sets $L_{T_u}(1)$ for $u \in L_T(k-1)$ are a partition of $L_T(k)$. By Lemma 5.3, we get

$$\sum_{w \in L_T(k)} |\Psi_T(r, w)| \widehat{f}(w) = \sum_{u \in L_T(k-1)} \sum_{w \in L_{T_u}(1)} |\Psi_T(r, w)| \widehat{f}(w)$$

$$= \sum_{u \in L_T(k-1)} |\Psi_T(r, u)| \sum_{w \in L_{T_u}(1)} |\Psi_T(u, w)| \widehat{f}(w).$$

Further, for every $u \in L_T(k-1)$ it holds that $L_{T_u}(1) \subset N_T(u)$, and for all $w \in L_{T_u}(1)$ we have

$$|\Psi_T(u, w)| = |\mathbb{P}_T[w \mid u] - \mathbb{P}_T[w \mid \overline{u}]| \leq \frac{\lambda_w}{1 + \lambda_w}.$$

Thus, by the strict clique dynamics condition, we get

$$\sum_{u \in L_T(k-1)} |\Psi_T(r, u)| \sum_{w \in L_{T_u}(1)} |\Psi_T(u, w)| \widehat{f}(w) < \sum_{u \in L_T(k-1)} |\Psi_T(r, u)| \sum_{w \in N_T(u)} \frac{\lambda_w}{1 + \lambda_w} \widehat{f}(w)$$

$$\leq (1 - \alpha) \sum_{u \in L_T(k-1)} |\Psi_T(r, u)| \widehat{f}(u).$$

By the induction hypothesis, we obtain

$$(1 - \alpha) \sum_{u \in L_T(k-1)} |\Psi_T(r, u)| \widehat{f}(u) \leq (1 - \alpha)(1 - \alpha)^{k-1} f(r) = (1 - \alpha)^k f(r),$$

which concludes the proof. $\square$

Now, we use this layerwise decay in the self-avoiding-walk tree to prove that Condition 1.3 is satisfied.

LEMMA 5.5. *Let (G, λ) be an instance of the multivariate hard-core model. If (G, λ) satisfies the strict clique dynamics condition for a function f and a constant α, then it also satisfies Condition* 1.3 *for $q = f$ and $C = \frac{1}{\alpha}$.*

Proof. Note that if (G, λ) satisfies the strict clique dynamics condition, the same holds for the instance $(G[S], \lambda[S])$ for the same function f and constant α.

Assume $G[S]$ is connected and let $G' = G[S]$. Further, let $T = \mathbb{T}(r, G')$ and $\widehat{S} = \bigcup_{v \in S} C_T(v)$, and define the function $\widehat{f} \colon \widehat{S} \to \mathbb{R}_{>0}$ as in Lemma 5.4. Recall that, by definition, $\Psi_{G'}(r, r) = 0$. By Lemma 5.2, we get

$$
\sum_{v \in S} |\Psi_{G'}(r, v)| f(v) = \sum_{v \in S \setminus \{r\}} |\Psi_{G'}(r, v)| f(v)
$$

$$
= \sum_{v \in S \setminus \{r\}} \left| \sum_{\widehat{v} \in C_T(v)} \Psi_T(r, \widehat{v}) \right| f(v)
$$

$$
\leq \sum_{v \in S \setminus \{r\}} \sum_{\widehat{v} \in C_T(v)} |\Psi_T(r, \widehat{v})| \widehat{f}(\widehat{v}).
$$

Note that the sets $C_T(v)$ for $v \in S \setminus \{r\}$ are a partition of $\widehat{S} \setminus \{r\}$. Recall that $L_T(k) \subset \widehat{S}$ denotes the vertices in T at layer $k \in \mathbb{N}$, and observe that the sets $L_T(k)$ for $k \in \mathbb{N}_{>0}$ are a partition of $\widehat{S} \setminus \{r\}$ as well. Thus, we have

$$
\sum_{v \in S \setminus \{r\}} \sum_{\widehat{v} \in C_T(v)} |\Psi_T(r, \widehat{v})| \widehat{f}(\widehat{v}) \leq \sum_{k \in \mathbb{N}_{>0}} \sum_{w \in L_T(k)} |\Psi_T(r, w)| \widehat{f}(w).
$$

By Lemma 5.4, we obtain the desired bound:

$$
\sum_{k \in \mathbb{N}_{>0}} \sum_{w \in L_T(k)} |\Psi_T(r, w)| \widehat{f}(w) \leq f(r) \sum_{k \in \mathbb{N}_{>0}} (1 - \alpha)^k = \frac{1}{\alpha} f(r).
$$

Now, assume $G[S]$ is not connected. Set G' to be the largest connected component in $G[S]$ that contains r and let $S' \subset S$ be the set of vertices in G'. The claim follows from applying the proof above to G' with vertex set S' and by observing that $\Psi_{G[S]}(r, v) = 0$ for all $v \in S \setminus S'$. $\square$

Lemma 1.5 immediately implies with Theorem 1.4 the following result for the mixing time of block dynamics under a strict clique dynamics condition.

COROLLARY 5.6. *Let (G, λ) be an instance of the multivariate hard-core model. Let Λ be a given disjoint clique cover for G of size m, and let $Z_{\max} = \max_{i \in [m]} \{Z(G[\Lambda_i])\}$. If (G, λ) satisfies the strict clique dynamics condition for a function f and a constant α, then the mixing time of the block dynamics $\mathcal{B} = \mathcal{B}(G, \lambda, \Lambda)$, starting from $\emptyset \in \mathcal{I}(G)$, is bounded by*

$$
\tau_{\mathcal{B}}^{(\emptyset)}(\varepsilon) \leq m^{O(1/\alpha^2)} Z_{\max}^{O(1/\alpha^2)} \ln\left(\frac{1}{\varepsilon}\right).
$$

Finally, Lemma 1.5 together with Lemma 3.5 implies the following result.

COROLLARY 5.7. *Let (G, λ) be an instance of the multivariate hard-core model that satisfies the strict clique dynamics condition for a function f and a constant α. For every $S \subseteq V$ it holds that $\beta_1\big(\Psi_{G[S]}\big) \leq \frac{1}{\alpha}$.*

6. The monoatomic hard-sphere model. We study the *grand canonical ensemble* of the monoatomic hard-sphere model in a d-dimensional finite cubic region $\mathbb{V} = [0, \ell]^d$ of Euclidean space with side length $\ell \in \mathbb{R}_{\geq 1}$. We write $|\mathbb{V}| = \ell^d$ for the

volume of $\mathbb{V}$. The hard-sphere model describes the distribution of identical particles, represented as d-dimensional balls in $\mathbb{V}$. This distribution is governed by a fugacity parameter $\lambda \in \mathbb{R}_{>0}$, describing the contribution of each particle to the chemical potential, and hard-core interactions between particles, meaning that no two particles are allowed to overlap. For simplicity, it is common to assume particles to have volume 1, meaning that their radius is $r = (1/v_d)^{1/d}$, where v_d denotes the volume of a unit sphere in d dimensions.

A probabilistic interpretation of grand canonical ensemble is that the centers of particles are distributed according to a Poisson point process on $\mathbb{V}$ with activity λ, conditioned on the fact that particles are nonoverlapping (i.e., each pair of distinct centers has distance at least $2r$). Note that this implies that particles are indistinguishable, meaning that exchanging the positions of two particles results in exactly the same configuration of the system. We aim for approximating the *grand canonical partition function*, which can be seen as the normalizing constant of the resulting distribution of system states. As a reminder, the partition function can formally be defined as

$$Z(\mathbb{V}, \lambda) = 1 + \sum_{k \in \mathbb{N}_{>0}} \frac{\lambda^k}{k!} \int_{\mathbb{V}^k} D\Big(x^{(1)}, \ldots, x^{(k)}\Big) \, \mathrm{d}\nu^{d \times k},$$

where

$$D\Big(x^{(1)}, \ldots, x^{(k)}\Big) = \begin{cases} 1 & \text{if } d\big(x^{(i)}, x^{(j)}\big) \geq 2r \text{ for all } i, j \in [k] \text{ with } i \neq j, \\ 0 & \text{otherwise} \end{cases}$$

and $\nu^{d \times k}$ is the Lebesgue measure on $\mathbb{R}^{d \times k}$.

6.1. Hard-core representation. To apply our result for block dynamics to the continuous hard-sphere model, we will approximate it by an instance of the hard-core model. The main idea of this discretization is to restrict the centers of spheres to vertices in an integer grid, while scaling the fugacity λ and the radius r appropriately. The resulting discrete hard-sphere model can easily be transformed into a hard-core instance. We proceed by formalizing the direct transformation from the continuous hard-sphere model instance to the discrete hard-core model.

Let $(\mathbb{V}, \lambda)$ be an instance of the continuous hard-sphere model with $\mathbb{V} = [0, \ell]^d$. Recall that we fixed the radius $r = (1/v_d)^{1/d}$. Let $\mathbb{G}(n) = \mathbb{Z}^d \cap [0, n)^d$ be a finite integer grid of side length $n \in \mathbb{N}_{>0}$. For any $\rho \in \mathbb{R}_{>0}$ such that $\rho\ell \in \mathbb{N}_{>0}$, the hard-core representation of $(\mathbb{V}, \lambda)$ with resolution ρ is a hard-core model (G_ρ, λ_ρ) with $G_\rho = (V_\rho, E_\rho)$ and
- there is a vertex $v_x \in V_\rho$ for each grid point $x \in \mathbb{G}(\rho\ell)$,
- there is an edge $(v_x, v_y) \in E_\rho$ for any pair of grid points $x, y \in \mathbb{G}(\rho\ell)$ with $x \neq y$ and $d(x, y) \leq 2\rho r$, and
- $\lambda_\rho = \rho^{-d}\lambda$.

Note that in the above definition $d(x, y)$ denotes the Euclidean distance.

We will use the following convergence result for the partition function of the hard-core representation in terms of the resolution ρ to approximate the hard-sphere partition function.

LEMMA 6.1. *Let $(\mathbb{V}, \lambda)$ be an instance of the continuous hard-sphere model in d dimensions. For each resolution $\rho \geq 2\sqrt{d}$ it holds that*

$$1 - \rho^{-1}\mathrm{e}^{\Theta(|\mathbb{V}| \ln |\mathbb{V}|)} \leq \frac{Z(\mathbb{V}, \lambda)}{Z(G_\rho, \lambda_\rho)} \leq 1 + \rho^{-1}\mathrm{e}^{\Theta(|\mathbb{V}| \ln |\mathbb{V}|)}.$$

Proof. Note that it suffices to bound the additive error $|Z(\mathbb{V}, \lambda) - Z(G_\rho, \lambda_\rho)|$. Because $Z(\mathbb{V}, \lambda) \geq 1$, this directly results in the desired multiplicative bound.

In order to obtain an additive bound, we start by transforming $Z(G_\rho, \lambda_\rho)$ to a form that is more similar to the form of $Z(\mathbb{V}, \lambda)$.

Let $\mathbb{V} = [0, \ell)^d$ and let $\varphi^{(\rho)} \colon \mathbb{G}(\rho\ell) \to \mathbb{V}$ with $(x_1, \ldots, x_d) \mapsto \varphi^{(\rho)}(x) = (x_1/\rho, \ldots, x_d/\rho)$. Note that, for all $x^{(i)}, x^{(j)} \in \mathbb{G}(\rho\ell)$ it holds that

$$d\left(x^{(i)}, x^{(j)}\right) \geq 2\rho r \;\leftrightarrow\; d\left(\varphi^{(\rho)}\left(x^{(i)}\right), \varphi^{(\rho)}\left(x^{(j)}\right)\right) \geq 2r.$$

Thus, we see that

$$\begin{aligned}
Z(G_\rho, \lambda_\rho) &= \sum_{I \in \mathcal{I}(G_\rho)} \lambda_\rho^{|I|} \\
&= 1 + \sum_{k \in \mathbb{N}_{>0}} \sum_{\substack{I \in \mathcal{I}(G_\rho) \\ |I| = k}} \lambda_\rho^k \\
&= 1 + \sum_{k \in \mathbb{N}_{>0}} \frac{\lambda_\rho^k}{k!} \sum_{\substack{\left(x^{(1)}, \ldots, x^{(k)}\right) \\ \in (\mathbb{G}(\rho\ell))^k}} D\left(\varphi^{(\rho)}\left(x^{(1)}\right), \ldots, \varphi^{(\rho)}\left(x^{(k)}\right)\right) \\
(6.1) \qquad &= 1 + \sum_{k \in \mathbb{N}_{>0}} \frac{\lambda^k}{k!} \sum_{\substack{\left(x^{(1)}, \ldots, x^{(k)}\right) \\ \in (\mathbb{G}(\rho\ell))^k}} \left(\frac{1}{\rho}\right)^{d \cdot k} D\left(\varphi^{(\rho)}\left(x^{(1)}\right), \ldots, \varphi^{(\rho)}\left(x^{(k)}\right)\right).
\end{aligned}$$

We continue by rewriting

$$\sum_{\substack{\left(x^{(1)}, \ldots, x^{(k)}\right) \\ \in (\mathbb{G}(\rho\ell))^k}} \left(\frac{1}{\rho}\right)^{d \cdot k} D\left(\varphi^{(\rho)}\left(x^{(1)}\right), \ldots, \varphi^{(\rho)}\left(x^{(k)}\right)\right)$$

for any fixed $k \in \mathbb{N}_{>0}$. Let $\varphi^{(\rho)}(\mathbb{G}(\rho\ell)) \subseteq \mathbb{V}$ denote the image of $\varphi^{(\rho)}$, and let $\Phi^{(\rho)} \colon \mathbb{V} \to \varphi^{(\rho)}(\mathbb{G}(\rho\ell))$ with

$$(x_1, \ldots, x_d) \mapsto \left(\frac{\lfloor \rho x_1 \rfloor}{\rho}, \ldots, \frac{\lfloor \rho x_d \rfloor}{\rho}\right).$$

Further, for all $k \in \mathbb{N}_{>0}$ and all $\left(x^{(1)}, \ldots, x^{(k)}\right) \in \left(\varphi^{(\rho)}(\mathbb{G}(\rho\ell))\right)^k$, let

$$\begin{aligned}
W_{x^{(1)}, \ldots, x^{(k)}}^{(\rho)} &= \left\{ \left(y^{(1)}, \ldots, y^{(k)}\right) \in \mathbb{V}^k \;\middle|\; \forall i \in [k] \colon \Phi^{(\rho)}\left(y^{(i)}\right) = x^{(i)} \right\} \\
&= \left(\left(\Phi^{(\rho)}\right)^{-1}\left(x^{(1)}\right)\right) \times \cdots \times \left(\left(\Phi^{(\rho)}\right)^{-1}\left(x^{(k)}\right)\right).
\end{aligned}$$

Note that the sets $W_{x^{(1)}, \ldots, x^{(k)}}^{(\rho)}$ partition $\mathbb{V}^k$ into $(d \times k)$-dimensional hypercubes of side length $1/\rho$. Thus, for all $\left(x^{(1)}, \ldots, x^{(k)}\right) \in \left(\varphi^{(\rho)}(\mathbb{G}(\rho\ell))\right)^k$, it holds that

$$\nu^{d \times k}\left(W_{x^{(1)}, \ldots, x^{(k)}}^{(\rho)}\right) = \left(\frac{1}{\rho}\right)^{d \cdot k}.$$

By this and by the definition of a Lebesgue integral for elementary functions, we obtain

$$
\sum_{\substack{(x^{(1)},\ldots,x^{(k)}) \\ \in (\mathbb{G}(\rho\ell))^k}} \left(\frac{1}{\rho}\right)^{d\cdot k} D\Big(\varphi^{(\rho)}\Big(x^{(1)}\Big),\ldots,\varphi^{(\rho)}\Big(x^{(k)}\Big)\Big)
$$

$$
= \sum_{\substack{(x^{(1)},\ldots,x^{(k)}) \\ \in (\mathbb{G}(\rho\ell))^k}} \nu^{d\times k}\Big(W^{(\rho)}_{\varphi^{(\rho)}(x^{(1)}),\ldots,\varphi^{(\rho)}(x^{(k)})}\Big) \cdot D\Big(\varphi^{(\rho)}\Big(x^{(1)}\Big),\ldots,\varphi^{(\rho)}\Big(x^{(k)}\Big)\Big)
$$

$$
= \sum_{\substack{(x^{(1)},\ldots,x^{(k)}) \\ \in \big(\varphi^{(\rho)}(\mathbb{G}(\rho\ell))\big)^k}} \nu^{d\times k}\Big(W^{(\rho)}_{x^{(1)},\ldots,x^{(k)}}\Big) \cdot D\Big(x^{(1)},\ldots,x^{(k)}\Big)
$$

$$
= \int_{\mathbb{V}^k} D\Big(\Phi^{(\rho)}\Big(x^{(1)}\Big),\ldots,\Phi^{(\rho)}\Big(x^{(k)}\Big)\Big)\, \mathrm{d}\nu^{d\times k}.
$$

Substituting this expression back into (6.1) yields

$$
Z(G_\rho,\lambda_\rho) = 1 + \sum_{k\in\mathbb{N}_{>0}} \frac{\lambda^k}{k!} \int_{\mathbb{V}^k} D\Big(\Phi^{(\rho)}\Big(x^{(1)}\Big),\ldots,\Phi^{(\rho)}\Big(x^{(k)}\Big)\Big)\, \mathrm{d}\nu^{d\times k}.
$$

We now express $|Z(\mathbb{V},\lambda) - Z(G_\rho,\lambda_\rho)|$ in terms of the absolute difference of the integrals for all $k\in\mathbb{N}_{>0}$. It holds that

$$
\left| \int_{\mathbb{V}^k} D\Big(x^{(1)},\ldots,x^{(k)}\Big)\, \mathrm{d}\nu^{d\times k} - \int_{\mathbb{V}^k} D\Big(\Phi^{(\rho)}\Big(x^{(1)}\Big),\ldots,\Phi^{(\rho)}\Big(x^{(k)}\Big)\Big)\, \mathrm{d}\nu^{d\times k} \right|
$$

$$
\leq \int_{\mathbb{V}^k} \left| D\Big(x^{(1)},\ldots,x^{(k)}\Big) - D\Big(\Phi^{(\rho)}\Big(x^{(1)}\Big),\ldots,\Phi^{(\rho)}\Big(x^{(k)}\Big)\Big) \right| \mathrm{d}\nu^{d\times k}.
$$

Let $N^{(\rho)} \subseteq \mathbb{V}^k$ be such that for all $(x^{(1)},\ldots,x^{(k)}) \in N^{(\rho)}$ we have that

$$
D\Big(x^{(1)},\ldots,x^{(k)}\Big) \neq D\Big(\Phi^{(\rho)}\Big(x^{(1)}\Big),\ldots,\Phi^{(\rho)}\Big(x^{(k)}\Big)\Big).
$$

As D is an indicator function, it holds that

$$
\int_{\mathbb{V}^k} \left| D\Big(x^{(1)},\ldots,x^{(k)}\Big) - D\Big(\Phi^{(\rho)}\Big(x^{(1)}\Big),\ldots,\Phi^{(\rho)}\Big(x^{(k)}\Big)\Big) \right| \mathrm{d}\nu^{d\times k} = \nu^{d\times k}\Big(N^{(\rho)}\Big).
$$

We construct a superset of $N^{(\rho)}$, for which we calculate the Lebesgue measure. First, note that $N^{(\rho)} = \emptyset$ for $k = 1$, as in this case $D(x^{(1)}) = D(\Phi^{(\rho)}(x^{(1)})) = 1$ for all $x^{(1)} \in \mathbb{V}$. Further, let $K = \big(\ell\sqrt{d}/(2r)\big)^d$. Note that, for all $k > K$, it holds that at least two particles have distance less than $2r$, meaning that such a configuration has always overlapping particles and $N^{(\rho)} = \emptyset$. We are left with considering $2 \leq k \leq K$.

We observe that, for all $(x^{(1)},\ldots,x^{(k)}) \in \mathbb{V}^k$ such that

$$
D\Big(x^{(1)},\ldots,x^{(k)}\Big) \neq D\Big(\Phi^{(\rho)}\Big(x^{(1)}\Big),\ldots,\Phi^{(\rho)}\Big(x^{(k)}\Big)\Big),
$$

there is a pair of points $x^{(i)}, x^{(j)}$ for $i,j \in [k]$ such that $i \neq j$ and

$$
d\Big(x^{(i)},x^{(j)}\Big) < 2r \leq d\Big(\Phi^{(\rho)}\Big(x^{(i)}\Big),\Phi^{(\rho)}\Big(x^{(j)}\Big)\Big) \ \text{ or }
$$

$$d\Big(x^{(i)}, x^{(j)}\Big) \geq 2r > d\Big(\Phi^{(\rho)}\Big(x^{(i)}\Big), \Phi^{(\rho)}\Big(x^{(j)}\Big)\Big).$$

As for every point $x^{(i)} \in \mathbb{V}$, it holds that

$$d\Big(x^{(i)}, \Phi^{(\rho)}\Big(x^{(i)}\Big)\Big) \leq \frac{\sqrt{d}}{\rho},$$

there is a pair of points $x^{(i)}, x^{(j)}$ for $i, j \in [k]$ such that $i \neq j$ and

$$\left|2r - d\Big(x^{(i)}, x^{(j)}\Big)\right| \leq 2\frac{\sqrt{d}}{\rho}.$$

For all $i, j \in [k]$ with $i \neq j$ let $S_{i,j}^{(\rho)} \subseteq \mathbb{V}^k$ be the set of points $\big(x^{(1)}, \ldots, x^{(k)}\big) \in \mathbb{V}^k$ such that this is the case. Then

$$\nu^{d \times k}\Big(N^{(\rho)}\Big) \leq \nu^{d \times k}\left(\bigcup_{1 \leq i < j \leq k} S_{i,j}^{(\rho)}\right) \leq \sum_{1 \leq i < j \leq k} \nu^{d \times k}\Big(S_{i,j}^{(\rho)}\Big).$$

By Fubini's theorem, noting that $S_{i,j}^{(\rho)}$ only depends on i and j, we get

$$\nu^{d \times k}\Big(S_{i,j}^{(\rho)}\Big) = \int_{\mathbb{V}^k} \mathbb{1}\left\{\left|2r - d\Big(x^{(i)}, x^{(j)}\Big)\right| \leq 2\frac{\sqrt{d}}{\rho}\right\} \mathrm{d}\nu^{d \times k}$$

$$= \ell^{d(k-2)} \int_{\mathbb{V}^2} \mathbb{1}\left\{\left|2r - d\Big(x^{(i)}, x^{(j)}\Big)\right| \leq 2\frac{\sqrt{d}}{\rho}\right\} \mathrm{d}\nu^{d \times 2}$$

$$\leq \ell^{d(k-1)} \cdot \left(\left(2r + 2\frac{\sqrt{d}}{\rho}\right)^d - \left(2r - 2\frac{\sqrt{d}}{\rho}\right)^d\right),$$

where the last equality comes from the fact that r was chosen as the radius of a ball of volume 1 in d dimensions. By the assumption $\rho \geq 2\sqrt{d}$ and the binomial theorem, we further bound

$$\left(2r + 2\frac{\sqrt{d}}{\rho}\right)^d - \left(2r - 2\frac{\sqrt{d}}{\rho}\right)^d = \sum_{i=0}^{d} 2 \cdot \mathbb{1}\{i \text{ is odd}\}\binom{d}{i}(2r)^{d-i}\left(2\frac{\sqrt{d}}{\rho}\right)^i$$

$$= \frac{2\sqrt{d}}{\rho}\sum_{i=1}^{d} 2 \cdot \mathbb{1}\{i \text{ is odd}\}\binom{d}{i}(2r)^{d-i}\left(2\frac{\sqrt{d}}{\rho}\right)^{i-1}$$

$$\leq \frac{2\sqrt{d}}{\rho}\sum_{i=1}^{d} 2 \cdot \mathbb{1}\{i \text{ is odd}\}\binom{d}{i}(2r)^{d-i}1^{i-1}$$

$$\leq 2\frac{2\sqrt{d}}{\rho}(2r + 1)^d.$$

Using this bound for $\nu^{d \times k}(S_{i,j}^{(\rho)})$, we obtain

$$\nu^{d \times k}\Big(N^{(\rho)}\Big) \leq k^2 \cdot 2 \cdot \ell^{d(k-1)} \cdot \frac{2\sqrt{d}}{\rho} \cdot (2r + 1)^d.$$

Thus, we get

$$|Z(\mathbb{V}, \lambda) - Z(G_\rho, \lambda_\rho)| \leq \sum_{k=2}^{K} \frac{\lambda^k}{k!} \nu^{d \times k}\left(N^{(\rho)}\right)$$

$$\leq \frac{1}{\rho} \sum_{k=2}^{K} \frac{\lambda^k}{k!} k^2 \cdot 4\ell^{d(k-1)} \cdot \sqrt{d} \cdot (2r+1)^d.$$

We simplify the bound further by

$$\frac{1}{\rho} \sum_{k=2}^{K} \frac{\lambda^k}{k!} k^2 \cdot 4\ell^{d(k-1)} \cdot \sqrt{d} \cdot (2r+1)^d \leq \frac{1}{\rho} K^2 \cdot 4\ell^{d(K-1)} \cdot \sqrt{d} \cdot (2r+1)^d \sum_{k=2}^{K} \frac{\lambda^k}{k!}$$

$$\leq \frac{1}{\rho} K^2 \cdot 4\ell^{d(K-1)} \cdot \sqrt{d} \cdot (2r+1)^d \mathrm{e}^\lambda,$$

where the last inequality follows from the Taylor expansion of e^x at 0.

Overall, we bound

$$|Z(\mathbb{V}, \lambda) - Z(G_\rho, \lambda_\rho)| \leq \frac{1}{\rho} K^2 \cdot 4\ell^{d(K-1)} \cdot \sqrt{d} \cdot (2r+1)^d \mathrm{e}^\lambda$$

$$\leq \frac{1}{\rho} \mathrm{e}^{\Theta\left(Kd\ln(\ell) + \ln(r+1) + \mathrm{e}^\lambda\right)}.$$

Observe that $r \in \mathrm{O}(1)$ and $\mathrm{e}^\lambda \in \mathrm{O}(1)$. Further, for $r = (1/v_d)^{1/d}$ it holds that $K \in \mathrm{O}\left(\ell^d\right)$. Thus we have

$$|Z(\mathbb{V}, \lambda) - Z(G_\rho, \lambda_\rho)| \leq \frac{1}{\rho} \mathrm{e}^{\Theta\left(\ell^d \ln(\ell^d)\right)} = \frac{1}{\rho} \mathrm{e}^{\Theta(|\mathbb{V}| \ln(|\mathbb{V}|))},$$

which concludes the proof. □

6.2. Approximation bound. We aim for applying Corollary 4.3 to the hard-core representation of the hard-sphere model. In order to do so, we need a bound on the maximum degree Δ_ρ of the graph G_ρ for any sufficiently large resolution ρ. Let $b_d(s)$ denote the number of integer grid points in a d-dimensional sphere of radius s centered at the origin. Note that the number of neighbors of a vertex $v_x \in V_\rho$ for any grid point $x \in \mathbb{G}(\rho\ell)$ is upper bounded by $b_d(2\rho r)$. We use the following bound on b_d.

LEMMA 6.2. *Let $\gamma \in (0, 1]$ and $s \in \mathbb{R}_{>0}$. For all $\rho \geq \left(2\sqrt{d}\right)^d/(\gamma s)$ it holds that $b_d(\rho s) \leq (1 + \gamma) \cdot v_d \cdot (\rho s)^d$.*

Proof. We start by considering a sphere of radius $\rho s + \sqrt{d}$ at the origin. Note that this enlarged sphere contains for each grid point $(x_1, \ldots, x_d)$ in the original sphere the cubic region $[x_1, x_1 + 1] \times \cdots \times [x_d, x_d + 1]$ of volume 1. Thus, the volume of the enlarged sphere is a trivial upper bound on the number of grid points in the original sphere.

Formally, we get

$$b_d(\rho s) \leq v_d \cdot \left(\rho s + \sqrt{d}\right)^d,$$

which we rewrite as

$$v_d \cdot \left(\rho s + \sqrt{d}\right)^d = v_d \cdot (\rho s)^d + v_d \cdot \sum_{i \in [d]} \binom{d}{i} (\rho s)^{d-i} \sqrt{d}^i.$$

Further, note that for our choice of ρ it holds that $\rho s \geq 1$. Thus, we get

$$v_d \cdot (\rho s)^d + v_d \cdot \sum_{i \in [d]} \binom{d}{i} (\rho s)^{d-i} \sqrt{d}^i \leq v_d \cdot (\rho s)^d + v_d \cdot (\rho s)^{d-1} \cdot 2^d \sqrt{d}^d$$

$$= v_d \cdot (\rho s)^d \cdot \left(1 + \frac{1}{\rho s}\left(2\sqrt{d}\right)^d\right).$$

We conclude the proof by noting that $\left(2\sqrt{d}\right)^d / \rho s \leq \gamma$. $\qquad\square$

As we fixed $r = (1/v_d)^{1/d}$ we can immediately conclude that for every $\gamma \in (0,1]$ there is some $\rho_\gamma \in \Theta(1/\gamma)$ such that for all $\rho \geq \rho_\gamma$ it holds that

$$\Delta_\rho \leq (1+\gamma)(2\rho)^d.$$

Finally, the following general lemma will help us to turn a sampling scheme for $\mu^{(G_\rho, \lambda_\rho)}$ into a randomized approximation of $Z(G_\rho, \lambda_\rho)$.

LEMMA 6.3 ([22, Lemma 13]). *Let (G, λ) be an instance of the multivariate hard-core model and let Λ be a clique cover of size m with $Z_{\max} = \max_{i \in [m]}\{Z(G_\rho[\Lambda_i])\}$. Further, for $i \in [m]$ let $V_i = V \setminus \bigcup_{j<i} \Lambda_j$. For every $\varepsilon \in (0,1]$ there are $s \in \Theta(m Z_{\max}/\varepsilon^2)$ and $\varepsilon_s \in \Theta(\varepsilon/(m Z_{\max}))$ such that a randomized ε-approximation of $Z(G, \lambda)$ can be computed by drawing s samples ε_s-approximately from $\mu^{(G[V_i])}$ for each $i \in [m]$.*

Note that sampling from $\mu^{(G[V_i])}$ means sampling from $\mu^{(G)}$ for $i = 0$ and ignoring all cliques $\{\Lambda_j\}_{j<i}$ for $i \geq 1$.

THEOREM 1.1. *Let $(\mathbb{V}, \lambda)$ be an instance of the continuous hard-sphere model with $\mathbb{V} = [0, \ell]^d$. If there is a $\delta \in (0,1]$ such that*

$$\lambda \leq (1-\delta)\frac{e}{2^d},$$

then for each $\varepsilon \in (0,1]$ there is a randomized ε-approximation of $Z(\mathbb{V}, \lambda)$ computable in time polynomial in $|\mathbb{V}|^{1/\delta^2}$ and $\frac{1}{\varepsilon}$.

Proof. Set $\gamma = \delta/2$ and $\varepsilon' = \varepsilon/3$. By combining Lemmas 1.2 and 6.2, we know that we can choose a resolution $\rho \in \Theta\left(e^{|\mathbb{V}| \ln |\mathbb{V}|}/(\varepsilon' \gamma)\right) = \Theta\left(e^{|\mathbb{V}| \ln |\mathbb{V}|}/(\varepsilon \gamma)\right)$ such that

$$(6.2) \qquad 1 - \varepsilon' \leq \frac{Z(\mathbb{V}, \lambda)}{Z(G_\rho, \lambda_\rho)} \leq 1 + \varepsilon' \text{ and}$$

$$(6.3) \qquad \Delta_\rho \leq (1+\gamma)(2\rho)^d.$$

Note that $(1-\varepsilon')^2 \geq 1-\varepsilon$ and $(1+\varepsilon')^2 \leq 1+\varepsilon$. Thus, (6.2) implies that it is sufficient to ε'-approximate $Z(G_\rho, \lambda_\rho)$. We start by arguing that we can apply Corollary 4.3 to $Z(G_\rho, \lambda_\rho)$. Then, we construct a disjoint clique cover and show that each step of the block dynamics based on that cover can be computed efficiently. Finally, we will use Lemma 6.3 to get the desired approximation.

To apply Corollary 4.3, we need to show that $\lambda_\rho \leq (1-\delta')\lambda_c(\Delta_\rho)$ for some $\delta' \in (0,1]$. To this end, we choose $\delta' = \delta/2$. Due to (6.3) we know that

$$\lambda_c(\Delta_\rho) = \frac{(\Delta_\rho - 1)^{\Delta_\rho - 1}}{(\Delta_\rho - 2)^{\Delta_\rho}} \geq \frac{\left((1+\gamma)(2\rho)^d - 1\right)^{(1+\gamma)(2\rho)^d - 1}}{\left((1+\gamma)(2\rho)^d - 2\right)^{(1+\gamma)(2\rho)^d}}.$$

Now, note that

$$\lambda \leq (1-\delta)\frac{e}{2^d} \leq \frac{1-\frac{\delta}{2}}{1+\frac{\delta}{2}}\frac{e}{2^d} = \frac{1-\delta'}{1+\gamma}\frac{e}{2^d} \leq (1-\delta')\rho^d \frac{\left(\left((1+\gamma)(2\rho)^d-1\right)\right)^{(1+\gamma)(2\rho)^d-1}}{\left(\left((1+\gamma)(2\rho)^d-2\right)\right)^{(1+\gamma)(2\rho)^d}},$$

where the last inequality comes from the fact that $x\frac{(x-1)^{x-1}}{(x-2)^x}$ converges to e from above as $x \to \infty$. Dividing by ρ^d yields $\lambda_\rho = \rho^{-d}\lambda \leq (1-\delta')\lambda_c(\Delta_\rho)$.

We now construct the disjoint clique cover that we are going to use. This is done by dividing the grid $\mathbb{G}(\rho\ell)$ into cubic regions of side length $a = \lfloor\frac{2\rho}{\sqrt{d}}v_d^{-1/d}\rfloor$. Formally, for a tuple $(i_1, \dots, i_d) \in \mathbb{N}^d$, let

$$\mathbb{H}_{i_1,\dots,i_d} = \{(x_1,\dots,x_d) \in \mathbb{G} \mid \forall j \in [d]\colon i_j a \leq x_j < (i_j+1)a\}.$$

Note that for every pair of grid points $x, y \in \mathbb{H}_{i_1,\dots,i_d}$ it holds that $d(x,y) < 2\rho v_d^{-1/d} = 2\rho r$. Thus, the set of vertices, corresponding to grid points in $\mathbb{H}_{i_1,\dots,i_d}$, form a clique in G_ρ. We obtain a clique cover Λ of size $m = |\Lambda| \in O\left((\rho\ell/a)^d\right) = O(|\mathbb{V}|)$. Further, it holds that

$$Z_{\max} \leq 1 + a^d\lambda_\rho = 1 + a^d\rho^{-d}\lambda \in O(1).$$

By Corollary 4.3, the block dynamics based on Λ have mixing time that is logarithmic in $1/\varepsilon_s$ for any sampling error $\varepsilon_s \in (0,1]$ and polynomial in $|\Lambda|$. The latter implies that the mixing time is also polynomials in $|\mathbb{V}|^{1/\delta^2}$.

We proceed by arguing that we can compute each step efficiently. Note that we cannot construct the graph explicitly, as it would be far too large for our choice of resolution ρ. However, by identifying each vertex by its corresponding grid point, deciding whether there is an edge between two vertices or if a vertex belongs to a certain clique can be done by comparing integers up to size $O(\rho\ell)$, which can be done in $O(\ln(\rho\ell)) = O(|\mathbb{V}|\ln|\mathbb{V}|)$. Choosing a clique from the clique cover can be done by choosing d integers up to size $O(\ell)$. Now, assume the current state of the block dynamics Markov chain is $I \in \mathcal{I}(G_\rho)$. For a given clique $\Lambda_i \in \Lambda$, we can sample from $\mu_{|\Lambda_i}(\ \cdot\ |\sigma_{|V\setminus\Lambda_i}^{(I)})$ by the following procedure.

(1) With probability $\frac{1}{Z(G_\rho[\Lambda_i])}$ set $I_+ = \emptyset$. Otherwise, draw $x \in \mathbb{H}_i$ uniformly at random, where $\mathbb{H}_i$ is the region of the grid corresponding to Λ_i, and set $I_+ = \{v_x\}$.

(2) If $I \cup I_+ \in \mathcal{I}(G_\rho)$, output I_+. Otherwise, restart from (1).

Simple calculations show that this rejection sampler results in the desired output distribution. Note that (1) involves computing $Z(G_\rho[\Lambda_i]) = 1+|\Lambda_i|\rho^{-d}\lambda$ and sampling d integers up to size $O(a)$, which can be done in $O(\ln(a)) = O(|\mathbb{V}|\ln|\mathbb{V}|)$. Moreover, we never reject I_+ in (2) if $I_+ = \emptyset$, which happens with probability at least $1/Z(G_\rho[\Lambda_i])$. Thus, the number of restarts is dominated by a geometric random variable with success probability $1/Z(G_\rho[\Lambda_i])$ and we almost surely require $O(Z(G_\rho[\Lambda_i])) = O(1)$ trials.

We now know that we can sample ε_s-approximately from $\mu^{(G_\rho, \lambda_\rho)}$ in time polynomial in $|\mathbb{V}|^{1/\delta^2}$ and $\ln(1/\varepsilon_s)$. Applying Lemma 6.3 proves the theorem. $\square$

REFERENCES

[1] V. L. Alev and L. C. Lau, *Improved analysis of higher order random walks and applications*, in Proceedings of STOC'20, 2020, pp. 1198–1211, https://doi.org/10.1145/3357713.3384317.

[2] N. Anari, V. Jain, F. Koehler, H. T. Pham, and T. Vuong, *Entropic Independence* II: *Optimal Sampling and Concentration via Restricted Modified Log-Sobolev Inequalities*, CoRR, abs/2111.03247, https://arxiv.org/abs/2111.03247, 2021.

[3] N. Anari, K. Liu, and S. O. Gharan, *Spectral independence in high-dimensional expanders and applications to the hardcore model*, SIAM J. Comput., (2021), pp. FOCS20–1–FOCS20–37, https://doi.org/10.1137/20M1367696.

[4] I. Bezáková, A. Galanis, L. A. Goldberg, and D. Štefankovič, *Inapproximability of the independent set polynomial in the complex plane*, in Proceedings of STOC'18, 2018, p. 1234–1240, https://doi.org/10.1145/3188745.3188788.

[5] A. Blanca, P. Caputo, A. Sinclair, and E. Vigoda, *Spatial mixing and nonlocal Markov chains*, Random Structures Algorithms, 55 (2019), pp. 584–614, https://doi.org/10.1002/rsa.20844.

[6] A. Blanca, Z. Chen, and E. Vigoda, *Swendsen-Wang dynamics for general graphs in the tree uniqueness region*, Random Structures Algorithms, 56 (2020), pp. 373–400, https://doi.org/10.1002/rsa.20858.

[7] T. Boublik, I. Nezbeda, and K. Hlavaty, *Statistical Thermodynamics of Simple Liquids and Their Mixtures*, Fundamental Studies in Engineering, Elsevier, Amsterdam, 1980.

[8] X. Chen, W. Feng, Y. Yin, and X. Zhang, *Rapid mixing of Glauber dynamics via spectral independence for all degrees*, in Proceedings of the 2021 IEEE 62nd Annual Symposium on Foundations of Computer Science, 2022, pp. 137–148, https://doi.org/10.1109/FOCS52979.2021.00022.

[9] Z. Chen, A. Galanis, D. Štefankovič, and E. Vigoda, *Rapid mixing for colorings via spectral independence*, in Proceedings of SODA'21, SIAM, Philadelphia, 2021, pp. 1548–1557.

[10] Z. Chen, K. Liu, and E. Vigoda, *Rapid mixing of Glauber dynamics up to uniqueness via contraction*, in Proceedings of FOCS20, 2020, pp. 1307–1318, https://doi.org/10.1109/FOCS46700.2020.00124.

[11] Z. Chen, K. Liu, and E. Vigoda, *Optimal mixing of Glauber dynamics: Entropy factorization via high-dimensional expansion*, in Proceedings of STOC'21, 2021, pp. 1537–1550, https://doi.org/10.1145/3406325.3451035.

[12] H.-T. Christoph et al., *Disagreement percolation for the hard-sphere model*, Electron. J. Probab., 24 (2019).

[13] M. Dyer, A. D. Flaxman, A. M. Frieze, and E. Vigoda, *Randomly coloring sparse random graphs with fewer colors than the maximum degree*, Random Structures Algorithms, 29 (2006), pp. 450–465, https://doi.org/10.1002/rsa.20129.

[14] M. Dyer, A. Frieze, and M. Jerrum, *On counting independent sets in sparse graphs*, SIAM J. Comput., 31 (2002), pp. 1527–1541, https://doi.org/10.1137/S0097539701383844.

[15] C. Efthymiou, *MCMC sampling colourings and independent sets of $g(n, d/n)$ near uniqueness threshold*, in Proceedings of SODA'14, 2014, pp. 305–316, https://doi.org/10.1137/1.9781611973402.22.

[16] C. Efthymiou, T. P. Hayes, D. Štefankovič, and E. Vigoda, *Sampling random colorings of sparse random graphs*, in Proceedings of SODA'18, 2018, pp. 1759–1771, https://doi.org/10.1137/1.9781611975031.115.

[17] W. Feng, H. Guo, Y. Yin, and C. Zhang, *Rapid mixing from spectral independence beyond the boolean domain*, in Proceedings of SODA'21, SIAM, Philadelphia, 2021, pp. 1558–1577.

[18] R. Fernández, A. Procacci, and B. Scoppola, *The analyticity region of the hard sphere gas. Improved bounds*, J. Stat. Phys., 128 (2007), pp. 1139–1143, https://doi.org/10.1007/s10955-007-9352-7.

[19] T. Friedrich, A. Göbel, M. Katzmann, M. Krejca, and M. Pappik, *Using Random Graphs to Sample Repulsive Gibbs Point Processes with Arbitrary-range Potentials*, preprint, arXiv:2204.01793, 2022.

[20] T. Friedrich, A. Göbel, M. Katzmann, M. S. Krejca, and M. Pappik, *Algorithms for Hard-constraint Point Processes via Discretization*, preprint, arXiv:2107.08848, 2021.

[21] T. Friedrich, A. Göbel, M. Krejca, and M. Pappik, *A spectral independence view on hard spheres via block dynamics*, in Proceedings of ICALP'21, 2021, pp. 66–1.

[22] T. Friedrich, A. Göbel, M. S. Krejca, and M. Pappik, *Polymer Dynamics via Cliques: New Conditions for Approximations*, CoRR, abs/2007.08293, http://arxiv.org/abs/2007.08293, 2020.

[23] A. Galanis, Q. Ge, D. Štefankovič, E. Vigoda, and L. Yang, *Improved inapproximability results for counting independent sets in the hard-core model*, Random Structures Algorithms, 45 (2014), pp. 78–110, https://doi.org/10.1002/rsa.20479.

[24] A. Galanis, L. A. Goldberg, and D. Štefankovič, *Inapproximability of the Independent Set Polynomial Below the Shearer Threshold*, CoRR, abs/1612.05832, http://arxiv.org/abs/1612.05832, 2016.

[25] D. Galvin and P. Tetali, *Slow mixing of Glauber dynamics for the hard-core model on regular bipartite graphs*, Random Structures Algorithms, 28 (2006), pp. 427–443, https://doi.org/10.1002/rsa.20094.

[26] H. Guo and M. Jerrum, *Perfect Simulation of the Hard Disks Model by Partial Rejection Sampling*, CoRR, abs/1801.07342, https://arxiv.org/abs/1801.07342v2, 2018.

[27] J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 4th ed., Academic Press, New York, 2013, https://doi.org/10.1016/B978-0-12-387032-2.00013-1.

[28] N. J. Harvey, P. Srivastava, and J. Vondrák, *Computing the independence polynomial: From the tree threshold down to the roots*, in Proceedings of SODA'18, 2018, pp. 1557–1576, https://doi.org/10.1137/1.9781611975031.102.

[29] T. P. Hayes and C. Moore, *Lower Bounds on the Critical Density in the Hard Disk Model via Optimized Metrics*, CoRR, abs/1407.1930, http://arxiv.org/abs/1407.1930, 2014.

[30] T. Helmuth, W. Perkins, and S. Petti, *Correlation Decay for Hard Spheres via Markov Chains*, CoRR, abs/2001.05323, https://arxiv.org/abs/2001.05323, 2020.

[31] M. Jenssen, F. Joos, and W. Perkins, *On the hard sphere model and sphere packings in high dimensions*, Forum Math. Sigma, 7 (2019), https://doi.org/10.1017/fms.2018.25.

[32] R. Kannan, M. W. Mahoney, and R. Montenegro, *Rapid mixing of several Markov chains for a hard-core model*, in Proceedings of ISAAC'03, 2003, pp. 663–675, https://doi.org/10.1007/978-3-540-24587-2_68.

[33] W. S. Kendall, *Perfect simulation for the area-interaction point process*, in Probability Towards 2000, Springer, New York, 1998, pp. 218–234, https://doi.org/10.1007/978-1-4612-2224-8_13.

[34] W. S. Kendall and J. Møller, *Perfect simulation using dominating processes on ordered spaces, with application to locally stable point processes*, Adv. Appl. Probab., 32 (2000), pp. 844–865, https://doi.org/10.1239/aap/1013540247.

[35] D. A. Levin and Y. Peres, *Markov Chains and Mixing Times*, AMS, Providence, RI, 2017.

[36] L. Li, P. Lu, and Y. Yin, *Correlation decay up to uniqueness in spin systems*, in Proceedings of SODA'13, 2013, pp. 67–84, https://doi.org/10.1137/1.9781611973105.5.

[37] F. Martinelli, *Lectures on Glauber dynamics for discrete spin models*, in Lectures on Probability Theory and Statistics, Springer, New York, 1999, pp. 93–191, https://doi.org/10.1007/978-3-540-48115-7_2.

[38] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *Equation of state calculations by fast computing machines*, J. Chemical Physics, 21 (1953), pp. 1087–1092, https://doi.org/10.1063/1.1699114.

[39] M. Michelen and W. Perkins, *Analyticity for Classical Gasses via Recursion*, CoRR, abs/2008.00972, http://arxiv.org/abs/2008.00972, 2020.

[40] M. Michelen and W. Perkins, *Potential-Weighted Connective Constants and Uniqueness of Gibbs Measures*, CoRR, abs/2109.01094, https://arxiv.org/abs/2109.01094, 2021.

[41] M. Michelen and W. Perkins, *Strong Spatial Mixing for Repulsive Point Processes*, CoRR, abs/2202.08753, https://arxiv.org/abs/2202.08753, 2022.

[42] S. B. Moka, S. Juneja, and M. R. H. Mandjes, *Analysis of perfect sampling methods for hard-sphere models*, SIGMETRICS Performance Evaluation Review, 45 (2018), p. 69–75, https://doi.org/10.1145/3199524.3199536.

[43] E. Mossel and A. Sly, *Gibbs rapidly samples colorings of $g(n, d/n)$*, Probab. Theory Related Fields, 148 (2010), pp. 37–69, https://doi.org/10.1007/s00440-009-0222-x.

[44] E. Mossel, D. Weitz, and N. Wormald, *On the hardness of sampling independent sets beyond the tree threshold*, Probab. Theory Related Fields, 143 (2009), pp. 401–439, https://doi.org/10.1007/s00440-007-0131-9.

[45] V. Patel and G. Regts, *Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials*, SIAM J. Comput., 46 (2017), pp. 1893–1919, https://doi.org/10.1137/16M1101003.

[46] M. D. Penrose, *Self-avoiding walks and trees in spread-out lattices*, J. Stat. Phys., 77 (1994), pp. 3–15, https://doi.org/10.1007/BF02186829.

[47] A. Sinclair, *Improved bounds for mixing rates of Markov chains and multicommodity flow*, Combin. Probab. Comput., 1 (1992), pp. 351–370, https://doi.org/10.1017/S0963548300000390.

[48] A. Sinclair, P. Srivastava, D. Štefankovič, and Y. Yin, *Spatial mixing and the connective constant: Optimal bounds*, Probab. Theory Related Fields, 168 (2017), pp. 153–197, https://doi.org/10.1007/s00440-016-0708-2.

[49] A. Sly, *Computational transition at the uniqueness threshold*, in Proceedings of FOCS'10, 2010, pp. 287–296, https://doi.org/10.1109/FOCS.2010.34.

[50] J. van den Berg and R. Brouwer, *Random sampling for the Monomer–Dimer model on a lattice*, J. Math. Phys., 41 (2000), pp. 1585–1597, https://doi.org/10.1063/1.533198.

[51] D. Weitz, *Combinatorial criteria for uniqueness of Gibbs measures*, Random Structures Algorithms, 27 (2005), pp. 445–475.

[52] D. Weitz, *Counting independent sets up to the tree threshold*, in Proceedings of STOC'06, 2006, pp. 140–149, https://doi.org/10.1145/1132516.1132538.

[53] B. Widom and J. S. Rowlinson, *New model for the study of liquid-vapor phase transitions*, J. Chem. Phys., 52 (1970), pp. 1670–1684, https://doi.org/10.1063/1.1673203.