Submitted to the Annals of Applied Probability
arXiv: arXiv:2001.05323

CORRELATION DECAY FOR HARD SPHERES VIA MARKOV CHAINS

By Tyler Helmuth*,
Will Perkins†,
and Samantha Petti‡

Durham University
University of Illinois at Chicago
Harvard University

We improve upon all known lower bounds on the critical fugacity and critical density of the hard sphere model in dimensions two and higher. As the dimension tends to infinity our improvements are by factors of 2 and 1.7, respectively. We make these improvements by utilizing techniques from theoretical computer science to show that a certain Markov chain for sampling from the hard sphere model mixes rapidly at low enough fugacities. We then prove an equivalence between optimal spatial and temporal mixing for hard spheres to deduce our results.

1. Introduction. For a fixed radius $r > 0$, the *hard sphere model* in a volume $\Lambda \subset \mathbb{R}^d$ at fugacity $\lambda \geq 0$ is a random point process $\mathbf{X}$ defined by conditioning a Poisson point process of intensity λ on Λ on the event that the points are at pairwise distance at least $2r$, or in other words, conditioning on the event that the points $\mathbf{X}$ are the centers of a packing of spheres of radius r. Conditioned on the number k of centers, the distribution is uniform over all sphere packings of Λ with k spheres. Note that by a 'sphere packing' we simply mean a configuration of non-overlapping spheres, not a 'close packing' which implies maximality.

The hard sphere model is a simple but fundamental model of monatomic gases. Its theoretical importance is in part due to the fact that it conjecturally possesses a crystalline phase [1, 5, 19]. Understanding the phase diagram of the model has presented a significant challenge even at the level of physics [3], and mathematical results are almost exclusively restricted to understanding the low-density (small λ) phase (see [30] for a notable exception). In particular, it is an open mathematical problem to prove the

*Supported in part by EPSRC grant EP/P003656/1.
†Supported in part by NSF Career award DMS-1847451.
‡Supported in part by NSF Graduate Research Fellowship DGE-1650044.
MSC 2010 subject classifications: Primary 82B21; secondary 60J05
Keywords and phrases: hard sphere model, Markov chains, correlation decay

existence of a phase transition in the hard sphere model. Not only is the model the most studied example of a broad class of Gibbs point processes, but it has played a starring role in the development of Markov chain Monte Carlo methods since the beginning: the Metropolis algorithm was first applied to the study of the two-dimensional hard sphere model [26].

We will give a more precise definition of the hard sphere model below, but for now we restrict our attention to aspects of the problem directly relevant to our results. The reader unfamiliar with the model may find [25] to be an inspiring introduction and broader overview. Without loss of generality, it will be convenient to choose the radius $r = r_d$ such that each sphere has volume one and $|\mathbf{X}|$, the number of spheres in the random packing, is also the volume covered by the packing.

The *critical fugacity* $\lambda_c(d)$ in dimension d is the supremum over λ such that the hard sphere model has a unique infinite volume limit in the sense of van Hove, i.e., such that the set of weak limit points of $\{\mu_{\Lambda,\lambda}\}_\Lambda$ as $\Lambda \to \mathbb{R}^d$ is a singleton set. If $\lambda_c(d) < \infty$, then the hard sphere model exhibits a *phase transition* at fugacity λ_c. When $d = 1$, $\lambda_c(d) = \infty$, but it is not known in any higher dimension whether or not $\lambda_c(d) < \infty$. It is believed that $\lambda_c(d)$ is finite in dimension 3 (and in some or all dimensions $d \geq 4$), while the case $d = 2$ is subtle and remains an active area of investigation even in the physics literature [3, 35].

Proving a lower bound on $\lambda_c(d)$ amounts to proving the absence of a phase transition for λ in an interval on the real line. Developing new and more powerful methods for proving the absence of a phase transition has been a central theme of statistical mechanics (e.g. [43, 10, 36, 7]), not only due to the interest in understanding phase diagrams but also because of a broad equivalence between the absence of phase transitions and important probabilistic and dynamical properties of finite and infinite systems [11]. In particular, the absence of a phase transition is related to mixing properties of natural dynamics to sample from finite volume Gibbs measures, i.e., the thermodynamics properties are connected to the performance of natural algorithms to sample from finite volume systems. While this connection has been made precise in the setting of lattice systems [34, 12, 38], it has only been shown under more restrictive conditions in the setting of Gibbs point processes like the hard sphere model (see the discussion below in Section 1.1).

Our main result is an improved lower bound on the critical fugacity $\lambda_c(d)$ that we prove by developing an unrestricted equivalence between finite volume Markov-chain mixing properties and correlation decay properties.

THEOREM 1.1. *For all $d \geq 2$, $\lambda_c(d) \geq 2^{1-d}$.*

The *density* of the hard sphere model on $\mathbb{R}^d$ at fugacity λ is

$$\rho(\lambda) = \liminf_{n\to\infty} \frac{1}{n}\mathbb{E}_{Q_n,\lambda}|\mathbf{X}|\,, \tag{1.1}$$

where Q_n is the d-dimensional cube of volume n centered at the origin, and the expectation is with respect to the hard sphere model on Q_n at fugacity λ. The use of liminf in (1.1) is necessary as *a priori* the limit is only known to exist for Lebesgue-a.e. value of λ. The *critical density* $\rho_c(d)$ of the hard sphere model is $\rho(\lambda_c(d))$ (or $\lim_{\lambda\to\infty}\rho(\lambda)$ if $\lambda_c = \infty$). That is, $\rho_c(d)$ is the limiting expected packing density of the hard sphere model at the critical fugacity $\lambda_c(d)$. By making use of Theorem 1.1 we can obtain an improved lower bound on the critical density.

THEOREM 1.2. *For all $d \geq 2$, $\rho_c(d) \geq \frac{2}{3\cdot 2^d}$. As the dimension d tends to infinity we have $\rho_c(d) \geq (.8526 + o_d(1))2^{-d}$.*

Before discussing how these results improve upon all previously known bounds, we briefly outline the proof of Theorems 1.1 and 1.2. At a high level this is done by adapting and combining three ingredients from the study of algorithms, probability theory, and combinatorics:

1. We analyze a Markov chain for sampling hard sphere configurations in a finite volume. By using techniques from theoretical computer science, namely path coupling with an optimized metric (e.g. [37]), we show that this Markov chain mixes rapidly at small enough fugacity. The conclusion is Theorem 3.2 below. Our main contribution here is to rigorously implement the idea that the hard sphere model behaves like a hard-core lattice gas on a finite graph with many triangles.
2. We establish a continuous analogue of the equivalence of spatial and temporal mixing from lattice spin systems (e.g. [34, 12]) to deduce exponential decay of correlations from our fast mixing results. For applications to bounds on the critical fugacity of the hard sphere model our main result here is Theorem 2.2. We also prove a full equivalence between spatial and temporal mixing, see Theorems 2.3 and 2.4. Our main contribution here is to prove an equivalence result for hard spheres that does not rely on the convergence of a cluster expansion; the importance of this is discussed below in Section 1.1.
3. We achieve the bounds on the critical density ρ_c in Theorem 1.2 by applying non-trivial lower bounds on the expected packing density of the hard sphere model [20].

After the presentation of some preliminaries in Section 2.1, we outline the first two of these steps in more detail in Sections 2.2 and 2.3. Before this we compare our results with those previously obtained in the literature.

1.1. *Previous results.* Historically, the main approach to proving the absence of a phase transition at low densities in the hard sphere model is to use the cluster expansion. This is a convergent power series expansion for the pressure of the hard sphere gas. The classical bound states that the cluster expansion converges for all complex λ with $|\lambda| \leq e^{-1}2^{-d}$, and thus $\lambda_c(d) > e^{-1}2^{-d}$. For a given dimension it may be possible to improve upon this result, e.g., in two dimensions, Fernández, Procacci, and Scoppola [13], proved the cluster expansion converges for $|\lambda| \leq .1277$. However, one does not expect to be able to improve the constant e^{-1} as the dimension d tends to infinity: the singularity determining the radius of convergence of the cluster expansion is known to lie on the negative real axis, and there is a compelling (but non-rigorous) argument that this singularity is at $-e^{-1}2^{-d}(1 + o_d(1))$, see [15]. It is known rigorously that the cluster expansion cannot have a radius of convergence greater than 2^{-d}, see [28, Remark 3.7].

To avoid the negative axis singularity a natural idea is to use techniques that do not require analyticity, e.g., probabilistic techniques which concern only positive fugacities λ. One approach in this direction was taken by Hofer-Temmel [18], who used disagreement percolation [36] and known bounds on the critical intensity of d-dimensional Poisson–Boolean percolation to prove lower bounds on the critical fugacity of the hard sphere model. In dimension 2, his bound is $\lambda_c(2) > .1367$. Hofer-Temmel's method and a bound based on the non-rigorous 'high-confidence' results of [2] for Poisson-Boolean percolation gives $\lambda_c(2) > .28175$ [18]. The asymptotics of the critical intensity of Poisson–Boolean percolation as $d \to \infty$ are known, and this gives a bound of $\lambda_c(d) \geq (1 + o(1))2^{-d}$, improving upon the cluster expansion bound by a factor e. Our probabilistic approach makes use of Markov chain mixing times, a well-honed tool, instead of disagreement percolation. For comparison, our bound $\lambda_c(d) \geq 2^{1-d}$ is an improvement of a factor 2 as $d \to \infty$, and of more than 3 compared to the rigorous results in dimension 2.

Finally, there has been work on developing exact sampling algorithms for the hard sphere model. Guo and Jerrum [16] showed that a partial rejection sampling algorithm is efficient in dimension 2 for $\lambda \leq .21027$ and Wellens improved this bound to $\lambda \leq .2344$ [40]. For an approach combining coupling from the past and rejection sampling, see [41].

While it appears we are the first to use Markov chains to estimate the crit-

ical fugacity of the hard sphere model, there have been previous works that obtain bounds on the critical density by showing that certain Markov chains for sampling a configuration of hard spheres mix rapidly. To lower bound the critical density these chains make use of the canonical ensemble, meaning the configurations consist of a fixed finite number of spheres in a finite volume. Results of this type include Kannan, Mahoney, and Montenegro who showed that a simple Markov chain for the canonical ensemble exhibits rapid mixing for densities $\rho < 2^{-1-d}$ [22], and Hayes and Moore who used an optimized metric to show that in dimension 2 this same Markov chain mixes rapidly at densities $\rho < .154$ [17]. The Markov chain studied in these papers moves spheres in a non-local way. Dynamics involving only local moves have been investigated by Diaconis, Lebeau and Michel as an application of a more general geometric framework [9]; these local dynamics are restricted to vanishing densities due to the existence of jammed configurations of arbitrarily low density, see [21].

To convert bounds on the critical fugacity to bounds on the critical density, we use the bound $\rho(\lambda) > \frac{\lambda}{1+2^d\lambda}$ where $\rho(\lambda)$ is the packing density of the hard sphere model at fugacity λ (see Section 5). Surprisingly, the estimates on ρ_c that result from combining this estimate with the high-dimensional bounds from [18] coincide with the high-dimensional bounds on ρ_c from [22]. Our results improve upon this by a factor of $4/3$. Our stronger result as $d \to \infty$ in Theorem 1.2, which is an improvement of roughly 1.7, makes use of a better estimate for $\rho(\lambda)$, see Section 5. Our results are also the best known in $d = 2$.

As mentioned above, our argument showing that rapid mixing implies exponential decay of correlations is based on the argument given in [12] for discrete spin systems on graphs. Previously there has been work relating mixing times and correlation decay for continuous-time birth-death chains for continuum particle systems with soft two-body potentials [4]. Later works allowed for hard core potentials [42, 6], but all of these results apply only in the low density regime, i.e., within the domain of convergence of the cluster expansion. For the reasons discussed above, it is essential for us to have a result that does not rely on the convergence of the cluster expansion. We achieve this by using combinatorial techniques in our proof of equivalence to avoid a low density hypothesis.

1.2. *Future directions.* One interesting benchmark for further progress on determining the uniqueness phase of the hard sphere model would be to obtain uniqueness for all $\rho \leq 2^{-d}$, the point at which the system no longer trivially (by a union bound) contains free volume. Passing this threshold ap-

pears to require new ideas. Another tool from computer science that may be applicable to the hard sphere model is Weitz's correlation decay method [39], although some adaptation will be necessary to deal with the continuous nature of space for the hard sphere model.[1]

While it should be straightforward to generalize our methods to more hard shapes that are not spheres, the adaptation to more general (soft) two-body potentials is less clear. A generalization that covers stable two-body potentials (the setting of, e.g., [29, 32]) would be interesting.

It is also worth remarking that the step in our arguments of obtaining density bounds from fugacity bounds is a challenging and interesting problem in itself. At very low fugacities one can use the cluster expansion to write a convergent power series formula for the density. Since, however, we are interested in going beyond the radius of convergence of the cluster expansion, this tool is not available. Developing alternative approaches to estimating the density, whether by analytic continuation, direct study of the Virial series, or other means, would be very interesting.

Lastly, there is of course the long-standing open problem of proving the existence of a phase transition for the hard sphere model.

2. Spatial and temporal mixing.

In this section we define the hard sphere model with boundary conditions, define the notions of strong spatial mixing and optimal temporal mixing, and reformulate our main results in terms of these notions.

2.1. *Hard spheres with boundary conditions.*

We begin by formally defining the hard sphere model in a bounded measurable volume $\Lambda \subset \mathbb{R}^d$. Recall that we write $r = r_d$ for the radius of a sphere of volume 1 in $\mathbb{R}^d$. Let

$$\Lambda_{\text{Int}} = \{x \in \Lambda : \text{dist}(x, \Lambda^c) \geq r\}.$$

The hard sphere model on volume Λ at fugacity $\lambda \geq 0$ with *free boundary conditions* is a Poisson point process of intensity λ on Λ_{Int} conditioned on the event that all points are at pairwise distance at least $2r$. In words, the hard sphere model arises by conditioning on the event that the points form the centers of a sphere packing in Λ with spheres of volume 1; we recall a sphere packing in a set A is any collection of pairwise disjoint open spheres that are entirely contained in A. Explicitly, the normalizing constant $Z_\Lambda(\lambda)$

[1] Subsequent to the posting of this article to the arXiv, one possible adaptation of Weitz's method to continuum particle systems has been found [27].

is

$$Z_\Lambda(\lambda) = \sum_{k\geq 0} \int_{\Lambda_{\text{Int}}^k} \frac{\lambda^k}{k!} \prod_{1\leq i<j\leq k} 1_{\|x_i - x_j\|\geq 2r} \prod_{i=1}^k dx_i \tag{2.1}$$

where dx_i is Lebesgue measure. We will denote the law of $\mathbf{X}$ by μ_Λ (the dependence on λ will be suppressed). The density of μ_Λ on Λ_{Int}^k with respect to Lebesgue measure is given by the integrand of (2.1) divided by the partition function $Z_\Lambda(\lambda)$. Note that the requirement that spheres lie entirely within Λ instead of just requiring the centers to lie in Λ makes no difference in the infinite volume limit, but it does have a regularizing effect in finite volume.

We will also be interested in the hard sphere model with boundary conditions τ. More precisely, we define $\tau \subseteq \Lambda_{\text{Int}}$ as a set of forbidden locations for centers. The hard sphere model on a volume Λ at fugacity $\lambda \geq 0$ with boundary conditions τ is a Poisson point process of intensity λ on $\Lambda_{\text{Int}} \setminus \tau$ conditioned on the event that all points are at pairwise distance at least $2r$. One possibility is that τ represents the volume blocked by a set of permanently fixed spheres: if Y is a set of centers and $\tau = \Lambda_{\text{Int}} \cap (\cup_{y\in Y} B_{2r}(y))$, then $\Lambda_{\text{Int}} \setminus \tau$ is the set of locations for centers that do not overlap with spheres defined by the centers in Y. Note τ need not have this form. The law of the hard sphere model on Λ with boundary condition τ will be denoted by μ_Λ^τ.

2.2. *Strong spatial mixing.* Let Ω_Λ be the set of all *configurations* for the hard sphere model on Λ, that is, the set of all finite point sets in Λ_{Int} whose pairwise distance is at least $2r$. Similarly, let Ω_Λ^τ be the set of configurations for the hard sphere model on Λ with boundary conditions τ. In particular, $\Omega_\Lambda = \Omega_\Lambda^\emptyset$.

For two probability measures μ_1 and μ_2 on Ω_Λ we let $\|\mu_1 - \mu_2\| = \|\mu_1 - \mu_2\|_{TV}$ denote their total variation distance. For $\Lambda' \subseteq \Lambda$ probability measures on Ω_Λ induce probability measures on $\Omega_{\Lambda'}$ by retaining only the points in Λ'. To control the resulting measures, we let $\|\mu_1 - \mu_2\|_{\Lambda'}$ denote the total variation distance between the pushforward of μ_1 and μ_2 to measures on configurations in Λ' under the projection map from Λ to Λ'. In particular, if $|\Lambda'| < 1$, then the only valid configuration is the empty set of centers and so $\|\mu_1 - \mu_2\|_{\Lambda'} = 0$. For $\Lambda \subset \mathbb{R}^d$ we denote its volume by $|\Lambda|$.

We can now define the strong spatial mixing property.

Definition 2.1. *The hard sphere model at fugacity λ exhibits* strong spatial mixing *(SSM) on $\mathbb{R}^d$ if there exist $\alpha, \beta > 0$ such that for all compact measurable $\Lambda' \subseteq \Lambda \subset \mathbb{R}^d$ and any pair of boundary conditions τ and τ',*

$$\|\mu_\Lambda^\tau - \mu_\Lambda^{\tau'}\|_{\Lambda'} \leq \beta \left|\Lambda'\right| \exp(-\alpha \cdot \text{dist}(\tau \triangle \tau', \Lambda')). \tag{2.2}$$

We define the strong spatial mixing threshold on $\mathbb{R}^d$ as

$$\lambda_{\mathrm{SSM}}(d) = \sup\{\lambda : \text{ SSM holds for } \lambda' < \lambda\}. \tag{2.3}$$

It is well-known that a much weaker spatial mixing condition implies uniqueness of infinite volume Gibbs measures (e.g. [33, 11]), and so $\lambda_c(d) \geq \lambda_{\mathrm{SSM}}(d)$. The inequality can in principle be strict; for example, it is expected that $\lambda_{\mathrm{SSM}}(2) < \lambda_c(2)$.

2.3. *Optimal temporal mixing.* Consider the following Markov chain on Ω_Λ^τ, called the *single-center dynamics*. Given a configuration $X_t \in \Omega_\Lambda^\tau$, form X_{t+1} as follows:

1. Pick $x \in \Lambda$ uniformly at random.
2. With probability $1/(1+\lambda)$, remove any $y \in X_t$ with $\mathrm{dist}(x, y) < r$; that is, let $X_{t+1} = X_t \setminus B_r(x)$.
3. With probability $\lambda/(1+\lambda)$, attempt to add a center at x. That is, let $X' = X_t \cup \{x\}$. If $X' \in \Omega_\Lambda^\tau$, then set $X_{t+1} = X'$; if not, then set $X_{t+1} = X_t$.

We show in Lemma 3.1 below that the stationary distribution of this Markov chain is indeed μ_Λ^τ.

Following [12], our notion of optimal temporal mixing for Markov chains in the next definition is essentially $O(n \log n)$ mixing for all regions Λ of volume n and all boundary conditions.

Definition 2.2. *Let $n = |\Lambda|$. The single-center dynamics for the hard sphere model on $\mathbb{R}^d$ has* optimal temporal mixing *at fugacity λ if there exist $b, c > 0$ so that for any compact measurable $\Lambda \subset \mathbb{R}^d$, any boundary condition τ, any $s > 0$, and any two instances (X_t) and (Y_t) of the single-center dynamics on Ω_Λ^τ,*

$$\|X_{\lfloor sn \rfloor} - Y_{\lfloor sn \rfloor}\|_{TV} \leq bne^{-cs}, \tag{2.4}$$

where, by an abuse of notation, the left-hand side means the total variation distance between the laws of $X_{\lfloor sn \rfloor}$ and $Y_{\lfloor sn \rfloor}$.

2.4. *New results.* Using the technique of coupling with an optimized metric from Vigoda's work on the discrete hard-core model on bounded-degree graphs [37], we establish optimal temporal mixing of the single-center dynamics for fugacities $\lambda < 2^{1-d}$.

Theorem 2.1. *For all $d \geq 2$ and all $\lambda < 2^{1-d}$, the single-center dynamics for the hard sphere model on $\mathbb{R}^d$ exhibits optimal temporal mixing.*

We then prove that optimal temporal mixing of the single-center dynamics implies strong spatial mixing.

THEOREM 2.2. *Fix $d \geq 2, \lambda > 0$. If the single-center dynamics has optimal temporal mixing on $\mathbb{R}^d$, then the hard sphere model on $\mathbb{R}^d$ exhibits strong spatial mixing.*

Together these theorems imply Theorem 1.1.

PROOF OF THEOREM 1.1. Theorems 2.1 and 2.2 together immediately imply that $\lambda_c(d) \geq \lambda_{\mathrm{SSM}}(d) \geq 2^{1-d}$, the first inequality by the remark following (2.3). □

The proof of Theorem 2.2 does not use anything specific about the single-center dynamics except that it performs updates within a randomly chosen ball of bounded radius. Another Markov chain with this property is the *heat-bath dynamics* with update radius $L > 0$. To define this chain, recall the ℓ-parallel set $A^{(\ell)}$ of $A \subset \mathbb{R}^d$ is

$$A^{(\ell)} = \{x \in \mathbb{R}^d : \mathrm{dist}(x, A) \leq \ell\}. \tag{2.5}$$

In particular, given our definition of Λ_{Int} above, we have $\Lambda = \Lambda_{\mathrm{Int}}^{(r)}$. To make one step of the heat-bath dynamics we pick a point $x \in \Lambda_{\mathrm{Int}}^{(L)}$ uniformly at random and then resample the centers in $B_L(x)$ subject to the boundary conditions induced by the other centers in the current configuration and τ. Optimal temporal mixing for the heat-bath dynamics also implies strong spatial mixing.

THEOREM 2.3. *Fix $d \geq 2, \lambda > 0, L > 0$. If the heat-bath dynamics with update radius L has optimal temporal mixing on $\mathbb{R}^d$, then the hard sphere model on $\mathbb{R}^d$ exhibits strong spatial mixing.*

The proof of this theorem is essentially identical to that of Theorem 2.2 (see Section 4.1), and hence will be omitted. We also prove a converse to Theorem 2.3: that strong spatial mixing implies that the heat-bath dynamics exhibit optimal temporal mixing, provided the update radius is sufficiently large (we define optimal temporal mixing for the heat-bath dynamics just as for the single-center dynamics).

THEOREM 2.4. *Fix $d \geq 2, \lambda > 0$. If the hard sphere model on $\mathbb{R}^d$ exhibits strong spatial mixing, then there is an $L_0 > 0$ such that for $L \geq L_0$ the heat-bath dynamics with update radius L exhibits optimal temporal mixing.*

2.5. *Notation and conventions.* We briefly collect some frequently used concepts. $B_\ell(x)$ denotes the open ball of radius ℓ centered at $x \in \mathbb{R}^d$, and $V_\ell = |B_\ell(x)|$ will denote the volume of this set. In particular, $V_r = 1$. More generally, $|A|$ will denote the Lebesgue measure of $A \subset \mathbb{R}^d$. For $\Lambda \subset \mathbb{R}^d$ the *ℓ-parallel set* $\Lambda^{(\ell)}$ of Λ is $\{x \in \mathbb{R}^d : d(x, \Lambda) \leq \ell\}$. By an abuse of notation, if B is a finite set, we will write $|B|$ for the cardinality of B.

3. A rapidly mixing Markov chain for the hard sphere model.

In this section we prove that the single-center dynamics for the hard sphere model at fugacities $\lambda < 2^{1-d}$ mixes rapidly. We begin by reviewing Markov chain mixing.

3.1. *Markov chain mixing.* Let Ω denote the state space of a discrete time Markov chain. Let $p^{(0)}$ be the initial probability distribution on Ω, and let $p^{(t)}$ be the distribution after t steps of the Markov chain. Suppose the chain has a unique stationary distribution μ. The mixing time of the chain is a worst-case estimate for the number of steps it takes the Markov chain to approach stationarity. More precisely,

Definition 3.1. *The mixing time of a Markov chain is*

$$t_{mix}(\varepsilon) = \sup_{p^{(0)} \in \mathcal{P}} \min \left\{ t : \|p^{(t)} - \mu\|_{TV} \leq \varepsilon \right\} \tag{3.1}$$

where $\mathcal{P}$ denotes the set of probability distributions on Ω.

A common approach to bounding the mixing time of a Markov chain is to construct a coupling. For our purposes, a coupling of two Markov chains $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ on Ω is a stochastic process $(X_t, Y_t)_{t \geq 0}$ with values in $\Omega \times \Omega$ such that the marginals $(X_t)_{t \geq 0}$ and $(Y_t)_{t \geq 0}$ are faithful copies of the Markov chains, and $X_{t+1} = Y_{t+1}$ whenever $X_t = Y_t$.

The path coupling theorem of Bubley and Dyer [8] says that constructing a coupling for single steps of the Markov chains from certain pairs of configurations in Ω is sufficient for establishing an upper bound on the mixing time. To use this approach, one must represent the state space Ω as the vertex set of a connected (finite or infinite) graph G_Ω with a function $\hat{D} \geq 1$ defined on the edges of G_Ω. The *path metric* D corresponding to $\hat{D}$ is the shortest path distance on the graph with edge weights given by $\hat{D}$, i.e.,

$$D(X, Y) = \inf_{\gamma \colon X \to Y} \left\{ \sum_{i=0}^{|\gamma|-1} \hat{D}(\gamma_i, \gamma_{i+1}) \right\}, \tag{3.2}$$

where the infimum is over nearest-neighbor paths $\gamma = (\gamma_0, \gamma_1, \ldots, \gamma_{|\gamma|})$ in G_Ω with $\gamma_0 = X$ and $\gamma_{|\gamma|} = Y$. The path coupling technique requires that $\hat{D}$ is a *pre-metric* on G_Ω, which by definition means that

$$\hat{D}(X, Y) = D(X, Y) \qquad \text{for all edges } \{X, Y\} \text{ of } G_\Omega.$$

To establish a rapid mixing regime for the single-center dynamics we will apply the version of Bubley and Dyer's path coupling theorem stated below. In the theorem and what follows, the diameter with respect to $D(\cdot, \cdot)$, of the graph on Ω is

$$\operatorname{diam}(\Omega) = \sup_{X,Y \in \Omega} D(X, Y). \tag{3.3}$$

THEOREM 3.1 ([23, Corollary 14.7]). *Suppose the state space Ω of a Markov chain is the vertex set of a connected graph G_Ω, and suppose that $\hat{D}$ is a pre-metric on G_Ω. Let D be the corresponding path metric.*

Suppose that for each edge $\{X_0, Y_0\}$ of G_Ω the following holds: if $p^{(0)}$ and $q^{(0)}$ are the distributions concentrated on the configurations X_0 and Y_0 respectively, then there exists a coupling (X_1, Y_1) of the distributions $p^{(1)}$ and $q^{(1)}$ such that

$$\mathbb{E}\left[D(X_1, Y_1)\right] \leq D(X_0, Y_0) e^{-\alpha},$$

where $\mathbb{E}$ is the expectation with respect to the Markov chain. Then

$$t_{\mathrm{mix}}(\varepsilon) \leq \left\lceil \frac{\log(\operatorname{diam}(\Omega)) + \log(1/\varepsilon)}{\alpha} \right\rceil.$$

REMARK 3.1. *[23, Corollary 14.7] concerns Markov chains on finite state spaces, but the proof applies essentially verbatim to our context.*

3.2. *Single-center dynamics.* We will use Theorem 3.1 to prove that the single-center dynamics introduced in Section 2.3 are rapidly mixing at fugacities $\lambda < 2^{1-d}$.

THEOREM 3.2. *Let $\Lambda \subset \mathbb{R}^d$ be compact and measurable, $n = |\Lambda|$, $\gamma \in (0, 1)$, and let $\lambda = (1-\gamma)2^{1-d}$. The mixing time of the single-sphere dynamics with fugacity λ on Ω_Λ^τ satisfies*

$$t_{mix}(\varepsilon) \leq \left\lceil \frac{4n(\log(2^{d+2}n) + \log(1/\varepsilon))}{\gamma} \right\rceil.$$

for all boundary conditions τ.

Before discussing the proof of this bound, we derive Theorem 2.1 from it.

PROOF OF THEOREM 2.1. Let Λ be a compact measurable subset of $\mathbb{R}^d$ of volume n. To show optimal temporal mixing with constants $b, c > 0$, it is enough to show that with an arbitrary initial distribution X_0, $\|X_{\lfloor sn \rfloor} - \mu_\Lambda^\tau\|_{TV} \leq \frac{b}{2} n e^{-cs}$, and then use the triangle inequality to bound $\|X_{\lfloor sn \rfloor} - Y_{\lfloor sn \rfloor}\|_{TV}$. In other words, setting $\varepsilon = \frac{b}{2} n e^{-cs}$, we want to show $\tau_{\text{mix}}(\varepsilon) \leq \lfloor sn \rfloor$. Taking $b = 2^{d+3}$ and $c = \gamma/4$ and applying Theorem 3.2 proves this. □

To establish rapid mixing for the single-center dynamics, we follow the approach of Vigoda for the discrete hard-core model on bounded degree graphs [37]. This approach makes use of an extended state space $\Omega^* \supseteq \Omega$. In our setting, let $\Omega_\Lambda^{\tau,*}$ be the collection of all sets of centers $X \subseteq \Lambda_{\text{Int}}$ such that each point in Λ is covered by at most two balls of radius r with a center in X, i.e.

$$X \in \Omega_\Lambda^{\tau,*} \iff \text{for all } x \in \Lambda, \quad |\{y \in X : \text{dist}(x, y) < r\}| \leq 2. \tag{3.4}$$

The purpose of this extended state space will become clear below when we introduce a pre-metric. Note that the boundary conditions τ play no role in the definition of $\Omega_\Lambda^{\tau,*}$. Next we extend our definition of the single-center dynamics to $\Omega_\Lambda^{\tau,*}$. At each step of the chain:

1. Pick $x \in \Lambda$ uniformly at random.
2. With probability $1/(1+\lambda)$, remove any $y \in X_t$ with $\text{dist}(x, y) \leq r$. That is, set $X_{t+1} = X_t \setminus B_r(x)$.
3. With probability $\lambda/(1+\lambda)$, attempt to add a center at x. Let $X' = X_t \cup \{x\}$. If $x \in \Lambda_{\text{Int}} \setminus \tau$ and $\text{dist}(x, X_t) \geq 2r$, then set $X_{t+1} = X'$. If not, set $X_{t+1} = X_t$. That is, we add a center at x if it locally satisfies the packing constraints and boundary conditions.

If $X_t \in \Omega_\Lambda^\tau$ then the chain will remain in Ω_Λ^τ and the dynamics agree with the definition given in Section 1. In addition, a Markov chain that starts in $\Omega_\Lambda^{\tau,*} \setminus \Omega_\Lambda^\tau$ will eventually reach Ω_Λ^τ. Since the chain has a unique invariant measure when considered on the state space Ω_Λ^τ, this shows the chain also has the same unique invariant measure on $\Omega_\Lambda^{\tau,*}$, and that the mixing time of the chain on $\Omega_\Lambda^{\tau,*}$ is an upper bound for the mixing time of the chain on Ω_Λ^τ.

Throughout the remainder of this section, we fix the dimension d, the region $\Lambda \subset \mathbb{R}^d$, and the boundary conditions τ. For simplicity we write $\Omega = \Omega_\Lambda^\tau$ and $\Omega^* = \Omega_\Lambda^{\tau,*}$.

LEMMA 3.1. *The stationary distribution of the single-center dynamics on Ω is the distribution of the hard sphere model on Ω.*

PROOF. Consider two distinct configurations $X, Y \in \Omega$. The transition density from X to Y is proportional to $\mathbf{1}_{|X \Delta Y|=1} \cdot \lambda^{|Y|}$. Suppose without loss of generality that $Y = X \cup \{x\}$. Let π denote the density of μ, and let $\pi_U(V)$ denote the transition density from state U to state V. Then $\pi(Y)/\pi(X) = \lambda$, and $\pi_X(Y)/\pi_Y(X) = \lambda$, and so the chain is reversible with respect to the hard sphere measure on Ω. □

Since the single-center dynamics are a Harris recurrent chain, the previous lemma implies that μ is the unique invariant measure for the dynamics on Ω, and that $p^{(t)} \to \mu$ for all initial distributions $p^{(0)}$, see, e.g., [31, Section 3.2].

3.3. *Proof of Theorem 3.2.* We begin with some preliminary definitions. For $X \in \Omega^*$ let

$$\Gamma(X) = (\Lambda \setminus \Lambda_{\mathrm{Int}}) \cup \tau \cup \left(\bigcup_{x \in X} B_{2r}(x) \right). \tag{3.5}$$

This is the 'blocked volume' of a configuration X where an additional center cannot be placed.

For $v \in \Lambda$ we write the ball $B_{2r}(v)$ as the disjoint union of the *occupied* (or *blocked*) set $O_X(v)$ and the *unoccupied* (or *free*) set $U_X(v)$,

$$O_X(v) = B_{2r}(v) \cap \Gamma(X), \qquad U_X(v) = B_{2r}(v) \setminus \Gamma(X). \tag{3.6}$$

We now use these notions to define a graph and pre-metric on Ω^*. For $X, Y \in \Omega^*$, we say that X and Y are adjacent ($X \sim Y$) if X has exactly one more center than Y, and all the centers in Y are also in X (or vice versa). That is, $X \sim Y$ if and only if $|X \Delta Y| = 1$. We will write G_{Ω^*} for the graph whose edges are pairs of adjacent vertices, and we define a pre-metric $\hat{D}(\cdot, \cdot)$ on this graph by

$$\hat{D}(X, X \cup \{v\}) = 2^d - c|O_X(v)|, \qquad c = \frac{\lambda 2^d}{2 + \lambda 2^d}. \tag{3.7}$$

Note that since c is in $[0, 1/2]$ for λ in $[0, 2^{1-d}]$, we have that $\hat{D}(X, Y)$ is in $[2^{d-1}, 2^d]$ for all edges $\{X, Y\}$ of G_{Ω^*}. Thus for each edge $\{X, Y\}$ we have $\hat{D}(X, Y) \geq 1$, and $\hat{D}$ is a premetric since any path from X to Y other than $\{X, Y\}$ has length at least 2^d under the corresponding path metric.

The pre-metric $\hat{D}$ is the continuous analogue of the pre-metric introduced by Vigoda in [37]. Defining the state space to be Ω^* rather than Ω affects the

metric D. Consider a simple example with free boundary conditions in which Λ is a ball of radius $3r/2$. Then $\Omega = \emptyset \cup \bigcup_{x\in\Lambda_{\text{Int}}} \{\{x\}\}$. For the state space Ω the graph of adjacent states is a star graph with center $\emptyset$, and so for non-empty distinct $X, Y \in \Omega$, $D(X,Y) = \hat{D}(X,\emptyset) + \hat{D}(Y,\emptyset) = 2^{d+1}$. In contrast, for the state space Ω^*, we have that $D(X,Y) \leq \hat{D}(X, X\cup Y) + \hat{D}(Y, X\cup Y) = 2^{d+1}(1-c)$. This is relevant in our proof when we bound the distance between a pair of configurations using the triangle inequality applied with a third configuration that is in $\Omega^* \setminus \Omega$ (see (3.15)).

To apply Theorem 3.1 we will couple adjacent configurations using the following coupling.

DEFINITION 3.2 (The identity coupling for the single-center dynamics). *The identity coupling for the single-center dynamics is defined as follows. If X_t and Y_t are separate instances of the single-center dynamics for μ_Λ^τ at time t, we couple them in a Markovian manner via the transition rule*

- *Choose a point $x \in \Lambda$ uniformly at random.*
- *With probability $1/(1+\lambda)$, in both X_t and Y_t delete any center in $B_r(x)$ to form X_{t+1} and Y_{t+1} respectively.*
- *With probability $\lambda/(1+\lambda)$, attempt to add a center at x in both X_t and Y_t.*

Consider $X, Y \in \Omega^*$ with $Y = X\cup\{v\}$. Let X' and Y' denote the resultant states after one step of the Markov chains coupled according to the above identity coupling, and let

$$\Delta = D(X', Y') - D(X, Y) \tag{3.8}$$

denote the random change in distance between configurations. The next lemma bounds the expectation of Δ.

LEMMA 3.2. *Let $X, Y \in \Omega^*$ such that $Y = X\cup\{v\}$. Let $\lambda = (1-\gamma)2^{1-d}$, with $\gamma \in (0,1)$. Then*

$$\mathbb{E}[\Delta] \leq \frac{2^d(2c-1)}{n(1+\lambda)} = -\frac{\gamma 2^d}{(2-\gamma)(1+\lambda)n} < 0. \tag{3.9}$$

PROOF. Let $Y = X\cup\{v\}$. The change in distance Δ is a random variable whose value is a function of the current configurations of the chains, the random point w chosen in a single step of the coupling, and whether or not the coupling tries to add or remove spheres. We begin with a case analysis of Δ. Throughout the proof we will use that $\hat{D}(X,Y) = D(X,Y)$ for edges $\{X,Y\}$.

1. Let A_1 be the event the center v is removed from Y, i.e., the chain removes spheres and w lies within distance r of v. The probability of this event is $1/(n(1+\lambda))$. After A_1 occurs, $X' = Y'$, and so $\Delta = -D(X,Y)$. It follows that

$$\mathbb{E}\left[\Delta \cdot 1_{A_1}\right] = -\frac{1}{n(1+\lambda)} D(X,Y) = -\frac{2^d - c|O_X(v)|}{n(1+\lambda)} \tag{3.10}$$

2. Let A_2 denote the event that a center is added to X but not Y. This occurs when w lies in $U_X(v)$ and the coupling attempts to add a sphere, as $U_X(v)$ is blocked in Y and not blocked in X. In this case we have $\Delta = D(X \cup \{w\}, Y) - D(X,Y)$. It follows that

$$\mathbb{E}\left[\Delta \cdot 1_{A_2}\right] = \frac{\lambda}{n(1+\lambda)} \int_{U_X(v)} \left(D(X \cup \{w\}, Y) - D(X,Y)\right) dw. \tag{3.11}$$

3. Let A_3 be the event that a new center w is added to both X and Y. Note that this event only occurs when $w \in \Lambda \setminus \Gamma(Y)$ and the coupling adds a center. In this case

$$\Delta = -c|\{x \in U_X(v) : \ x \text{ is blocked by the new center } w\}|.$$

For $x \in U_X(v)$, let A_3^x be the event that x becomes blocked by the new center, i.e., that $X' = X \cup \{w\}$, $Y' = Y \cup \{w\}$ and $x \in O_{X \cup \{w\}}(v)$. In order for the event A_3^x to occur, it must be the case that $w \in B_{2r}(x) \setminus \Gamma(Y)$. Hence

$$\begin{aligned}
\mathbb{E}\left[\Delta \cdot 1_{A_3}\right] &= \mathbb{E}\left[\int_{U_x(v)} -c 1_{A_3^x}\, dx\right] \\
&= -\frac{c\lambda}{n(1+\lambda)} \int_{U_x(v)} \int_{\Lambda} 1_{w \in B_{2r}(x) \setminus \Gamma(Y)}\, dw\, dx \\
&= -\frac{c\lambda}{n(1+\lambda)} \int_{U_X(v)} |B_{2r}(x) \setminus \Gamma(Y)|\, dx
\end{aligned} \tag{3.12}$$

4. Let A_4 be the event that at least one center is removed in both X and Y, and v is not removed. Let S_w be the set of centers removed; since $w \notin B_r(v)$ we have $S_w = X \cap B_r(w) = Y \cap B_r(w)$. In this case,

$$\Delta = c|\{x \in O_X(v) : \ x \text{ is no longer blocked after } S_w \text{ is removed}\}|.$$

For $x \in O_X(v)$, let A_4^x be the event that $X' = X \setminus S_w$, $Y' = Y \setminus S_w$, and $x \in U_{X \setminus S_w}(v)$. If A_4^x occurs there is a center $b_x \in X$ that is the

closest center to x that blocks x. In particular, $b_x \in S_w$, and hence $w \in B_r(b_x)$. Using this observation we obtain

$$\begin{aligned}\mathbb{E}\left[\Delta \cdot 1_{A_4}\right] &= \mathbb{E}\left[\int_{O_X(v)} c 1_{A_4^x}\, dx\right] \\ &\leq \frac{c}{n(1+\lambda)} \int_\Lambda \int_{O_X(v)} 1_{w \in B_r(b_x)}\, dx\, dw \\ &= \frac{c\,|O_X(v)|}{n(1+\lambda)}. \end{aligned} \tag{3.13}$$

The events A_1, A_2, A_3, and A_4 are mutually exclusive and exhaustive, so

$$\mathbb{E}[\Delta] = \mathbb{E}\left[\Delta \cdot \sum_{i=1}^{4} 1_{A_i}\right]. \tag{3.14}$$

To derive an upper bound on $\mathbb{E}[\Delta]$ we first need to estimate the integrand in (3.11). We will use the triangle inequality with the configurations $Y \cup \{w\}$, $X \cup \{w\}$, and Y. Temporarily deferring the justification of the use of the triangle inequality, note that since $c \geq 0$, $D(Y \cup \{w\}, X \cup \{w\}) \leq D(Y, X)$. Further, by definition, $D(Y \cup \{w\}, Y) = 2^d - c|B_{2r}(w) \cap \Gamma(Y)|$. Hence by the triangle inequality

$$\begin{aligned} D(X \cup \{w\}, Y) - D(X, Y) &\leq D(Y \cup \{w\}, X \cup \{w\}) + D(Y \cup \{w\}, Y) - D(X, Y) \\ &\leq 2^d - c|B_{2r}(w) \cap \Gamma(Y)|. \end{aligned} \tag{3.15}$$

To justify this use of the triangle inequality we must establish that $X \cup \{v, w\} \in \Omega^*$. Note that no point of Λ is covered by three balls of radius r whose centers are in Y because $Y \in \Omega^*$. No point that is covered by $B_r(w)$ is covered by $B_r(u)$ for some $u \in X$ since w is added to X by the Markov chain. It follows that no point in Λ is covered three times by $Y \cup \{w\}$, i.e., $Y \cup \{w\} \in \Omega^*$.

Inserting the estimates given in (3.10)–(3.13) into (3.14) we obtain

$$\begin{aligned}\mathbb{E}\left[\Delta\right] &\leq \frac{1}{n(1+\lambda)} \Bigg(-(2^d - c|O_X(v)|) + \lambda \int_{U_X(v)} (2^d - c|B_{2r}(w) \cap \Gamma(Y)|)\, dw \\ &\qquad - c\lambda \int_{U_X(v)} |B_{2r}(x) \setminus \Gamma(Y)|\, dx + c|O_X(v)| \Bigg) \\ &= \frac{1}{n(1+\lambda)} \left(-2^d + 2c|O_X(v)| + \lambda 2^d (1-c) |U_X(v)| \right), \end{aligned}$$

where the last line follows from $|B_{2r}(x) \cap \Gamma(Y)| + |B_{2r}(x) \setminus \Gamma(Y)| = 2^d$. Since $|U_X(v)| + |O_X(v)| = 2^d$ and $2c = \lambda 2^d(1-c)$, it follows that

$$\mathbb{E}[\Delta] \le \frac{2^d(2c-1)}{n(1+\lambda)} = -\frac{\gamma 2^d}{(2-\gamma)(1+\lambda)n}. \qquad \square$$

Now we deduce Theorem 3.2 from Theorem 3.1.

PROOF OF THEOREM 3.2. First we bound the diameter of Ω^* with respect to $D(\cdot,\cdot)$. Note that if $X \in \Omega^*$ then $|X| \le 2n$ since each ball covers one unit of volume and each point cannot be covered more than twice. Recall that we defined the graph G_{Ω^*} by putting an edge between configurations that differ by exactly one sphere location. It follows that the combinatorial diameter of G_{Ω^*} is bounded above by $4n$. For two adjacent states X and Y, $D(X,Y) = \hat{D}(X,Y) \le 2^d$, and hence $\operatorname{diam}(\Omega^*) \le n2^{d+2}$. Note also that G_{Ω^*} is connected since there is a path from any configuration to the empty configuration.

Next we find a suitable value for α in the statement of Theorem 3.1. Let $X_0 = X$ and $Y_0 = X \cup \{v\}$. Applying Lemma 3.2 we obtain

$$\begin{aligned}
\mathbb{E}[D(X_1,Y_1)] &= D(X_0,Y_0)\left(1 + \frac{\mathbb{E}[\Delta(X_0,Y_0)]}{D(X_0,Y_0)}\right) \\
&\le D(X_0,Y_0)\left(1 - \frac{\gamma}{n(2-\gamma)(1+\lambda)}\right) \\
&\le D(X_0,Y_0)\exp\left[-\frac{\gamma}{n(2-\gamma)(1+\lambda)}\right] \\
&\le D(X_0,Y_0)e^{-\frac{\gamma}{4n}}.
\end{aligned}$$

The first inequality used that $\mathbb{E}[\Delta] < 0$ and $D(X_0,Y_0) \le 2^d$, and the last used that $1+\lambda \le 2$. Since $\hat{D}$ is a pre-metric, we have verified the hypotheses of Theorem 3.1 with $\alpha = \gamma/4n$, which gives Theorem 3.2. $\square$

4. Spatial and temporal mixing. In this section we prove Theorems 2.2 and 2.4 following the approach of Dyer, Sinclair, Vigoda, and Weitz [12] who proved similar results for the discrete hard-core model on the integer lattice $\mathbb{Z}^d$. At the heart of this technique is the idea of disagreement percolation, bounding the distance that a disagreement between two copies of a Markov chain can typically spread in a fixed number of steps. This idea appeared in [34] in the context of spatial and temporal mixing with further refinements and applications due to van den Berg [36].

The main complication in the continuous setting is that there is a richer variety of domains and boundary conditions to consider. To handle this we will need the following lemma about the volume of parallel sets in Euclidean space.

LEMMA 4.1 (Fradelizi–Marsiglietti [14]). *Suppose $L \geq r$, then*

$$|\Lambda^{(L)}| \leq \frac{L^d}{r^d} |\Lambda^{(r)}| \,.$$

In particular, for $L \geq r$ we have $|\Lambda^{(L)}_{\text{Int}}| \leq \frac{L^d}{r^d}|\Lambda|$.

PROOF. For $\Lambda, B \subset \mathbb{R}^d$ compact with B convex, Fradelizi and Marsiglietti [14, Proposition 2.1] proved that the function $|s\Lambda + B| - s^d|\Lambda|$ is non-decreasing and continuous as a function of s on $\mathbb{R}_+$, where $s\Lambda + B$ is the Minkowski sum of $s\Lambda$ and B. In particular,

$$|\Lambda + B_L(0)| - |\Lambda| \leq \left|\frac{L}{r}\Lambda + B_L(0)\right| - \left(\frac{L}{r}\right)^d |\Lambda| \,,$$

where we have used the continuity in L to obtain the result for the open ball $B_L(0)$ instead of its closure. Since $L \geq r$ this implies

$$|\Lambda + B_L(0)| \leq \left|\frac{L}{r}\Lambda + B_L(0)\right| = \left(\frac{L}{r}\right)^d |\Lambda + B_r(0)| \,,$$

and the first claim follows since $\Lambda^{(L)} = \Lambda + B_L(0)$ and $\Lambda^{(r)} = \Lambda + B_r(0)$. The second claim follows since $\Lambda = \Lambda^r_{\text{Int}}$. □

4.1. *From temporal to spatial mixing.* The following lemma bounds how fast a disagreement between two copies of the single-center dynamics can spread. This is a continuum variant of [12, Lemma 3.1].

For $\Lambda' \subset \Lambda \subset \mathbb{R}^d$ we write $X_t\,[\Lambda']$ to denote the projection of $X_t \subset \Lambda_{\text{Int}}$ to the set Λ'_{Int}. In words, $X_t[\Lambda']$ is the set of centers of spheres in X_t that are entirely contained in Λ'.

LEMMA 4.2. *Let X_t and Y_t be two copies of the single-center dynamics for the hard sphere model on Λ with boundary conditions τ_X and τ_Y and initial conditions $X_0, Y_0 \in \Omega^{\tau,*}_\Lambda$. Suppose both $X_0 \triangle Y_0$ and $\tau_X \triangle \tau_Y$ are contained in $A \subset \Lambda$. Let $B \subset \Lambda$ with $\operatorname{dist}(A, B_{\text{Int}}) = s > 0$. Then for all positive $\eta \leq \frac{s}{e^2 r \cdot 4^{d+1}}$, under the identity coupling we have*

$$\Pr\left[X_{\lfloor \eta n \rfloor}[B] \neq Y_{\lfloor \eta n \rfloor}[B]\right] \leq |B| \cdot e^{-s/(4r)}, \tag{4.1}$$

where $n = |\Lambda|$.

PROOF OF LEMMA 4.2. We couple X_t and Y_t via the identity coupling. Say t' is the smallest t so that $X_t[B] \neq Y_t[B]$. That is, there is a center in B_{Int} in one configuration but not the other. Since removing a center will not create a disagreement, at step t' exactly one of the Markov chains must add a center at some $w \in B_{\mathrm{Int}}$. In order for the update point w to be added to only one of the Markov chains, it must be that $B_{2r}(w)$ contains a point y of disagreement, meaning that y is a center in one of the configurations but not the other.

Proceeding further, if $X_{\lfloor \eta n \rfloor}[B] \neq Y_{\lfloor \eta n \rfloor}[B]$ then there must be a connected (overlapping) chain of balls of radius $2r$ joining A to B_{Int} with the property that there is a point of disagreement in each update ball. In particular, the balls must be ordered in time to propagate a disagreement forward. We call such a chain of balls an *ordered chain*. With $s = \operatorname{dist}(A, B_{\mathrm{Int}})$, there must be an ordered chain of at least $m = \lceil \frac{s}{4r} \rceil$ balls connecting A to B_{Int}.

In any ordered chain each ball must intersect the last ball added to the chain, so the probability of extending a chain of balls of radius $2r$ by one ball is at most $4^d/n$. The probability of forming an ordered chain of ℓ balls of radius $2r$ with the final ball centered in B_{Int} is thus at most

$$\binom{\lfloor \eta n \rfloor}{\ell} \frac{|B_{\mathrm{Int}}|}{n} \left(\frac{4^d}{n}\right)^{\ell-1}, \tag{4.2}$$

where we have neglected the constraint that a disagreement must be created in each update ball. This upper bounds the probability of a disagreement in B at time $\lfloor \eta n \rfloor$ by

$$\frac{|B_{\mathrm{Int}}|}{4^d} \sum_{\ell=m}^{\lfloor \eta n \rfloor} \binom{\lfloor \eta n \rfloor}{\ell} \left(\frac{4^d}{n}\right)^{\ell} \leq \frac{|B_{\mathrm{Int}}|}{4^d} \sum_{\ell=m}^{\infty} \left(\frac{e\eta 4^d}{\ell}\right)^{\ell} \leq \frac{|B_{\mathrm{Int}}|}{4^d} \sum_{\ell=m}^{\infty} \left(\frac{s}{4re\ell}\right)^{\ell}.$$

The first inequality used $\binom{M}{\ell} \leq (eM/\ell)^{\ell}$ and the second used the hypothesis on η, i.e., $\eta < \frac{s}{e^2 r \cdot 4^{d+1}}$. The ratio of consecutive terms in the summation is at most $1/e$, so the entire series is bounded by twice the first term. This gives an upper bound of

$$\frac{2\,|B_{\mathrm{Int}}|}{4^d} \left(\frac{s}{4erm}\right)^m \leq \frac{2\,|B_{\mathrm{Int}}|}{4^d} e^{-s/(4r)} \leq |B| e^{-s/(4r)},$$

where the first inequality is due to the definition of m as a ceiling. □

REMARK 4.1. *An inspection of the preceding proof reveals that it also applies to the situation in which the boundary conditions $\tau_X = \tau_X(t)$ and $\tau_Y = \tau_Y(t)$ change in time, provided $\tau_X(t) \triangle \tau_Y(t) \subset A$ for all t. In this*

situation the configurations are in $\Omega_\Lambda^{\tau,*}$ as they may not satisfy the boundary conditions.

The next lemma shows that optimal temporal mixing implies what is called *projected optimal mixing* [12, Lemma 4.1]. Recall the definition of $\|\cdot\|_{\Lambda'}$ from Section 2.2.

LEMMA 4.3. *If the single-center dynamics has optimal temporal mixing on $\mathbb{R}^d$ with constants $b, c > 0$ then there exist constants $b', c' > 0$ such that, for any compact measurable $\Lambda \subset \mathbb{R}^d$, any boundary condition τ, any two instances X_t and Y_t of the dynamics on Ω_Λ^τ, and any measurable $\Lambda' \subset \Lambda$, we have that*

$$\|X_{\lfloor \eta n \rfloor} - Y_{\lfloor \eta n \rfloor}\|_{\Lambda'} \leq b' \left|\Lambda'\right| e^{-c'\eta}$$

for any $\eta > e^{-2}4^{-d-1}$, where $n = |\Lambda|$. The same conclusion also holds if $X_0, Y_0 \in \Omega_\Lambda^{\tau,*}$, as long as $X_0[A_R], Y_0[A_R] \in \Omega_{A_R}$, with A_R as defined below in the proof.

PROOF. Fix $\eta > e^{-2}4^{-d-1}$ and let $R = \eta e^2 r \cdot 4^{d+1}$. Define A_R to be

$$A_R = \{x \in \Lambda : \operatorname{dist}(x, \Lambda'_{\text{Int}}) \leq R\}.$$

If $|\Lambda'| < 1$, then the total variation distance is zero since both $X_t[\Lambda']$ and $Y_t[\Lambda']$ are the empty set since no spheres fit inside Λ'. Hence we may assume that $|\Lambda'| \geq 1$, and our assumption on η implies that $|A_R| \geq |A_r| = |\Lambda'| \geq 1$.

The proof proceeds by defining auxiliary Markov chains X_t^R and Y_t^R on A_R that imitate X_t and Y_t closely, and then using the triangle inequality:

(4.3)
$$\|X_{\lfloor \eta n \rfloor} - Y_{\lfloor \eta n \rfloor}\|_{\Lambda'} \leq \|X_{\lfloor \eta n \rfloor} - X^R_{\lfloor \eta n \rfloor}\|_{\Lambda'} + \|X^R_{\lfloor \eta n \rfloor} - Y^R_{\lfloor \eta n \rfloor}\|_{\Lambda'} + \|Y^R_{\lfloor \eta n \rfloor} - Y_{\lfloor \eta n \rfloor}\|_{\Lambda'}.$$

The definition of A_R will ensure that it is unlikely that information can pass from outside A_R to Λ', which will ensure the first and third terms are small. The second term will be handled by the optimal temporal mixing hypothesis.

In detail, we define the Markov chains X_t^R and Y_t^R to be empty outside of $(A_R)_{\text{Int}}$ for all t, and to agree with X_0 and Y_0 respectively inside $(A_R)_{\text{Int}}$ at $t = 0$. The two chains have the same dynamics:

- Uniformly select an update point x from Λ.
- If $x \notin A_R$, do nothing.
- Otherwise perform an update of the chain, with the configuration outside A_R held empty as a boundary condition. Formally, the boundary condition for this update is $\tau_R = \Lambda \setminus (A_R)_{\text{Int}}$.

X_t^R and Y_t^R are lazy dynamics on A_R: with probability $1 - |A_R|/n$ nothing occurs, otherwise an update on A_R is performed.

We couple X_t^R with X_t by a variant of the identity coupling: if X_t updates at a point outside A_R then X_t^R does nothing, otherwise attempt the same update. The projections of X_t and X_t^R to Λ' are both copies of the hard sphere model on Λ' with boundary conditions and initial conditions that only differ outside $(A_R)_{\mathrm{Int}}$, see the third bullet above. As a result, Lemma 4.2 and Remark 4.1 imply that

$$\|X_{\lfloor \eta n \rfloor} - X^R_{\lfloor \eta n \rfloor}\|_{\Lambda'} \leq |\Lambda'| e^{-R/(4r)} = |\Lambda'| e^{-\eta e^2 4^d}.$$

The application of Lemma 4.2 is valid by the definition of R and A_R, i.e., that $\operatorname{dist}(\Lambda'_{\mathrm{Int}}, A_R^c) > R$. In particular, this holds even if $X_0 \in \Omega_\Lambda^{\tau,*}$. Exactly the same reasoning and bound apply to $\|Y^R_{\lfloor \eta n \rfloor} - Y_{\lfloor \eta n \rfloor}\|_{\Lambda'}$.

As the configurations of X_t^R and Y_t^R agree outside A_R and $X_0[A_R], Y_0[A_R] \in \Omega_{A_R}$, the hypothesis of optimal temporal mixing applies. Hence the second term of (4.3) is small provided the chain takes enough steps. There is probability $|A_R|/n$ that the update point lies in A_R. So in ηn steps, we expect $\eta |A_R|$ updates to occur in A_R. By a Chernoff bound at least $\eta |A_R|/2$ updates occur in A_R with probability at least $1 - e^{-\eta |A_R|/8} \geq 1 - e^{-\eta/8}$ since $|A_R| \geq 1$. This gives

$$\|X^R_{\lfloor \eta n \rfloor} - Y^R_{\lfloor \eta n \rfloor}\|_{\Lambda'} \leq \|X^R_{\lfloor \eta n \rfloor} - Y^R_{\lfloor \eta n \rfloor}\|_{A_R} \leq b\, |A_R|\, e^{-c\eta/2} + 2e^{-\eta/8}.$$

The first inequality has used that the total variation distance is weakly decreasing when projecting to subsets. For the second we have applied the definition of optimal temporal mixing and used a union bound to ensure both $X^R_{\lfloor \eta n \rfloor}$ and $Y^R_{\lfloor \eta n \rfloor}$ have taken $\eta |A_R|/2$ steps. Putting these bounds together with (4.3), we have

$$\begin{aligned} \|X_{\lfloor \eta n \rfloor} - Y_{\lfloor \eta n \rfloor}\|_{\Lambda'} &\leq 2|\Lambda'| e^{-\eta e^2 4^d} + b\, |A_R|\, e^{-c\eta/2} + 2e^{-\eta/8} \\ &\leq 2|\Lambda'| e^{-\eta e^2 4^d} + b(R/r)^d e^{-c\eta/4} |\Lambda'| e^{-c\eta/4} + 2e^{-\eta/8} \end{aligned}$$

since Lemma 4.1 implies $|A_R| \leq (R/r)^d |\Lambda'|$. With

$$b' = \sup_{\eta \geq e^{-2} 4^{-d-1}} b(\eta e^2 4^{d+1})^d e^{-c\eta/4} + 4 \quad \text{and} \quad c' = \min\{c/4, 1/8\},$$

this proves the claim. □

Using these lemmas we prove Theorem 2.2. Our proof follows that of [12, Theorem 2.3].

PROOF OF THEOREM 2.2. Fix λ and d, and suppose that the single-center dynamics for the hard sphere model on $\mathbb{R}^d$ at fugacity λ exhibits optimal temporal mixing with constants b, c. Let $\Lambda \subset \mathbb{R}^d$ be compact and measurable, and suppose τ, τ' are two boundary conditions on Λ. Let $\Lambda' \subset \Lambda$ be measurable, and let $s = \operatorname{dist}(\tau \triangle \tau', \Lambda'_{\text{Int}})$.

Let Z_t be a copy of the single-center dynamics with stationary distribution μ_Λ^τ and let Z'_t be a copy of the dynamics with stationary distribution $\mu_\Lambda^{\tau'}$, and take both initial conditions to be the same sample from $\mu_\Lambda^{\tau'}$. In particular Z'_t is distributed as $\mu_\Lambda^{\tau'}$ for all t (and thus $Z'_t \in \Omega_\Lambda^{\tau'}$). On the other hand, we only know $Z_0 \in \Omega_\Lambda^{\tau,*}$ since the initial condition might violate the boundary condition τ.

We have, by the triangle inequality,

$$\|\mu_\Lambda^\tau - \mu_\Lambda^{\tau'}\|_{\Lambda'} = \|\mu_\Lambda^\tau - Z'_t\|_{\Lambda'} \le \|\mu_\Lambda^\tau - Z_t\|_{\Lambda'} + \|Z_t - Z'_t\|_{\Lambda'} \,,$$

for any choice of t. From Lemma 4.3, we have projected optimal mixing, and so if we take

$$t = \left\lfloor \frac{sn}{e^2 r \cdot 4^{d+1}} \right\rfloor \,,$$

we have

$$\|\mu_\Lambda^\tau - Z_t\|_{\Lambda'} \le b'|\Lambda'| e^{-c's/(e^2 r 4^{d+1})} \,.$$

We can apply Lemma 4.3 even though $Z_0 \in \Omega_\Lambda^{\tau,*}$ since $Z_0[A_R] \in \Omega_{A_R}$ with A_R as defined in the proof of Lemma 4.3.

The centers of Z_t and Z'_t outside of Λ' determine the boundary conditions of the projected chain restricted to Λ'. The symmetric difference of these boundary conditions are contained in $(\tau \triangle \tau')^{(2r)}$. Therefore, by Remark 4.1 our choice of t allows us to apply Lemma 4.2, which gives

$$\|Z_t - Z'_t\|_{\Lambda'} \le |\Lambda'| e^{-(s-2r)/(4r)} \,.$$

Setting $\beta = b' + e^{1/2}$ and $\alpha = \min\{c'/(e^2 r 4^{d+1}), 1/(4r)\}$ and putting these bounds together gives

$$\|\mu_\Lambda^\tau - \mu_\Lambda^{\tau'}\|_{\Lambda'} \le b'|\Lambda'| e^{-c's/(e^2 r 4^{d+1})} + |\Lambda'| e^{-s/(4r)} \le \beta |\Lambda'| e^{-\alpha s} \,. \qquad \square$$

4.2. *From spatial to temporal mixing.* Here we will show that strong spatial mixing implies that the heat-bath dynamics with radius $L \ge L_0(d, \alpha, \beta)$ exhibits optimal temporal mixing (Theorem 2.4). Along with Theorem 2.3, this shows that strong spatial mixing and optimal temporal mixing of the heat bath dynamics are essentially equivalent.

PROOF OF THEOREM 2.4. Assume the hard-sphere model on $\mathbb{R}^d$ exhibits strong spatial mixing with constants α and β. We will prove optimal temporal mixing for the heat-bath dynamics with update radius $L = Kr$, for K to be chosen large enough in the course of the proof.

We construct a path coupling using Hamming distance. That is, $D(X, Y) = |X \triangle Y|$, the number of centers in the symmetric difference of X and Y. If $|\Lambda| = n$, then at most n centers can fit in a valid configuration, and so the diameter of Ω_Λ^τ under Hamming distance is at most $2n$.

Suppose X_t and Y_t are two copies of the radius-L heat-bath chain on Ω_Λ^τ, with $X_0 = Y_0 \cup \{u\}$. Again we use an identity coupling to couple the chains: we choose the same update ball in each chain; if the boundary conditions agree, we make the same update. If the boundary conditions disagree, then we will choose a specific coupling detailed below.

Let $\Delta = D(X_1, Y_1) - D(X_0, Y_0)$ under this coupling. If x is the random center of the update ball and $u \in B_L(x)$, then the boundary conditions agree and so with probability 1, $X_1 = Y_1$, and so $\Delta = -1$. This occurs with probability

$$\Pr[u \in B_L(x)] = \frac{|B_L(u)|}{|\Lambda_{\text{Int}}^{(L)}|} = \frac{K^d}{N},$$

where we set $N = |\Lambda_{\text{Int}}^{(L)}|$.

If $u \notin B_{L+2r}(x)$, then again the boundary conditions of the update ball agree, and so with probability 1 we will have $X_1 = Y_1 \cup \{u\}$ and so $\Delta = 0$.

Finally, if $u \in B_{L+2r}(x) \setminus B_L(x)$, the boundary conditions of the update ball differ by the presence of u, and so the Hamming distance may increase. We bound the probability that u is in this width $2r$ boundary of the update ball:

$$\begin{aligned}\Pr[u \in B_{L+2r}(x) \setminus B_L(x)] &= \frac{|(B_{L+2r}(u) \cap \Lambda^{(L)}) \setminus B_L(u)|}{|\Lambda_{\text{Int}}^{(L)}|} \\ &\leq \frac{|B_{L+2r}(u) \setminus B_L(u)|}{|\Lambda_{\text{Int}}^{(L)}|} \\ &= \frac{(K+2)^d - K^d}{N}.\end{aligned}$$

Next we bound the expected increase in Hamming distance in this case under a specific coupling.

Fix $x \in \Lambda_{\text{Int}}^{(L)}$ so that $u \in B_{L+2r}(x) \setminus B_L(x)$. Let τ_X be the boundary condition on $B_L(x)$ induced by X_0 and let τ_Y be the boundary condition

induced by Y_0. In particular, $\tau_X \triangle \tau_Y \subseteq B_{2r}(u)$. Set $t = r(K/8d)^{1/d} - 2r$, and let $A = \{y \in B_L(x) : \mathrm{dist}(y, u) \le t\}$ and $\overline{A} = B_L(x) \setminus A$. We will bound the change in Hamming distance by considering A and $\overline{A}$ separately; the coupling will be chosen to control the change on $\overline{A}$.

The increase in Hamming distance can be written as the sum of the increase in Hamming distance restricted to spheres that intersect A plus the increase in Hamming distance restricted to the configuration in $\overline{A}$. An upper bound on the increase in Hamming distance for spheres intersecting A is twice the maximum number of centers possible in a valid configuration, which we can upper bound by $2V_{2r+t}$.

We now turn to $\overline{A}$. We can bound the total variation distance between $\mu^{\tau_X}_{B_L(x)}$ and $\mu^{\tau_Y}_{B_L(x)}$ restricted to $\overline{A}$ using the strong spatial mixing assumption:

$$\begin{aligned}\|\mu^{\tau_X}_{B_{Kr}(x)} - \mu^{\tau_Y}_{B_{Kr}(x)}\|_{\overline{A}} &\le \beta|\overline{A}^{(r)}|e^{-\alpha \mathrm{dist}(\tau_X \triangle \tau_Y, \overline{A})} \\ &\le \beta|\overline{A}^{(r)}|e^{-\alpha \mathrm{dist}(B_{2r}(u), \overline{A})} \\ &\le \beta(K+1)^d e^{-\alpha(t-2r)}\,.\end{aligned}$$

Therefore, there exists a coupling of X_1, Y_1 so that X_1 and Y_1 disagree within $\overline{A}$ with probability at most $\beta(K+1)^d e^{-\alpha(t-2r)}$. An upper bound on the increase in Hamming distance restricted to $\overline{A}$ is twice the maximum number of centers that can be placed in $\overline{A}^{(r)}$, which is $2(K+1)^d$. Under this coupling we can therefore bound the expected change in Hamming distance by

$$\mathbb{E}\left[\Delta \middle| u \in B_{(K+2)r}(x) \setminus B_{Kr}(x)\right] \le 2V_{2r+t} + 2\beta(K+1)^{2d} e^{-\alpha(t-2r)}\,.$$

Putting this together yields that the expected change in Hamming distance is at most

$$\mathbb{E}\left[\Delta\right] \le -\frac{K^d}{N} + \frac{(K+2)^d - K^d}{N}\left(2V_{2r+t} + 2\beta(K+1)^{2d} e^{-\alpha(t-2r)}\right)\,.$$

Now since $2V_{t+r} = K/4d$ and $(K+2)^d - K^d \le 2d(K+2)^{d-1}$, we have

$$\mathbb{E}\left[\Delta\right] \le -\frac{1}{N}\left[K^d - 2d(K+2)^{d-1}\left(\frac{K}{4d} + 2\beta(K+1)^{2d} e^{-\alpha(r(K/8d)^{1/d} - 3r)}\right)\right]\,,$$

and choosing K large enough as a function of d, α, β we can ensure that

$$\mathbb{E}[\Delta] \le -\frac{K^d}{3}\frac{1}{N} \le -\frac{1}{3n}$$

since $N = |\Lambda^{(L)}_{\mathrm{Int}}| \le K^d n$ by Lemma 4.1. Then combining this bound and the diameter bound, Theorem 3.1 gives optimal temporal mixing. □

5. Bounds on the critical density. In this section we prove Theorem 1.2; this requires two preparatory results. Recall that

$$\rho(\lambda) = \liminf_{n\to\infty} \frac{1}{n}\mathbb{E}_{Q_n,\lambda}|\mathbf{X}|\,,$$

where Q_n is the box of volume n centered at the origin in $\mathbb{R}^d$. We first give an easy lower bound on $\rho(\lambda)$. This is closely related to an inequality of Lieb [24] that applies to the canonical ensemble.

LEMMA 5.1. *For all d and all $\lambda \geq 0$,*

$$\rho(\lambda) \geq \frac{\lambda}{1+2^d\lambda}\,.$$

PROOF. Let $\rho_\Lambda(\lambda) = \frac{1}{|\Lambda|}\mathbb{E}_{\Lambda,\lambda}|\mathbf{X}|$ be the expected density of the hard sphere model on Λ with free boundary conditions, and let

$$F_\Lambda(\lambda) = \frac{\mathbb{E}_{\Lambda,\lambda}|\{y \in \Lambda_{\mathrm{Int}} : \mathrm{dist}(y,\mathbf{X}) > 2r\}|}{|\Lambda|}$$

be the expected *free volume* fraction of Λ. A short calculation gives the identity $\rho_\Lambda(\lambda) = \lambda F_\Lambda(\lambda)$ for all bounded measurable Λ of positive volume [20, Lemma 7]. Further,

$$F_\Lambda(\lambda) \geq |\Lambda_{\mathrm{Int}}|/|\Lambda| - 2^d\rho_\Lambda(\lambda)\,,$$

since each center in $\mathbf{X}$ can block at most volume 2^d. With a little algebra, this implies

$$\rho_\Lambda(\lambda) \geq \frac{|\Lambda_{\mathrm{Int}}|}{|\Lambda|}\frac{\lambda}{1+\lambda 2^d}.$$

Applying this bound to $\Lambda = Q_n$ and taking a limit gives the lemma. □

We will also require the following bound on $\rho(\lambda)$.

THEOREM 5.1 ([20, Proof of Theorem 2]). *For all $d \geq 2$ and all $\lambda > 0$,*

$$\rho(\lambda) \geq \inf_z \max\left\{\lambda e^{-z}, z2^{-d}e^{-2\lambda 3^{d/2}}\right\}.$$

In particular if $\lambda = c2^{-d}$, we have $\rho(\lambda) \geq (1+o_d(1))W(c)2^{-d}$ where $W(\cdot)$ is the Lambert-W function, i.e. the inverse of $f(W) = We^W$.

PROOF OF THEOREM 1.2. To prove the first statement in Theorem 1.2, we combine Lemma 5.1 and Theorem 1.1 to get

$$\rho_c \geq \frac{2^{1-d}}{1+2^d 2^{1-d}} = \frac{2}{3\cdot 2^d}\,,$$

To prove the second statement in Theorem 1.2, we use Theorem 5.1 and the bound $\lambda_c(d) \geq 2^{1-d}$, to obtain

$$\rho_c(d) \geq (1+o_d(1))W(2)2^{-d} \geq (.8526+o_d(1))\cdot 2^{-d}\,,$$

as $d \to \infty$. □

Acknowledgements. We thank Arnaud Marsiglietti for pointing us towards Lemma 4.1, and the anonymous referees for many helpful comments and suggestions.

References.

[1] B. J. Alder and T. E. Wainwright. Studies in molecular dynamics. ii. behavior of a small number of elastic spheres. *The Journal of Chemical Physics*, 33(5):1439–1451, 1960.

[2] P. Balister, B. Bollobás, and M. Walters. Continuum percolation with steps in the square or the disc. *Random Structures & Algorithms*, 26(4):392–403, 2005.

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

[4] L. Bertini, N. Cancrini, and F. Cesi. The spectral gap for a Glauber-type dynamics in a continuous gas. *Annales de l'Institut Henri Poincare B: Probability and Statistics*, 38(1):91–108, 2002.

[5] X. Blanc and M. Lewin. The crystallization conjecture: a review. *EMS Surveys in Mathematical Sciences*, 2(2):255–306, 2015.

[6] A.-S. Boudou, P. Caputo, P. Dai Pra, and G. Posta. Spectral gap estimates for interacting particle systems via a Bochner-type identity. *Journal of Functional Analysis*, 232(1):222–258, 2006.

[7] D. C. Brydges. A short course on cluster expansions. *Les Houches*, (PART I), 1984.

[8] R. Bubley and M. Dyer. Path coupling: A technique for proving rapid mixing in Markov chains. In *Proceedings 38th Annual Symposium on Foundations of Computer Science*, pages 223–231. IEEE, 1997.

[9] P. Diaconis, G. Lebeau, and L. Michel. Geometric analysis for the Metropolis algorithm on Lipschitz domains. *Inventiones Mathematicae*, 185(2):239–281, 2011.

[10] R. L. Dobrushin. Definition of random variables by conditional distributions. *Teoriya Veroyatnostei i ee Primeneniya*, 15(3):469–497, 1970.

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

[12] M. Dyer, A. Sinclair, E. Vigoda, and D. Weitz. Mixing in time and space for lattice spin systems: A combinatorial view. *Random Structures & Algorithms*, 24(4):461–479, 2004.

[13] R. Fernández, A. Procacci, and B. Scoppola. The analyticity region of the hard sphere gas. Improved bounds. *J. Stat. Phys.*, 5:1139–1143, 2007.

[14] M. Fradelizi and A. Marsiglietti. On the analogue of the concavity of entropy power in the Brunn–Minkowski theory. *Advances in Applied Mathematics*, 57:1–20, 2014.
[15] H. Frisch, N. Rivier, and D. Wyler. Classical hard-sphere fluid in infinitely many dimensions. *Physical Review Letters*, 54(19):2061, 1985.
[16] H. Guo and M. Jerrum. Perfect simulation of the hard disks model by partial rejection sampling. *Annales de l'Institut Henri Poincaré D*, 8(2):159–177, 2021.
[17] T. P. Hayes and C. Moore. Lower bounds on the critical density in the hard disk model via optimized metrics. *arXiv preprint arXiv:1407.1930*, 2014.
[18] C. Hofer-Temmel. Disagreement percolation for the hard-sphere model. *Electronic Journal of Probability*, 24, 2019.
[19] W. G. Hoover and F. H. Ree. Melting transition and communal entropy for hard spheres. *The Journal of Chemical Physics*, 49(8):3609–3617, 1968.
[20] M. Jenssen, F. Joos, and W. Perkins. On the hard sphere model and sphere packings in high dimensions. In *Forum of Mathematics, Sigma*, volume 7. Cambridge University Press, 2019.
[21] M. Kahle. Sparse locally-jammed disk packings. *Annals of Combinatorics*, 16(4):773–780, 2012.
[22] R. Kannan, M. W. Mahoney, and R. Montenegro. Rapid mixing of several Markov chains for a hard-core model. In *Algorithms and computation*, volume 2906 of *Lecture Notes in Comput. Sci.*, pages 663–675. Springer, Berlin, 2003.
[23] D. A. Levin and Y. Peres. *Markov Chains and Mixing Times*, volume 107. American Mathematical Soc., 2017.
[24] E. Lieb. New method in the theory of imperfect gases and liquids. *Journal of Mathematical Physics*, 4(5):671–678, 1963.
[25] H. Löwen. Fun with hard spheres. In *Statistical Physics and Spatial Statistics*, volume 554, pages 295–331. Springer, 2000.
[26] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.
[27] M. Michelen and W. Perkins. Analyticity for classical gasses via recursion. *arXiv preprint arXiv:2008.00972*, 2020.
[28] T. X. Nguyen and R. Fernández. Convergence of cluster and virial expansions for repulsive classical gases. *Journal of Statistical Physics*, 179:448–484, 2020.
[29] O. Penrose. Convergence of fugacity expansions for fluids and lattice gases. *Journal of Mathematical Physics*, 4(10):1312–1320, 1963.
[30] T. Richthammer. Lower bound on the mean square displacement of particles in the hard disk model. *Communications in Mathematical Physics*, 345(3):1077–1099, 2016.
[31] G. O. Roberts and J. S. Rosenthal. General state space Markov chains and MCMC algorithms. *Probability Surveys*, 1:20–71, 2004.
[32] D. Ruelle. Correlation functions of classical gases. *Annals of Physics*, 25:109–120, 1963.
[33] D. Ruelle. *Statistical mechanics: Rigorous results*. World Scientific, 1999.
[34] D. W. Stroock and B. Zegarlinski. The logarithmic Sobolev inequality for discrete spin systems on a lattice. *Communications in Mathematical Physics*, 149(1):175–193, 1992.
[35] A. L. Thorneywork, J. L. Abbott, D. G. Aarts, and R. P. Dullens. Two-dimensional melting of colloidal hard spheres. *Physical Review Letters*, 118(15):158001, 2017.
[36] J. van den Berg. A uniqueness condition for Gibbs measures, with applications to the 2-dimensional Ising antiferromagnet. *Communications in Mathematical Physics*, 152(1):161–166, 1993.

[37] E. Vigoda. A note on the Glauber dynamics for sampling independent sets. *The Electronic Journal of Combinatorics*, 8(1):8, 2001.

[38] D. Weitz. Combinatorial criteria for uniqueness of gibbs measures. *Random Structures & Algorithms*, 27(4):445–475, 2005.

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

[40] J. Wellens. A note on partial rejection sampling for the hard disks model in the plane. *arXiv preprint arXiv:1808.03367*, 2018.

[41] D. B. Wilson. How to couple from the past using a read-once source of randomness. *Random Structures & Algorithms*, 16(1):85–113, 2000.

[42] L. Wu. Estimate of spectral gap for continuous gas. *Annales de l'Institut Henri Poincare B: Probability and Statistics*, 40(4):387–409, 2004.

[43] C.-N. Yang and T.-D. Lee. Statistical theory of equations of state and phase transitions. i. theory of condensation. *Physical Review*, 87(3):404, 1952.

Department of Mathematical Sciences
Durham University
Mathematical Sciences & Computer Science Building
Upper Mountjoy Campus
Stockton Road
Durham University, DH1 3LE
E-mail: tyler.helmuth@durham.ac.uk

Department of Mathematics,
Statistics, and Computer Science
University of Illinois at Chicago
851 S. Morgan St.
Chicago, IL 60607
E-mail: math@willperkins.org

NSF-Simons Center for Mathematical and Statistical Analysis of Biology
Harvard University
16 Divinity Avenue
Cambridge, MA 02138
E-mail: spetti@fas.harvard.edu