The Journal of Chemical Physics

RESEARCH ARTICLE | AUGUST 15 2024

Hole statistics of equilibrium 2D and 3D hard-sphere crystals

Haina Wang ; David A. Huse ; Salvatore Torquato

Check for updates

J. Chem. Phys. 161, 074106 (2024)
https://doi.org/10.1063/5.0228208

View Online
Export Citation

25 August 2024 14:56:01

Nanotechnology & Materials Science
Optics & Photonics
Impedance Analysis
Scanning Probe Microscopy
Sensors
Failure Analysis & Semiconductors



Unlock the Full Spectrum.
From DC to 8.5 GHz.
Your Application. Measured.
Find out more
Zurich Instruments

AIP Publishing

The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

Hole statistics of equilibrium 2D and 3D hard-sphere crystals

Cite as: J. Chem. Phys. 161, 074106 (2024); doi: 10.1063/5.0228208
Submitted: 11 July 2024 • Accepted: 1 August 2024 •
Published Online: 15 August 2024

View Online Export Citation CrossMark

Haina Wang,[1] David A. Huse,[2] and Salvatore Torquato[3,a)]

AFFILIATIONS

[1] Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA
[2] Department of Physics, Princeton University, Princeton, New Jersey 08544, USA
[3] Department of Chemistry, Department of Physics, Princeton Institute of Materials, and Program in Applied and Computational Mathematics, Princeton University, Princeton, New Jersey 08544, USA

[a)] **Author to whom correspondence should be addressed:** torquato@princeton.edu

ABSTRACT

The probability of finding a spherical "hole" of a given radius r contains crucial structural information about many-body systems. Such hole statistics, including the void conditional nearest-neighbor probability functions $G_V(r)$, have been well studied for hard-sphere fluids in d-dimensional Euclidean space $\mathbb{R}^d$. However, little is known about these functions for hard-sphere crystals for values of r beyond the hard-sphere diameter, as large holes are extremely rare in crystal phases. To overcome these computational challenges, we introduce a biased-sampling scheme that accurately determines hole statistics for equilibrium hard spheres on ranges of r that far extend those that could be previously explored. We discover that $G_V(r)$ in crystal and hexatic states exhibits oscillations whose amplitudes increase rapidly with the packing fraction, which stands in contrast to $G_V(r)$ that monotonically increases with r for fluid states. The oscillations in $G_V(r)$ for 2D crystals are strongly correlated with the local orientational order metric in the vicinity of the holes, and variations in $G_V(r)$ for 3D states indicate a transition between tetrahedral and octahedral holes, demonstrating the power of $G_V(r)$ as a probe of local coordination geometry. To further study the statistics of interparticle spacing in hard-sphere systems, we compute the local packing fraction distribution $f(\phi_l)$ of Delaunay cells and find that, for $d \leq 3$, the excess kurtosis of $f(\phi_l)$ switches sign at a certain transitional global packing fraction. Our accurate methods to access hole statistics in hard-sphere crystals at the challenging intermediate length scales reported here can be applied to understand the important problem of solvation and hydrophobicity in water at such length scales.

Published under an exclusive license by AIP Publishing. https://doi.org/10.1063/5.0228208

25 August 2024 14:56:01

I. INTRODUCTION

The venerable idealized hard-sphere model has been fruitfully employed to fundamentally understand the structure and bulk properties of a broad range of realistic systems with strong short-ranged repulsive interactions, including crystals and liquids of noble gases, glasses, granular media, and biological systems.[1–3] The determination of the equilibrium phase diagrams for two-dimensional (2D) hard-disk and three-dimensional (3D) hard-sphere systems has been an active area of study since the pioneering studies of Alder and Wainwright,[4,5] who presented numerical evidence for entropy-driven disorder–order phase transitions. The 3D hard-sphere system undergoes a first-order liquid–solid phase transition, with the corresponding freezing and melting packing fractions $\phi_f = 0.494$ and $\phi_m = 0.545$, respectively.[6] The densest packing fraction for the 3D hard-sphere crystal is given by $\phi_J = \pi/\sqrt{18}$, which is achieved by the uncountably infinite stacking variants of hexagonal close-packed layers (e.g., fcc and hcp). There exist strong numerical evidence that the fcc lattice has the highest entropy among all possible stacking variants.[7,8] For the equilibrium 2D hard-disk system, recent studies[9–11] show that its phase diagram contains a fluid phase, a hexatic phase with short-range translational order and quasi-long-range orientational order, and a triangle-lattice crystal phase with the densest possible packing fraction $\phi_J = \pi/\sqrt{12}$. The freezing point for the hard-disk system, which corresponds to the transition between the fluid and the hexatic phases, occurs at $\phi_f = 0.69$.[6,10] A continuous phase transition between the hexatic and the crystal phases occurs at $\phi_h = 0.72$.[11] In this work, we focus on the "hole" statistics of 2D and 3D hard-sphere crystals.

Published under an exclusive license by AIP Publishing

A key microstructure descriptor used to probe the hole statistics in many-body systems is given by the "void" nearest-neighbor functions.[2,12,13] These functions, including the void probability density function $H_V(r)$ and the associated hole probability and conditional probability functions $E_V(r)$ and $G_V(r)$, respectively, have been extensively studied for hard spheres along the disordered (fluid) branch.[2,12–15] The void nearest-neighbor conditional probability distribution function $G_V(r)$ is defined such that $\rho s_1(r) G_V(r) dr$ gives the probability of finding a particle at a distance between r and $r + dr$, given that one has found a hole of radius r centered at a random point, where $s_1(r)$ is the surface area of a sphere with radius r.[2,12] It is known that $G_V(r)$, $H_V(r)$, and $E_V(r)$ (defined in Sec. II) can be expressed as integrals over all the n-particle correlation functions $g_2, g_3, \ldots$.[2,12] and contain crucial information about equilibrium and nonequilibrium properties, such as the pressure and the excess chemical potential of equilibrium hard-sphere systems,[2,16] as well as the degree to which a many-body system is out of equilibrium.[17] Moreover, these functions play an important role in the theoretical study of liquids and amorphous matter,[13,15,18,19] granular flows,[20] stellar dynamics,[21] disordered photonic materials,[22,23] and covering and quantizer problems in discrete geometry.[24] Because of these theoretical and practical interests associated with the nearest-neighbor functions, it is highly desirable to precisely determine them for nontrivial models of interacting many-body systems, including the hard-sphere model.

For one-dimensional (1D) equilibrium hard rods, $H_V(r), E_V(r)$ and $G_V(r)$ are known analytically at all packing fractions ϕ.[13,25] Exact solutions for these functions are not available for equilibrium hard spheres in two and higher dimensions, but can be very well approximated along the fluid (disordered) branch via scaled particle theory due to Reiss *et al.*[12] These authors showed that $G_V(r)$ for a hard-sphere fluid is a monotonically increasing function with r and its asymptotic for large r is given by

$$G_V(r) = a_0 + \frac{a_1}{r} + \frac{a_2}{r^2} + \cdots + \frac{a_{d-1}}{r^{d-1}} + \mathcal{O}\left(\frac{1}{r^d}\right), \quad r \gg \mathcal{D}, \quad (1)$$

where $\mathcal{D}$ is the sphere diameter and a_i are coefficients that depend on ϕ. Equation (1) provides a good approximation at small and intermediate r with $r \geq \mathcal{D}/2$ as well.[12,14] Somewhat more accurate expressions for a_i and thus $G_V(r)$ were obtained by Torquato *et al.*[13] Torquato further showed that the form (1) is still valid in the metastable extension of the liquid branch of hard-sphere systems.[15]

Compared to disordered states, relatively less is known about the void nearest-neighbor functions for crystal states. While theoretical and simulation studies have been performed to determine concentrations and free energies of vacancies in 2D and 3D equilibrium hard-sphere crystals,[26–30] they were mainly concerned with the void nearest-neighbor functions at the specific distance $r = \mathcal{D}$, which determines the thermodynamic properties of hard-sphere systems. Studying nearest-neighbor statistics in crystals for a large range of r up to $r = 2\mathcal{D}$ is challenging because large holes are extremely rare at high packing fractions. The standard method of computing these functions, which involves sampling nearest-neighbor distances from particle centers to randomly placed test points, becomes inaccurate with greater than 100% error when $r \sim 0.7\mathcal{D}$ for 2D and 3D crystals near melting, at which $E_V(r) \sim 10^{-6}$.[17,31] Furthermore, simulations in the grand canonical (GC) ensemble are required to efficiently study hole statistics, as the GC ensemble allows particles to be added and removed to create or migrate large holes. However, standard Monte Carlo (MC) simulations in the grand canonical ensemble are not feasible in the crystal phase, since particle additions result in a very high rejection rate.

In this work, we overcome these computational challenges by introducing a biased-sampling scheme for equilibrium hard spheres in the first three spatial dimensions that enables accurate and efficient sampling of extremely rare large holes. Our algorithm is based on the Monte Carlo simulations of hard spheres and a test point that repels its nearest hard-sphere neighbor via a biasing potential to create large holes. We apply our algorithm to compute void nearest-neighbor functions on ranges of r beyond the hard-sphere diameter, up to rare holes that occur with probabilities at least five orders of magnitude smaller than the sensitivity limit of the aforementioned unbiased approach, which samples nearest-neighbor distances to random test points. For 2D crystals, hexatic phases, and dense fluids with $\phi \leq 0.73$, we are able to compute $G_V(r)$ within ±3% errors up to $r = 1.95\mathcal{D}$ and ±5% errors up to $r = 2\mathcal{D}$, which is around 2.5 times larger than the ranges on which one can compute these functions accurately via the unbiased approach. At these packing fractions, the sensitivity of hole sampling are improved by more than 40 orders of magnitude compared to the unbiased scheme and holes with probability down to $E_V(r) \sim 10^{-48}$ can been detected. For 2D crystals with $0.74 \leq \phi \leq 0.85$ and 3D crystals with $0.54 \leq \phi \leq 0.60$, we are able to accurately determine $G_V(r)$ up to at least $1.1\mathcal{D}$ within 300 h on 200 parallel threads of 2.8 GHz Intel CPUs, which improves the computational time by at least five orders of magnitude compared to the unbiased method for all the values of ϕ that we study. This capability enables us to investigate the relationship between variations in the void nearest-neighbor functions and the local coordination geometry of large holes in crystals, which could not be done in previous studies due to the limited ranges of the hole statistics. To further characterize the distribution of interparticle spacing in hard-sphere systems, we also study the local packing fraction distribution $f(\phi_l)$ of Delaunay cells as a function of the global packing fraction in the first three spatial dimensions.

We validate our methodology by comparing our numerical nearest-neighbor statistics for 1D hard rods with the known analytical expressions,[2] and find that excellent agreement up to $r = 2\mathcal{D}$ can be achieved with a small system size of $\langle N \rangle = 200$. Importantly, for 2D and 3D hard-sphere crystals, we find for the first time that that the conditional void probability function $G_V(r)$ exhibits oscillations whose amplitudes increase rapidly with the packing fraction. The function $G_V(r)$ is also oscillatory for 2D hard spheres in the hexatic phase. On the other hand, $G_V(r)$ for disordered fluid states increases monotonically with r, even for high-density fluids near freezing. We discover that oscillations in $G_V(r)$ for the 2D crystal are strongly correlated with the Nelson–Halperin orientational order metric ψ_6[32] in the vicinity of the holes. In particular, the minima in $G_V(r)$ correspond to "stable" hole sizes that preserve the local hexatic order of the crystal, i.e., holes created by removing 1, 3, 7, ... particles in the triangle lattice. For 3D hard-sphere fcc crystals, we observe a shoulder in the first peak of $G_V(r)$, which indicates the transition between tetrahedral and octahedral holes. These findings demonstrate the power of $G_V(r)$ for probing the compatibility of the crystal structures with spherical cavities of given sizes, which is

25 August 2024 14:56:01

Published under an exclusive license by AIP Publishing

closely related to the local coordination geometry of such holes. We find that in each of the first three space dimensions, the excess kurtosis of the local packing fraction distribution of Delaunay cells $f(\phi_l)$ increases with the global packing fraction ϕ and switches from negative to positive at a certain dimension-dependent transition packing fraction ϕ^*. The values of ϕ^* lie in the crystal phases for $d = 2$ and 3. This switch from a platykurtic to a leptokurtic behavior in $f(\phi_l)$ as ϕ increases reflects a transition between weakly correlated displacements of hard particles at lower ϕ to highly correlated collective displacements at higher ϕ.

In Sec. II, we provide preliminary definitions and background. In Sec. III, we describe our biased-sampling scheme to compute hole statistics of equilibrium systems. Section IV presents the results for the hole statistics of equilibrium hard spheres obtained via the sampling algorithm. In Sec. V, we study the local packing distribution of Delaunay cells. We provide conclusive remarks in Sec. VI.

II. DEFINITIONS AND PRELIMINARIES

In this section, we introduce some fundamental concepts on void nearest-neighbor functions for classical many-body systems, as well as their relations to important thermodynamic properties. We also provide exact expressions for the void nearest-neighbor functions for 1D hard-rod fluids,[18] as well as accurate approximations of these functions for 2D hard-sphere systems along the disordered branch.[15]

Consider a many-body system in d-dimensional Euclidean space $\mathbb{R}^d$. The void nearest-neighbor probability density function $H_V(r)$ is defined as[13]

$$H_V(r)dr = \text{probability that at an arbitrary located point in the system, the nearest particle center lies at a distance between } r \text{ and } r + dr. \tag{2}$$

The associated complementary hole probability function, also called the void exclusion probability function $E_V(r)$, is given by[13]

$$E_V(r) = 1 - \int_0^r H_V(r')dr' = \text{probability of finding a spherical region of radius } r \text{ centered at some arbitrary point void of particle centers.} \tag{3}$$

This definition is often given succinctly as the probability of finding a hole of radius r. Finally, the associated conditional probability function $G_V(r)$ is defined as[13]

$$G_V(r) = \frac{H_V(r)}{\rho s_1(r) E_V(r)} = \frac{-1}{\rho s_1(r)} \frac{d \ln E_V(r)}{dr}, \tag{4}$$

where $s_1(r) = d\pi^{d/2} r^{d-1}/\Gamma(1 + d/2)$ is the surface area of a d-dimensional sphere of radius r. The function $G_V(r)$ has the interpretation that $\rho s_1(r) G_V(r) dr$ is the conditional probability that, given a spherical cavity of radius r empty of particle centers, there exist particle centers in the spherical shell of volume $s_1(r)dr$ encompassing the cavity.

The void nearest-neighbor functions are related to crucial thermodynamic quantities. For example, the reduced pressure of an equilibrium monodisperse hard-sphere system at packing fraction ϕ is given by[2,18]

$$\frac{p}{\rho k_B T} = G_V(\infty) = 1 + 2^{d-1} \phi G_V(\mathcal{D}), \tag{5}$$

where $\rho = \phi / v_1(\mathcal{D}/2)$ is the number density, $v_1(r) = \pi^{d/2} r^d / \Gamma(1 + d/2)$ is the volume of a d-dimensional sphere of radius r, and T is the temperature. Furthermore, the excess chemical potential μ' of an equilibrium hard-sphere system is given by[2,16]

$$\mu' = \mu - \mu_{\text{id}} = -k_B T \ln [E_V(\mathcal{D})], \tag{6}$$

where $\mu_{\text{id}} = k_B T \ln(\rho \Lambda^d)$ is the chemical potential of the ideal gas at density ρ and Λ is the de Broglie wavelength.

For any monodisperse hard-sphere packing, whether in equilibrium or not, the void nearest-neighbor functions on the range $[0, \mathcal{D}/2]$ are exactly given by[12,13]

$$H_V(r) = \rho s_1(r), \quad r \in [0, \mathcal{D}/2], \tag{7a}$$

$$E_V(r) = 1 - \rho v_1(r), \quad r \in [0, \mathcal{D}/2], \tag{7b}$$

$$G_V(r) = \frac{1}{1 - \rho v_1(r)}, \quad r \in [0, \mathcal{D}/2]. \tag{7c}$$

For equilibrium 1D hard rods, the void nearest-neighbor functions for all $r \geq \mathcal{D}/2$ are known exactly,[18,25]

$$H_V(r) = 2\phi \exp\left(\frac{-2\phi(r/\mathcal{D} - 1/2)}{1 - \phi}\right), \quad r \geq \mathcal{D}/2, \tag{8a}$$

$$E_V(r) = (1 - \phi) \exp\left(\frac{-2\phi(r/\mathcal{D} - 1/2)}{1 - \phi}\right), \quad r \geq \mathcal{D}/2, \tag{8b}$$

$$G_V(r) = \frac{1}{1 - \phi}, \quad r \geq \mathcal{D}/2. \tag{8c}$$

For 2D and 3D equilibrium hard spheres, it is not possible to obtain exact expressions for the void nearest-neighbor functions for $r \geq \mathcal{D}/2$. However, accurate approximations for these functions along the disordered branch have been obtained by Torquato.[15] In particular, for $d = 2$, we have

$$G_V(r) = a_0 + \frac{a_1}{r}, \quad r \geq \mathcal{D}/2, \tag{9}$$

where the coefficients a_0 and a_1 are given by[15]

$$a_0 = \begin{cases} \dfrac{1 + 0.128\phi}{(1 - \phi)^2}, & 0 \leq \phi \leq \phi_f, \\ 2 g_f(1) \dfrac{\phi_c - \phi_f}{\phi_c - \phi} - \dfrac{1}{1 - \phi}, & \phi_f \leq \phi \leq \phi_c, \end{cases} \tag{10a}$$

$$a_1 = \begin{cases} \dfrac{-0.564\phi}{(1 - \phi)^2}, & 0 \leq \phi \leq \phi_f, \\ -g_f(1) \dfrac{\phi_c - \phi_f}{\phi_c - \phi} + \dfrac{1}{1 - \phi}, & \phi_f \leq \phi \leq \phi_c, \end{cases} \tag{10b}$$

where $\phi_f = 0.69$, $\phi_c = 0.82$, and $g_f(1) = (1 - 0.436\phi_f)/(1 - \phi_f)^2$.

25 August 2024 14:56:01

Published under an exclusive license by AIP Publishing

III. BIASED-SAMPLING SCHEME FOR HOLE STATISTICS

Here, we introduce the biased-sampling algorithm that enables one to accurately compute void nearest-neighbor functions for hard-sphere crystals and high-density fluids for radial distances beyond the sphere diameter, i.e., up to rare holes that occur with probabilities at least five orders of magnitude lower than the sensitivity limit of the unbiased approach. Motivated by the work of Zhang *et al.*[22] to create large holes in particle systems, this algorithm utilizes a test point that interacts with hard spheres via a biasing potential, thereby creating holes that are much larger than those observable in standard simulations. For clarity, we first describe the algorithm applied to the canonical ensemble, before extending it to the grand canonical ensemble. As is given in Sec. IV, finite size effects are significantly reduced in the latter ensemble.

A. Canonical ensemble

We consider a configuration $\mathbf{r}^N$ of N hard spheres of diameter $\mathcal{D}$ in $\mathbb{R}^d$ in a simulation box under periodic boundary conditions of fixed volume V. To sample the void nearest-neighbor probability distribution function $H_V(r)$, we introduce one additional test point at position vector $\mathbf{t}$, which interacts with its nearest hard-sphere neighbor via a biasing potential $u_i(r)$, where $r = \min_{\mathbf{r}\in\mathbf{r}^N}|\mathbf{r}-\mathbf{t}|$ is the distance from $\mathbf{t}$ to its nearest hard-sphere center. The potential $u_i(r)$ on the interval $[0, R)$ is subject to iterative optimization to sample hole radii in this range, where $R \le 2\mathcal{D}$ is the largest hole radius that we are interested in sampling at a given ϕ. In the initial iteration $(i = 0)$, we simply use $u_0(r) = 0$, i.e., there is no initial interaction between the test point and the hard spheres. However, in subsequent iterations $(i \ge 1)$, $u_i(r)$ is repulsive on $r \in (\mathcal{D}/2, R)$ due to the updating rule (12) described in the following. In each iteration i, Monte Carlo (MC) simulations are performed at the dimensionless temperature $k_BT = 1/\beta = 1$ in the canonical ensemble. It should be noted that both the hard spheres and the test point undergo attempted MC displacements.

Once equilibrium is reached, we use the equilibrated configurations to compute the probability distribution function of the nearest-neighbor distance from the test point to its nearest hard-sphere center, denoted as $p_i(r)$. To adequately sample this function, a configuration snapshot is taken every 20 sweeps for each simulation trajectory consisting of 2×10^5 sweeps. The resulting $p_i(r)$ is obtained from a binned histogram of nearest-neighbor distances to the test point with bin size $\delta = 0.01\,\mathcal{D}$, averaged over all snapshots in 200 independent simulation trajectories. Compared to the standard unbiased method, the probability density of holes of radius r in biased simulations is preferentially enhanced by a Boltzmann factor $\exp[-\beta u_i(r)]$. To compensate for this bias, one must divide $p_i(r)$ by this Boltzmann weighting factor to recover the unbiased $H_V(r)$. Thus, the estimation of $H_V(r)$ in iteration i is given by

$$H_V^i(r) = \frac{p_i(r)\exp[\beta u_i(r)]}{\int_{\mathbb{R}} p_i(r)\exp[\beta u_i(r)]dr}, \tag{11}$$

where the denominator is a normalization factor. It should be noted that Boltzmann reweighting factors have been used in numerous previous studies that employ biased-sampling techniques.[27,33,34]

In the next iteration, the biasing potential is updated binwise as

$$\beta u_{i+1}(r) = \beta u_i(r) + \ln[p_i(r)\,\mathcal{D}] + C_i \tag{12}$$

on the range of r, where $p_i(r)\,\mathcal{D} > 10^{-3}$. For r values such that $p_i(r)\,\mathcal{D} \le 10^{-3}$, we assume that the statistics for $p_i(r)$ is not sufficiently accurate to properly update $u_i(r)$, and thus a linear extrapolation is used up to $r = 2\,\mathcal{D}$. For $r > 2\,\mathcal{D}$, we set $u_i(r) = \infty$, i.e., holes with radii larger than $2\,\mathcal{D}$ are forbidden.

We choose to use an adaptive biasing potential (12), rather than a fixed potential, so that all the nearest-neighbor distances in the range $(0, R)$ can be adequately uniformly sampled. To facilitate comparison of $u_i(r)$ across iterations, the constants C_i are chosen such that $\beta u_i(r)$ at the first bin vanishes for any i. It should be noted that the values of these constants have no effect on the simulation results. The iterations are repeated until $p_i(r)$ is nearly uniform within the range of interest $(0, R)$, i.e.,

$$\frac{1}{R}\int_0^R (p_i(r)\,\mathcal{D})^2 dr - \left(\frac{1}{R}\int_0^R p_i(r)\,\mathcal{D}dr\right)^2 \le 0.05, \tag{13a}$$

and

$$p_i(r) > 0.8/R, \quad \forall r \in (0, R). \tag{13b}$$

The convergence criterion (13b) implies that given $R \le 2\,\mathcal{D}$, there are at least 8000 counts in each bin. We remark that analogs of this condition have been used in various other studies that utilize biased-sampling techniques.[27,35]

Because the time τ for the system to reach an equilibrium steady state (in terms of the number of MC sweeps) increases with ϕ, we choose R such that the hole statistics within $[0, R)$ can be well equilibrated within 2×10^6 sweeps; see Sec. IV C for details. The system is taken to be well equilibrated if $G_V(r)$ obtained from configurations 10^5 sweeps apart differ within 3% at all bins on $[0, R)$. For 2D and 3D crystals, the magnitude of $-\beta u_i(R)$ in the final iteration is on the order of 10^2.

B. Grand canonical ensemble

In contrast to the canonical ensemble, numerical simulation in the grand canonical ensemble facilitates the creation and migration of large holes through the addition and removal of particles, thereby significantly reducing finite size effects compared to the canonical ensemble, in which an interstitial has to be created for each vacancy. In standard grand canonical simulations without the biasing potential, particle additions at random positions experience a high rejection rate. By contrast, in the biased-sampling scheme, a large hole centered at $\mathbf{t}$ can be easily created as a result of the biasing potential. Therefore, new particles inserted in the vicinity of the test point have a higher probability to be accepted.

In light of this observation, we propose the following method to sample the grand canonical ensemble with chemical potential μ associated with the prescribed mean packing fraction ϕ. The initial configurations of our simulations are perfect crystals at ϕ. The simulation boxes of volume V under periodic boundary conditions are rhombic in two dimensions and cubic in three dimensions, which are commensurate with 2D triangle-lattice and 3D fcc Bravais lattice vectors, respectively. In an attempted particle addition, a hard-sphere particle is inserted, the center of which is randomly drawn from the uniform distribution on the spherical region of radius $\lambda - \mathcal{D}$ centered at $\mathbf{t}$, where $\lambda > \mathcal{D}$ is the nearest-neighbor distance to the test point before the addition. This is to ensure that the newly added particle becomes the nearest neighbor to the test

25 August 2024 14:56:01

Published under an exclusive license by AIP Publishing

point and that it does not overlap with any existing hard spheres. If $\lambda \le \mathcal{D}$, no particle addition is attempted. In attempted particle deletions, the nearest hard-sphere neighbor of the test point is removed if its distance to $\mathbf{t}$ is smaller than $\lambda' - \mathcal{D}$, where $\lambda' > \mathcal{D}$ is the second nearest-neighbor distance to the test point prior to the deletion. If $\lambda' \le \mathcal{D}$, no particle deletion is attempted.

To maintain detailed balance, the attempted additions and deletions are accepted with probabilities,

$$P_{\text{add}} = \min\left[1, \exp\left(-\beta\Delta\Phi + \beta\mu'\right)\rho_0 v_1(\lambda - \mathcal{D})\right], \tag{14a}$$

$$P_{\text{del}} = \min\left[1, \frac{\exp\left(-\beta\Delta\Phi - \beta\mu'\right)}{\rho_0 v_1(\lambda' - \mathcal{D})}\right], \tag{14b}$$

where $\Delta\Phi$ is the change in the biasing potential energy caused by the addition or deletion, $\rho_0 = \phi/v_1(\mathcal{D}/2)$ is the mean number density corresponding to the prescribed mean packing fraction, and μ' is the chemical potential in excess of the ideal gas contribution. In order to relax the full system after an addition or deletion, 50 sweeps of attempted particle displacements are done between each attempted addition or deletion. Additions and deletions are attempted randomly with equal probabilities. The iterative computations of $H_V^i(r)$ proceed in the same way as in Sec. III A.

Prior to executing the biased-sampling algorithm in the grand canonical ensemble, one must first find the excess chemical potential μ'. The chemical potentials for equilibrium hard-sphere systems in one, two, and three dimensions have been reported in various previous studies across packing fractions.[13,28,29,36] For 1D hard rods, μ' is exactly known from Eqs. (6) and (8b), i.e.,

$$\beta\mu'_{d=1}(\phi) = -\ln\left[E_V(\mathcal{D})\right] = -\ln(1-\phi) + \frac{\phi}{1-\phi}. \tag{15}$$

For 2D hard-disk and 3D hard-sphere crystals, we use the approximation obtained by Stillinger *et al.*,[36]

$$\begin{aligned}\beta\mu'_d(\phi) &= -d\,\ln\left(1 - \frac{\phi}{\phi_J}\right) + \frac{\beta p(\phi; d)}{\rho} \\ &= -d\,\ln\left(1 - \frac{\phi}{\phi_J}\right) + \frac{d}{1 - \phi/\phi_J(d)},\end{aligned} \tag{16}$$

where $\phi_J(2) = \pi/\sqrt{12}$ and $\phi_J(3) = \pi/\sqrt{18}$ are the close-packing fractions of the 2D triangle lattice and the 3D fcc crystals, respectively, and $\beta p(\phi, d)/\rho = d/(1 - \phi/\phi_J(d))$ is the reduced pressure approximated via free-volume theory.[37] It should be noted that Eq. (16) agrees very well with the numerical results for μ' via density functional theory[28] and kinetic Monte Carlo[29] calculations.

It is not known if Eq. (16) provides an accurate approximation for the excess chemical potential in the 2D hexatic phase. Thus, in this case, we numerically determine μ' following the gradual insertion method developed by Mon and Griffiths,[38] which is an extension of the well-established Widom insertion method[39] to treat high-density systems in which direct particle insertions experience a high rejection rate. In this method, a pseudo-hard sphere at position vector $\mathbf{p}$ is inserted into a configuration $\mathbf{r}^N$ of N hard spheres, and it interacts with the existing hard-sphere *centers* via a pseudo-hard-sphere potential $v(r;\,\sigma)$, where σ is the effective hard-sphere diameter. This system is equilibrated using Monte Carlo simulations, while σ is increased gradually from 0 to $\mathcal{D}$ in the canonical ensemble. The excess chemical potential is given by

$$\mu' = \int_0^{\mathcal{D}} \left\langle \frac{d\Psi(\sigma)}{d\sigma} \right\rangle_{N+\sigma} d\sigma, \tag{17}$$

where $\langle x \rangle_{N+\sigma}$ is the mean of quantity x in the canonical ensemble of N hard spheres and the additional pseudo hard sphere with pair interaction $v(r;\,\sigma)$ and

$$\Psi(\sigma) = \Sigma_{\mathbf{r}_i \in \mathbf{r}^N} v(|\mathbf{r}_i - \mathbf{p}|; \sigma) \tag{18}$$

is the potential energy experienced by the pseudo hard sphere. In this work, we use the following pseudo-hard-sphere potential:

$$\beta v(r;\sigma) = \left(\frac{4}{(r/\sigma)^{50}}\right)\left(\frac{-\operatorname{atan}\left(\frac{r/\sigma - 1}{0.001}\right)}{\pi} + \frac{1}{2}\right). \tag{19}$$

Equation (19) has the same repulsive part (up to scaling) as the well-known generalized Lennard-Jones potential,[40] but has the additional advantage that $v(r;\,\sigma)$ decreases to zero rapidly due to the presence of the arctangent function. The integral (17) is evaluated via the trapezoidal rule using 20 equal-sized bins, while the quantity $d\Psi(\sigma)/d\sigma$ is computed via automatic differentiation.[41]

IV. RESULTS FOR THE HOLE STATISTICS OF EQUILIBRIUM HARD SPHERES

Here, we present the results for the void nearest-neighbor functions obtained using our biased-sampling scheme described in Sec. III. Since $H_V(r), E_V(r)$ and $G_V(r)$ can be easily converted to one another via Eqs. (3) and (4), here we report only the results for $G_V(r)$ due to its direct relationship to the reduced pressure (5) and the fact that plots of $G_V(r)$ are more visually intuitive than those of $H_V(r)$ and $E_V(r)$.[42]

A. 1D hard-rod fluids

To demonstrate the validity and accuracy of our biased-sampling scheme, we apply it to compute the void nearest-neighbor functions for equilibrium 1D hard-rod fluids, whose functional forms in the thermodynamic limit are exactly known, as given by Eqs. (7) and (8). Figure 1 shows the results for $G_V(r)$ for the 1D hard-rod fluid at prescribed packing fraction $\phi = 0.95$. As shown in Fig. 1(a), the standard sampling approach in the canonical ensemble requires a large number of particles ($N = 10\,000$) to achieve reasonable accuracy (3% error) on the range $[0, \mathcal{D}]$. It should be noted that all the simulated $G_V(r)$ increase sharply at a certain hole radius r^*, which indicates that in standard sampling, one cannot detect holes with radii larger than r^*. By contrast, Fig. 1(b) shows that with biased-sampling in the canonical ensemble, one can accurately determine hole statistics for larger hole radii, up to $r = 1.5\mathcal{D}$, using much smaller system sizes compared to standard simulations. Due to the finite-size effect caused by fixed N and V in the canonical ensemble, $G_V(r)$ obtained via such simulations increase with r in the range $r \ge \mathcal{D}/2$. However, as N increases, the corresponding $G_V(r)$ approaches the expected value $20 = 1/(1-\phi)$ in the thermodynamic limit. Figure 1(c) shows that the biased-sampling scheme

25 August 2024 14:56:01

Published under an exclusive license by AIP Publishing

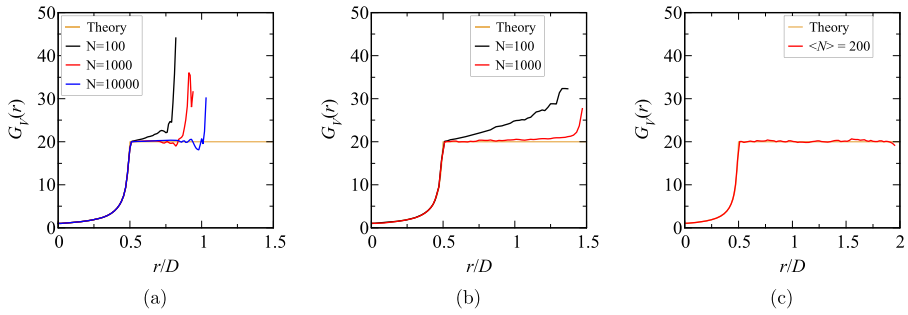


FIG. 1. Conditional void probability function $G_V(r)$ for equilibrium 1D hard-rod fluid at $\phi = 0.95$ computed via various methods. (a) $G_V(r)$ computed via the standard sampling method in the canonical ensemble. (b) $G_V(r)$ computed via biased sampling in the canonical ensemble. (c) $G_V(r)$ computed with biased sampling in the grand canonical ensemble.

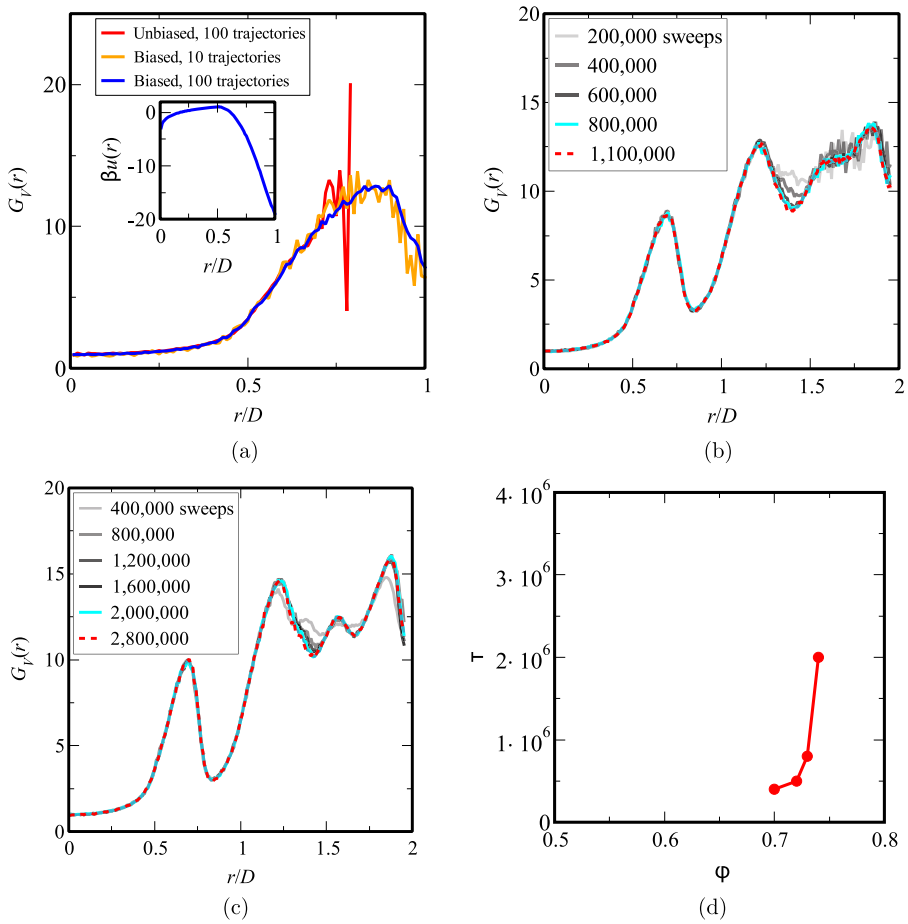


FIG. 2. (a) Comparison of $G_V(r)$ computed via the unbiased and biased-sampling approaches for the 2D equilibrium hard-disk crystal at $\phi = 0.73$ in the *canonical* ensemble. The inset shows the corresponding biasing potential. (b)–(c) The evolution of $G_V(r)$ for the 2D equilibrium hard-disk crystal at the prescribed mean packing fractions (b) $\phi = 0.73$ and (c) $\phi = 0.74$ under the biased-sampling scheme in the *grand canonical* ensemble. (d) The equilibration time τ for $G_V(r)$ in the range $[0, 1.5\mathcal{D})$ at $\phi = 0.7, 0.72, 0.73$, and 0.74 in the grand canonical ensemble.

25 August 2024 14:56:01

Published under an exclusive license by AIP Publishing

in the grand canonical ensemble is able to very accurately determine $G_V(r)$ up to $r = 1.95\mathcal{D}$ using merely $\langle N \rangle = 200$ particles. Compared to biased sampling in the canonical ensemble, the grand canonical ensemble further improves the sensitivity and accuracy of hole sampling because holes can be readily created via the deletion of particles. Importantly, due to the strong finite-size effect associated with the canonical ensemble, we subsequently exclusively use the grand canonical ensemble in our large-scale calculations for 2D and 3D crystal states.

B. Computational complexity for crystal states

To ascertain the accuracy and efficiency of our biased-sampling method to study the hole statistics of 2D and 3D crystal states, we first perform a preliminary study by computing $G_V(r)$ for the 2D equilibrium hard-disk crystal at $\phi = 0.73$ in the *canonical* ensemble using both the biased and the standard (unbiased) methods, as shown in Fig. 2(a). Here, the biasing potential (plotted in the inset) is determined via the iterative procedure in Sec. III, such that $p_i(r)$ is approximately uniform on the range $[0, \mathcal{D}]$ [see Eq. (13)]. Each simulation trajectory, in which 2500 configurations with $N = 256$ particles are generated, takes around 30 min to complete on a 2.8 GHz Intel CPU. We find that with 100 trajectories, $G_V(r)$ via the standard unbiased method has significant noise at $r \sim 0.75\mathcal{D}$, and no hole with radii larger than $0.80\mathcal{D}$ is found. This limitation is due to the fact that $E_V(0.8\mathcal{D}) \sim 10^{-5}$, which is the smallest probability that standard simulations on this scale can detect. By contrast, the biased-sampling method reaches comparative precision with the unbiased method at $r \sim 0.75\mathcal{D}$ with as few as ten trajectories and can detect larger holes with radii up to $\mathcal{D}$. For 100 trajectories, the unbiased method produces accurate hole statistics on the range of interest $[0, \mathcal{D}]$. Notably, according to our biased-sampling result, $E_V(\mathcal{D}) \sim 10^{-10}$, which is 10^{-5} of the sensitivity limit of the unbiased method. Thus, it would require the unbiased method on the order of 10^5 h to reach comparative accuracy at $r \sim \mathcal{D}$ with the biased-sampling method. We remark that in this preliminary investigation, the canonical ensemble is used merely to facilitate the comparison between biased and unbiased algorithms. In our subsequent large-scale calculations (200 trajectories of 10^4 configurations), we exclusively use biased-sampling in the grand canonical ensemble.

Prior to performing large-scale simulations for 2D and 3D crystal states, we first determine the ranges $[0, R)$ on which one can accurately sample the hole statistics at given values of ϕ within 2×10^6 sweeps, which takes ~24 h for ~250 particles in two dimensions. To determine R, we estimate the equilibration time τ required to achieve the convergence condition in Sec. III using the biased-sampling scheme in the grand canonical ensemble for 2D crystals slightly above the hexatic–solid transition packing fraction $\phi_h = 0.72$.[11] Figures 2(b) and 2(c) show the evolution of $G_V(r)$ for the 2D crystal at $\phi = 0.73$ and 0.74, respectively, under their corresponding final biasing potentials, starting from the perfect triangle-lattice configurations at the prescribed mean packing fractions in rhombic simulation cells with $N = 256$. At $\phi = 0.73$, $G_V(r)$ in the range $[0, 1.25\mathcal{D})$ reaches a steady state within the first 200 000 sweeps. The range $[1.25\mathcal{D}, 1.95\mathcal{D})$ is equilibrated within the first 800 000 sweeps. The significant increase in the equilibration time as r increases is due to the low transition rate between the first, second, and third stable hole sizes, with $r \sim \mathcal{D}, 1.5\mathcal{D}$ and $2\mathcal{D}$, respectively. While the creation of large holes is favored energetically due to the biasing potential, the dynamics of such processes is slow, as particles in the vicinity of the test point must undergo large displacements compared to the mean displacement magnitudes in the unbiased equilibrium crystal. At $\phi = 0.74$, we find that 2×10^6 sweeps are required to equilibrate $G_V(r)$ up to $1.5\mathcal{D}$. Due to the dramatic increase in equilibration time as ϕ increases, as shown in Fig. 2(d), we choose decreasing values of R as ϕ increases, and our choices of R at each 2D and 3D packing fraction are presented in Table I. Except in the case $d = 2$, $\phi = 0.73$, we always discard the data between $1.95\mathcal{D}$ and $2\mathcal{D}$ because the imposed infinite biasing potential at $r = 2\mathcal{D}$ means that no holes with radii large than $2\mathcal{D}$ can be created, which causes an anomalous sharp increase in $G_V(r)$ just before $2\mathcal{D}$. However, in just one case $d = 2, \phi = 0.73$, we set the hard wall in the biasing potential to be at $2.1\mathcal{D}$ rather than $2\mathcal{D}$, so that the positions of the local extrema in $G_V(r)$ up to $2\mathcal{D}$ can be accurately determined. We select this packing fraction since it lies fully in the crystal phase and its corresponding $G_V(r)$ can be readily computed in 2×10^6 sweeps. We find that in this case, errors in $G_V(r)$ on $[0, 2\mathcal{R}]$ are within 5%. Notably, $G_V(r)$ in the range $[0, 1.1\mathcal{D})$ is equilibrated within the first 200 000 sweeps for all 2D cases with $\phi \leq 0.85$ and 3D cases with $\phi \leq 0.60$. The total computation time for all biased-sampling calculations is ~300 h under the respective final biasing potentials with 200 parallel threads. In addition, the computation resources needed to evolve the biasing potential prior to convergence is 3–10 times of the final simulation time. We remark that at any ϕ value presented in Table I, the biased-sampling scheme improves the sensitivity limit and thus the computational time for reaching comparable accuracy, by at least five orders of magnitude compared to the unbiased method.

C. 2D hard-disk states

Here, we present the results for the void nearest-neighbor functions for 2D equilibrium hard-disk crystals, as well as those for a high-density liquid and a hexatic state. Furthermore, we study the

TABLE I. Upper radius below which $G_V(r)$ can be determined within ±3% error in 2×10^6 sweeps.

ϕ for 2D states	$R/\mathcal{D}$	ϕ for 3D states	$R/\mathcal{D}$
≤0.72	1.95	0.54	1.5
0.73	1.95; errors are within 5% with $R = 2\mathcal{D}$.	0.55, 0.60	1.1
0.74	1.5		
0.75	1.3		
0.76, 0.78, 0.85	1.1		

25 August 2024 14:56:01

Published under an exclusive license by AIP Publishing

correlation between large-r behaviors of $G_V(r)$ and the local coordination geometry of holes, which could not be done in previous studies due to the absence of accurate hole statistics for $r \geq \mathcal{D}$.

1. Crystal states

Figure 3(a) shows the simulated results of $G_V(r)$ for the 2D hard-disk crystal at the prescribed mean packing fraction $\phi = 0.73$, which is slightly above the packing fraction $\phi_h = 0.72$ corresponding to the second-order phase transition between the hexatic phase and the crystal phase.[11] The grand-canonical biased-sampling scheme enables us to determine accurate $G_V(r)$ within ±5% error up to $r = 2\mathcal{D}$. We note that $E_V(2\mathcal{D}) \sim 10^{-48}$, which is more than 40 orders of magnitude smaller than the sensitivity limit of the unbiased method described in Sec. I. Importantly, we find, for the first time, that the simulated $G_V(r)$ for the crystal state exhibits significant oscillations. Such oscillations persist on a large range of r up to $r = 2\mathcal{D}$, and the period of the oscillations is ~$0.5\mathcal{D}$. Figure 3(b) plots the final biasing potential $u(r)$ used in the grand-canonical biased-sampling scheme to compute the void nearest-neighbor functions for the 2D equilibrium hard-disk crystal at $\phi = 0.73$. It should be noted that $u(r)$ is an increasing function with r on $[0, \mathcal{D}/2]$ because $H_V(r)$ increases on this range; see (7a). Thus, $u(r)$ must increase with r to sample these small holes uniformly. However, $u(r)$ for r larger than the disk radius is strongly repulsive, so that large holes up to $r = 2\mathcal{D}$ can be adequately sampled. We remark that, up to an additive constant, the functional form of $u(r)$ is very close to that of $\ln[H_V(r)]$ on the range of interest $[0, 2\mathcal{D}]$, as expected from the updating rule (12) and the convergence criterion (13). Indeed, the difference between $u(r)$ and $\ln[H_V(r)]$ is within 2% for all r in the

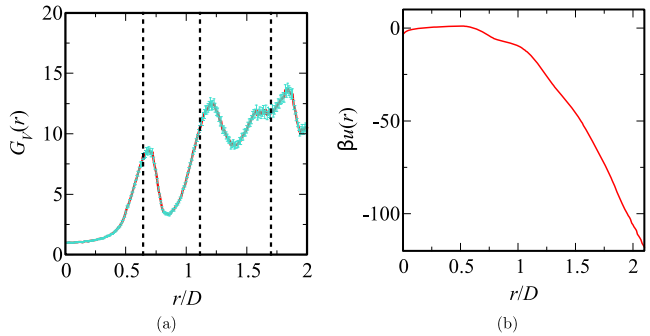


FIG. 3. (a) Function $G_V(r)$ for the 2D equilibrium hard-disk crystal at $\phi = 0.73$, $\phi/\phi_J = 0.805$ obtained from the grand-canonical biased-sampling scheme. The vertical lines at $r = 0.6435\mathcal{D}$, $1.115\mathcal{D}$ and $1.702\mathcal{D}$ correspond to the radii of the circumscribed circles of the idealized holes shown in Figs. 5(a)–5(c), which provide lower bounds of the first three maxima in $G_V(r)$. (b) The final biasing potential $u(r)$ used in the biased-sampling scheme in the grand canonical ensemble to compute the void nearest-neighbor functions for the 2D equilibrium hard-disk crystal at $\phi = 0.73$.

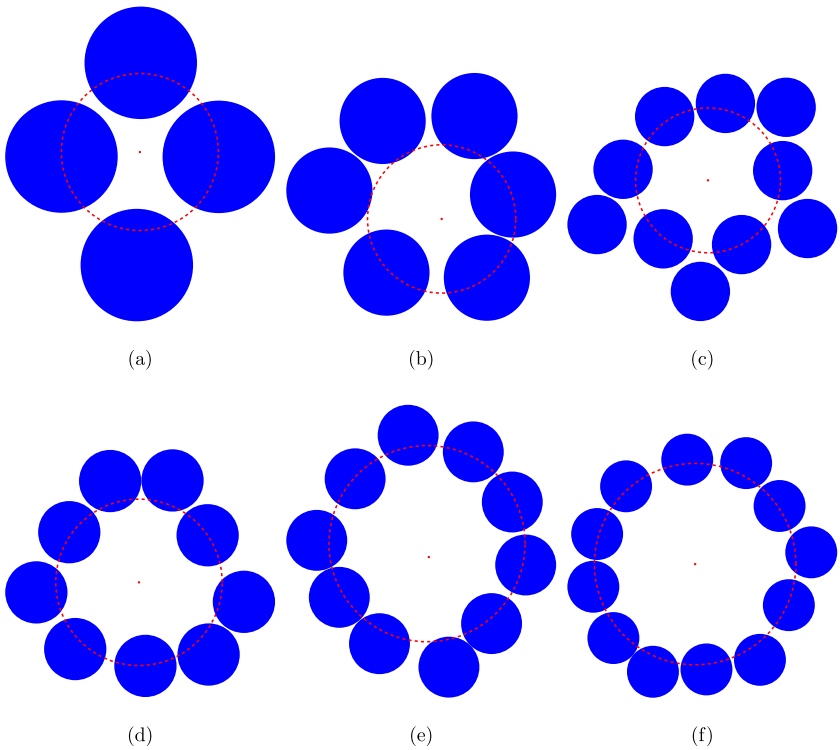


FIG. 4. Local coordination configurations of circular-like holes in an equilibrium hard-disk crystal at $\phi = 0.73$. (a) A hole of radius $r = 0.69\mathcal{D}$, corresponding to a local maximum in $G_V(r)$. (b) A hole of radius $r = 0.86\mathcal{D}$, corresponding to a local minimum in $G_V(r)$. (c) A hole of radius $r = 1.20\mathcal{D}$, corresponding to a local maximum in $G_V(r)$. (d) A hole of radius $r = 1.37\mathcal{D}$, corresponding to a local minimum in $G_V(r)$. (e) A hole of radius $r = 1.61\mathcal{D}$, corresponding to a shoulder in $G_V(r)$. (f) A hole of radius $r = 1.95\mathcal{D}$, corresponding to a local minimum in $G_V(r)$.

25 August 2024 14:56:01

Published under an exclusive license by AIP Publishing

range of interest. Thus, the oscillations in $u(r)$ imply that $\ln[H_V(r)]$ also contains such oscillations.

To investigate the origin of the oscillations in $G_V(r)$ for crystals, we study the local coordination geometry of circular-like holes in the 2D crystal configurations. We exclusively consider 2D circular-like holes, because $G_V(r)$ explicitly samples d-dimensional spherical holes. Indeed, we find that $G_V(r)$ is relatively insensitive to aspherical holes. Here, we define the coordination number Z of a hole of radius r as the number of disks that overlap with the circle of radius r centered at the test point. Figure 4 shows the representative coordination geometries in the equilibrium hard-disk crystal at $\phi = 0.73$. The first local maximum in $G_V(r)$, occurring at $r = 0.69\mathcal{D}$, corresponds to four-coordinated holes [Fig. 4(a)]. As expected, this hole radius is approximately half of the diagonal length of a square with side length $\mathcal{D}$. The first local minimum in $G_V(r)$ at $r = 0.86\mathcal{D}$ corresponds to hexagonal holes with $Z = 6$ [Fig. 4(b)]. Figure 4(c) shows that the second local maximum in $G_V(r)$, occurring at $r = 1.20\mathcal{D}$, corresponds to a significantly distorted hexagonal hole, i.e., the hexagon formed by the centers of the nearest six hard disks to the test point is distinctly different from a regular hexagon. Figures 4(d) and 4(f) show that the second and third local minima in $G_V(r)$, at $r = 1.37\mathcal{D}$ and $r = 1.95\mathcal{D}$, respectively, correspond to 9- and 12-coordinated holes that are created by removing three and seven particles in the triangle lattices, respectively. We also find that the shoulder of $G_V(r)$ at $r \sim 1.65\mathcal{D}$ corresponds to ten-coordinated holes [Fig. 4(e)]. These observations suggest that the minima in $G_V(r)$ corresponding to "stable" hole radii are compatible with the local configuration of the triangle-lattice crystal, whereas maxima in $G_V(r)$ correspond to hole sizes that lead to significant distortion of the triangle lattice and represent the transition state between stable hole sizes.

To explain quantitatively the positions of the maxima in $G_V(r)$, we compare the local coordination geometry of the stable holes in equilibrium hard-disk systems at $\phi = 0.73$ with the corresponding idealized holes in a perfect triangle lattice, as shown in Fig. 5. The idealized triangular (3-coordinated) and hexagonal (6-coordinated) holes have well-defined radii of $r/\mathcal{D} = 0.6435$ and 1.115, respectively, which are the radii of the circumscribed circles of the polygons formed by the disk centers. The 9- and 12-coordinated holes have two characteristic length scales, given by the radii of the inscribed and the circumscribed circles. We find that the radii of the circumscribed circles of the 3-, 6-, and 9-coordinated idealized holes provide good lower bounds of the first three maxima in $G_V(r)$, respectively; see Fig. 3(a). These lower bounds are expected since to create holes slightly larger than the circumscribed circles of the idealized holes, all particles surrounding the hole must be simultaneously displaced away from the hole center. Such significant distortions of the crystal geometry result in maxima in $G_V(r)$. While we do not report data for $G_V(r)$ beyond $2\mathcal{D}$, we expect another local maximum of $G_V(r)$ to occur just above the radius of the circumscribed circle of the idealized 12-coordinated hole at $r = 2.229\mathcal{D}$. On the other hand, the local minima in $G_V(r)$, which indicate stable circular-like hole sizes, occur at r values smaller than the radii for the corresponding idealized holes shown in Fig. 5.[43] This observation is consistent with the well-known effect that a vacancy in an equilibrium crystal has asymmetric fluctuations in the radius of the largest spherical hole that it can contain. The circular-like vacancies tend to compress rather than expand compared to their sizes in the perfect crystal.[26,27]

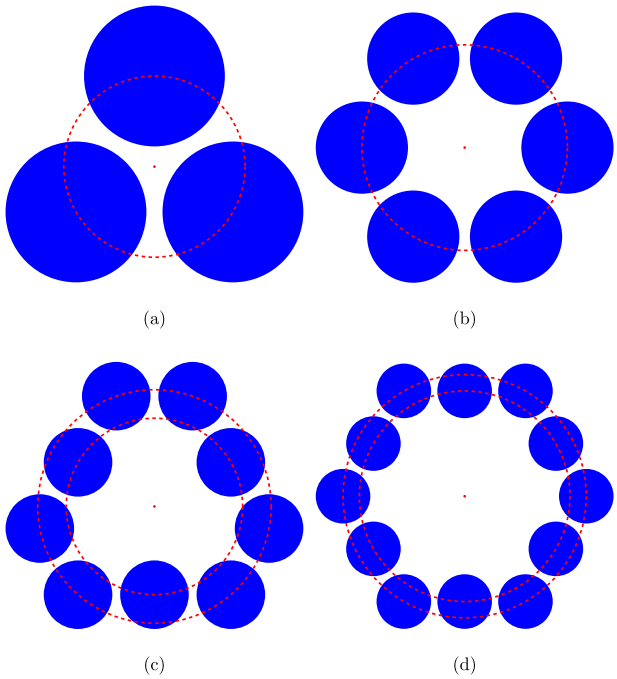


FIG. 5. Local coordination configurations of circular-like holes in a perfect hard-disk crystal at $\phi = 0.73$. (a) A triangular hole of radius $r = 0.6435\mathcal{D}$, which is the covering radius of the perfect crystal. (b) A hexagonal hole of radius $r = 1.115\mathcal{D}$. (c) A 9-coordinated hole with the radius of the inscribed circle $r = 1.287\mathcal{D}$ and radius of the circumscribed circle $r = 1.702\mathcal{D}$. (d) A 12-coordinated hole of radius of the inscribed circle $r = 1.931\mathcal{D}$ and radius of the circumscribed circle $r = 2.229\mathcal{D}$.

Compression of the vacancies increases the free volumes of the remaining particles, thereby increasing the entropy of the crystal. Furthermore, as particles surrounding the vacancies undergo asymmetric displacements, the vacancies become aspherical, which again decreases the radius of the largest spherical hole contained in the vacancy.

We expect the effect of holes on the local crystalline configuration to be reflected in the local orientational order in the vicinity of the holes. To study the correlation between the oscillations in $G_V(r)$ and local orientational order, we compute the bond-orientational order parameter ψ_n,[44] which measures the n-fold orientational order in a 2D many-body system, given by

$$\psi_n = \left| \frac{1}{M} \sum_{j=1}^{M} \exp\left(in\theta_j\right) \right|, \tag{20}$$

where the sum is over M nearest neighbors $\mathbf{r}_j$ of the test particle at $\mathbf{t}$, θ_j is the angle between the vector $\mathbf{r}_j - \mathbf{t}$ and an arbitrary but fixed reference vector.[32,45] In this work, we take $M = 18$, a choice that considers hard disks in the first 2–3 coordination shells of the test point. The red curve in Fig. 6 shows ψ_6 for the 2D hard-disk crystal at $\phi = 0.73$ as a function of the hole radius r. It is clear that ψ_6 in the vicinity of the test point exhibits oscillating behaviors as r increases, with a minimum at $r = 0.86\mathcal{D}$ and a maximum at $r = \mathcal{D}$. This shows that the hexatic orientational order of the triangle lattice is preserved with triangular holes and hexagonal holes, but is

25 August 2024 14:56:01

Published under an exclusive license by AIP Publishing

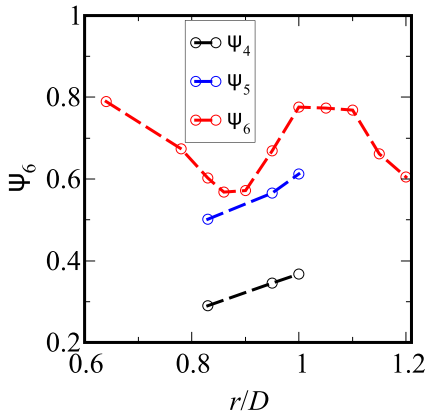


FIG. 6. Local orientational order metrics ψ_4, ψ_5, and ψ_6 (20) for the nearest 18 particles to the test point in a 2D hard disk crystal at $\phi = 0.73$.

disrupted at hole radii that lie between these two "stable" local hole geometries. Figure 6 also shows ψ_4 and ψ_5 with $M = 18$ at hole sizes near $r = 0.86\mathcal{D}$. We observe that while ψ_6 has a local minimum for these hole sizes, it is still larger than ψ_4 and ψ_5. Thus, while these hole sizes may be 4- or 5-coordinated [e.g., as shown in Fig. 4(b)], the second and third coordination shells still possess hexatic orientational order to a larger extent than four- or five-fold orientational order.

Figure 7(a) shows $G_V(r)$ for 2D crystals at various packing fractions. The corresponding hole probability functions $E_V(r)$ are shown in Fig. 12(a) in the Appendix. As ϕ increases, the amplitudes of the oscillations grow rapidly and the first maximum of $G_V(r)$ occurs at decreasing r values. As ϕ approaches the jamming packing fraction $\phi_J = \pi/\sqrt{12}$ for the triangle lattice, the position of the first peak approaches the so-called covering radius $r_c(\phi_J) = \mathcal{D}/\sqrt{3} = 0.577\,\mathcal{D}$,[24] which is the distance from the centroid of the triangular hole to the center of any hard disk on the vertex of the triangular hole. It is known that at $\phi = \phi_J$, $G_V(r)$ develops a pole of order one at $r_c(\phi_J)$.[46] At all 2D packing fractions that we study, we find that the covering radius $r_c(\phi)$ provides a lower bound of the position of the first peak r_m in $G_V(r)$, as shown in Fig. 7(b). In addition, as $\phi \to \phi_J$, the second peak of $G_V(r)$ approaches $G_V(r)$ for a hexagonal hole created by removing one particle in the prefect triangle lattice [dark blue curve shown in Fig. 7(a)], the expression of which is the same as $G_V(r)$ for the densest honeycomb packing of hard disks in the range $\mathcal{D}/\sqrt{3} \le r \le \mathcal{D}$, given in Ref. 24.

2. Dense fluid and hexatic states

Because oscillations in $G_V(r)$ are already significant for 2D crystals at packing fractions well below jamming, we investigate whether they also appear at lower packing fractions, i.e., in dense fluid and hexatic phases. Figures 8(a) and 8(b) shows $G_V(r)$ for the 2D equilibrium hard-disk system in the dense fluid phase at $\phi = 0.670$ and the hexatic phase at $\phi = 0.702$, respectively. It should be noted that in the former case, ϕ is just below the first-order fluid–hexatic transition packing fraction $\phi_f = 0.69$.[6,10] Figure 8(a) also shows the asymptotic expression for $G_V(r)$ along the disordered branch predicted from Eqs. (9) and (10). It is obvious that the simulated $G_V(r)$ for the fluid state is monotonically increasing with r and agrees closely with the asymptotic expression. We remark that the monotonicity of $G_V(r)$ is in strong contrast with the functional form of the radial distribution functions $g_2(r)$ for dense hard-disk fluids, which possess strong oscillations up to $r = 4\mathcal{D}$.[47] On the other hand, Fig. 8(b) shows that $G_V(r)$ for the hexatic state no longer monotonically increases with r. The existence of oscillations in $G_V(r)$ is consistent with the statement that the hexatic phase is orientationally ordered. Figure 8(b) also shows the asymptotic expression for $G_V(r)$ extrapolated from the disordered branch in Eqs. (9) and (10). It is noteworthy that for most values of r on the interval $(\mathcal{D}, 1.95\,\mathcal{D})$, the simulated $G_V(r)$ is smaller than the disordered-branch asymptotic expression. The simulated value of $G_V(\mathcal{D})$ is also smaller than the value predicted via Eqs. (9) and (10). These differences clearly demonstrate that the hexatic phase cannot be viewed as an extrapolation of the disordered branch. Using Eq. (5), we deduce that the

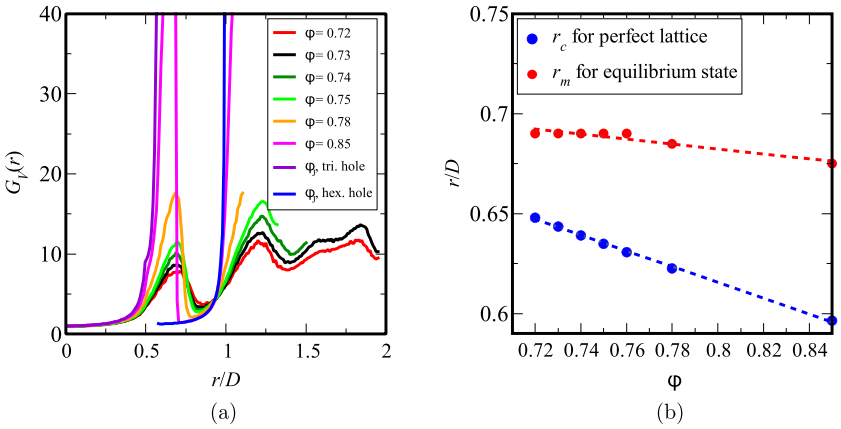


FIG. 7. (a) Function $G_V(r)$ for 2D hard-disk crystals at different packing fractions obtained from the grand-canonical biased-sampling scheme. (b) Plots of the covering radius $r_c(\phi)$ of the perfect triangle lattice and the position of the first peak in $G_V(r)$ for equilibrium hard-disk crystals $r_m(\phi)$.

25 August 2024 14:56:01

Published under an exclusive license by AIP Publishing

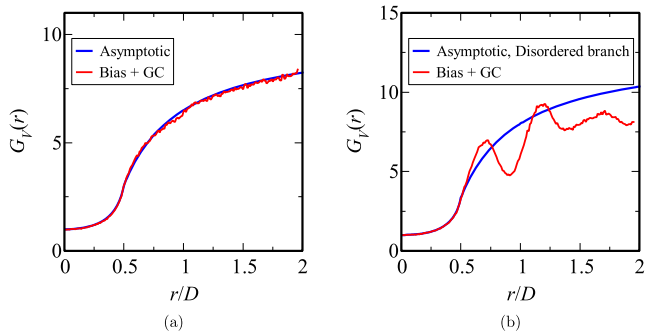


FIG. 8. (a) Function $G_V(r)$ for 2D hard-disk fluid at $\phi = 0.670$ obtained from the grand-canonical biased-sampling scheme, as well as the asymptotic expression via Eqs. (9) and (10). (b) The function $G_V(r)$ for 2D hard-disk hexatic state at $\phi = 0.702$ obtained from the grand-canonical biased-sampling scheme, as well as the asymptotic expression extrapolated from the disordered branch [Eqs. (9) and (10)]. The significant differences between the two curves show that the hexatic phase cannot be viewed as an extrapolation of the disordered branch.

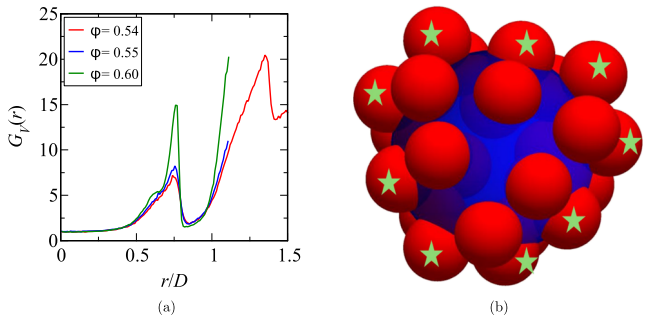


FIG. 9. (a) Function $G_V(r)$ for the 3D hard-sphere systems at $\phi = 0.54, 0.55$ and 0.60 obtained via biased sampling in the grand canonical ensemble. (b) Local coordination configuration of a hole (indicated with a blue sphere) in an equilibrium hard-sphere system at $\phi = 0.54$ of radius $r = 1.47\mathcal{D}$, approximately at the second minimum in $G_V(r)$. The nine hard-sphere particles along a great circle of this hole are indicated with stars.

pressure of the hexatic phase is lower than the disordered metastable state at the same packing fraction.

In summary, $G_V(r)$ for 2D hard disks monotonically increases with r for fluid states, is oscillatory with approximately symmetric peaks for hexatic phases, and is oscillatory with highly asymmetric peaks for crystals. These significant differences in $G_V(r)$ are consistent with the well-known fact that the translational and orientational correlation functions behave differently for the three phases.[10,11] In contrast to commonly used translational[48] and orientational[32] order metrics, which are real-valued numbers that describe global properties of many-body systems averaged over multiple length scales, $G_V(r)$ provides detailed scale-dependent information on the variation of the hole probability with the hole radius r.

D. 3D FCC hard-sphere crystals

Figure 9(a) shows $G_V(r)$ for the equilibrium 3D hard-sphere fcc-lattice crystal at three different packing fractions. The corresponding hole probability functions $E_V(r)$ are shows in Fig. 12(b) in the Appendix. The cases $\phi = 0.55$ and 0.60 are slightly above the melting packing fraction $\phi_m = 0.545$.[6] The case $\phi = 0.54$, slightly below the melting point, is chosen to facilitate an efficient and accurate calculation of $G_V(r)$ up to $r = 1.5\mathcal{D}$, thereby providing preliminary qualitative information about hole statistics for 3D crystals near melting. For $\phi = 0.54$, the proportion of the liquid phase is expected to be negligible with $\langle N \rangle = 256$ particles used in our study since ϕ is very close to ϕ_m.[49] Indeed, we have not observed phase separation at $\phi = 0.54$ in our simulations.

As is the case for the 2D hard-disk crystal, we find significant oscillations in $G_V(r)$ with amplitudes that rapidly increase with ϕ, and the period of the oscillations is approximately $\mathcal{D}/2$. The first minimum of $G_V(r)$ corresponds to a vacancy created by the removal of one particle in the crystal. The coordination number of such holes is $Z = 12$. Figure 9(b) shows a representative local coordination configuration of a hole of radius $r = 1.47\mathcal{D}$ in the hard-sphere system at $\phi = 0.54$, i.e., approximately at the second minimum in $G_V(r)$ at $r = 1.42\mathcal{D}$. Such holes have $Z = 24$ and correspond to cavities in the fcc crystal created by removing four particles that form the vertices of a tetrahedral hole. It should be noted that in a perfect close-packed fcc lattice, there are 25 particles that touch a tetrahedral four-particle cluster. In equilibrium hard-sphere crystals, Z is slightly lower than 25 due to the fact that the 25-coordinated cavity created from the perfect fcc is highly aspherical, and they are distorted in the equilibrium states to yield nearly spherical holes. In particular, we find that there are usually nine hard-sphere particles fitted along a great circle of these holes, indicated with stars shown in Fig. 9(b). These 9-coordinated big circles are analogous to the 2D local coordination configuration shown in Fig. 4(e). It should be noted that the second maximum and minimum in $G_V(r)$, as well as the tetrahedron-like holes, are features that are almost certainly qualitatively correct for the crystal phase near melting. Indeed, the tetrahedron-like holes have also been observed in simulations at $\phi = 0.55$ and 0.60. However, we do not include statistics for these holes shown in Fig. 9(a) because $G_V(r)$ near $r = 1.5\mathcal{D}$ did not converge within 2×10^6 sweeps at these packing fractions.

In contrast to $G_V(r)$ for 2D hard-disk crystals [Figs. 3(a) and 3(c)], we observe in Fig. 9 that for the fcc crystals, the first peak of $G_V(r)$ has a shoulder at $r \sim 0.65\mathcal{D}$, which becomes more apparent as ϕ increases. This is due to the fact that the 2D triangle-lattice packing contains only triangular holes, whereas the 3D fcc packing contains both tetrahedral and octahedral holes. In a perfect fcc lattice with packing fraction ϕ, the radius of the tetrahedral holes is given by

$$r_t(\phi) = \sqrt{\frac{3}{8}}\left(\frac{\phi_J}{\phi}\right)^{1/3}\mathcal{D}, \tag{21}$$

where $\phi_J = \pi/\sqrt{18}$ is the packing fraction for the fcc close packing. Thus, we have $r_t(0.55) = 0.676\mathcal{D}$ and $r_t(0.60) = 0.657\mathcal{D}$, which are approximately the hole radii where the shoulders in $G_V(r)$ appear in Fig. 9. Thus, the shoulders represent the transition between smaller holes that may be derived from either tetrahedral or octahedral holes in the fcc lattice and larger holes that are exclusively derived from octahedral holes. More precisely, in the equilibrium crystal, a spherical cavity of radius $r > r_t(\phi)$ is much more likely to be found within an octahedral hole than in a distorted tetrahedral hole. On the other hand, both tetrahedral and octahedral holes can accommodate spherical cavities of radii smaller than $r_t(\phi)$ without significantly

25 August 2024 14:56:01

Published under an exclusive license by AIP Publishing

distorting the crystal structure. It should also be noted that analogous to the 2D case, $r_t(\phi)$ provides a lower bound of the position of the first peak in $G_V(r)$.

V. LOCAL PACKING FRACTION DISTRIBUTION OF DELAUNAY CELLS

Here, to further characterize the distribution of interparticle spacing in configurations of hard-sphere systems, we study the local packing fraction distribution of Delaunay cells as a function of the global packing fraction ϕ in one, two, and three dimensions. For a point configuration $\mathbf{r}^N$ in $\mathbb{R}^d$, a Delaunay tessellation is a d-dimensional triangulation $DT(\mathbf{r}^N)$ such that no point in $\mathbf{r}^N$ is inside the circum-hypersphere of any d-simplex in $DT(\mathbf{r}^N)$.[50] Here, we use point configurations derived from the sphere centers in equilibrium d-dimensional hard-sphere systems in the canonical ensemble with $N \sim 500$ particles (without the test point) to create the corresponding Delaunay tessellations. We then compute the local packing fraction ϕ_l, i.e., the fraction of the particle phase in each Delaunay cell, and determine the distribution of ϕ_l over all the Delaunay cells in 200 independent configurations, denoted by $f(\phi_l)$. Figure 10(a) shows $f(\phi_l)$ for the equilibrium 2D hard-disk crystal at global packing fraction $\phi = 0.85$. Apparently, $f(\phi_l)$ is asymmetric about ϕ, as it resembles a Gaussian for $\phi_l < \phi$ but decays more rapidly for $\phi_l > \phi$. This is due to the fact that $\phi = 0.85$ is close to the close packing fraction $\phi_f = \pi/\sqrt{12} = 0.9069$, which is an upper bound of ϕ_l.

To more closely investigate the non-Gaussian behaviors of the local packing fraction distribution across ϕ and d, we show in Fig. 10(b) the excess kurtosis of $f(\phi_l)$ for equilibrium 1D high-density hard-rod fluids as well as for equilibrium 2D and 3D hard-sphere crystals. The kurtosis of a random variable X with mean μ and standard deviation σ is given by

$$\mathrm{Kurt}\,(X) = \mathrm{E}\left[\left(\frac{X-\mu}{\sigma}\right)^4\right], \tag{22}$$

and the excess kurtosis is defined as $\mathrm{Kurt}(X) - 3$.[51] A distribution with negative excess kurtosis is known as platykurtic, indicating that it contains fewer or less extreme outliers than the normal distribution of the same mean and standard deviation.[51] Interestingly,

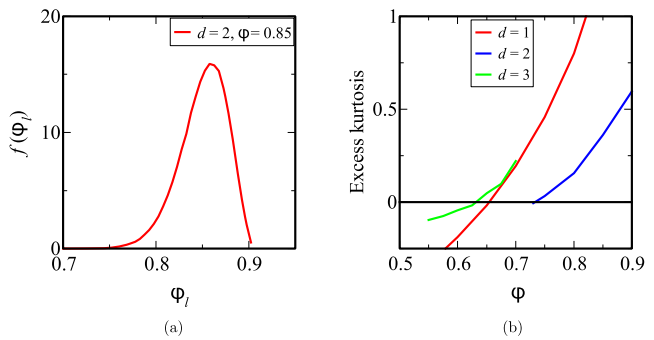


FIG. 10. (a) Probability distribution $f(\phi_l)$ of the local packing fraction of Delaunay cells in the 2D hard-disk crystal at $\phi = 0.85$. (b) The excess kurtosis of $f(\phi_l)$ for high-density hard-sphere systems as a function of ϕ in one, two, and three dimensions.

Fig. 10(b) shows that in each of the first three dimensions, the excess kurtosis increases with ϕ and switches from negative to positive at a certain dimension-dependent transition packing fraction ϕ^*. Indeed, we observe that for ϕ away from jamming, where the excess kurtosis is negative, the shape of $f(\phi_l)$ is approximately symmetric and Gaussian-like. Thus, it is expected that the distribution is platykurtic at low ϕ because $f(\phi_l)$ has a finite support $[0, \phi_J]$, and its tails on both ends must decay faster than a Gaussian. However, at high ϕ close to jamming, $f(\phi_l)$ is asymmetric and has a heavier tail toward smaller ϕ_l, indicating that the hard spheres undergo collective motion to creating regions of significantly lower local packing fractions than ϕ.

For the 2D state, $\phi^* = 0.73$ is slightly above the packing fraction for the hexatic-crystal transition. For the 3D state, $\phi^* = 0.625$ lies between the melting packing fraction and the fcc close-packing fraction. Furthermore, we find that as the dimensionality increases, the magnitudes of the excess kurtosis are closer to zero, i.e., $f(\phi_l)$ becomes closer to the Gaussian distribution. The variation of the excess kurtosis with ϕ is also less steep in higher dimensions. This reflects the decorrelation principle, which states that unconstrained spatial correlations vanish asymptotically for pair distances beyond the hard-core diameter in the high-dimensional limit.[52] As the particles become increasingly decorrelated in higher dimensions, $f(\phi_l)$ is expected to approach a Gaussian as a result of the central limit theorem.

VI. DISCUSSION AND CONCLUSIONS

In summary, we have introduced a biased-sampling scheme that accurately determines void nearest-neighbor functions, i.e., the hole statistics, for dense hard-sphere crystals and liquids on ranges of the hole radius r that far extends the sizes that could be previously explored. Using this algorithm, we find that for 2D hexatic and crystal phases, $G_V(r)$ exhibits oscillations with amplitudes that rapidly increase with the packing fraction. Such oscillations stand in contrast with the 2D fluid states, for which $G_V(r)$ monotonically increases with r up to the fluid-hexatic transition point. For 3D crystals, $G_V(r)$ is also non-monotonic, and a shoulder develops its first peak as ϕ increases, indicating the transition between tetrahedral to octahedral holes in the fcc crystal. We find that the oscillations in $G_V(r)$ for order states are strongly correlated with the local orientation order. For both 2D and 3D systems, minima of $G_V(r)$ correspond to stable hole whose local coordination geometry is compatible with the crystal structure, whereas the maxima of $G_V(r)$ correspond to geometrically strained holes that significantly disturb the local crystal structure. We have also studied the local packing fraction distribution $f(\phi_l)$ of Delaunay cells and find that the excess kurtosis of $f(\phi_l)$ increases with increasing ϕ, and changes sign at a dimension-dependent transition packing fraction ϕ^*. This switch from a platykurtic to a leptokurtic behavior in $f(\phi_l)$ as ϕ increases indicates the transition between predominantly weakly correlated displacements of hard spheres at lower ϕ to significant collective displacements at higher ϕ. It is likely that ϕ^* lies in the crystal phase also for certain higher dimensions $d \geq 4$.

It is important to remark on the relationship between the void nearest-neighbor functions and the n-body correlation function g_n for hard-sphere systems, which has been originally explored by

25 August 2024 14:56:01

Published under an exclusive license by AIP Publishing

Reiss *et al.*[12] The exact series representation of $G_V(r)$ in terms of g_n is given by[18]

$$G_V(r) = \lim_{|\mathbf{x}-\mathbf{r}_1|\to r} \frac{1}{\rho E_V(r)} \times \left[\rho + \sum_{k=1}^{\infty} \frac{(-1)^k \rho^{k+1}}{k!} \int g_{k+1}(\mathbf{r}^{k+1}) \prod_{i=2}^{k+1} m(|\mathbf{x}-\mathbf{r}_i|;r) d\mathbf{r}_i\right], \tag{23}$$

where

$$m(y;r) = \begin{cases} 1, & y < r, \\ 0, & y \geq r, \end{cases} \tag{24}$$

and the series representation of $E_V(r)$ is given by

$$E_V(r) = 1 + \Sigma_{k=1}^{\infty} (-1)^k \frac{\rho^k}{k!} \int g_k(\mathbf{r}_1, \ldots \mathbf{r}_k) v_k^{\text{int}} \times (\mathbf{r}_1, \ldots \mathbf{r}_k; r) d\mathbf{r}_1 \ldots d\mathbf{r}_k, \tag{25}$$

where $v_k^{\text{int}}(\mathbf{r}_1, \ldots \mathbf{r}_k; r)$ is the intersection volume of k equal spheres of radius r centered at positions $\mathbf{r}_1, \ldots, \mathbf{r}_k$. Because $G_V(r)$ involves an integral of $g_2(r)$ in Eq. (23), any singularity in $g_2(r)$ translates to a corresponding singularity in $G_V(r)$ of a higher order. Indeed, Reiss *et al.*[12] explicitly demonstrated that $G_V(r)$ is smoother than $g_2(r)$ in certain ranges because $g_2(r)$ on $[\mathcal{D}, 2\mathcal{D}/\sqrt{3}]$ can be exactly expressed in terms of the first two derivatives of $G_V(r)$ on $[\mathcal{D}/2, \mathcal{D}/\sqrt{3}]$. These authors also show that there exist singularities in $G_V(r)$ at $r = \mathcal{D}/2, \mathcal{D}/\sqrt{3}, \ldots$, i.e., the hole radii at which $g_2, g_3, \ldots$ just become non-vanishing in (23).[12] Moreover, in light of the fact that Stillinger[53] has shown that $g_2(r)$ for 2D and 3D hard spheres contain singularities for all $r > \mathcal{D}$ at positive number density,[54] $G_V(r)$ must also be nonanalytic for all $r \geq \mathcal{D}/2$, with dense singularities of higher orders than the corresponding ones in $g_2(r)$.[55]

Figures 11(a) and 11(b) show the first and second derivatives $G_V'(r/\mathcal{D})$ and $G_V'(r/\mathcal{D})$, respectively, of $G_V(r/\mathcal{D})$ for the 2D hard-disk crystal at $\phi = 0.85$. The first derivative $G_V'(r/\mathcal{D})$ drops sharply to -4×10^3 $r = 0.70\mathcal{D}$. The magnitudes of the second derivative in the vicinity of this radius reach as large as 3×10^5. Thus, for hard-sphere packings close to jamming, $G_V(r)$ can be considered to be nearly discontinuous in the first and higher derivatives. As ϕ approaches ϕ_J, $G_V(r)$ becomes truly discontinuous at the covering radius for the close-packing triangle lattice $r_c(\phi_J) = 0.577\mathcal{D}$.

Apart from the spike, we find that $G_V'(r/\mathcal{D})$ steeply increases at $r/\mathcal{D} = 0.58$ and has a plateau on the range $[0.60, 0.67]$. Small holes of $r \leq 0.58\mathcal{D}$ can be fitted into the triangular holes of the 2D crystal without disturbing the local hexatic symmetry. Thus, increasing r implies that the test point must be placed closer to the centroids of the triangular holes. The increase in $G_V'(r/\mathcal{D})$ due to the effect of test-point positions can be predicted from the expression of $G_V(r)$ for the perfect triangle lattice,[24] which exhibits rapid increases in all derivatives in $G_V(r)$ up to the singularity at the covering radius. By contrast, for the range $[0.60, 0.67]$, the key contribution to variations in $G_V(r)$ is the expansion of the triangular holes. Similarly, the increase in $G_V'(r/\mathcal{D})$ at $r = \mathcal{D}$ and the plateau beyond this radius are due to the position of the test point toward the centroids of the hexagonal holes and the expansion of the hexagonal holes, respectively. We further remark that except for the spikes, $G_V'(r/\mathcal{D})$ exhibits less significant oscillations than $G_V(r)$ itself, and the second derivative $G_V''(r/\mathcal{D})$ [curvature of $G_V(r)$] is rather constant in r. These behaviors suggest that the discontinuities in the first and second derivatives of $G_V(r)$ are primarily due to the abrupt transition between triangular and hexatic stable holes at $r \sim 0.70\mathcal{D}$, and $G_V(r)$ values away from this transitional radius can be well approximated by polynomials or rational functions.

Our investigation of hard spheres facilitates the study of structural and bulk properties of realistic materials. While in actual crystalline materials, the interactions between the constituent particles are different from hard-sphere interactions, we expect similar oscillatory behavior of $G_V(r)$ to occur, since the hard-sphere system is a general idealized model for systems with strong short-ranged

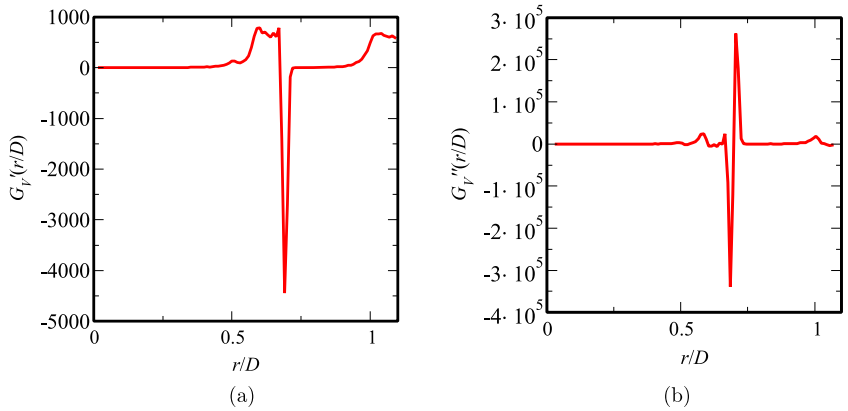


FIG. 11. (a) First and (b) second derivatives $G_V'(r/\mathcal{D})$ and $G_V''(r/\mathcal{D})$, respectively, of $G_V(r/\mathcal{D})$ for the 2D hard-disk crystal at $\phi = 0.85$. The first derivative increases steeply at $r/\mathcal{D} = 0.58$ and 1.0, and plateaus in the ranges $[0.60, 0.67]$ and $[1, 1.2]$. Spikes with very large magnitudes are observed in both derivatives, indicating that $G_V(r)$ is nearly discontinuous at $r \sim 0.70\mathcal{D}$. Except for the spike, the second derivative ([curvature of $G_V(r)$]) is relatively flat.

25 August 2024 14:56:01

Published under an exclusive license by AIP Publishing

repulsive interactions. Furthermore, our investigation of the local density distribution can also be generalized to systems with non-hard-core interactions, such as the well-known Lennard-Jones or Gaussian-core models.[56] For such systems, ϕ_l is not well-defined, but one can study the distribution of the inverse volumes of the Delaunay cells, and we expect that the kurtosis of this distribution to increase with the global number density ρ.

Finally, we remark that because a hole in a many-body system can be considered as a "solute" particle in a "solvent" of particles,[12] the accurate methods that we have developed to access solvation thermodynamics in hard-sphere crystals at the challenging intermediate length scales reported here can be applied to understand the important problem of solvation and hydrophobicity in water at such length scales.[57–62] The hydrophobic solubility has been extensively studied in the framework of scaled particle theory,[58,59,63] and hard-sphere models have often been used as idealized reference systems to understand the formation of cavities in water.[64,65] Importantly, both hard-sphere fluids and liquid water have the property that the probability of observing N particle centers within molecular-sized volumes of the fluids are governed by the Gaussian distribution.[65] Thus, the Pratt–Chandler theory of hydrophobicity,[66] based on the Gaussian-field model for fluids,[64] applies to both systems. While analytical expressions for the hydrophobic chemical potential are available in the limit of small and large solute sizes,[59,63] the hydrophobicity of intermediate-sized solute particles remains a challenging problem. Our findings highlight that the solvation thermodynamics on the intermediate scale in dense, structured fluid or solid solvents cannot be described by a simple interpolation between small- and large-scale behaviors, as is sometimes assumed.[63] For example, the geometric compatibility of spherical cavities with the crystal structures is manifested by the oscillations in $G_V(r)$ that cannot exist in any simple interpolation between the small- and large-r behaviors of $G_V(r)$. The ability to predict void probabilities (and, by extension, solvation thermodynamics) at intermediate length scales that match the numerical results in our work is a stringent new test for theories of condensed matter, e.g., classical density functional theory.

While the radial void nearest-neighbor functions [such as $G_V(r)$] are essential in the study of thermodynamic properties of hard-sphere systems (see Sec. II), they do not explicitly reveal detailed statistics for aspherical holes. In equilibrium crystals, it is possible to create aspherical vacancies. For example, removing two neighboring particles in 2D and 3D crystals generates an elongated vacancy, whereas removing three neighboring particles in a 3D crystal generates a flat vacancy. In such cases, $G_V(r)$ samples spherical void regions contained within such vacancies, and information about the shape of larger aspherical vacancies may be hidden in its higher derivatives. However, one could probe the statistics of aspherical holes by considering the analogous vector-dependent function $G_V(\mathbf{r})$, which is proportional to the conditional probability of finding a particle at $\mathbf{r}$ from the test point, given that no particle center is found on the ray $\{\lambda\mathbf{r} : \lambda \in [0,1)\}$. For example, in two dimensions, one can define $G_V(\mathbf{r}) = G_V(r,\theta)$ such that $\rho G_V(r,\theta) r dr d\theta$ is the conditional probability of finding a particle center at polar coordinates (r,θ), given that no particle center is found in the sector defined by the region $\{(r',\theta') : r' \in [0,r), \theta' \in (\theta - d\theta/2, \theta + d\theta/2)\}$. Precise determination of the vector-dependent hole statistics can enable the study of solubility in molecular solid solutions in which the solutes are typically aspherical molecules.[67]

ACKNOWLEDGMENTS

H.W. and S.T. were supported by National Science Foundation Grant No. CBET-2133179. D.A.H. was supported by National Science Foundation QLCI Grant No. OMA-2120757.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Haina Wang: Data curation (lead); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (lead); Validation (equal); Writing – review & editing (equal). **David A. Huse**: Conceptualization (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Validation (equal); Writing – review & editing (equal). **Salvatore Torquato**: Conceptualization (lead); Formal analysis (equal); Investigation (equal); Methodology (equal); Validation (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

APPENDIX: HOLE PROBABILITIES

Figure 12 shows plots of the hole probability functions $E_V(r)$ for 2D hard-disk crystals and 3D hard-sphere systems. They correspond to the functions $G_V(r)$ shown in Figs. 7(a) and 9(a), respectively, as obtained via relation (4). It should be noted that at higher packing fractions, $E_V(r)$ exhibits a rapid change in its first derivative $E_V'(r)$ at $r \sim 0.70\mathcal{D}$ in the 2D case and at $r \sim 0.80\mathcal{D}$ in the 3D case. In two dimensions, the sharp change in $E_V'(r)$ represents the abrupt transition between unstable distorted triangular holes to stable hexagonal holes created by monovacancies. Similarly, in three dimensions, this behavior represents the transition between distorted octahedral holes to 12-coordinated monovacancies.

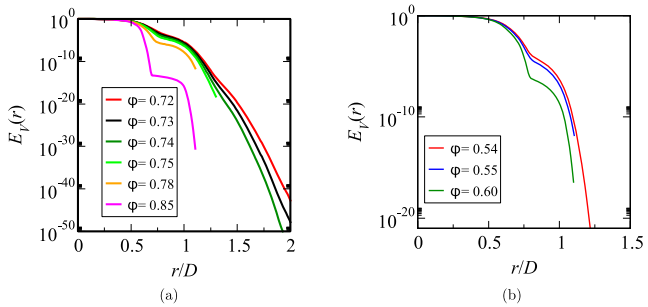


FIG. 12. (a) Hole probability function $E_V(r)$ for (a) 2D hard-disk crystals and (b) 3D hard-sphere systems at different packing fractions obtained from the grand-canonical biased-sampling scheme.

25 August 2024 14:56:01

Published under an exclusive license by AIP Publishing

REFERENCES

[1] J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic Press, New York, 1986).
[2] S. Torquato, *Random Heterogeneous Materials: Microstructure and Macroscopic Properties* (Springer-Verlag, New York, 2002).
[3] S. Torquato and F. H. Stillinger, "Jammed hard-particle packings: From Kepler to Bernal and beyond," Rev. Mod. Phys. **82**, 2633 (2010).
[4] B. J. Alder and T. E. Wainwright, "Phase transition for a hard sphere system," J. Chem. Phys. **27**, 1208–1209 (1957).
[5] B. J. Alder and T. E. Wainwright, "Phase transition in elastic disks," Phys. Rev. **127**, 359–361 (1962).
[6] W. G. Hoover and F. H. Ree, "Melting transition and communal entropy for hard spheres," J. Chem. Phys. **49**, 3609–3617 (1968).
[7] P. G. Bolhuis, D. Frenkel, S.-C. Mau, and D. A. Huse, "Entropy difference between crystal phases," Nature **388**, 235–236 (1997).
[8] S. C. Mau and D. A. Huse, "Stacking entropy of hard-sphere crystals," Phys. Rev. E **59**, 4396–4401 (1999).
[9] C. H. Mak, "Large-scale simulations of the two-dimensional melting of hard disks," Phys. Rev. E **73**, 065104 (2006).
[10] J. J. Kozak, J. Brzezinski, and S. A. Rice, "A conjecture concerning the symmetries of planar nets and the hard disk freezing transition," J. Phys. Chem. B **112**, 16059–16069 (2008).
[11] E. P. Bernard and W. Krauth, "Two-step melting in two dimensions: First-order liquid-hexatic transition," Phys. Rev. Lett. **107**, 155704 (2011).
[12] H. Reiss, H. L. Frisch, and J. L. Lebowitz, "Statistical mechanics of rigid spheres," J. Chem. Phys. **31**, 369–380 (1959).
[13] S. Torquato, B. Lu, and J. Rubinstein, "Nearest-neighbour distribution function for systems on interacting particles," J. Phys. A: Math. Gen. **23**, L103–L107 (1990).
[14] E. Helfand, H. Reiss, H. L. Frisch, and J. L. Lebowitz, "Scaled particle theory of fluids," J. Chem. Phys. **33**, 1379–1385 (1960).
[15] S. Torquato, "Mean nearest-neighbor distance in random packings of hard D-dimensional spheres," Phys. Rev. Lett. **74**, 2156–2159 (1995).
[16] S. Torquato and S. B. Lee, "Computer simulations of nearest-neighbor distribution functions and related quantities for hard-sphere systems," Physica A **167**, 361–383 (1990).
[17] H. Wang and S. Torquato, "Equilibrium states corresponding to targeted hyperuniform nonequilibrium pair statistics," Soft Matter **19**, 550–564 (2023).
[18] S. Torquato, B. Lu, and J. Rubinstein, "Nearest-neighbor distribution functions in many-body systems," Phys. Rev. A **41**, 2059–2075 (1990).
[19] M. D. Rintoul, S. Torquato, and G. Tarjus, "Nearest-neighbor statistics in a one-dimensional random sequential adsorption process," Phys. Rev. E **53**, 450–457 (1996).
[20] M. Larcher and J. T. Jenkins, "Segregation and mixture profiles in dense, inclined flows of two types of spheres," Phys. Fluids **25**, 113301 (2013).
[21] S. Chandrasekhar, "Stochastic problems in physics and astronomy," Rev. Mod. Phys. **15**, 1–89 (1943).
[22] G. Zhang, F. H. Stillinger, and S. Torquato, "Can exotic disordered 'stealthy' particle configurations tolerate arbitrarily large holes?," Soft Matter **13**, 6197–6207 (2017).
[23] M. A. Klatt, P. J. Steinhardt, and S. Torquato, "Wave propagation and band tails of two-dimensional disordered systems in the thermodynamic limit," Proc. Natl. Acad. Sci. U. S. A. **119**, e2213633119 (2022).
[24] S. Torquato, "Reformulation of the covering and quantizer problems as ground states of interacting particles," Phys. Rev. E **82**, 056109 (2010).
[25] Z. W. Salsburg, R. W. Zwanzig, and J. G. Kirkwood, "Molecular distribution functions in a one-dimensional fluid," J. Chem. Phys. **21**, 1098–1107 (1953).
[26] M. A. Bates and D. Frenkel, "Influence of vacancies on the melting transition of hard disks in two dimensions," Phys. Rev. E **61**, 5223–5227 (2000).
[27] S. Pronk and D. Frenkel, "Point defects in hard-sphere crystals," J. Phys. Chem. B **105**, 6722–6727 (2001).
[28] R. Roth, K. Mecke, and M. Oettel, "Communication: Fundamental measure theory for hard disks: Fluid and solid," J. Chem. Phys. **136**, 081101 (2012).
[29] E. A. Ustinov, "Thermodynamics and simulation of hard-sphere fluid and solid: Kinetic Monte Carlo method versus standard Metropolis scheme," J. Chem. Phys. **146**, 034110 (2017).
[30] A. Purohit, A. J. Schultz, S. G. Moustafa, J. R. Errington, and D. A. Kofke, "Free energy and concentration of crystalline vacancies by molecular simulation," Mol. Phys. **116**, 3027–3041 (2018).
[31] Because $H_V(r)$ for hard-sphere systems decay asymptotically as $\exp(-r^d)$,[16] it is not possible to significantly improve upon this computational limit by simply scaling up computational resources.
[32] D. R. Nelson and B. I. Halperin, "Dislocation-mediated melting in two dimensions," Phys. Rev. B **19**, 2457–2484 (1979).
[33] P. L. Geissler and D. Chandler, "Importance sampling and theory of nonequilibrium solvation dynamics in water," J. Chem. Phys. **113**, 9759–9765 (2000).
[34] B. G. Keller and P. G. Bolhuis, "Dynamical reweighting for biased rare event simulations," Annu. Rev. Phys. Chem. **75**, 137–162 (2024).
[35] M. Heying and D. S. Corti, "Scaled particle theory revisited: New conditions and improved predictions of the properties of the hard sphere fluid," J. Phys. Chem. B **108**, 19756–19768 (2004).
[36] F. H. Stillinger, Jr., Z. W. Salsburg, R. L. Kornegay, and R. L. Kornegay, "Rigid disks at high density," J. Chem. Phys. **43**, 932–943 (1965).
[37] Z. W. Salsburg and W. W. Wood, "Equation of state of classical hard spheres at high density," J. Chem. Phys. **37**, 798–804 (1962).
[38] K. K. Mon and R. B. Griffiths, "Chemical potential by gradual insertion of a particle in Monte Carlo simulation," Phys. Rev. A **31**, 956–959 (1985).
[39] B. Widom, "Potential-distribution theory and the statistical mechanics of fluids," J. Phys. Chem. **86**, 869–872 (1982).
[40] J. Jover, A. J. Haslam, A. Galindo, G. Jackson, and E. A. Müller, "Pseudo hard-sphere potential for use in continuous molecular-dynamics simulation of spherical and chain molecules," J. Chem. Phys. **137**, 144505 (2012).
[41] J. Revels, M. Lubin, and T. Papamarkou, "Forward-mode automatic differentiation in Julia," arXiv:1607.07892 [cs.MS] (2016).
[42] H. Wang, F. H. Stillinger, and S. Torquato, "Sensitivity of pair statistics on pair potentials in many-body systems," J. Chem. Phys. **153**, 124106 (2020).
[43] For 9- and 12-coordinated holes, the hole radius is taken to be the average of the radii of the inscribed and the circumscribed circles.
[44] D. Levine and P. J. Steinhardt, "Quasicrystals. I. Definition and structure," Phys. Rev. B **34**, 596–616 (1986).
[45] The special case of ψ_6 was first proposed by Nelson and Halperin.
[46] T. M. Middlemas and S. Torquato, "Nearest-neighbor functions for disordered stealthy hyperuniform many-particle systems," J. Stat. Mech.: Theory Exp. **2020**, 103302.
[47] F. Moučka and I. Nezbeda, "Detection and characterization of structural changes in the hard-disk fluid under freezing and melting conditions," Phys. Rev. Lett. **94**, 040601 (2005).
[48] S. Torquato, G. Zhang, and F. H. Stillinger, "Ensemble theory for stealthy hyperuniform disordered ground states," Phys. Rev. X **5**, 021020 (2015).
[49] It would be useful for a future study to examine substantially larger systems to test this expectation.
[50] D. F. Watson, "Computing the n-dimensional Delaunay tessellation with application to Voronoi polytopes," Comput. J. **24**, 167–172 (1981).
[51] K. P. Balanda and H. L. Macgillivray, "Kurtosis: A critical review," Am. Stat. **42**, 111–119 (1988).
[52] S. Torquato and F. H. Stillinger, "New conjectural lower bounds on the optimal density of sphere packings," Exp. Math. **15**, 307–331 (2006).
[53] F. H. Stillinger, "Pair distribution in the classical rigid disk and sphere systems," J. Comput. Phys. **7**, 367–384 (1971).
[54] Here, the condition of strictly positive density is important to show the dense singularities in $g_2(r)$, because the virial expansion does not show singularities for all r.
[55] We remark that Reiss *et al.*[12] were not aware of the result in Ref. 53 that $g_2(r)$ is nonanalytic at all $r > \mathcal{D}$.
[56] F. H. Stillinger and T. A. Weber, "Gaussian core model in two dimensions. I. Melting transition," J. Chem. Phys. **74**, 4015–4019 (1981).
[57] H. J. Axon and W. Hume-Rothery, "The lattice spacings of solid solutions of different elements in aluminium," Proc. R. Soc. A **193**(1032), 1–24 (1948).

25 August 2024 14:56:01

Published under an exclusive license by AIP Publishing

[58]F. H. Stillinger, "Structure in aqueous solutions of nonpolar solutes from the standpoint of scaled-particle theory," J. Solution Chem. **2**, 141–158 (1973).
[59]K. W. Lum, D. W. Chandler, and J. D. Weeks, "Hydrophobicity at small and large length scales," J. Phys. Chem. B **103**, 4570–4577 (1999).
[60]Y. Zhang, J. Evans, and S. Yang, "The prediction of solid solubility of alloys: Developments and applications of Hume-Rothery's rules," J. Cryst. Phys. Chem. **1**, 103–119 (2010).
[61]E. Xi, V. Venkateshwaran, L. Li, N. Rego, A. J. Patel, and S. Garde, "Hydrophobicity of proteins and nanostructured solutes is governed by topographical and chemical context," Proc. Natl. Acad. Sci. U. S. A. **114**, 13345–13350 (2017).
[62]M. A. Klatt, J. Kim, T. E. Gartner III, S. Torquato, and S. Torquato, "Local number fluctuations in ordered and disordered phases of water across temperatures: Higher-order moments and degrees of tetrahedrality," J. Chem. Phys. **160**, 204502 (2024).
[63]H. S. Ashbaugh and L. R. Pratt, "Colloquium: Scaled particle theory and the length scales of hydrophobicity," Rev. Mod. Phys. **78**, 159–178 (2006).
[64]D. Chandler, "Gaussian field model of fluids with an application to polymeric fluids," Phys. Rev. E **48**, 2898–2905 (1993).
[65]G. E. Crooks and D. Chandler, "Gaussian statistics of the hard-sphere fluid," Phys. Rev. E **56**, 4217–4221 (1997).
[66]L. R. Pratt and D. Chandler, "Theory of the hydrophobic effect," J. Chem. Phys. **67**, 3683–3704 (1977).
[67]M. Lusi, "A rough guide to molecular solid solutions: Design, synthesis and characterization of mixed crystals," CrystEngComm **20**, 7042–7052 (2018).

25 August 2024 14:56:01

Published under an exclusive license by AIP Publishing