

Characterization of void space in polydisperse sphere packings: Applications to hard-sphere packings and to protein structure analysis

Moumita Maiti¹, Arun Lakshminarayanan^{1,3,a}, and Srikanth Sastry^{1,2,b}

¹ Theoretical Sciences Unit, JNCASR, Jakkur, Bangalore 560065, India

² TIFR Centre for Interdisciplinary Sciences, 21 Brundavan Colony, Narsingi, Hyderabad 500075, India

³ Indian Institute of Technology, Kharagpur, West Bengal 721302, India

Received 20 April 2012 and Received in final form 27 October 2012

Published online: 28 January 2013 – © EDP Sciences / Società Italiana di Fisica / Springer-Verlag 2013

Abstract. The implementation of a method for the exact evaluation of the volume and surface area of cavities and free volumes in polydisperse sphere packings is described. The generalization of an algorithm for Voronoi tessellation by Tanemura *et al.* is presented, employing the radical plane construction, as a part of the method. We employ this method to calculate the equation of state for monodisperse and polydisperse hard-sphere fluids, crystals, and for the metastable amorphous branch up to random close packing or jamming densities. We compute the distribution of free volumes, and compare with previous results employing a heuristic definition of free volume. We show the efficacy of the method for analyzing protein structure, by computing various quantities such as the distribution of sizes of buried cavities and pockets, the scaling of solvent accessible area to the corresponding occupied volume, the composition of residues lining cavities, etc.

1 Introduction

The geometry of sphere packings is of interest in understanding diverse physical and chemical phenomena. From the understanding of possible crystal structures, to the structure of fluids, granular material, to the structure of biomolecules, especially proteins, the analysis of the geometry of sphere packings finds wide ranging applications. A specific problem of interest is the characterization of empty or “void” space in sphere packings, and the interface between empty and occupied space; examples arise in the study of the thermodynamics of hard-sphere fluids and solids [1], modeling flow through porous media [2] and random heterogeneous materials [3], and the characterization of the solvent accessible surface of proteins [4–6], and of voids and pockets [7].

A number of methods has been developed over the years to calculate the occupied volume and the surface area of packings of spheres, especially in the context of analyzing protein structure [4–16], but methods that take account also of the connectivity of void space, important, *e.g.* in detecting buried cavities in proteins, are relatively rare [12–15]. In [15], an algorithm was presented for the

detection of disconnected components of void space (or *cavities*), and for the exact computation of their volumes and surface areas (subject only to machine precision, unlike, *e.g.*, estimates based on Monte Carlo sampling), for monodisperse and polydisperse packings of overlapping spheres. The method is based on the Voronoi construction, for polydisperse spheres, its generalization in the form of the radical plane construction, which tessellates space into polyhedra (Voronoi cells) consisting of points that are closest to a given sphere than any other (while the distance in question is the distance from the sphere center in the case of monodisperse spheres, for the radical plane construction, it is the tangent distance to the sphere; see below for details). The vertices of the Voronoi cells are equidistant from four spheres for arbitrary sphere packings, whose centres define *Delaunay* tetrahedra or simplices. The algorithm identifies subsets of Delaunay tetrahedra that contain the void space corresponding to individual cavities, and calculates their volumes and surface areas by further dividing individual Delaunay simplices into sub-simplices such that one needs only to calculate the intersection volume of a sphere with the corresponding sub-simplex. In [16], this method was extended to calculate the free volumes of individual spheres, defined as the volume swept by the centre of the sphere without overlapping with any of the other spheres, which are kept fixed in their positions. Employing this extension, the free-volume

^a Present address: eBay Inc, 2065 Hamilton Ave, San Jose CA 95125 USA.

^b e-mail: sastry@tifrh.res.in

distributions for the monodisperse hard-sphere fluid, and the equation of state from the free volumes, were evaluated.

In the present work, we consider the analysis of void space in polydisperse sphere packings, with applications to free volume in hard-sphere systems, and the analysis of cavities, pockets, and other structural features in proteins. To this end, we first present a generalization of the algorithm for the Voronoi construction due to Tanemura *et al.* [17], to perform the radical plane construction [18] for polydisperse sphere packings.

The paper is organized as follows: In sect. 2, we present the details of the methods used, including the new algorithm for the radical plane construction. Section 3 contains a discussion of and results concerning free volumes and surface areas for monodisperse and polydisperse hard-sphere systems. Section 4 contains a discussion and results concerning the analysis of protein structures. Section 5 contains a summary of results and conclusions.

2 Methods

In this section, we present the details of the algorithm used to calculate the volume and surface areas of cavities in a packing of spheres of unequal size, and define the terminology used in the rest of the paper. As described in detail in [15], the algorithm (here described for packings of spheres of unequal size) consists in the following steps:

- 1) *Radical plane construction:* For a set of spheres i of radius r_i , the radical plane construction divides the space into non-overlapping polyhedra VP_i such that for any point \mathbf{x} within VP_i the tangent distance to sphere i is smaller than to any other sphere j , or

$$d^2(\mathbf{x}, \mathbf{x}_i) - r_i^2 < d^2(\mathbf{x}, \mathbf{x}_j) - r_j^2, \quad j \neq i, \quad (1)$$

where \mathbf{x}_i is the location of the sphere centre i , and d^2 are the squared Euclidean distances between \mathbf{x} and \mathbf{x}_i . For any pair of spheres, the condition of equal tangent distance defines a plane (illustrated in fig. 1), and VP_i are convex polyhedra defined by the intersection of such planes separating pairs of spheres that are termed as geometric neighbors.

The usual Voronoi construction (wherein VP_i are termed Voronoi polyhedra) is a special case that arises when the radii of all the spheres are treated as equal. We will refer to the polyhedra VP_i as Voronoi polyhedra also in the polydisperse case. In the general case (*i.e.*, barring sphere configurations possessing special symmetries as in crystalline ordered structures), the vertices of Voronoi polyhedra have equal tangent distance from four (in three dimensions) surrounding spheres. (Where no confusion arises, we use “distance” as a short hand for tangent distance, and phrases such as “ \mathbf{x} is closer to i than j ” are used as a shorthand for “the tangent distance from \mathbf{x} to the sphere i is smaller than the tangent distance to sphere j ”.) Tetrahedra with these sphere centres as vertices are termed *Delaunay tetrahedra* (DT) or simplices and are dual to

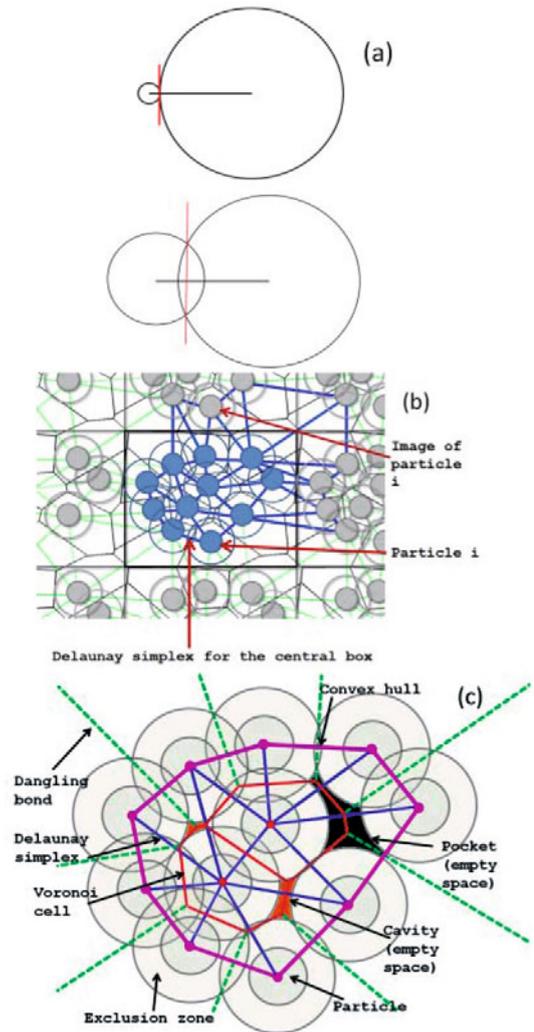


Fig. 1. Illustration of the radical plane construction in two dimensions. (a) The dividing surface defined by the condition of equal tangent distance is a plane, for select cases. For overlapping spheres, the dividing plane passes through the circle of intersection which is obvious from the expression for the tangent distance, when applied to a point on the circle of intersection, and for spheres in contact, through the point of contact. (b) The radical plane construction with the application of periodic boundary conditions or replicated images. The spheres in the central box are replicated in all directions. In this case, the Delaunay simplices cover the system volume. (c) The radical plane construction for a union of spheres with open boundary conditions. In this case, the union of Delaunay simplices defines the convex hull. We have marked an example each of a pocket and a cavity in this illustration (see discussion in sect. 4).

the Voronoi vertices. Similarly, the edges of Voronoi polyhedra are equidistant from triplets of spheres, and are dual to faces of DTs, and as already mentioned, faces of Voronoi polyhedra are equidistant from pairs of spheres, and are dual to edges of the DTs. Both the Voronoi and Delaunay polyhedra divide space into non-overlapping regions, or *tessellate* space. We describe below the generalization of the algorithm for

the Voronoi construction due to Tanemura *et al.* [17] to perform the radical plane construction.

- 2) *Identification of cavities:* After obtaining Voronoi polyhedra, the subset of Voronoi vertices that are “in the void” are identified by the condition that their squared tangent distance is greater than zero. Similarly, Voronoi edges that are in the void are identified by the condition that the point along the edge that is closest to a Delaunay face has a squared tangent distance greater than zero. Such a point is either the point of intersection of the edge with the face (when the two Voronoi vertices that the edge terminates at are on either side of the face) or the closer of the Voronoi vertices in cases where both the Voronoi vertices lie on the same side of the face. Having identified such vertices and edges, one identifies subsets of vertices that are in the void that are connected by edges in the void. As described in [19,20], each set of such connected vertices corresponds to a cavity.
- 3) *Identification of Delaunay tetrahedra enclosing cavities:* The set of Delaunay tetrahedra that are dual to the vertices that define a given cavity fully enclose the cavity. Thus, the cavity volume and surface area are obtained by calculating the void space that is present in the union of such Delaunay tetrahedra, which are considered one at a time.
- 4) *Calculation of cavity volume and surface area contained within each Delaunay tetrahedron:* In order to calculate the void space contained in a Delaunay tetrahedron, one must consider the complement of the intersection of the DT with the union of the spheres defining the Delaunay tetrahedron. Rather than consider such an intersection volume, a procedure is adopted where each Delaunay simplex is divided into 24 subsimplices, such that each subsimplex has as its vertices (a) a sphere centre (total of 4), (b) the midpoint of a Delaunay edge (three per sphere), (c) the intersection point of a Voronoi edge with the plane of a Delaunay face (two per sphere per Delaunay edge), and (d) the Voronoi vertex. Within each such subsimplex, it is sufficient to consider the intersection of the subsimplex with the sphere whose centre defines one of its vertices. The intersection volume, and the surface area of the sphere contained within the subsimplex can be calculated exactly [15]. Adding up the contributions from individual subsimplices of each Delaunay tetrahedron relevant for a given cavity yield the volume and surface area of individual cavities.

In considering the hard-sphere system, one analyzes the geometric problem defined by unions of “exclusion spheres” whose radius equals the sum of radii of the spheres present in the system, and a new sphere that may be inserted into the system. The free volume of a given sphere then corresponds to the volume of a cavity containing the coordinates of the sphere centre, that is obtained by removing the sphere in question from the collection of exclusion spheres whose radii is defined using the radius of the removed sphere. The geometric problem to consider

therefore is that of calculating the volume and surface area of a cavity that is generated by the removal of one of the spheres.

The extension of the above procedure to evaluate the free volumes, described for equal sized spheres in [16], carries over straightforwardly for the polydisperse case, with the generalization of the two theorems proved in [16] to the polydisperse case. These generalizations are stated and proved below.

We end this summary by noting that if one employs periodic boundary conditions, as done in simulations of the hard-sphere system that we consider later, the Delaunay tetrahedra cover the space (*i.e.* the sum of the volume of all the Delaunay tetrahedra equals the volume of the system that is subjected to periodic boundary conditions), whereas if one considers a union of spheres subject to open boundary conditions, as the case of proteins that we consider later, the union of Delaunay tetrahedra define a convex polyhedron termed the convex hull. In order to evaluate the external surface area in the latter case (as we need to in order to compute the solvent accessible surface area for proteins), we find it convenient to apply periodic boundary conditions in this case as well, with a suitably chosen but arbitrary containing volume. With this procedure, the external surface can be easily calculated as the surface area of the cavity that lies outside the union of spheres. These cases are illustrated in fig. 1.

2.1 Generalized algorithm for Voronoi construction

The generalization of the Voronoi tessellation algorithm by Tanemura *et al.* [17] for the radical plane is described below. As mentioned earlier, the Voronoi construction, obtained by the procedure of Tanemura *et al.* can be viewed as the radical plane construction (defined by eq. (1) above) for equal sized spheres. There are, however, situations peculiar to the radical plane construction which are not encountered in the usual Voronoi construction. While in the Voronoi construction the bisector plane between two points is always situated between the two points, the centre of a sphere may, in the radical plane construction, lie in the half-space of the other sphere (the easiest way to visualize this is to note that the radical plane of intersecting spheres is the plane containing the circle of intersection. Depending on the radii of and separation between the two spheres, the centres of both the spheres may lie on the same side of the plane of intersection). Also, it is possible that the generalized Voronoi cell of a given sphere, resulting from the radical plane construction, has zero volume. This situation arises when (and only when) the sphere in question is contained entirely within another. In such an event, of course, one may as well discard such a sphere as it in no way affects the analysis of the void space properties. This is the procedure followed here. The results presented below do not make any assumption about the location of the radical plane with respect to the sphere centres (unless stated), although the accompanying illustrations are drawn to indicate typical situations.

Two modifications are made to the original algorithm: i) In [17] one loops through each sphere in the system and constructs Delaunay tetrahedra that enclose the sphere. Such a procedure is redundant as each DT is encountered four times, and further, construction of the first DT each time is expensive. In the algorithm here, this redundancy is eliminated. ii) The procedure for constructing the first DT in [17] starts by identifying a sphere that is guaranteed to be a geometric neighbor (*i.e.* one that shares a Voronoi cell face with the central sphere). The generalization of the procedure for polydisperse spheres is valid only under restricted conditions. Hence, finding the first geometric neighbor to a given sphere is more involved in the present case.

The generalized Voronoi construction, based on Theorems I-IV to be proved subsequently, involves the following steps:

Step 1. Find if any sphere is completely contained within any of the other spheres present. Remove all such spheres from the configuration to be analyzed.

Step 2. Find a sphere i such that the centre of i lies in its own (generalized) Voronoi cell. This means that for all pairs of spheres i and k , the centre of i lies in the i half-space, as defined by the radical plane for the pair i and k . If such a sphere is not found (it appears extremely unlikely that such a sphere cannot be found for any sphere packing, but we have not attempted to prove that such is the case in general), step 3(b) below is implemented.

Step 3. If a sphere is found whose centre lies in its own Voronoi cell, a) find sphere j such that the $\text{CMSTD}(i, j)$ (CMSTD stands for the “common minimum squared tangent distance” for the spheres indicated in brackets. STD below stands for “squared tangent distance”) is smaller than $\text{CMSTD}(i, k) \forall k \neq i, j$. Then, i, j are geometric neighbors. Else, b) considering all pairs of spheres i, l , find a pair i, j such that the STD to i, j of the point $P \equiv \text{PCMSTD}(i, j)$ (PCMSTD stands for the coordinates of the point that has the CMSTD for the spheres indicated in brackets) is smaller than its STD to $k \forall k \neq i, j$. Then, i, j are geometric neighbors.

Step 4. With i, j from step 3, find sphere k such that the $\text{CMSTD}(i, j, k)$ is smaller than for any other triplet of spheres. Then, i, j, k are geometric neighbors.

Step 5. With i, j, k from step 4, find sphere l such that $\text{CMSTD}(i, j, k, l)$ is smaller than for any other quadruplet of spheres. Then, i, j, k, l form a DT. Add the DT found to the list DTLIST as the first entry. Set the count of the number of DTs found, γ , to 1.

For each DT encountered in steps 5, 6, define the count of the number of adjacent DTs for a given DT, DTN . Set $\text{DTN}(1) = 0$. For each face β of each DT α found in steps 5, 6, define the count $l(\alpha, \beta)$ which is equal to the number of DTs sharing the face. Set $l(1, \beta) = 1$.

Step 6. In this step, looping through DTLIST , for each DT α such that $\text{DTN}(\alpha) \neq 4$, and for each face β such that $l(\alpha, \beta) < 2$ (which means that the face is shared by less than two DTs), construct a new DT that shares the face β , as follows:

Let the face β be formed by spheres i, j, k . Define a coordinate system such that i, j, k lie in the x, y plane, and the fourth sphere l which completes the DT has a negative z coordinate. Consider all spheres which lie in the positive z half-space. Among these, find sphere m such that the z coordinate of the PCMSTD of i, j, k, m has the lowest value. Then, i, j, k, m form a DT.

When all DTs α in DTLIST have $\text{DTN}(\alpha) = 4$, equivalently all $l(\alpha, \beta) = 2$, the tessellation is complete.

In the cases (*e.g.* proteins) where open boundary conditions are used, some of the DTs α will have faces β that form part of the convex hull, and which are identified by the fact that $l(\alpha, \beta) = 1$ at the termination of the tessellation. In this case, the termination condition will be that no new DTs are added after an iteration looping through all the DTs that have $l(\alpha, \beta) = 1$.

We present below proofs of four theorems (which are generalizations of ones by Tanemura *et al.* [17]) which must be valid for the above steps to result in a generalized Voronoi tessellation for the given configuration of spheres. While these results are necessary to justify the procedure we describe, they are not necessary for understanding the method, and may be skipped.

Theorem I. *For a given sphere i such that its centre lies inside its own Voronoi cell, the STD of its centre to i is smaller than to any other sphere, if $\text{CMSTD}(i, j)$ is smaller than $\text{CMSTD}(i, k) \forall k \neq i, j$, then i and j are contiguous.*

Proof. The premise is $\text{CMSTD}(i, j) < \text{CMSTD}(i, k) \forall k \neq i, j$. For any two spheres, the set of points which have the same squared tangent distance (STD) to both i and j lie on a plane, the radical plane, that is perpendicular to the line ij , joining spheres i and j . The intersection point P of the radical plane and the line ij thus has the “common minimum squared tangent distance” (CMSTD) to spheres i and j . To prove that i, j are contiguous, it is sufficient (but not necessary) to show that the STD to point P from any other sphere $k \neq i, j$ is never smaller than the CMSTD of i and j .

Assume that a sphere k exists such that the STD from P to k is smaller than to i, j . This implies that if one draws the radical plane between spheres i and k , P lies in the k half-space (see illustration in fig. 2). Since the centre of the sphere i and P lie in different half-spaces defined by the radical plane of i, k , the line joining i and P must intersect the radical plane of i, k . Let P' be the point of intersection. Let Q be the $\text{PCMSTD}(i, k)$. By definition of the radical plane, *i.e.* iQP' is a right-handed triangle and hence $\text{STD}(i, Q) < \text{STD}(i, P')$. Further, as P' lies between i and P , and hence the distance from i to P' is smaller than the distance from i to P . Hence, $\text{STD}(i, P') < \text{STD}(i, P)$, and we have proved $\text{STD}(i, Q) < \text{STD}(i, P')$. This leads to the conclusion that $\text{CMSTD}(i, k) = \text{STD}(i, Q) < \text{CMSTD}(i, j)$ which contradicts the premise. The theorem is hence proved.

Theorem II. *Consider spheres i, j, k with i, j as in Theorem I, such that $\text{CMSTD}(i, j, k)$ (the set of points which satisfy $\text{STD}(i, j) = \text{STD}(i, k) = \text{STD}(j, k)$) is smaller*

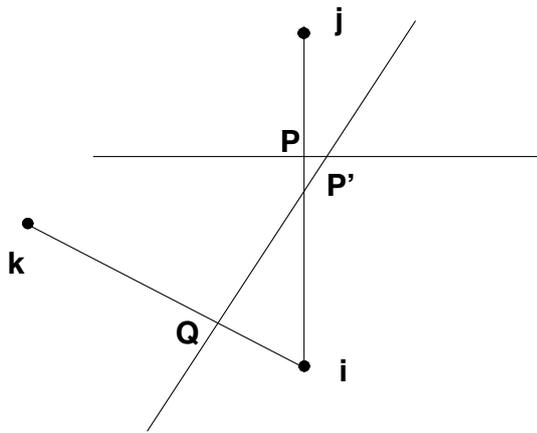


Fig. 2. Illustration for Theorem I. The line passing through P , perpendicular to the line joining i, j represents the radical plane i, j . The line passing through Q is the radical plane i, k . P' is the intersecting point of the radical plane i, k and the line joining i, j .

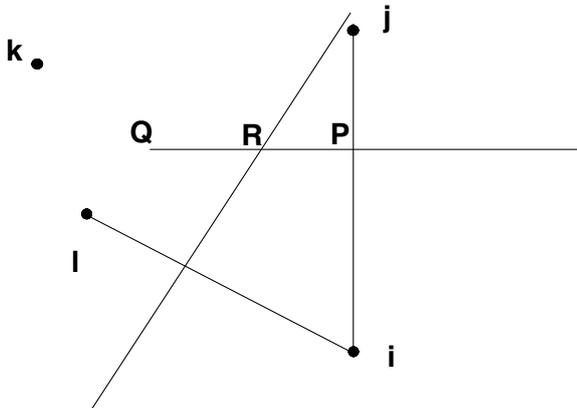


Fig. 3. Illustration for Theorem II. The line passing through the point R which intersects the line i, l represents the radical plane i, l . The line perpendicular to the line i, j which passes through the point P represents the radical plane i, j . The point Q is in the l half-space.

than $CMSTD(i, j, l) \forall l \neq (i, j, k)$. Then, i, j, k are contiguous.

Proof. The premise is $CMSTD(i, j, k) < CMSTD(i, j, l) \forall l \neq i, j, k$. For any triplet, the set of points of equal STD to all three spheres describes a line. In order for i, j, k to be contiguous, it is sufficient (but not necessary) to show that the STD of Q ($Q \equiv PCMSTD(i, j, k)$) to any other sphere l is greater than the STD to i, j, k , i.e. $STD(l, Q) > CMSTD(i, j, k)$.

To show this, assume the contrary, i.e. that for some l , $STD(l, Q) < CMSTD(i, j, k)$. Define the plane containing the sphere centres i, j, k as the x, y plane, and the direction i, j as the y axis (see illustration in fig. 3). The radical plane of i, j then defines the x, z plane, and the point Q is on the x axis as it is on the radical plane i, j . Since we assume that the STD from Q to l is smaller than $CMSTD(i, j, k)$, if we construct the radical plane of i, l , Q is in the l half-space. The point $P \equiv PCMSTD(i, j)$ is

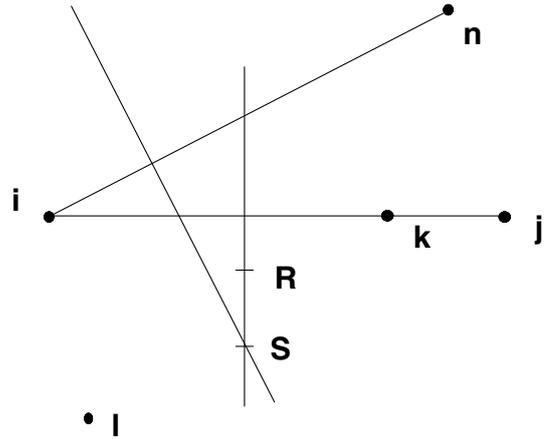


Fig. 4. Illustration for Theorem IV. i, j and k are in a plane perpendicular to the plane of the paper. R is a point on $CMSTD(i, j, k, m)$ which is also on the radical plane i, j . The line passing through S which intersects the radical plane i, n represents the radical plane i, n .

in the i half-space from Theorem I. Let R be the point of intersection of the line PQ and the line defined by the intersection of the radical plane of (i, l) with the x, y plane. R clearly lies between P and Q as P and Q belong to different half-spaces. By definition, R is a point on the line of equal STD to i, j, l . Since the distance PR is less than the distance PQ , $STD(i, R) < STD(i, Q) = CMSTD(i, j, k)$. Since any point on the radical plane has the same STD, $STD(j, R) = STD(i, R) = STD(l, R)$. Hence $R \equiv PCMSTD(i, j, l)$. This implies that $CMSTD(i, j, l) < CMSTD(i, j, k)$, which contradicts the premise. Then the theorem is proved.

Theorem III. Consider spheres (i, j, k, l) with i, j, k as in Theorems I and II, and $CMSTD(i, j, k, l)$ being the smallest of all quadruplets of spheres. Then i, j, k, l are contiguous, i.e. they form a DT.

Let $P \equiv PCMSTD(i, j)$, $Q \equiv PCMSTD(i, j, k)$. Let i, j, k define the x, y plane, and i, j define the y axis. Let $R \equiv PCMSTD(i, j, k, l)$.

Let us assume that there is some sphere m such that $STD(m, R) < CMSTD(i, j, k, l)$. Then, defining the radical plane of i, m , R belongs to the m half-space. From Theorem II, Q belongs to the i half-space. Let S be the intersection point of QR with the radical plane i, m . S must lie between Q and R . Hence, $STD(i, S) < STD(i, R)$. By definition, S is the $PCMSTD(i, j, k, m)$ which implies that $CMSTD(i, j, k, m) < CMSTD(i, j, k, l)$ which contradicts the premise. The theorem is hence proved.

Theorem IV. Consider the DT i, j, k, l . Let i, j, k define the x, y plane and let l be in the negative z half-space. Consider m such that m is in the positive z half-space, and $PCMSTD(i, j, k, m)$ has the smaller z coordinates than the $PCMSTD$ for any i, j, k, n where n is in the positive z half-space. Then i, j, k, m form a DT.

Proof. The premise is that z -coordinate of $PCMSTD(i, j, k, m)$ is smaller than $PCMSTD(i, j, k, n)$. Let $R \equiv PCMSTD(i, j, k, m)$. R lies along the line of

equal STD to i, j, k (see illustration in fig. 4), which is parallel to the z axis. l is in the negative z and n is in the positive z half-space. In order to form the DT i, j, k, n , assume that there is a sphere n such that $\text{STD}(n, R) < \text{CMSTD}(i, j, k, m)$. Then, if one constructs the radical plane of i, n , R lies in the n half-space. Let S be the intersection of the radical plane of i, n with the line of equal STD from i, j, k . By definition, $S \equiv \text{PCMSTD}(i, j, k, n)$. Since n is in the positive z half-space, and the radical plane i, n is perpendicular to the line joining i and n , from a point on the radical plane, to reach a point in the n half-space with the same x, y coordinates, one must increase z . Since S and R have the same x, y coordinates, in order to reach R from S , one must move along the direction of increasing z coordinate. This implies that S has a lower z coordinate than R , if R is in the n half-space which is false by assumption. Hence the theorem is proved.

Finally, we state below the two theorems relevant to the calculation of free volumes along with the corresponding proofs.

Theorem F-I. *If a sphere is removed from the system, only the Voronoi polyhedra of its geometric neighbors must be re-tessellated.*

Proof. If a sphere is removed, points inside the Voronoi cell of that sphere will become points contained in the modified Voronoi cells of other spheres. We need to prove that the spheres whose Voronoi cells are modified are the geometric neighbors of the removed sphere.

Consider the Voronoi polyhedron V_j of an arbitrary sphere j . Let the sphere i be removed. Let a point P belonging to the Voronoi polyhedron V_i of i , now belong to the redefined Voronoi polyhedron V_j' of j . In order for a point P that does not belong to V_j to belong to the redefined Voronoi polyhedron V_j' , the removal of i must modify the boundary of V_j . Since the boundaries of V_j are formed by radical planes between j and one of its geometric neighbors, the removal of i can only modify the boundaries of V_j if i is a geometric neighbor of j . Thus the spheres whose Voronoi polyhedra change as a result of the removal of a sphere are its geometric neighbors.

Theorem F-II. *Pairs of geometric neighbours of a sphere that share a Voronoi face continue to do so after the sphere is removed.*

Proof. Consider spheres k and k' which are geometric neighbours of sphere i that is removed, and share a common Voronoi face. Any point on the common face has a smaller tangent distance to spheres k and k' than any other atom in the system. Clearly the removal of any other sphere (including i) will not change this fact, verifying the theorem.

2.2 Tessellation with replicated images

As described above, for systems with open boundary conditions, the Delaunay tessellation does not tessellate the system volume (which may arbitrarily be defined so as to enclose the union of spheres under consideration). There

are however circumstances where it is desirable to do so. In order to address such cases, we adopt here the approach of replicating the spheres along $\pm x, \pm y, \pm z$ directions. Thus, for a system that is contained in a volume of dimension L_x, L_y, L_z , with N spheres, we consider an extended system of dimension $(3L_x \times 3L_y \times 3L_z)$, containing $27N$ spheres. However, tessellation is performed only to obtain DTs that involve the original N spheres. In addition to cases such as proteins, with open boundary conditions, such a procedure is also useful in considering systems where one employs periodic boundary conditions, under circumstances where the heterogeneity of the system is large enough that the minimum image convention normally used along with periodic boundary conditions fails to produce a meaningful tessellation.

However, with such a procedure, the computations involved in step 6 of the algorithm, involving a search over all spheres that are prospective geometric neighbors, can become very expensive. This step involves the determination of the z coordinate of the PCMSTD for a triplet of spheres along with a candidate fourth sphere. Even while we consider only DTs that contain at least one of the original N spheres as a vertex, it is useful to have a termination condition for the search for a fourth sphere that shares an unpaired DT face. Here we describe such a criterion. We note that the stopping criterion defined here is applicable in general, although motivated here by the procedure of replicated images.

Let us consider a DT face i, j, k which defines the x, y plane, with an existing fourth neighbor l in the negative z half-space. In order to find an m , in the positive z half-space such that i, j, k, m defines a DT, we compute the z coordinate of the PCMSTD for i, j, k, m and iterate till we find m with the lowest z . In doing so, we iterate through a list of m which are ordered in increasing order of the distance $|x_m^{\vec{}}|$ ($x_m^{\vec{}}$ is the vector distance) from one of the DT vertices, say i (see illustration in fig. 5(a)).

We denote the PCMSTD of the face i, j, k is by \vec{c}_3 and the normal to the face by \hat{n} . The PCMSTD of i, j, k, m , \vec{c}_4 , is then

$$\vec{c}_4 = \vec{c}_3 + z\hat{n}, \quad (2)$$

where z is the z coordinate of the PCMSTD of i, j, k, m (see fig. 5 for illustration). Using the condition of equal tangent distance $((\vec{x}_i - \vec{c}_4)^2 - R_i^2 = (\vec{x}_m - \vec{c}_4)^2 - R_m^2)$ and eq. (2), z can be written as follows:

$$z = \frac{x_m^2 + R_i^2 - R_m^2 - 2c_3|x_m^{\vec{}}|\sin\phi}{2|x_m^{\vec{}}|\cos\phi}. \quad (3)$$

We note in passing that the use of the above expression rather than solving simultaneous equations for \vec{c}_4 also reduces the computational cost considerably.

As we iterate through candidate neighbors m , at any given point, we have the current minimum z value z_{curr} . We wish to define a criterion such that, given z_{curr} , we need not search further when the distance $|x_m^{\vec{}}|$ of a candidate neighbor m exceeds a threshold value x_{cut} . To determine this condition, we note that for a given distance $|x_m^{\vec{}}|$, there is a minimum value of z depending on the value

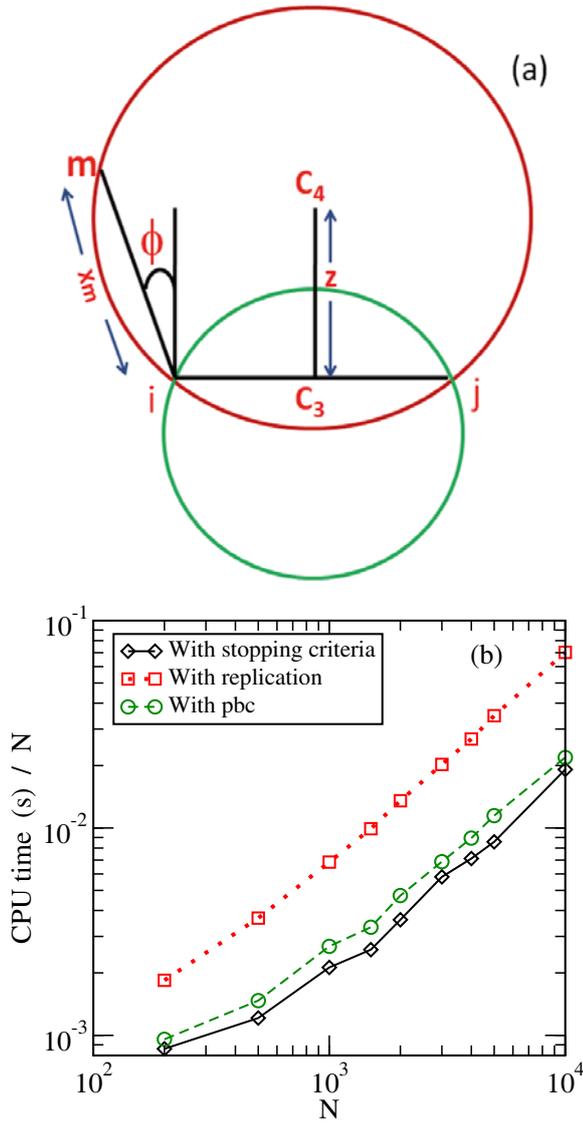


Fig. 5. (a) Illustration of the geometry considered for defining the stopping criterion for adding new Delaunay tetrahedra. (b) Comparison of the CPU time of tessellation for a set of homogeneous configurations of different system size. The CPU time with periodic boundary condition with the minimum image convention (circles) is shown along with the case of replicated images without (squares), and with (diamonds) the stopping criterion. In the latter cases, we use the expression given in the text for the z coordinate of the PCMSTD. (see text).

of ϕ . Minimizing z with respect to ϕ , we get the condition

$$\sin \phi = \frac{2c_3 |x_m^z|}{x_m^2 + R_i^2 - R_m^2}. \quad (4)$$

Using this condition, and the current value of z_{curr} in eq. (2), we get an expression for x_{cut} as follows (note that this condition depends on R_m . In order for the expression to hold for all the available candidates m , we use the minimum value of the available radii):

$$x_{\text{cut}} = \sqrt{(c_3^2 + z_{\text{curr}}^2) + \sqrt{(c_3^2 + z_{\text{curr}}^2) - (R_i^2 - R_m^2)}}. \quad (5)$$

When $|x_m^z| > x_{\text{cut}}$, we terminate the search for m . As shown in fig. 5 (configurations of number density 0.5 of N random points with $N = 200$ to 20,000 were used as test cases), the use of the termination criterion considerably reduces the cost of the tessellation.

3 The hard-sphere system

The hard-sphere system has been studied extensively [21–28] over the decades as an idealized model system that nevertheless displays many essential condensed phase phenomena. Of particular interest has been the use of this system to understand aspects of amorphous phase structure [23] and jamming [24–28] and other aspects of granular matter [29–34].

In the present work, we calculate the hard-sphere equation of state for monodisperse and bi-disperse hard-sphere fluids, and nearly jammed configurations, as well as the crystal in the monodisperse phase. We note that the term “free volume” is defined variously in the literature, but we use the definition of refs. [1, 16] which lead to the expression for the equation of state we mention below. For some of these cases, we also calculate the free-volume distribution. The free-volume distributions and the equation of state using free volumes and surface areas for the monodisperse hard-sphere system was previously computed in [16].

To this end, we first perform Monte Carlo simulations for a range of packing fractions for monodisperse hard-sphere systems, and a 50 : 50 binary mixture of spheres of unequal size, with the ratio of (additive) diameters $\sigma_2/\sigma_1 = 1.4$. The number of particles is $N = 2000$. Equilibration runs between $100 \times \tau$ and $500 \times \tau$ are performed for each independent run, where the relaxation time τ is estimated by the decay of a density-density correlation function. For each state point (density), we perform simulations for 20–25 independent initializations. In each case, production runs are performed for about 2000τ and we analyze about 2000 configurations for each independent run. The durations for runs in the crystalline phase, in the FCC lattice, are much shorter since short equilibration are sufficient in these cases.

The pressure of the monodisperse hard-sphere system is given in terms of the free volumes and surface areas by [1, 16]

$$\frac{P}{k_B T} = \rho + \frac{\rho \sigma}{2D} \left\langle \frac{s_f}{v_f} \right\rangle, \quad (6)$$

where k_B is the Boltzmann constant and T is the temperature, σ is the diameter of a hard sphere, ρ is number density and D is dimensionality of the system. As shown in [35], the corresponding relation for polydisperse hard-sphere systems is

$$\frac{P}{k_B T} = \sum_i \rho_i + \sum_i \frac{\rho_i \sigma_{ii}}{2D} \left\langle \frac{s_f^i}{v_f^i} \right\rangle, \quad (7)$$

where the index i labels components with sizes σ_{ii} and a partial number density of ρ_i . We use these relations in computing the equation of state (EOS).

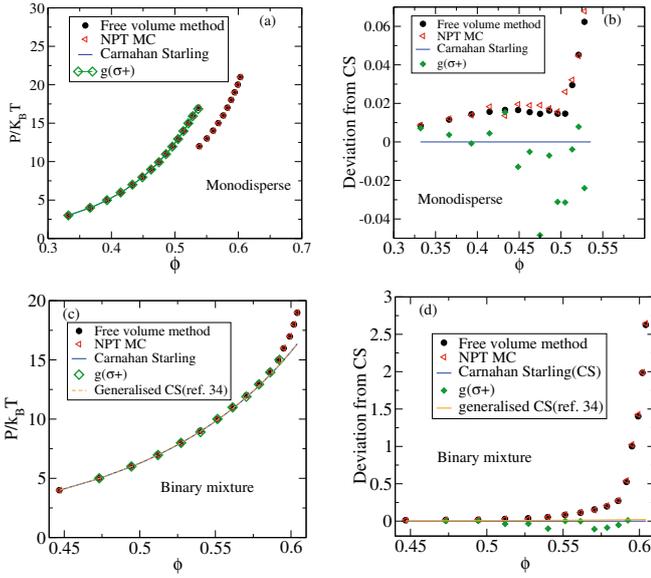


Fig. 6. (a) Equation of state of the monodisperse hard-sphere fluid and crystal, obtained from i) free volumes and surface areas (circles), ii) densities calculated in NPT Monte Carlo simulations (squares), and iii) the contact value of $g(r)$ (diamonds). The Carnahan-Starling equation of state (in the unit of $\frac{P}{k_B T}$) for the monodisperse hard-sphere fluid for the same methods. The deviations become pronounced at high densities. The freezing density is indicated by a vertical line. (c) The equation of state for a polydisperse (binary mixture) hard-sphere fluid. Details as in panel (a) with the Mansoori-Carnahan-Starling shown (line) for reference, along with the generalised CS discussed in ref. [36]. (d) Deviations from the Mansoori-Carnahan-Starling equation of state.

In fig. 6(a), the EOS of the monodisperse hard-sphere fluid and crystal, obtained from eq. (6) is shown. For comparison, we show the EOS obtained from i) constant pressure simulations (NPT) where the average density for a range of pressures is calculated, ii) the contact value of the pair correlation function $g(r)$ using the formula

$$\frac{P}{\rho k_B T} = 1 + \frac{2\pi\rho}{3} \sigma^3 g(\sigma^+), \quad (8)$$

and iii) the Carnahan-Starling (CS) EOS [37]. In fig. 6(b) we show the comparison between these various evaluations as deviations from the CS EOS. As expected, the estimation based on the contact value of the $g(r)$ is not very accurate, but the other methods agree reasonably well, and indicate that the measured pressure is in excess of the CS EOS at high densities. At densities above the freezing packing fraction of 0.4945, we ensure that we sample the metastable liquid by calculating the pair correlation function and the global orientational order parameter Q_6 [38]. Above the packing fraction 0.525, we observe spontaneous crystallization of the fluid samples and thus do not report them here. In fig. 6(c) we show the EOS (with other methods as for monodisperse spheres) for the binary mixture fluid. As a reference, we use the Mansoori-Carnahan-Starling (MCS) EOS [39]. The deviations from

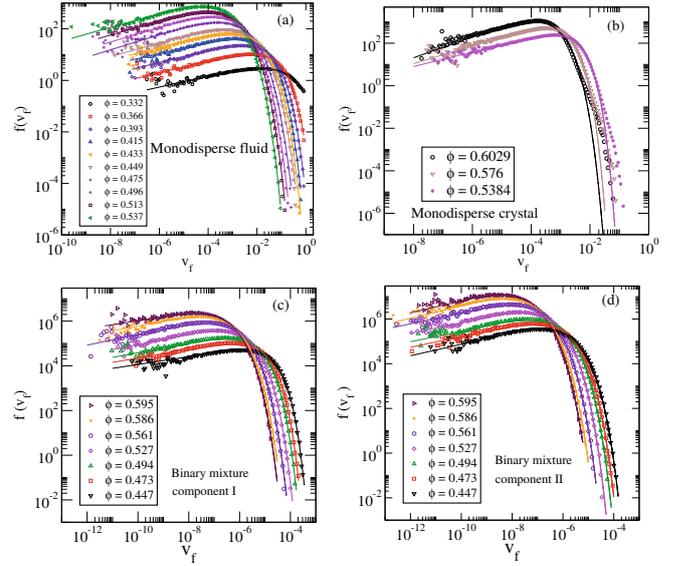


Fig. 7. Free-volume distributions for (a) the monodisperse hard-sphere fluid, (b) the monodisperse hard-sphere crystal, (c) the smaller component of the binary hard-sphere mixture, and (d) the larger component of the binary hard-sphere mixture. In each case, the solid lines are fits to the generalized gamma distribution (eq. (9)). We see that eq. (9) provides a good description of the free-volume distributions for the fluids, whereas systematic deviations are seen for the crystal.

the MCS EOS shown in fig. 6(d) show stronger deviations from the MCS EOS than in the monodisperse case.

Next we compute the full distribution of free volumes. In fig. 7(a) we show the free-volume distribution for the monodisperse fluid. The free-volume peak moves to smaller value with increasing packing fraction as expected. The numerical data are fit to the generalized gamma distribution (also considered in [16] and references therein, and [40])

$$f(v_f) = \frac{\gamma\beta^{\frac{\alpha+1}{\gamma}}}{\Gamma\left(\frac{\alpha+1}{\gamma}\right)} v_f^\alpha \exp(-\beta v_f^\gamma), \quad (9)$$

which are shown as solid lines in the figure. We see that the fits are excellent at all densities. The free-volume distributions for the monodisperse crystal are shown in fig. 7(b). Fits of the numerical data to the generalized gamma distribution are not very good in this case, with significant deviations in the tails. Free-volume distributions for the binary mixture fluid for the two components are shown in fig. 7(c) and (d), along with the fits to the generalized gamma distribution. The fits for the bi-disperse hard-sphere fluid are also very good.

The fit parameters α , β and γ of eq. (9) are plotted for the monodisperse fluid and crystal phases in fig. 8(a), (b) and (c). We see that the parameters α and γ are slowly varying, with α values varying non-monotonically in the range of 0.22 to 0.3, and γ values approaching a nearly constant value of 0.5 at high densities in the fluid phase. In fig. 8(d), (e) and (f) we report the fit parameters for the

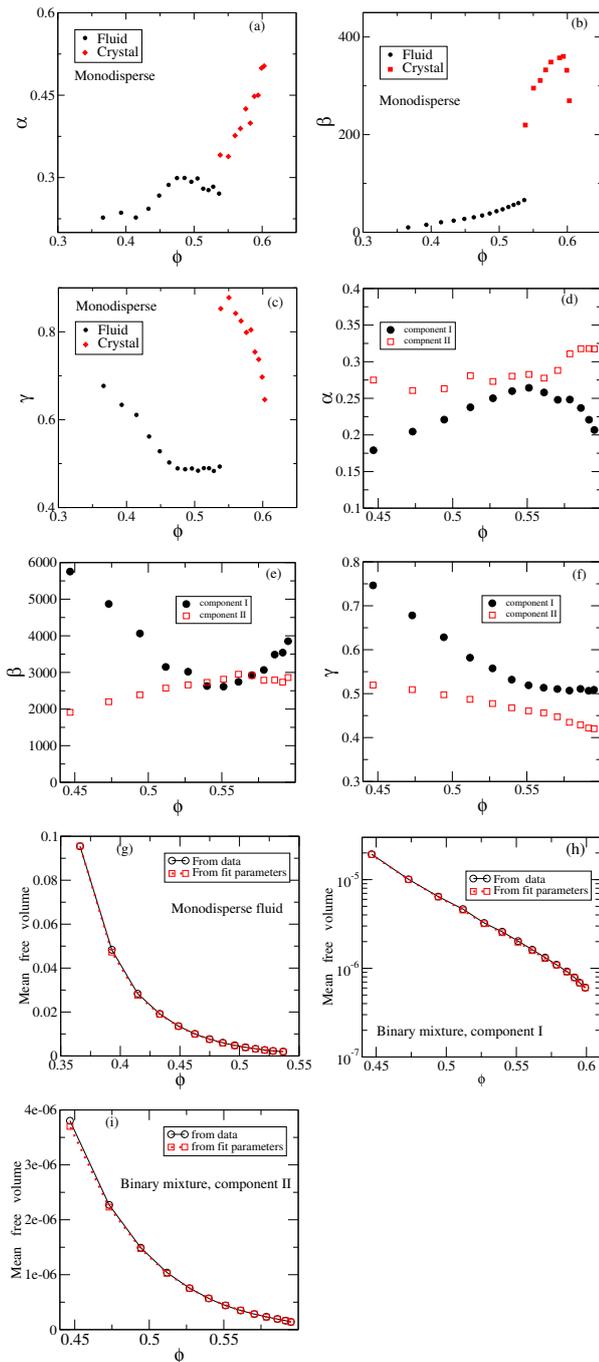


Fig. 8. Parameters α , β and γ of eq. (9). (a) α (b) β , and (c) γ , for the monodisperse hard-sphere fluid and crystal. (d) α (e) β , and (f) γ , for the two components of the binary hard-sphere fluid. The average free volumes for the (g) monodisperse fluid, (h) smaller, and (i) larger components of the binary mixture.

bi-disperse case. We compare the mean free volumes obtained from fit parameters and numerical data in fig. 8(g) for monodisperse fluid and in fig. 8(h) and (i) for the bi-disperse case. The comparison shows that the fits faithfully reproduce the behavior of the mean values.

We have also studied the free-volume distributions of amorphous configurations at higher densities, close to ran-

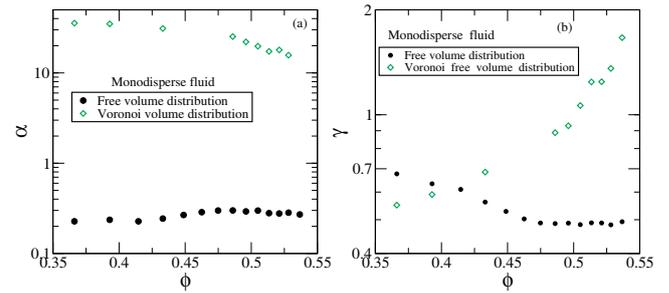


Fig. 9. Parameters α , and γ for *Voronoi free volumes* fitted to eq. (9), with corresponding values from free-volume distributions shown in fig. 8. Values of (a) α and (b) γ from [40] and the present work for Voronoi free volumes compare well, but are significantly different both in values and trends from the corresponding values for free volumes.

dom close packing or the jamming point. These configurations show deviations from the behavior displayed by fluid configurations near and below the freezing density. A detailed discussion of this case is taken up elsewhere [41]. In [40], distributions of Voronoi free volumes were considered, defined to be the volume of the Voronoi cells of spheres, subtracting the average volume of the Voronoi cells at close packing. These distributions were fitted to the generalized gamma distribution in eq. (9). We report in fig. 9(a) and (b) the parameters α and γ from [40] along with our evaluations for the monodisperse fluid, which are seen to be in good agreement (in the notation we use, the parameters m and δ in [40] are given by $m = (\alpha + 1)\gamma$, $\delta = \gamma$). However, we note that by comparison, the corresponding fit parameters for the free volume show different values, as well as an opposite trend with density. The use of the Voronoi free volume *vs.* the free volume in different situations thus bears further scrutiny.

4 Protein structure analysis

The analysis of the structure of proteins in the native state is of great interest as it forms the basis of understanding the functioning of the proteins [4–11, 42]. Structure based ligand design requires detailed characterization of the exterior of the protein, including the geometry and composition of pockets [7]. Further, the nature of the packing of residues in proteins may also shed light on the mechanisms of protein folding [43–47]. Specifically, the statistics and role of pockets and cavities has been addressed in this context [44, 45]. In order to demonstrate the utility of our methods to the analysis of protein structure, we compute various quantities that have been discussed in the literature, some of which have already been computed and well characterized. Towards this end, we analyze up to 800 proteins from the protein data bank, of sequence length ranging from 20 to 4500, most of the analyzed proteins being in the range of 100 to 1000. For these proteins, we compute the occupied volume, surface area, and the composition of the surface residues, the volumes, surface areas and com-

position of cavities and pockets, and the connectivity of residues lining cavities.

The volume of a protein is defined as that contained within the solvent accessible (SA) surface [4], which we refer to as the SA volume. The SA volume, and the SA surface area are computed by using a tessellation with replicated images as described above. The SA volume is the complement of the volume of the largest cavity that corresponds to the exterior of the protein, and the SA surface area is the surface area of the largest (exterior) cavity. We do not consider in this work the evaluation of the molecular surface volume [4]. For these calculations, we consider the geometric problem of spheres located at atomic coordinates of size equal to their van der Waals radius plus that of a water molecule (1.4 \AA) treated as the probe. The buried cavities are all other cavities that are identified, excluding the exterior “cavity” (see fig. 1). In order to define pockets, we use the physical definition that a pocket is a concave region on the surface of the protein such that the empty space defining the pocket is separated from the exterior by a “neck” that is narrower than the largest sphere that may be accommodated by the pocket. In order to identify such pockets, we use the following procedure:

A tessellation of the protein is performed which yields a set of Delaunay tetrahedra, the external surface of whose union defines the convex hull. The DT at the periphery are identified by the fact that at least one of the edges of each such DT is not shared by a second DT that together encloses the DT face. Such “dangling edges” are identified and used in the procedure to identify pockets (see the illustration fig. 1(c)). We identify the set of Voronoi vertices and edges that are in the void, and by analyzing the connectivity of such vertices through edges that are in the void, identify the cavities as corresponding to a connected set of vertices in the void, none of which are connected to a dangling edge. The cavities are excluded from further analysis. Next, we perform an iterative procedure, starting with each of the dangling edges. Considering the vertex to which a dangling edge is connected, we test whether the vertex is contained within the DT it corresponds to. If not, the dangling edge and the vertex are eliminated from consideration, and the other edges connected to the vertex are now added to the list of dangling edges. The procedure terminates when each of the dangling edges is connected to a vertex that is contained within the DT it corresponds to. Each dangling edge defines the mouth of a pocket, and in principle a given pocket could have more than one such mouth. We next perform a cluster analysis to identify the set of connected vertices and edges among the remaining set. Each connected set of vertices then corresponds to a pocket.

Now we describe the results from our analysis. Figure 10(a) shows the number of pockets, cavities, and pockets plus cavities plotted against the SA volume. Figure 10(b) shows the sum of the volumes of pockets plus cavities, pockets, and cavities *vs.* SA volume. Figure 10(c) shows the SA surface area *vs.* SA volume. The linear correlation between them was also observed in a previous

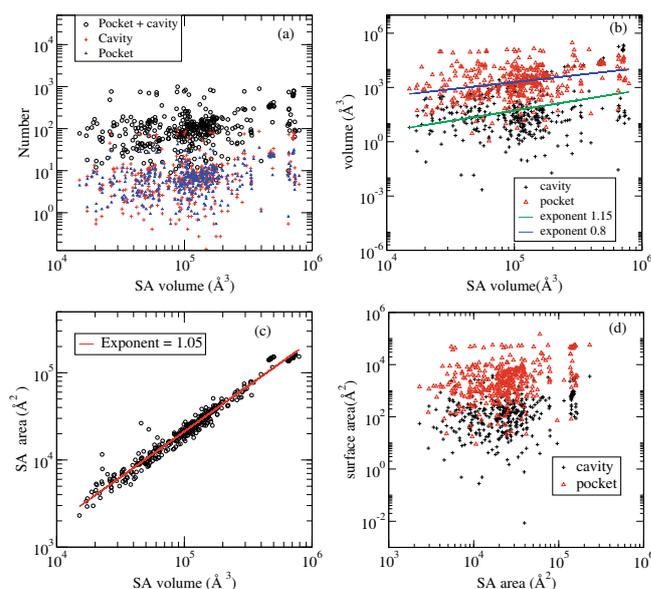


Fig. 10. The number and volumes of cavities and pockets in proteins as a function of SA volume and surface area. (a) Number of pockets and cavities *vs.* SA volume. (b) Volume of pockets and cavities *vs.* SA volume. (c) SA surface area *vs.* SA volume. (d) Pockets and cavities area *vs.* SA area. See text for the procedure used.

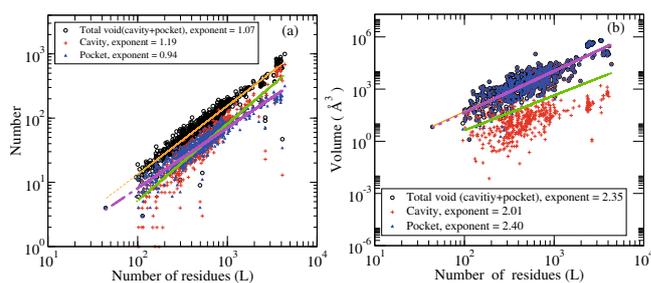


Fig. 11. (a) Number of cavities and pockets *vs.* the sequence length of proteins. (b) Volume of cavities and pockets *vs.* the sequence length of proteins.

study [44]. Figure 10(d) shows the sum of the surface areas of pockets, and cavities *vs.* the SA surface area.

Figure 11(a) shows the number of cavities and pockets *vs.* the sequence length of the proteins. As found in [44], we find a linear relationship between the sequence length of the proteins and the number of pockets and cavities. Figure 11(b) shows the sum of the volumes of cavities and pockets *vs.* the sequence length of the proteins. Unlike the observation in [44] that the volume of the cavities and pockets does not correlate with the sequence length, however, we find that there is a reasonable (but poorer than the number in fig. 11(a)) correlation, with the volumes scaling as L^2 .

In fig. 12(a) we show the size distribution of pockets *plus* cavities. This distribution was considered by Liang and Dill in [44], who found that the frequency of sizes scaled with the size as a power law, with an exponent of -1.67 . Liang and Dill analyzed this result in terms of the statistics expected for a random packing of spheres. The

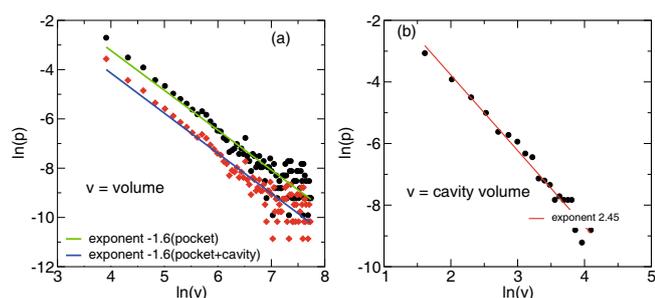


Fig. 12. (a) Size distribution of pockets + cavities, pockets. (b) Size distribution of cavities.

expected exponent for three dimensions is -1.5 , which therefore makes the analogy not fully conclusive. In this context, it is also interesting to inquire whether the buried cavities (which have a direct analogy with cavities in sphere packings) might have a different scaling. To study this, we calculate the size distribution separately for pockets and cavities, which are shown in fig. 12(b) and (a), respectively. Interestingly, we find that the size distribution of pockets shows essentially the same behavior as the distribution of pockets *plus* cavities, whereas the size distribution of cavities, while still being a power law distribution, is rather different, with an exponent of ~ -2.5 . This observation, and its implications for the nature of packing observed in proteins, is described elsewhere [48].

In fig. 13(a) we show the fraction of residues on the surface *vs.* interior residues as a function of the sequence length. Over a broad range of sizes, these fractions remain roughly constant. This is consistent with the fact that the surface area varies roughly linearly with the volume of proteins, as shown in fig. 11(c) and observed in [44]. We show in fig. 13(b) the fraction of residues on the surface that forms a part of a pocket. Interestingly, this fraction varies very strongly with sequence length, with nearly 80% of the surface residues being a part of a pocket for larger proteins. In fig. 13(c), we show similarly the fraction of interior residues that line a cavity. The fraction is roughly constant over a broad range of protein lengths. The inset of fig. 13(c) shows a similar correlation of the number of cavity residues and interior residues with protein length. In fig. 14(a) we show the fraction of hydrophobic, polar and charged residues as a function of sequence length. These fractions, except for increases in scatter, do not depend on the sequence length. With this as reference, we consider the fraction of hydrophobic residues lining cavities and pockets, as a fraction of the number of hydrophobic residues in the entire protein in fig. 14(b), and similarly the fraction of polar and charged residues in fig. 14(c) and (d). We find that compared to the full protein, cavities have a pronouncedly increased fraction of hydrophobic residues, and a decreased fraction of charged residues. The fraction of polar residues also is lower than the full protein, but the decrease is not as marked as the charged residues. The opposite trends are observed for pockets. Such an observation supports the *hydrophobic core collapse* model of protein folding [49]. In fig. 15 we show information regarding the secondary structure and connectivity of residues

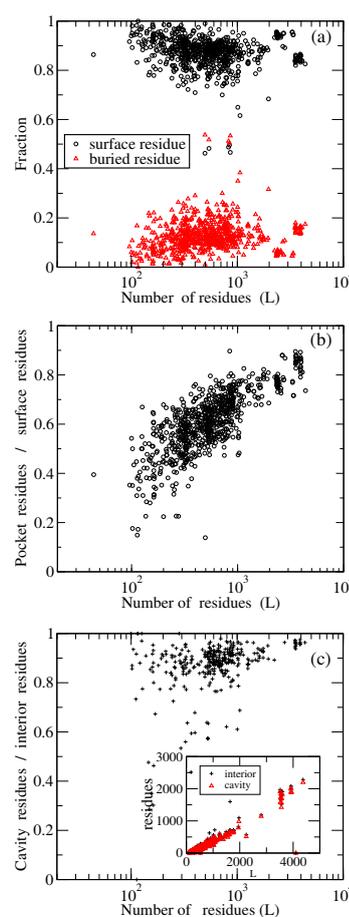


Fig. 13. (a) Fraction of surface *vs.* buried residues. (b) Fraction of residues lining pockets among surface residues. (c) Fraction of residues lining cavities among interior residues. Inset shows total number of interior and cavity residues *vs.* total protein residues.

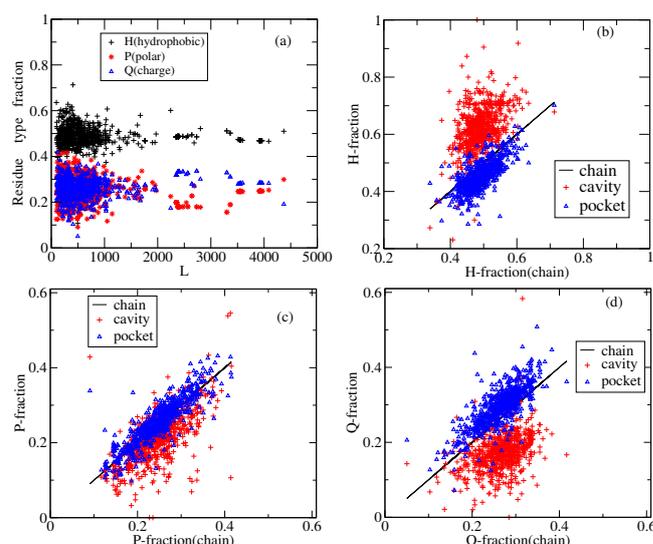


Fig. 14. (a) Fraction of hydrophobic residues out of all residues in the proteins. (b) Fraction of hydrophobic, (c) polar and (d) charged residues among residues lining cavities, and pockets.

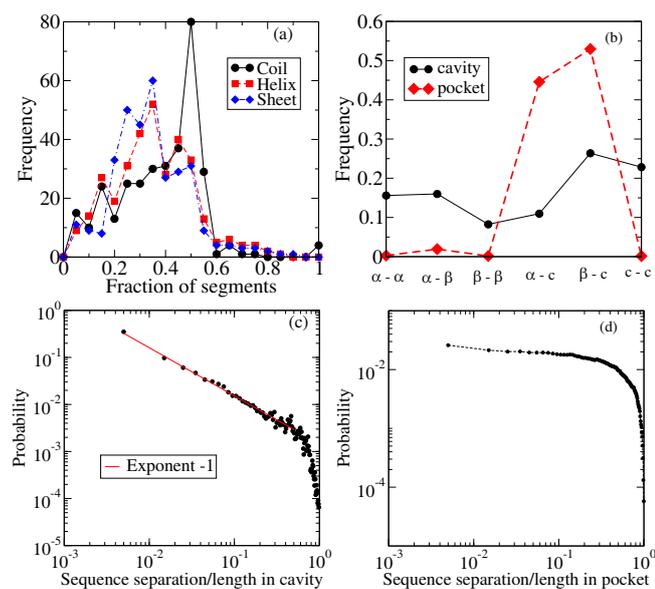


Fig. 15. Secondary structure and connectivity of residues lining cavities. (a) Percentage of α helix or β sheet secondary structure (S), and those which are part of coils (C). (b) Distribution of contact types (see text for the definition of contact type) in cavities and pockets. (c) Distribution of sequence separation between neighbor residues normalized by protein length lining cavities. The distribution is power law with exponent -1 . (d) Distribution of sequence separation between neighbor residues lining pockets.

lining cavities. In fig. 15(a) we show the percentage of α helix, β -sheet and coil for the whole protein. In fig. 15(b) we show the distribution of contact types between neighbor residues lining cavities and pockets. We represent α helix, β sheet and coil segments by α , β , and c . There are six possible contacts α - α , α - β , β - β , α - c , and c - c . We find that helix-helix, helix-sheet and sheet-sheet contacts are present in cavities. Surprisingly they are almost absent in pockets. Previous studies [47] reported that helix formation takes place in the early stages of protein folding. We have shown in fig. 14(b) that cavities are rich in hydrophobic residues, and it is of interest to ask what role hydrophobic collapse and secondary structure formation in the protein folding pathway play in the formation of cavities. As a characterization of the structure of the cavities, we study the sequence separation [47] between neighboring residues lining the cavities. We observe power law distribution of these separations (normalized to the protein length), as shown in fig. 15(c), whereas the separations between neighbors lining pockets (fig. 15(d)) show a much flatter distribution, with a cutoff imposed by the length of the proteins.

5 Conclusions

In this manuscript we discussed an algorithm for characterizing the void space in polydisperse sphere packings. We presented an algorithm for carrying out the radical

plane construction, by generalizing a previously developed method for the Voronoi construction. We employed the methodology to study monodisperse and polydisperse hard-sphere fluids and crystals, obtaining the equation of state, and the free-volume distributions. The free-volume distributions for polydisperse hard-sphere packings have not been evaluated before. We also applied the methods presented here to the analysis of protein structures. The presented results indicate that the methods we discuss can be employed very effectively in obtaining valuable information regarding the structure of proteins.

We wish to thank Dr. Niels Ellegaard for very useful discussions related to the use of replicated images in our tessellation algorithm. The invaluable help of Shiladitya Sengupta and Vishwas Vasisht in the preparation of the manuscript and their comments are gratefully acknowledged.

References

1. R.J. Speedy, *J. Chem. Soc. Faraday Trans II* **76**, 693 (1980).
2. A.E. Scheidegger, *The Physics of Flow through Porous Media* (University Toronto Press, Toronto, 1974).
3. S. Torquato, *Random Heterogeneous Materials: Microstructure and Macroscopic Properties* (Springer-Verlag, 2002).
4. B. Lee, F.M. Richards, *J. Mol. Biol.* **55**, 379 (1971).
5. F.M. Richards, *Methods Enzymol.* **115**, 440 (1985).
6. J. Finney, *J. Mol. Biol.* **96**, 721 (1975).
7. J. Liang, H. Edelsbrunner, C. Woodward, *Protein Sci.* **7**, 1884 (1998).
8. T.H. Connolly, *J. Appl. Cryst.* **16**, 548 (1983).
9. L.R. Dodd, D.N. Theodorou, *Mol. Phys.* **72**, 1313 (1991).
10. G. Perrot *et al.*, *J. Comput. Chem.* **13**, 1 (1992).
11. A. Poupon, *Curr. Opin. Struct. Biol.* **14**, 233 (2004).
12. H. Edelsbrunner, M. Facello, J. Liang, *Discrete Appl. Math.* **88**, 83 (1998).
13. J. Liang, H. Edelsbrunner, P. Fu, P.V. Sudhakar, S. Subramaniam, *Proteins: Struct., Funct., Genet.* **33**, 1 (1998).
14. J. Liang, H. Edelsbrunner, P. Fu, P.V. Sudhakar, S. Subramaniam, *Proteins: Struct., Funct., Genet.* **33**, 18 (1998).
15. S. Sastry, D.S. Corti, P.G. Debenedetti, F.H. Stillinger, *Phys. Rev. E* **56**, 5524 (1997).
16. S. Sastry, T.M. Truskett, P.G. Debenedetti, S. Torquato, F.H. Stillinger, *Mol. Phys.* **95**, 280 (1998).
17. M. Tanemura, T. Ogawa, N. Ogita, *J. Comput. Phys.* **51**, 191 (1983).
18. B.J. Gellatly, J.L. Finney, *J. Non-Cryst. Solids* **50**, 313 (1981).
19. A.R. Kerstein, *J. Phys. A* **16**, 3071 (1983).
20. S.C. van der Marck, *Phys. Rev. Lett.* **77**, 1785 (1996).
21. B.J. Alder, T.E. Wainwright, *Phys. Rev. Lett.* **18**, 988 (1967).
22. H. Reiss, H.L. Frisch, J.L. Lebowitz, *J. Chem. Phys.* **31**, 369 (1959).
23. J.D. Bernal, J. Mason, *Nature* **188**, 910 (1960).
24. F.H. Stillinger, E.A. DiMarzio, R.L. Kornegay, *J. Chem. Phys.* **40**, 1564 (1964).

25. S. Torquato, F.H. Stillinger, *J. Phys. Chem.* **105**, 11849 (2001).
26. S. Torquato, F.H. Stillinger, *J. App. Phys.* **102**, 093511 (2007).
27. S. Torquato, F.H. Stillinger, *Rev. Mod. Phys.* **82**, 2633 (2010).
28. G. Parisi, F. Zamponi, *Rev. Mod. Phys.* **82**, 789 (2010).
29. S. Slotterback, M. Toiya, L. Goff, J.F. Douglas, W. Losert, *Phys. Rev. Lett.* **101**, 258001 (2008).
30. T. Aste, G.W. Delaney, T. Di Matteo, *AIP Conf. Proc.* **1227**, 157 (2009).
31. G.E. Schröder-Turk *et al.*, *EPL* **90**, 34001 (2010).
32. J.G. Puckett, F. Lechenault, K.E. Daniels, *Phys. Rev. E* **83**, 041301 (2011).
33. K.A. Newhall, I. Jorjadze, E. Vanden-Eijnden, J. Brujic, *Soft Matter* **7**, 11518 (2011).
34. S.-C. Zhao, S. Sidle, H.L. Swinney, M. Schröter, *EPL* **97**, 34004 (2012).
35. D.S. Corti, R.K. Bowles, *Mol. Phys.* **96**, 1623 (1999).
36. H. Hansen-Goose, R. Roth, *J. Chem. Phys.* **124**, 154506 (2006).
37. N.F. Carnahan, K.E. Starling, *J. Chem. Phys.* **51**, 635 (1969).
38. P.J. Steinhardt, D.R. Nelson, M. Ronchetti, *Phys. Rev. B* **28**, 784 (1983).
39. G.A. Mansoori, N.F. Carnahan, K.E. Starling, T.W. Leland, jr., *J. Chem. Phys.* **54**, 1523 (1971).
40. V.S. Kumar, V. Kumaran, *J. Chem. Phys.* **123**, 114501 (2005).
41. M. Maiti, S. Sastry, preprint.
42. S.J. Hubbard, P. Argos, *J. Mol. Biol.* **261**, 289 (1996).
43. B. Honig, *J. Mol. Biol.* **293**, 283 (1999).
44. J. Liang, K.A. Dill, *Biophys. J.* **81**, 751 (2001).
45. M. Bueno, N. Cremades, J. Luis Neira, J. Sancho, *J. Mol. Biol.* **358**, 701 (2006).
46. J.R. Banavar, A. Maritan, *Rev. Mod. Phys.* **75**, 23 (2003).
47. C. Micheletti, J.R. Banavar, A. Maritan, F. Seno, *Phys. Rev. Lett.* **82**, 3372 (1999).
48. M. Maiti, S. Sastry, preprint.
49. H.S. Chan, S. Bromberg, K.A. Dill, *Philos. Trans. R. Soc. London Ser. B., Biol. Sci.* **348**, 61 (1995).