Voronoi tessellation-based algorithm for determining rigorously defined classical and generalized geometric pore size distributions

Samarth Agrawal,[1,2,*] Sandra Galmarini,[1,†] and Martin Kröger[2,3,‡]

[1]*Laboratory for Building Energy Materials and Components, Swiss Federal Laboratories for Science and Technology, Empa, Überlandstrasse 129, 8600 Dübendorf, Switzerland*

[2]*Polymer Physics, Department of Materials, ETH Zurich, CH-8093 Zurich, Switzerland*

[3]*Magnetism and Interface Physics, Department of Materials, ETH Zurich, CH-8093 Zurich, Switzerland*

(Received 25 September 2022; accepted 6 January 2023; published 25 January 2023)

The geometric pore size distribution (PSD) $P(r)$ as function of pore radius r is an important characteristic of porous structures, including particle-based systems, because it allows us to analyze adsorption behavior, the strength of materials, etc. Multiple definitions and corresponding algorithms, particularly in the context of computational approaches, exist that aim at calculating a PSD, often without mentioning the employed definition and therefore leading to qualitatively very different and apparently incompatible results. Here, we analyze the differences between the PSDs introduced by Torquato *et al.* and the more widely accepted one provided by Gelb and Gubbins, here denoted as T-PSD and G-PSD, respectively, and provide rigorous mathematical definitions that allow us to quantify the qualitative differences. We then extend G-PSD to incorporate the ideas of coating, which is significant for nanoparticle-based systems, and of finite probe particles, which is crucial to micro and mesoporous particles. We derive how the extended and classical versions are interrelated and how to calculate them properly. We next analyze various numerical approaches used to calculate classical G-PSDs and may be used to calculate the generalized G-PSD. To this end, we propose a simple yet sufficiently complicated benchmark for which we calculate the different PSDs analytically. This approach allows us to completely rule out a recently proposed algorithm based on radical Voronoi tessellation. Instead, we find and prove that the output of a grid-free classical Voronoi tessellation, namely, the properties of its triangulated faces, can be used to formulate an algorithm, which is capable of calculating the generalized G-PSD for a system of monodisperse spherical particles (or points) to any precision, using analytical expressions. The Voronoi-based algorithm developed and provided here has optimal scaling behavior and outperforms grid-based approaches.

DOI: 10.1103/PhysRevE.107.015307

I. INTRODUCTION

The characterization of the void region in porous materials is an important aspect for understanding their behavior and for engineering purposes. The role of the void region has been studied in the context of glassy systems [1–3], mechanical behavior of cellular solids [4], hydrophobicity in amino acids [5], phase transition in two-dimensional (2D) colloids [6], strength of concrete [7], etc. The void regions are characterized in terms of quantities such as solvent accessible surface area (SASA) [8,9] and pore size distribution (PSD) [10–12]. The latter is an important parameter for designing mesoporous and microporous materials which can be used for engineering applications such as gas storage [13–15] and separation of different components [16]. Experimentally, a PSD can be estimated using mercury porosimetry [11,12], however this method is intrusive and could potentially damage the sample during the measurement process. To avoid this issue, methods have been developed which extract a PSD from adsorption isotherms [17–19]. These methods typically assume the geometry of pores and shape of the distribution before fitting with the experimental data to estimate a PSD. The major drawback of such approaches is the assumption of pore geometry, and even state-of-the-art methods based on nonlocal density-functional theory [20,21] have shown to produce spurious results for some cases [22]. Therefore, an alternative approach wherein a PSD is directly calculated from the microstructure of the material serves as a viable solution. The microstructure in question can be either generated using nonintrusive high-resolution microscopy techniques [23–25] or via computational methods [26]. The mathematical definition of a pore size distribution, however, is not unique [27–29]. Definitions provided by Torquato *et al.* [30] and by Gelb *et al.* [31] differ significantly. While the former is based on the nearest-neighbor statistics, the latter derives the concept of a pore from the total accessible volume to a probe of given radius.

In the present work we focus on the definition provided by Gelb and Gubbins [31], also referred to as geometric pore size distribution $P(r)$, and denoted as classical G-PSD in the remainder of this paper because it provides results consistent with indirect methods such as Barrett-Joyner-Halenda (BJH) [17] and nonlocal density-functional theory (NLDFT) [20]

*samarth.agrawal@empa.ch
†sandra.galmarini@empa.ch
‡mk@mat.ethz.ch

for simple geometries such as cylinders, spheres while also providing meaningful results for more complex structures such as porous glasses [31]. It was demonstrated with extraordinary clarity that the T-PSDs defined by Torquato *et al.* [30] should not be used if a physically meaningful PSD is targeted [28]. Most literature available in the context of PSD, including computational approaches, consider the adsorbate gas molecule (used for studying the adsorption isotherm), to be a point sized sphere (probe radius $r_p = 0$). However, for the case of micro- and mesoporous materials such as metal organic frameworks, aerogels, etc. [32], the size of the adsorbate gas molecule is comparable to the size of cavities present in the adsorbent material and therefore has a significant effect on the adsorption isotherms and derived quantities such as SASA and PSDs [33]. Pinheiro *et al.* [34] have studied the probe-size dependent G-PSD for computationally generated Zeolite structures, but have not provided the equivalent mathematical expression for the Monte Carlo scheme used. For the class of nanoparticles and their porous composites the PSD is affected by a temperature- or solvent quality-depending coating thickness [35,36]. Such stabilizing coatings usually contain thermoresponsive polymers [37].

Here, we derive the analytical expression for probe radius r_p and coating-thickness r_c-dependent generalized G-PSD $P(r; r_p|r_c)$. We examine how it changes qualitatively for a simple benchmark case, for which the analytic solution is derived as well. We consider both finite coating thickness r_c and probe particle radius r_p mainly for the reason that they can be treated using identical methods. The G-PSD $P(r; r_p|0)$ for a naked system (uncoated, $r_c = 0$) with finite probe size r_p can be mapped onto the problem $P(r; 0|r_c)$ of a r_c-coated system with vanishing probe radius. The G-PSD for the coated system, in turn, is the $P(r)$ for the modified, r_c-coated original material, which we denote by $P(r|r_c)$. There are several important points to realize: (i) the G-PSD for a coated system is qualitatively different from the G-PSD for the original system and does not correspond to a simply r-shifted $P(r)$; (ii) the generalized $P(r; r_p|r_c)$ can be obtained from $P(r|r_{\text{eff}})$ with an effective coating thickness

$$r_{\text{eff}} = r_c + r_p, \tag{1}$$

as we going to explain; (iii) the two versions of T-PSDs introduced by Torquato *et al.* can be expressed in terms of a shifted $P(0|r)$ and are therefore qualitatively completely different to $P(r|0)$, as we will discuss in detail.

Different definitions of the PSD can be expressed in terms of the void volume accessible to a full r sphere in an r_{eff}-coated system denoted by

$$V(r|r_{\text{eff}}), \quad (r \geqslant 0, r_{\text{eff}} \geqslant -r_\circ).$$

This volume is significantly larger and more difficult to estimate than the volume accessible to the center of an r sphere; $V(r|r_{\text{eff}})$ is the most fundamental quantity in this work. The generalized G-PSD is expressed in terms of $V(r|r_{\text{eff}})$ in Eq. (4) below. All PSDs treated in this work and related quantities can be derived from $V(r|r_{\text{eff}})$, as we show later. However, it should be noted that this does not imply that an algorithm that determines a PSD must focus on calculating $V(r|r_{\text{eff}})$. Negatively coated particles appear in connection with the T-PSDs and the classical G-PSD requires knowledge of $V(r|0)$ only. Examples for the classical case of $V(r|0)$ and its geometrical construction for a 2D system composed of $N = 20$ circular material particles of radius $r_\circ = 0.1$ are shown in Fig. 1. The centers of the circular pores of radius r are located in a region whose points are further away than r from the surface of the material. To visualize this region, we show the euclidean distance maps (EDM) in the first column. Each point in an EDM map carries the distance to the closest center of a material circle. The Voronoi tessellation is an EDM map where only those points are shown as Voronoi edges that have equal distance to more than a single material circle. Each Voronoi cell surrounding a material circle then represents all points that are closer to this material circle than to any other material circle.

We revisit and eventually generalize existing algorithms, and provide a new one that is particularly suited for the calculation of the generalized $P(r; r_p|r_c) = P(r - r_p|r_{\text{eff}})$ for monodisperse systems composed of spherical particles of radius $r_\circ$. This gives rise to a second relevant radius we use later, the radius of the effectively coated particle,

$$r_s = r_\circ + r_{\text{eff}}. \tag{2}$$

The calculation of G-PSD is not straightforward, and various approaches have been suggested. For all such methods, the distribution is obtained at finite resolution: using M random insertions or M grid nodes. Algorithms that aim to fill the void space based on pixelated images of the material, have been suggested previously [38,39]. Various Monte Carlo schemes have also been proposed, including Bhattacharya *et al.* [40], wherein each Monte Carlo step is formulated as a constrained maximization problem and found it to be a computationally cheap. However, given the nature of optimization algorithms it is susceptible to locating local extrema as opposed to optimal solutions. Pinheiro *et al.* [34] suggested a semi-analytical solution to this problem by utilizing a radical Voronoi partition of the porous system. However, as we explain later, this method is actually erroneous and not suitable for calculating the G-PSD. Given its extensive use to characterize porous materials, the development of a fast algorithm which provides correct results is imperative. We adapt the existing numerical methods to generate probe size dependent G-PSD, and also present a Voronoi-based numerical method which can be used to calculate the same specifically for monodisperse systems composed of spherical particles. Furthermore, we compare the accuracy of the numerical algorithms presented for some simple configurations.

The definition of G-PSD and its generalization, as well as its connection to accessible volumes is provided in Sec. II. Existing numerical methods are revisited in Sec. III, along with a complete description of the new proposed Voronoi-based $O(M)$ algorithm for calculating the G-PSD for two-dimensional (2D) systems. Section IV presents our recommended benchmark, for which we derive analytic expressions in Sec. IV A. The benchmark serves to rule out the radial Voronoi approach in Sec. IV B. The differences between T-PSDs and the G-PSD are worked out in Sec. IV C. The convergence behavior and computational efficiency of the Voronoi-based algorithm is discussed in Sec. IV D, for the benchmark and also for a more "realistic" system, where computational approaches are the only means to estimate

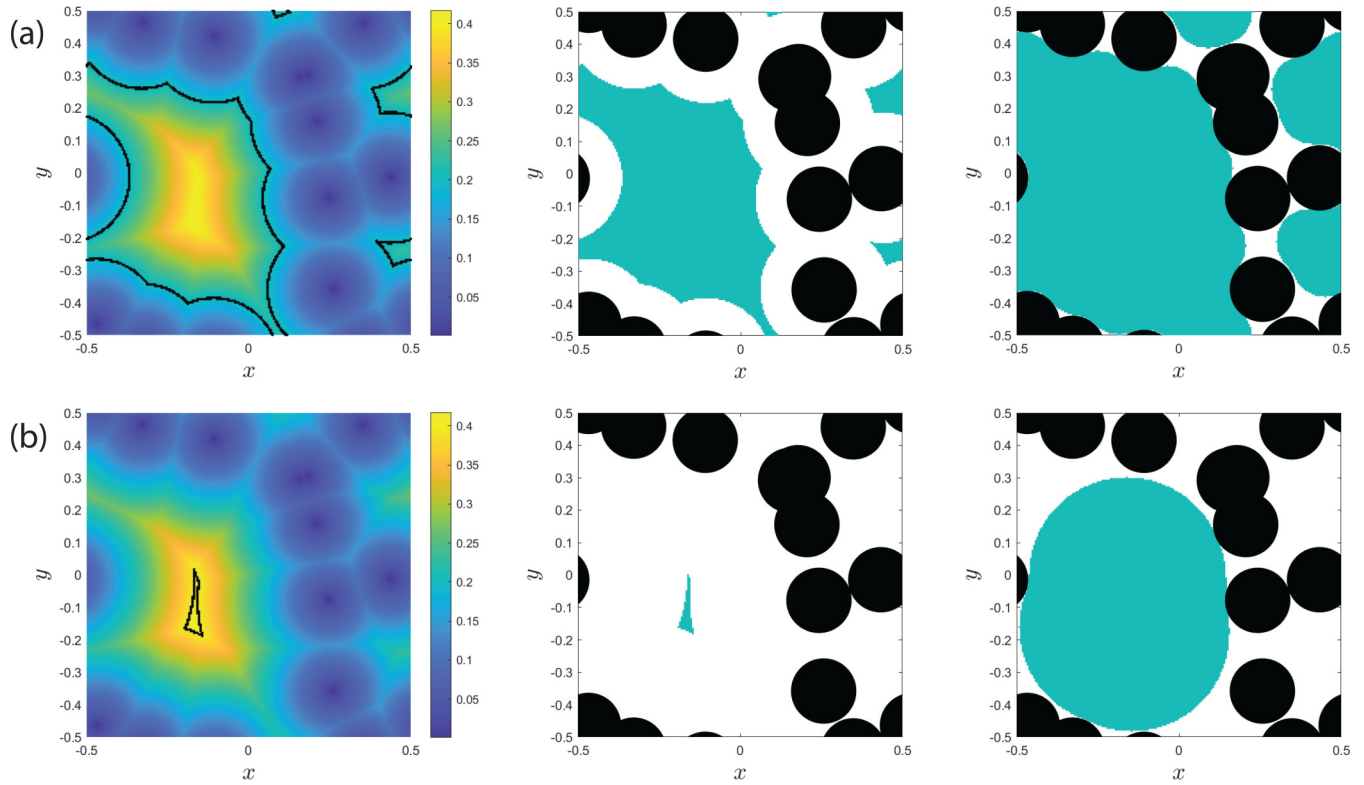

FIG. 1. Multidisk setup. Construction, shape and size of the volume $V(r) = V(r|0)$ for a 2D system consisting of 20 noncoated circles of radius $r_\circ = 0.1$ (filled circles) in a periodic box of side length $L = 1$ for two selected cases: (a) $r = 0.1$, and (b) $r = 0.3$. The first column shows the euclidean distance map (EDM, colored) constructed from the particles center positions (dark blue), as well as a single contour line (black) at altitude $r + r_\circ$. The second column highlights the region (green) enclosed by the EDM contour line; it carries all possible center coordinates of circular pores of radius r. The last column shows the volume $V(r|0)$ (green).

the PSD. Conclusions based on these results are provided in Sec. V.

II. GEOMETRIC PORE SIZE DISTRIBUTIONS G-PSD

For an arbitrary porous system in D dimensions, the space can be divided into r_c-coated material and void regions. Here, we denote by r_c-coated material the expanded material, whose outer surface is at distance r_c away from the original surface; $r_c = 0$ corresponds to the classical naked case. These coatings may overlap with themselves, and may also be regarded as hypothetical coatings for the purpose of this paper. Then, the size of a spherical pore for a given point $\mathbf{p}$ inside the void space is defined by the largest sphere that resides completely in the void region and contains this point $\mathbf{p}$ [40]. Such a D-dimensional sphere is termed as largest enclosing sphere (LES) and is represented as $r(\mathbf{p}|r_c)$, where r is its radius which is a function of the chosen point $\mathbf{p}$ subject to the condition that the material is r_c coated.

Classical G-PSD. The classical G-PSD, $P(r)$, is the probability density function of $r(\mathbf{p}) = r(\mathbf{p}|0)$ of the LES for uniformly chosen $\mathbf{p}$ in the void region of the naked material. It has been shown that $P(r)$ is defined as [31]

$$P(r) = -\frac{V'(r)}{V(0)} \quad (r \geqslant 0), \tag{3}$$

wherein $V(r)$ denotes the accessible pore volume, which is the total volume of the void region that is within reach to spheres of radius r, and $V'(r)$ represents the derivative with respect to r. The denominator $V(0)$ is then the total volume of the void region as the entirety of the void region is accessible to a point-sized sphere. As seen readily, $P(r)$ is normalized because, as $V(\infty) = 0$ for any finitely sized—or periodic—system, Eq. (3) implies $\int_0^\infty P(r)dr = 1$.

Generalized G-PSD. For the case of micro- and mesoporous materials such as metal organic frameworks, aerogels, etc. [32], the size of the adsorbate gas molecule (used for studying the adsorption isotherm) is comparable to the size of cavities present in the adsorbent material [33]. The total pore volume that is accessible experimentally to finitely sized probe particles is not necessarily the entirety of the void region, and hence the concept of the size of the pore at a point $\mathbf{p}$ is redefined as—the largest sphere that resides in the void region and completely contains the probe sphere centered at $\mathbf{p}$ with radius r_p. Such a sphere is hereby termed as the largest probe enclosing sphere (LPES) and is represented by $r(\mathbf{p}; r_p|r_c)$ because it is a function of the chosen point $\mathbf{p}$ as well as of the spherical probe radius r_p, within the r_c-coated material. A schematic figure showing the LES $r(\mathbf{p}) = r(\mathbf{p}|0)$ and LPES $r(\mathbf{p}; r_p) = r(\mathbf{p}; r_p|0)$ for a given $\mathbf{p}$ and r_p, in an arbitrary naked landscape, are respectively shown in Figs. 2(a) and 2(b). The dashed lines around the material particles in Fig. 2(d) to form the surface of the hypothetically r_p-coated material, and the region marked by the dashed circle, of radius $r - r_p$ inside the LPES is the total volume accessible to the center position $\mathbf{p}$ of the finitely sized probe particle, while the

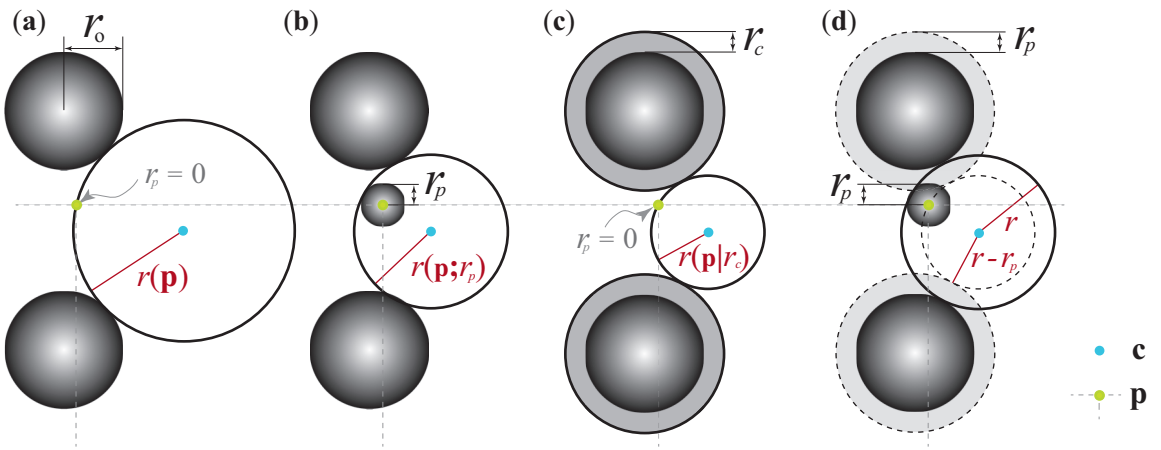

FIG. 2. Enclosing spheres for a material with just two solid 3D spheres. (a) Largest enclosing sphere (LES, black empty circle, actually representing a sphere) of radius $r(\mathbf{p})$, centered at $\mathbf{c}$ (cyan), containing the probe point particle centered at $\mathbf{p}$ (green). (b) Largest probe enclosing sphere (LPES, black empty circle) of radius $r(\mathbf{p};r_p)$ centered at another $\mathbf{c}$, containing the probe sphere (radius $r_p > 0$) centered at $\mathbf{p}$. (c) LES of radius $r(\mathbf{p}|r_c)$ within the r_c-coated material. (d) A hypothetical material shell of size r_p can be used to calculate $r(\mathbf{p};r_p) - r_p = r(\mathbf{p}|r_p)$, cf. Eq. (4), and to reduce the LPES problem to the setup shown in panel (c), where r_c is replaced by an effective radius $r_{\text{eff}} = r_c + r_p$, cf. Eq. (1). The distance $R(\mathbf{c})$ between $\mathbf{c}$ and the center of the nearest material particle (radius $r_\circ$) is different in panels (a)–(c). For sufficiently large probe particles, there are regions in the void space (coating overlap) that cannot host a finitely sized probe particle. Note that the surface of the probe particle must not touch the LES or LPES in general, if there are more than two material spheres. This will become obvious later below.

volume accessible to the probe particle itself also includes the portions of the hypothetical coating.

We can find the expression for the generalized G-PSD, $P(r;r_p|r_c)$, by reducing it to a class of problem treated by Eq. (3). The point $\mathbf{p}$ cannot be chosen within a distance of r_p from the r_c-coated material region to prevent the probe particle from overlapping. Hence, we can operate in a modified system wherein all the material elements are hypothetically coated with an additional thickness of r_p. For the r_c-coated system, the volume accessible to a sphere of radius r is denoted by $V(r|r_c)$. Therefore the entirety of the void region of the $r_{\text{eff}} \equiv r_c + r_p$-coated material is represented by $V(0|r_{\text{eff}})$; any possible probe particle center resides within this volume. Any LES which contains the center of the probe particle is then effectively the LPES, as extension into the imaginary coating is permissible. Therefore the problem of finding the LPES is now reduced to the original problem of finding the LES in the coated system. Figure 2 illustrates the idea of finding the LES and the LPES via hypothetical coating, while Fig. 3 shows the change in the EDM for finite r_{eff}. Based on these considerations, the generalized G-PSD is

$$P(r;r_p|r_c) = -\frac{V'(r-r_p|r_{\text{eff}})}{V(0|r_{\text{eff}})} \quad (r \geqslant r_p), \tag{4}$$

where the prime denotes a derivative with respect to r, and where we recall that $V(r|r_{\text{eff}})$ is the fundamental void volume in the r_{eff}-coated system that is accessible to a full r sphere.

A special case is the naked version (3) since $V(r) = V(r|0)$. If we write Eq. (4) as

$$P(r+r_p;r_p|r_c) = -\frac{V'(r|r_{\text{eff}})}{V(0|r_{\text{eff}})} \quad (r \geqslant 0), \tag{5}$$

with the effective coating $r_{\text{eff}} = r_c + r_p$, assumes the same form as Eq. (1), therefore indicating that G-PSD depends on r and r_{eff} only. It is therefore in every case sufficient to study either the case $P(r;0|r_c)$ of point-like probe particles in the original material with variable coating, or the case $P(r;r_p|0)$ of r_p-sized probe particles in the naked system, to obtain the full $P(r;r_p|r_c)$. If the material consists of monodisperse spheres of equal radius $r_\circ$, these particles can also be regarded as $r_\circ$-coated points. For the special case of point-like probe particles in naked material, Eq. (4) reduces to Eq. (3).

With Eq. (4) at hand, the cumulative distribution, $P_{\text{cum}}(r;r_p|r_c) = \int_{r_p}^{\infty} P(\rho;r_p|r_c)d\rho$ corresponding to the generalized G-PSD is

$$P_{\text{cum}}(r+r_p;r_p|r_c) = 1 - \frac{V(r|r_{\text{eff}})}{V(0|r_{\text{eff}})} \quad (r \geqslant 0), \tag{6}$$

where $V(0|r_{\text{eff}})$ is the total void volume in the effectively or hypothetically r_{eff}-coated system. As seen the proper limits are $P_{\text{cum}}(r_p;r_p|r_c) = 0$ and $\lim_{r\to\infty} P_{\text{cum}}(r;r_p|r_c) = 1$, for any r_c and any r_p, confirming that $P(r;r_p|r_c)$ is indeed a properly normalized probability density. The mean pore radius as function of r_p and r_c is

$$\langle r\rangle = \int_{r_p}^{\infty} rP(r;r_p|r_c)dr = r_p + \frac{\int_0^{\infty} V(r|r_{\text{eff}})dr}{V(0|r_{\text{eff}})}. \tag{7}$$

These expressions allow us to calculate the G-PSD from V, which is the easier quantity for analytic calculations. We now calculate the G-PSD for some regular shaped geometries. In case of a r_{eff}-coated cylinder of radius $r_\bigcirc > r_{\text{eff}}$ and length L, we have $V(r|r_{\text{eff}}) = \pi(r_\bigcirc - r_{\text{eff}})^2 L$ for $r \leqslant r_\bigcirc - r_{\text{eff}}$, and $V(r|r_{\text{eff}}) = 0$ otherwise. According to Eq. (4), $P(r;r_p|r_c) = \delta(r_\bigcirc - r_c - r)$ has a single dirac-distributed peak as long as $r_{\text{eff}} \leqslant r_\bigcirc$ and vanishes otherwise. This examples highlights the fact that $r_{\text{eff}} = r_c + r_p$ can be freely distributed over probe radius and coating layer thickness, and how this splitting affects the G-PSD.

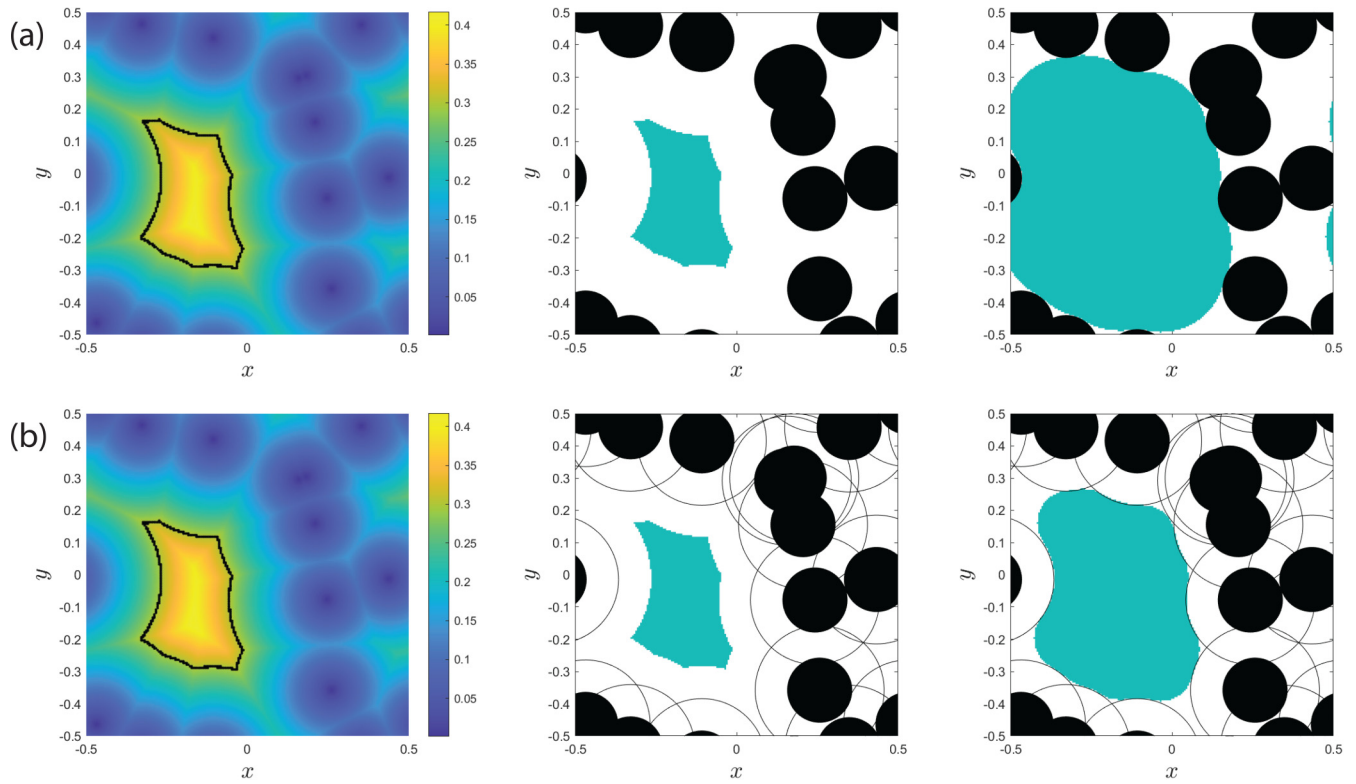

FIG. 3. Multidisk setup. Construction, shape and size of the volume $V(r|r_{\text{eff}})$ for the system treated already by Fig. 1, here for two selected cases which share their sum $r + r_{\text{eff}} = 0.2$: (a) $r = 0.2$, $r_{\text{eff}} = 0$, (b) $r = r_{\text{eff}} = r_c + r_p = 0.1$ (for example $r_c = 0.1$ and $r_p = 0$ or $r_c = 0$ and $r_p = 0.1$). The first column shows the euclidean distance map (EDM, colored) constructed from the particles center positions (dark blue), as well as a single contour line (black) at altitude $r + r_s$. The second column highlights the region (green) enclosed by the EDM contour line; it carries all possible center coordinates of circular pores of radius r. The black circles in panel (b) mark the outer surface of r_{eff}-coated particles. The last column shows the volume $V(r|r_{\text{eff}})$ (green), i.e., $V(0.2, 0)$ in panel (a) and $V(0.1, 0.1)$ in panel (b). All variants of PSD's are encoded in the two-parametric $V(r|r_{\text{eff}})$. The numerical result for $V(r|r_{\text{eff}})$ over the whole range of semipositive r and r_{eff} for this setup is shown in Fig. 6, along with the corresponding results for the generalized G-PSD.

On the other hand, for a hollow cube, $P(r; r_p|r_c)$ is very sensitive to r_p and r_c, as the fraction of void region accessible to the probe in the neighborhood of the corners changes considerably with r_p and r_c. Some cases for which all the quantities $V(r|r_{\text{eff}})$ and $P(r; r_p|r_c)$ can be calculated analytically, are collected in Appendix. For more complicated systems, such as those generated using methods such as molecular dynamics, calculating the geometric G-PSD analytically is impossible [31,41] and numerical methods must be employed. In the following section, we briefly describe existing numerical methods that can, or cannot be used to calculate geometric G-PSD $P(r; r_p|r_c)$ and also propose a new method based on Voronoi tessellation that can be used for systems made of monodisperse spherical particles.

As discussed in detail, we can analyze the case of point-sized probe particle in r_{eff}-coated material instead of considering nonvanishing r_p and r_c separately. To this end we finally define, using Eqs. (5) and (6),

$$P_{\text{cum}}(r|r_{\text{eff}}) \equiv P_{\text{cum}}(r; 0|r_{\text{eff}}) = 1 - \frac{V(r|r_{\text{eff}})}{V(0|r_{\text{eff}})} \quad (r \geqslant 0),$$

$$P(r|r_{\text{eff}}) \equiv \frac{dP_{\text{cum}}(r|r_{\text{eff}})}{dr} = -\frac{V'(r|r_{\text{eff}})}{V(0|r_{\text{eff}})} \quad (r \geqslant 0), \quad (8)$$

keeping in mind that these quantities contain the generalized G-PSD and cumulative G-PSD as function of r_c and r_p via Eqs. (5) and (6). To be clear, the generalized G-PSD is obtained from the two-parametric $V(r|r_{\text{eff}})$ via

$$P(r; r_p|r_c) = P(r - r_p|r_{\text{eff}}), \quad (9)$$

$$P_{\text{cum}}(r; r_p|r_c) = P_{\text{cum}}(r - r_p|r_{\text{eff}}). \quad (10)$$

With the help of $P_{\text{cum}}(r|r_{\text{eff}})$ the mean pore radius can also be written as $\langle r \rangle = r_p + \int_0^\infty [1 - P_{\text{cum}}(r|r_{\text{eff}})]\, dr$.

T-PSDs. Using unchanged notation, we can provide a mathematical definition of the two T-PSD versions suggested by Torquato *et al.* [30], which are only applicable for the special case of $r_c = r_p = r_{\text{eff}} = 0$ (for a derivation see Sec. IV C; to avoid any confusion we label them by $\mathcal{P}$ instead of P). There is on one hand a center-based T_c-PSD, $\mathcal{P}^c(r) = d\mathcal{P}^c_{\text{cum}}(r)/dr$ with

$$\mathcal{P}^c_{\text{cum}}(r) = 1 - \frac{V(0|r - r_\circ)}{V(0|-r_\circ)} \quad (r \geqslant 0), \quad (11)$$

and on the other a surface-based T_s-PSD, defined via $\mathcal{P}^s_{\text{cum}}(r) = \mathcal{P}^c_{\text{cum}}(r + r_\circ)$ for $r \geqslant 0$. The qualitative differences between the three PSDs (G versus center-based T_c-PSD and surface-based T_s-PSD) are highlighted by the two examples

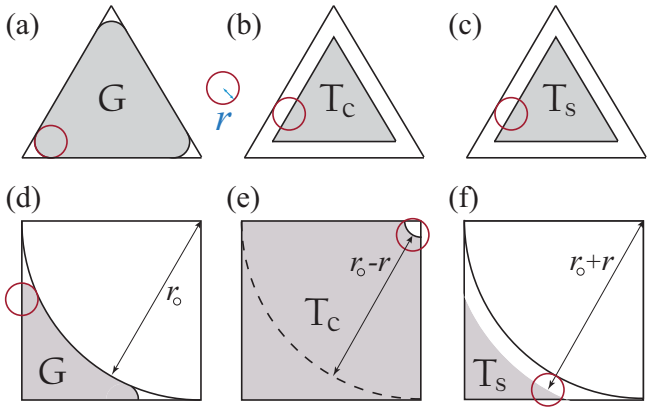

FIG. 4. Relevant PSD volumes for the triangle and our benchmark setup. Drastic differences between the relevant (gray-shaded) areas used to define the G-PSD $P(r)$ of Gelb & Gubbins (G) [31] as well as the T_c-PSD and T_s-PSD of Torquato *et al.* [30]. While for the T_c-PSD (T_s-PSD) the red circle touches at least one center (one surface) of a material particle, for the G-PSD it touches at least two surfaces of material particles. (a)–(c) Naked triangle serving as material. Its edges can be thought of consisting of infinitely many circles of radius $r_\circ = 0$. (d)–(f) Our naked benchmark setup, consisting of a single material particle of radius $r_\circ$ subject to periodic boundary conditions, i.e., tangent to its periodic image. (a) Shaded area $V(r|0)$ accessible to a circle of radius r, relevant for the G-PSD. (b), (c) Shaded area $V(0|r - r_\circ) = V(0|r)$ enclosed by r-coated triangle, relevant for both T_c-PSD and T_s-PSD, because $r_\circ = 0$ for the triangle. (d) One quarter of the area $V(r|0)$ (gray-shaded) accessible to a circle of radius r, relevant for the G-PSD. (e) One quarter of the area $V(0|r - r_\circ)$ (gray-shaded) created by the material circle coated by a shell of thickness $r - r_\circ$ (for this particular example, negatively coated), relevant for the T_c-PSD. (f) One quarter of the area $V(0|r)$ (gray-shaded) created by the material circle coated by a shell of thickness r, relevant for the T_s-PSD. While the T_c- and T_s-PSDs are both trivially calculated at minor computational effort, the efficient calculation of the G-PSD still poses a challenge, as these simple examples might already suggest.

provided in Fig. 4. While the first example [Figs. 4(a)–4(c)] is a simple triangle whose edges consist of infinitely many material points ($r_\circ = 0$), the second example [Figs. 4(d)–4(f)] will become our benchmark case composed of a single material point subject to periodic boundary conditions, to be discussed in detail in Sec. IV. It is interesting to note that the T-PSDs are linear in r for empty cylinders and spheres, while the G-PSD to which we focus in this work considers a spherical inclusion of radius $r_\bigcirc$ to be a well-defined pore with unique radius $r_\bigcirc$. As this Fig. 4 attempts to explain by showing the relevant volumes entering the PSD definitions, the versions produce qualitatively different pore size distributions (Sec. IV C).

Having provided proper mathematical definitions of the existing PSDs, explained their differences, and added the definition of the generalized G-PSD, it is an apparently straightforward exercise to calculate all these quantities numerically, e.g., by evaluating the two-parametric function $V(r|r_{\text{eff}})$ defined in Sec. I. Many previous works have reported PSD curves, however, a careful inspection of the literature shows that most so-called PSD curves show something that is either not well-defined, or as easy to compute as the T-PSDs, while the calculation of the most useful G-PSD is much more difficult to compute without making approximations or introducing a finite grid resolution. After revisiting possible approaches we derive an algorithm that is approximation- and grid-free, efficient as it finds the coordinates of the LPES semi-analytically, and exhibits the best possible $O(M)$ scaling.

III. EXISTING NUMERICAL METHODS FOR GEOMETRIC AND GENERALIZED GEOMETRIC G-PSDs

Most numerical methods proposed in literature focus on determining the G-PSD $P(r) = P(r|0)$ for the naked material via a Monte Carlo scheme wherein large number of $\mathbf{p}$ is chosen uniformly in the void region and the corresponding LES is determined to build the histogram and subsequently, the approximate distribution. Past works have mostly focused on finding fast and accurate algorithms to determine the LES. We briefly explain the existing methods, mention how they can be adapted to calculate LPES or the generalized $P(r|r_{\text{eff}})$ based on LES for the coated material, and propose a new algorithm for determining $P(r|r_{\text{eff}})$ for a material consisting of N monodisperse spheres in Sec. IV.

A. Brute force scan using a grid

The possibly most straightforward approach to directly estimate $P(r|r_{\text{eff}})$ rather than $V(r|r_{\text{eff}})$ for a fixed coating thickness r_{eff} at high memory and computational cost is by considering a regular lattice of M coordinates $\{\mathbf{p}_i\}$ in the void space of the r_{eff}-coated structure, with an initially zeroed LES field $r(\mathbf{p}_i|r_{\text{eff}})$. Note here already that the LES $r(\mathbf{p}|r_{\text{eff}})$ for the coated system is not identical to the LPES $r(\mathbf{p};r_{\text{eff}})$ for the naked system, and that this difference matters, if one is not only interested in the classical $P(r)$. For each of the lattice points $\mathbf{c} \in \{\mathbf{p}_i\}$ all the nodes residing within the volume of the largest sphere centered at $\mathbf{c}$, with radius $R(\mathbf{c}|r_{\text{eff}})$, are visited and the r values of all those nodes are updated to $\max(r, R(\mathbf{c}|r_{\text{eff}}))$. $R(\mathbf{c}|r_{\text{eff}})$ is determined by the distance between $\mathbf{c}$ and coated material surface next to $\mathbf{c}$. For a substrate consisting of N polydisperse spheres of radius $r_\circ^k$, centered at $\mathbf{x}^k$, the radius $R(\mathbf{c}|r_{\text{eff}})$ of the largest sphere at $\mathbf{c}$ fully residing in the void is given by $R(\mathbf{c}|r_{\text{eff}}) = R(\mathbf{c}) - r_{\text{eff}}$ with the usual smallest distance

$$R(\mathbf{c}) = \min_{k=1,\ldots,N} \left\{ |\mathbf{c} - \mathbf{x}^k| - r_\circ^k \right\}, \tag{12}$$

where $|\cdots|$ represents the Euclidean distance between two points. For other types of substrates, $R(\mathbf{c})$ can be calculated from the shape of the material surface, and $R(\mathbf{c}) \geqslant r_{\text{eff}}$ by construction. At the end of this procedure one is left with a list of size M whose entries are $r(\mathbf{p}_i|r_{\text{eff}})$ values. The histogram of all these r values is $P(r|r_{\text{eff}})$ multiplied by a constant, and this constant is simply given by $P_{\text{cum}}(\infty|r_{\text{eff}}) = 1$.

The accuracy of the method depends on the resolution of the lattice and eventually requires large amounts of memory to store the details of the lattice. If an upper limit for a pore radius is determined in advance, memory can be saved and resolution increased upon applying the grid-based approach to regions of a size smaller than a multiple of the limiting

pore radius. The lattice can also be displaced and the procedure repeated to increase the resolution. If the grid spacing is δ, the computational effort of the grid-based approach is $\approx M \times N$ for the generation of the R grid, and $\approx M(\langle r\rangle/\delta)^D$ for the accumulation of the D-dimensional r grid, for each r_{eff}. If $M \gg N$, the overall computational effort is therefore somewhere between $O(M)$ and $O(M^2)$, and the details of the PSD can only be obtained up to finite resolution δ.

B. Constrained local numerical optimization

The problem of finding the LES radius and its center position $\mathbf{c}$ for a given $\mathbf{p}$ and effective coating thickness r_{eff} can also be formulated as a nonlinear constrained optimization problem, using $R(\mathbf{c})$ from the previous section, as

$$\begin{aligned} &\text{maximize} \quad R(\mathbf{c}|r_{\text{eff}}) = R(\mathbf{c}) - r_{\text{eff}}, \\ &\text{constraint} \quad R(\mathbf{c}|r_{\text{eff}}) \geqslant |\mathbf{c}-\mathbf{p}|, \end{aligned} \tag{13}$$

wherein the constraint guarantees that point probe particle is completely covered by the R sphere and does not overlap with the coating. A main difference with the grid-based approach is that $\mathbf{c}$ and $\mathbf{p}$ must both not reside on a grid, and $\mathbf{c}$ can reside anywhere within the void space. As for most off-lattice problems of this kind, this system of equation and inequality cannot be strictly solved using a nonlinear programming module, as the landscape can be highly nontrivial. While the grid-based approach cannot miss the constrained extremum within the given resolution, the off-lattice version does not guarantee a certain resolution, but can potentially lead to correct results at infinitely large resolution. The optimization procedure is repeated M times to result in a G-PSD $P(r|r_{\text{eff}})$ that has a resolution comparable with the grid-based approach. The $P(r; r_p|r_c)$ is then obtained from $P(r|r_{\text{eff}})$ via Eq. (10). As opposed to the grid-based approach, $R(\mathbf{c})$ cannot be calculated once and stored but must also be calculated each time it is required. If the average step size of the solver is denoted by δ, the computational effort is $MN\langle r\rangle/\delta$.

C. Radical Voronoi tessellation

Voronoi tessellation techniques partition the space using a distance metric. In the standard method, given a set of N points $\{\mathbf{x}^k\}$, the physical space can be divided into regions based on the Euclidean distance from each point. In the current context, the space is partitioned into N regions each corresponding to one of the points. These regions are also known as Voronoi cells. The Voronoi cell of point i contains the set of all positions, $\mathbf{r}$, which satisfy

$$\forall_{j=1,\ldots,N} \quad |\mathbf{r}-\mathbf{x}^i| \leqslant |\mathbf{r}-\mathbf{x}^j|. \tag{14}$$

Such a partition of space results in cells shaped as irregular polyhedra, with the faces, edges and vertices being sets of points that are equidistant with respect to adjacent particle centers. To find the LES of yet unknown radius $r(\mathbf{p})$ centered at some $\mathbf{c}$, Pinheiro *et al.* [34] proposed to consider, in addition, the Voronoi tessellation for the $(N+1)$ points $(\mathbf{p}, \mathbf{x}^1, \ldots, \mathbf{x}^N)$. Then, the cell corresponding to the point $\mathbf{p}$, hereby denoted as $\mathbf{V}_p$, termed the "ghost cell," contains the set of all coordinates $\mathbf{r}$ which satisfy

$$\forall_{j=1,\ldots,N} \quad |\mathbf{r}-\mathbf{p}| \leqslant |\mathbf{r}-\mathbf{x}^j|. \tag{15}$$

For the case of naked point particles ($r_\circ^i = r_c = 0$), and point probe particles ($r_p = 0$), the center $\mathbf{c}$ of the largest included sphere that contains $\mathbf{p}$ then resides on either the vertices of the N- or $(N+1)$-Voronoi meshes; the vertex with the largest $R(\mathbf{c})$ serves as the solution to the LES problem. Given this property, it was further proposed [34] that the radical Voronoi tessellation could be adapted to find the center and radius of the LPES if both the material and the probe particle are composed of spheres with finite radii, i.e., for $r_\circ^i > 0$ and $r_p > 0$. In the case of radical Voronoi tessellation, the space is partitioned according to modified Eq. (14):

$$\forall_{j=1,\ldots,N} \quad |\mathbf{r}-\mathbf{x}^i|^2 - \left(r_\circ^i\right)^2 \leqslant |\mathbf{r}-\mathbf{x}^j|^2 - \left(r_\circ^j\right)^2, \tag{16}$$

where $r_\circ^i$ corresponding to $\mathbf{p}$ is r_p, when the radical Voronoi is applied to $N+1$ points via Eq. (15). The metric $|\mathbf{r}-\mathbf{x}^i|^2 - (r_\circ^i)^2$ is essentially square of the tangential length. It was suggested that this metric is used instead of a more intuitive $|\mathbf{r}-\mathbf{x}^i| - r_\circ^i$ because it results in irregular polyhedra much like the standard Voronoi tessellation. Using this approach, the center $\mathbf{c}$ of the LPES is then assumed to lie either on one of the radical Voronoi vertices of the ghost cell, or on one of the radical Voronoi vertices using N spheres. We can, however, show that this approach exhibits a major deficiency in methodology and is not suitable for all but the simplest geometries. In particular, one can show analytically, that the center of the LPES cannot coincide with the vertices of the ghost cell.

Let $\mathbf{c}_p$ be a vertex of $\mathbf{V}_p$, then as per the definition of radical Voronoi tessellation, it satisfies

$$|\mathbf{c}_p-\mathbf{p}|^2 - r_p^2 = |\mathbf{c}_p-\mathbf{x}^i|^2 - r_\circ^2, \tag{17}$$

where i is index of all particles which share the vertex $\mathbf{c}_p$, and where we assumed monodisperse material particles, $r_\circ^i = r_\circ$, only to shorten the following expressions. Then, the LPES radius $r(\mathbf{p}; r_p) = R(\mathbf{c}_p)$ must be large enough to completely cover the probe particle. Mathematically, these conditions can be represented as $R(\mathbf{c}_p) - r_p \geqslant |\mathbf{c}_p - \mathbf{p}|$ and $R(\mathbf{c}_p) + r_\circ \leqslant |\mathbf{c}_p - \mathbf{x}^i|$, where we basically reiterated Eqs. (12) and (13). These inequalities imply

$$0 \leqslant |\mathbf{c}_p-\mathbf{x}^i| - r_\circ - |\mathbf{c}_p-\mathbf{p}| - r_p. \tag{18}$$

Upon multiplying Eq. (18) with the positive $|\mathbf{c}_p-\mathbf{x}^i| + r_\circ$ one obtains with the help of Eq. (17)

$$\begin{aligned} 0 &\leqslant |\mathbf{c}_p-\mathbf{x}^i|^2 - r_\circ^2 - (|\mathbf{c}_p-\mathbf{p}| + r_p)(|\mathbf{c}_p-\mathbf{x}^i| + r_\circ) \\ &= |\mathbf{c}_p-\mathbf{p}|^2 - r_p^2 - (|\mathbf{c}_p-\mathbf{p}| + r_p)(|\mathbf{c}_p-\mathbf{x}^i| + r_\circ). \end{aligned} \tag{19}$$

Upon dividing by the positive $|\mathbf{c}_p-\mathbf{p}| + r_p$, this inequality implies

$$|\mathbf{c}_p-\mathbf{p}| - r_p \geqslant |\mathbf{c}_p-\mathbf{x}^i| + r_\circ. \tag{20}$$

This last inequality, however, negates inequality (18) and is only possible when $r_\circ = r_p = 0$ and $|\mathbf{c}_p-\mathbf{p}| = |\mathbf{c}_p-\mathbf{x}^i|$. The latter equality is ensured by the radical Voronoi construction, if the former is fulfilled. As the radical Voronoi reduces to the classical Voronoi construction for $r_\circ = r_p = 0$, the radical Voronoi construction can simply not be used for calculating the classical or generalized G-PSD. We are going to calculate the radical Voronoi vertices for the benchmark used within

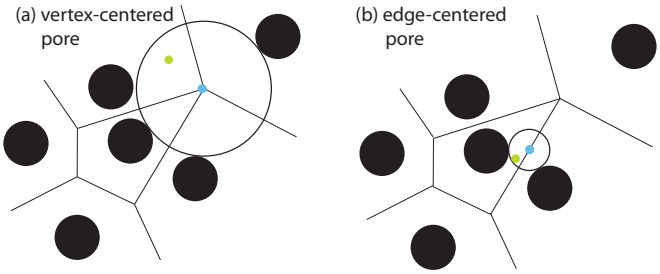

FIG. 5. The Voronoi-based algorithm locates the center **c** (blue point) and radius of the pore (black open circle) for a given point **p** (green) either on (a) a vertex or (b) an edge of the Voronoi diagram (black straight lines) corresponding to the material circles (black solid circles). We calculate the position of **c** analytically, $t \in \{0, 1\}$ for vertex-centered pores, and $t \in (0, 1)$ for edge-centered pores in Sec. III D.

this work in Sec. IV B to confirm that they do not lead to the correct $R(\mathbf{c}|r_{\text{eff}})$.

D. Voronoi-based global analytical optimization

We have already seen that the radical Voronoi construction cannot be used to calculate the G-PSD, except where it reduces to the classical Voronoi, i.e., for the case of point-like probe and point-like, uncoated material particles. However, the information contained in the Voronoi network of the N material particles alone can be used devise an algorithm that greatly improves over the suitably adapted classical approaches (Sec. III B).

Finding the center **c** of the LES or LPES for a given **p** can be simplified by considering the fact that they must touch at least two spheres at the same time (Fig. 5). In the case of $D = 2$ dimensions, this amounts to the center **c** residing on an edge of any of the N Voronoi cells, and on any of the Voronoi faces in case of $D = 3$. Instead of randomly walking through continuous D-dimensional space as done within constrained optimization in its original form, we can restrict the algorithm to inspect a number of relevant $(D-1)$-dimensional faces. For each such face, the optimum solution satisfying the constraints can be solved analytically not only in 2D but also in three dimensions (3D). Each face gives rise to a candidate pore radius; the largest among those candidate pore radii is the pore radius for the given **p**.

Here, we are limiting ourselves to derive and describe the 2D algorithm that returns the generalized G-PSD $P(r|r_{\text{eff}})$ for a probe circle (radius r_p) in a system of eventually overlapping monodisperse material circles (radius $r_\circ$, coating thickness r_c). The corresponding algorithm for the 3D case can be treated in an analogous fashion. Once $P(r; r_{\text{eff}})$ has been computed, $P(r; r_c|r_p)$ is immediately available via Eq. (10). As a first step, we make use of an existing algorithm to calculate, for each material particle center $\mathbf{x} \in \{\mathbf{x}^k\}$, the vertex coordinates of its Voronoi cell. Each Voronoi cell can be uniquely subdivided into nonoverlapping adjacent triangles that all meet at **x**. Importantly, as the $T = O(N)$ triangles corresponding to the N material circles are all known from the Voronoi construction at minor $O(N)$ effort, to subsequently calculate the G-PSD using $M \gg T$ probe particle insertions efficiently, and with proper statistics, without searching for each insertion point **p** all possible candidate edges, we can visit each of the T triangular regions and its finite number of candidate edges just once. To this end we randomly pick a known but triangle-dependent number m_Δ of points **p** within each triangle $\Delta \in \{1, \ldots, T\}$, such that m_j equals M times the void area covered by the triangle Δ within the system of r_{eff}-coated material circles, divided by the total void area $V(0|r_{\text{eff}})$ available in the r_{eff}-coated system. This implies that, if a given triangle does not have any overlap with the void space, no **p** have to be placed in this region because this triangle does not contribute to $V(r|r_{\text{eff}})$. For each of the m_Δ different positions **p** located within a triangle Δ (whose three vertices are respectively at the center position **x** of the material particle to which it belongs and at two adjacent vertices of the Voronoi face belonging to **x**; the two relative vectors $\mathbf{A}_\Delta$ and $\mathbf{B}_\Delta$ point from **x** to the other two vertices), we have to calculate a center **c** and radius r of the LES that includes the point **p** and does not overlap with the r_{eff}-coated material circles. Once calculated, the G-PSD $P(r|r_{\text{eff}})$ is the probability density constructed from the M different r values. There are now two equally important ingredients of our proposed algorithm left to be described, that are key to the speed of execution and optimal scaling of CPU time $\propto M$: (i) the generation of static subsets E_Δ of potentially relevant Voronoi edges for the given triangle Δ, and (ii) an analytic expression for a candidate r_e value (and its $\mathbf{c}_e$) for the given Voronoi edge $e \in E_\Delta$ and point **p**. The final pore radius for **p** is then given by the largest candidate r_e with $e \in E_\Delta$. (i) The generation of static subsets $E_\Delta \subset \{1, \ldots, T\}$ is useful as the number of pore-relevant edges for points **p** located inside any of the T triangles is usually much smaller than both m_Δ and N. A list E_Δ is obtained by first calculating the shortest distances $\{s_e\}$ between the triangle Δ and all the $T/2$ different Voronoi edges e (connecting $\mathbf{A}_e$ with $\mathbf{B}_e$). All those edges e for which $\max(|\mathbf{A}_e|, |\mathbf{B}_e|) - r_s \geqslant s_e$ are added to E_Δ. Because $\max(|\mathbf{A}_e|, |\mathbf{B}_e|) - r_s$ is finite, the computational effort is at most $O(N)$ and therefore does not affect the scaling $\propto M$ for sufficiently large M. (ii) The given Voronoi edge $e \in E_\Delta$ is a line connecting two Voronoi vertices at $\mathbf{A} \equiv \mathbf{A}_e$ and $\mathbf{B} \equiv \mathbf{B}_e$, that form a triangle with the corresponding material center at **x**. Hence $\mathbf{c}_e - \mathbf{x} = \mathbf{A} + t\mathbf{C}$ with $\mathbf{C} \equiv \mathbf{B} - \mathbf{A}$ and yet unknown $t \in [0, 1]$. The pore radius in the r_{eff}-coated system is given by the maximum $R(\mathbf{c}|r_{\text{eff}}) = |\mathbf{c} - \mathbf{x}| - r_s = c - r_s$ subject to the constraint $R(\mathbf{c}|r_{\text{eff}}) \geqslant |\mathbf{c} - \mathbf{p}| = |\mathbf{A} - \mathbf{P} + t\mathbf{C}|$, where $c \equiv |\mathbf{c}_e - \mathbf{x}|$ and $\mathbf{P} \equiv \mathbf{p} - \mathbf{x}$. Since $R(c|r_{\text{eff}}) \geqslant 0$, maximizing $R^2(c|r_{\text{eff}})$ presents an optimal solution for $R(c|r_{\text{eff}})$ as well. We observe that $\partial^2[r_s + R(c|r_{\text{eff}})]^2/\partial c^2 = 2 > 0$, which indicates that $R^2(c|r_{\text{eff}})$ and by implication $R(c|r_{\text{eff}})$ has a local minimum only and no local maximum. Hence, for any given interval $t \in [t_1, t_2]$, the maximum lies on either of the endpoints provided where the above inequality is satisfied. Consequently, we first test one of the two extremal values $t \in \{0, 1\}$ corresponding to the edge terminals [Fig. 5(a)]. If $|\mathbf{A} - \mathbf{P}| \geqslant |\mathbf{B} - \mathbf{P}|$, we set $t_0 = 1$, otherwise we set $t_0 = 0$. If $|\mathbf{A} + t_0\mathbf{C}| - r_s \geqslant |\mathbf{A} - \mathbf{P} + t_0\mathbf{C}|$, we are done. The candidate pore radius is $r_e = |\mathbf{A} + t_0\mathbf{C}| - r_s$ and the corresponding pore center is at $\mathbf{c}_e = \mathbf{x} + \mathbf{A} + t_0\mathbf{C}$ because the two additional candidates to be derived next cannot lead to a larger r_e anymore. The so far unused $1 - t_0$ we have to keep in mind. If the two terminals did not yet give rise to a final candidate pore radius r_e, we need to continue and calculate at most the two solutions $t_\pm$ to the effectively quadratic equation $c - r_s = |\mathbf{A} - \mathbf{P} + t\mathbf{C}|$

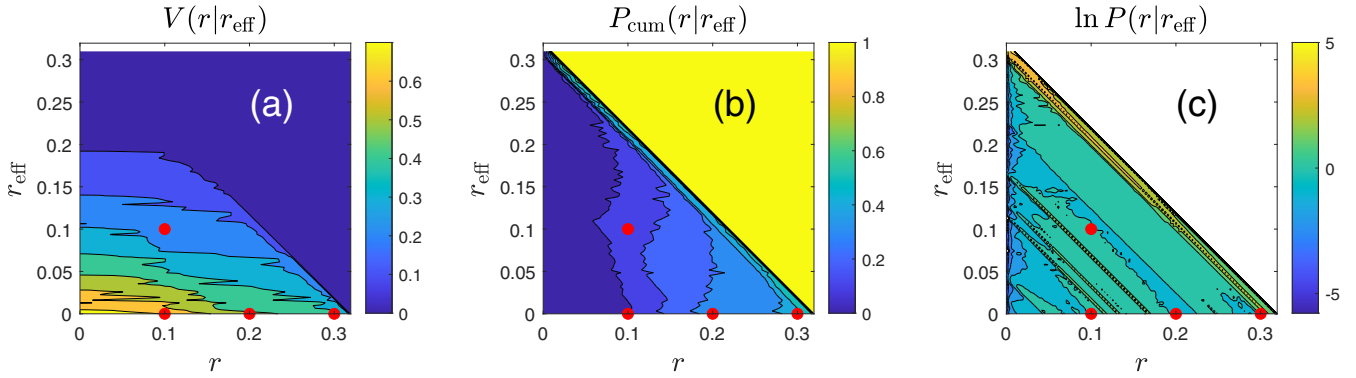

FIG. 6. Multidisk setup. Results from the presented algorithm for the configuration inspected in Figs. 1 and 3. The former figures provided visual information on how to obtain 4 selected points (marked by red bullets) in the present quantitative plots (a) $V(r|r_{\rm eff})$, (b) the cumulative G-PSD $P_{\rm cum}(r|r_{\rm eff})$, and (c) the G-PSD $P(r|r_{\rm eff})$ defined in Eq. (8). Beyond a certain $r + r_{\rm eff}$ (the radius of the LES), there is no contour line anymore in the left column of Fig. 1, and therefore vanishing $V(r|r_{\rm eff}) = 0$, $P(r|r_{\rm eff}) = 0$, and $P_{\rm cum}(r|r_{\rm eff}) = 1$. For the uncoated system ($r_c = 0$), the classical PSD $P(r)$ for point-like probe particles ($r_p = 0$) is contained in panel (c) on the x axis, i.e., at $r_{\rm eff} = r_c + r_p = 0$.

over the domain $t_\pm \in (0, 1)$ [Fig. 5(b)]. Some basic algebraic operations lead to $t_\pm = a \pm (a^2 + b)^{1/2}$, with

$$a = \frac{1}{2}\frac{(2\mathbf{P}\cdot\mathbf{A} + r_s^2 - P^2)\mathbf{P}\cdot\mathbf{C} - 2r_s^2\mathbf{A}\cdot\mathbf{C}}{(r_sC)^2 - (\mathbf{P}\cdot\mathbf{C})^2},$$
$$b = \frac{1}{4}\frac{(2\mathbf{P}\cdot\mathbf{A} + r_s^2 - P^2)^2 - (2r_sA)^2}{(r_sC)^2 - (\mathbf{P}\cdot\mathbf{C})^2}, \qquad (21)$$

where $A = |\mathbf{A}|, C = |\mathbf{C}|$, and $P = |\mathbf{P}|$. These expressions for a and b in Eq. (21) that determine the three remaining candidate t in the set $\{1 - t_0, t_-, t_+\}$ can be derived most conveniently by first subtracting $c^2 + r_s^2$ from both sides of the squared equation $c - r_s = |\mathbf{c} - \mathbf{p}|$, and then squaring it again. Out of these three t, the $t_* \in [0, 1]$ leading to the largest $c_* = |\mathbf{A} + t_*\mathbf{C}|$ solves the problem, i.e., the candidate pore radius is $r_e = c_* - r_s$ and the candidate pore center is located at $\mathbf{c}_e = \mathbf{x} + \mathbf{A} + t_*\mathbf{C}$. To summarize, this algorithm relies on the Voronoi network for a given configuration. It ensures that we find the LES ($r_p = 0$, $r_{\rm eff} = r_c$) or LPES ($r_c = 0$, $r_{\rm eff} = r_p$), i.e., the global optimum of the constrained optimization problem stated in Eq. (13). Each of the analytically calculated r values, whose ensemble with M members gives rise to a probability distribution G-PSD $\mathrm{P}(r|r_{\rm eff})$, has computational effort $O(1)$. The number of required calculations for the G-PSD is therefore strictly proportional to M, apart from the relatively cheap Voronoi and list E_j constructions that have to be done only once at $O(N)$ effort for all subsequent choices of system parameters $r_\circ$, r_c, and r_p. As opposed to the other mentioned strategies, the algorithm proposed here is unaffected by a mean pore radius.

The 2D code just described is part of our Supplemental Material [42]. The slightly more complicated analytic expressions for the 3D case we are planning to publish along with a 2D + 3D code elsewhere, as mentioned in the conclusion.

IV. RESULTS AND DISCUSSION

As we have exemplified the definition of the two-parametric volume $V(r|r_{\rm eff})$ in Figs. 1 and 3 by inspecting just four different $(r|r_{\rm eff})$ pairs, we start with completing this numerical example first, before we enter a more fundamental discussion that makes, apart from the development of the algorithm presented in Sec. III D, the core of our results. To this end, Fig. 6 displays $V(r|r_{\rm eff})$ as well as $P_{\rm cum}(r|r_{\rm eff})$ and $P(r|r_{\rm eff})$ versus r and $r_{\rm eff}$ for the configuration shown as filled black circles in Figs. 1 and 3.

The remainder of this section has two parts: In the first part we develop the analytic expression for the generalized G-PSD (which includes T-PSDs as special cases) of an extremely simple, but not too simple, configuration because it allows us to test G-PSD algorithms rigorously. The strategy turns out useful in falsifying the radical Voronoi approach. In the second part we present a new G-PSD algorithm, which improves over the local numerical optimization approach because it allows us to make sure that the global optimum is found exactly, and also efficiently, as we provide analytic expressions that solely require the coordinates of the N-Voronoi network.

A. Benchmark with analytically known G-PSD

We consider a simple geometry wherein the pore size distribution can be calculated analytically, and therefore serves as a validation case for the algorithms to be tested. The idea for this benchmark is derived from simple reasoning—if the phase of a given setup is flipped then the algorithm should be able to correctly compute the geometric pore size distribution of the "negative material." A figure of the setup is shown in Fig. 7. While the setup is essentially two dimensional, it is trivially extended to three dimensions. We therefore keep the symbol V for areas within the remainder of this paper.

To calculate the generalized G-PSD $P(r; r_p|r_c)$ which contains the classical G-PSD $P(r)$ for point-like particles in the naked benchmark setup as a special case, we need to derive an expression for $V(r|r_{\rm eff})$, the void area accessible to circles of radius r in the $r_{\rm eff}$-coated setup for the only nontrivial case of $r_s = r_\circ + r_{\rm eff} \geqslant 0$. Given the symmetry of the benchmark setup, it is sufficient to focus on calculating $V(r|r_{\rm eff})/8$, the shaded area in Fig. 8. It corresponds to the area accessible to a sphere of radius r in the $r_{\rm eff}$-coated configuration. One selected circle of radius r is shown, with its center coordinate at x on the y axis marking its outermost position in the void space. The cylinder is gray, the actual coating is blueish, the

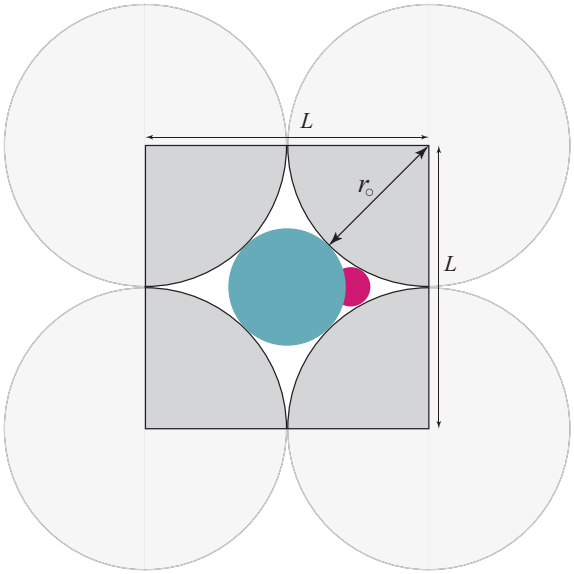

FIG. 7. Naked (uncoated) benchmark setup. The setup consists of identical circles of radius $r_\circ = 1$ or, equivalently, infinitely long and aligned cylinders, that just touch each other ($L = 2r_\circ$). The colored filled circles represent two possible largest circles that can be inscribed within the (white) void space enclosed by the large circles. The blue circle with radius $r_{\max}$ (24), represents the circle with largest possible radius that can be inscribed within this setup. This situation is realized in a periodic simulation cell (black square) of side length L, carrying a single circle of radius $r_\circ$.

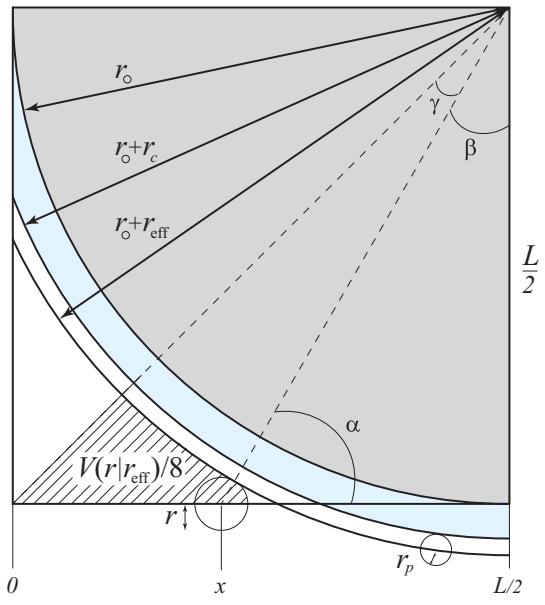

FIG. 8. Benchmark geometry. First quadrant of the r_c-coated benchmark setup (Fig. 7). Here we introduce radius r of the circle centered at $(x, 0)$, radius r_p of the probe particle, and some interrelated angles α, β, and γ. The shaded area $V(r|r_{\text{eff}})$ is the area that needs to be calculated as function of r, for given effective shell thickness, $r_{\text{eff}} = r_c + r_p$ to obtain the generalized G-PSD. The classical G-PSD is adsorbed by the special case $r_{\text{eff}} = 0$.

additional virtual coating of thickness r_p required to perform the calculation for probe particles of radius r_p is white. Some angles α, β, and γ have been added to help writing down the equations. One has $\beta + \gamma = \pi/4$ and $\alpha + \beta = \pi/2$, moreover $\tan\beta = (r_\circ - x)/r_\circ$ and $\sin\beta = (r_\circ - x)/(r_\circ + r_{\text{eff}} + r)$. This implies

$$x(r) = r_\circ - \sqrt{(r_s + r)^2 - r_\circ^2}, \tag{22}$$

$$\beta(r) = \tan^{-1}\left[1 - \frac{x(r)}{r_\circ}\right], \tag{23}$$

where we have re-introduced $r_s = r_\circ + r_{\text{eff}}$, Eq. (2). The largest possible circle that can be placed in the void space is centered at the origin with radius

$$r_{\max} = \sqrt{2}\, r_\circ - r_s. \tag{24}$$

The shaded area can be seen as a difference between two areas—a large triangle with base length x and height $r_\circ$ and the fraction $\gamma/2\pi$ of the circle of radius $r_\circ + r_{\text{eff}} = r_s$—plus a fraction $\alpha/2\pi$ of the r circle. Putting this together gives

$$V(r|r_{\text{eff}}) = 4\big[r_\circ x(r) - r_s^2\gamma(r) + r^2\alpha(r)\big]\Theta(r_{\max} - r), \tag{25}$$

with $\alpha(r) = \pi/2 - \beta(r)$, $\gamma(r) = \alpha(r) - \pi/4$, and Θ denotes the Heaviside step function. One has $V(r_{\max}|r_{\text{eff}}) = \pi r_{\max}^2$, as the largest pore of radius $r_{\max}$ is located in the center of the setup, where it occupies a large fraction of the void space. To express the G-PSD, we need the total area of the void space in the r_{eff}-coated setup which is already known from Eq. (25) as

$$V(0|r_{\text{eff}}) = 4\big[r_\circ x(0) - r_s^2\gamma(0)\big], \tag{26}$$

where r_{eff} is again hidden in r_s, $x(0)$ and $\gamma(0)$ via Eq. (2). By now, we have calculated all ingredients to write down the final results for the benchmark setup, with and without coating, for point-like and finitely sized probe particles. Point of departure is the final cumulative G-PSD for a point-like probe particle in the r_{eff}-coated material:

$$P_{\text{cum}}(r|r_{\text{eff}}) = 1 - \frac{r^2\alpha(r) + r_\circ x(r) - r_s^2\gamma(r)}{r_\circ x(0) - r_s^2\gamma(0)}\Theta(r_{\max} - r), \tag{27}$$

with $\alpha(r)$, $\gamma(r)$ given in terms of $x(r)$ and $\beta(r)$ specified by Eqs. (22) and (23), and $r_{\max}$ given by Eq. (24). The classical cumulative G-PSD is $P_{\text{cum}}(r) = P_{\text{cum}}(r|0)$, which amounts to replacing r_s by $r_\circ$ in the derived expression. This yields also the simple $V(0) = V(0|0) = (4 - \pi)r_\circ^2$.

It is important to recall once again the meaning of $r_{\text{eff}} = r_p + r_c$ and the identities $P(r|r_{\text{eff}}) = P(r + r_p; r_p|r_c) = dP_{\text{cum}}(r|r_{\text{eff}})/dr$ as stated in Eq. (8). The general result (27) contains three special cases: (i) the classical G-PSD for point-like probe particles in the naked setup, $P(r) = P(r|0)$ for $r \geqslant 0$, (ii) the generalized G-PSD for a finitely sized probe particle of radius $r_p > 0$ in the naked setup, $P(r; r_p|0) = P(r - r_p|r_p)$ for $r \geqslant r_p$, and (iii) the generalized G-PSD for a point probe in the r_c-coated setup, $P(r; 0|r_c) = P(r|r_c)$ for $r \geqslant 0$. Additionally the analytical expression for the mean pore radius $\langle r\rangle$ can be derived from Eq. (7). We are plotting $\langle r\rangle$ for various cases in Fig. 9.

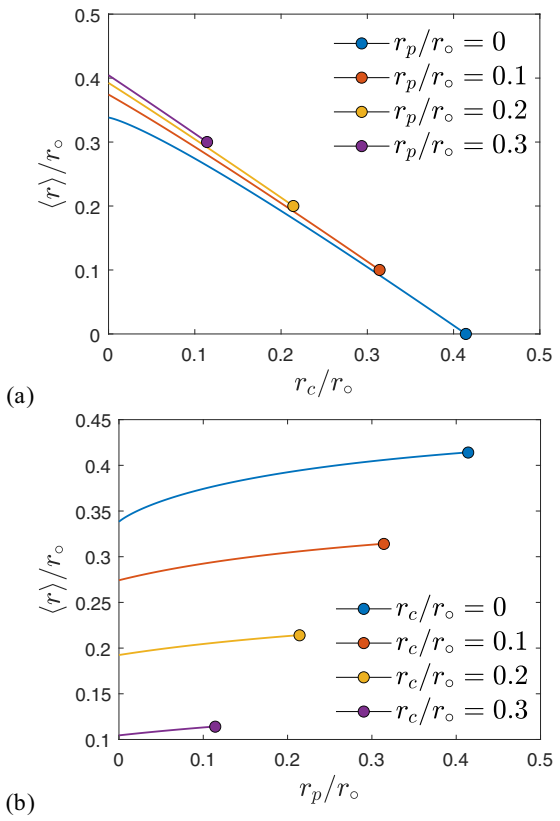

FIG. 9. Benchmark (analytic results). Mean pore radius $\langle r\rangle$ for the benchmark configuration, obtained analytically. (a) $\langle r\rangle$ versus r_c for various probe particle radii r_p. (b) $\langle r\rangle$ versus r_c for various coating thickness values r_c. The sum $r_{\text{eff}} = r_c + r_p$ is limited by $r_{\max}$ for $r_{\text{eff}} = 0$, which is $(\sqrt{2}-1)r_\circ \approx 0.41\, r_\circ$.

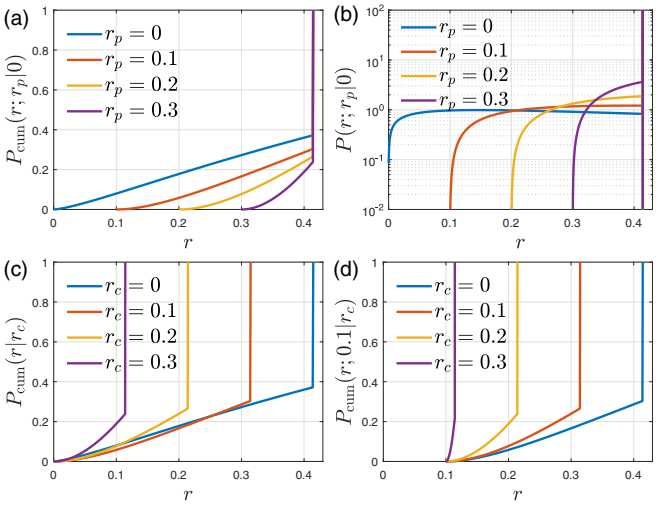

FIG. 10. Benchmark. Analytic distribution functions for the benchmark, Eq. (27), all exactly reproduced with the numerical scheme of Sec. III D. (a) $P_{\text{cum}}(r;r_p|0) = P_{\text{cum}}(r-r_p|r_p)$ for various probe radii r_p in the naked setup. (b) G-PSD $P(r;r_p|0)$ corresponding to the cumulative G-PSD shown in panel (a). (c) $P_{\text{cum}}(r;0|r_c) = P_{\text{cum}}(r|r_c)$ for point-like probe particle at various coating thickness values r_c. (d) $P_{\text{cum}}(r;r_p|r_c) = P_{\text{cum}}(r-r_p|r_p+r_c)$ for probe particles of radius $r_p = 0.1$ at various coating thickness values r_c.

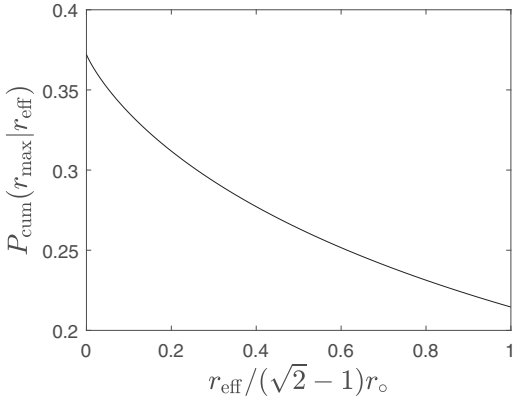

FIG. 11. Benchmark. The largest pore radius $r_{\max}$ is covering a significant portion $V(r_{\max}|r_{\text{eff}})$ of the total pore volume $V(0|r_{\text{eff}})$ for the r_{eff}-coated benchmark setup, leading to a jump in the cumulative G-PSD $P_{\text{cum}}(r|r_{\text{eff}})$ from the value $P_{\text{cum}}(r_{\max}|r_{\text{eff}})$ to unity for $r > r_{\max}$. The effect of r_{eff} on this value is displayed here because it is an important characteristics of the whole benchmark G-PSD. Because r_{eff} can only take values between 0 and $(\sqrt{2}-1)r_\circ$ for the benchmark (Fig. 8), we here plotted versus the scaled a dimensionless r_{eff} that resides on the interval [0,1].

The $P_{\text{cum}}(r;r_p|0)$ (Fig. 10) starts being nonzero at $r = r_p$, then increases monotonically until it jumps from its value $P_{\text{cum}}(r_{\max}|r_{\text{eff}})$ to unity at $r > r_{\max}$. There are two values that are particularly useful when comparing a numerical implementation of the G-PSD with the analytical solution. The magnitude of $P_{\text{cum}}(r_{\max}|r_{\text{eff}})$ (Fig. 11), and the mean pore radius (Fig. 9). For $r_{\text{eff}} = 0$, we obtain $P_{\text{cum}}(r_{\max}) = 2[2-(2-\sqrt{2})\pi]/(4-\pi) \approx 0.37208$ and $\langle r\rangle \approx 0.3383\, r_\circ$.

As there are no spurious local minima, the analytic results are fully consistent with the solution to the benchmark problem of aforementioned numerical algorithms (constrained local optimization, grid-based brute force), according to the chosen resolution of the numerical methods.

To summarize, in this section we have derived and presented analytic expressions for the interrelated $V(r|r_{\text{eff}})$, $P(r|r_{\text{eff}})$, $P_{\text{cum}}(r|r_{\text{eff}})$ for the benchmark with probe particles of arbitrary radius r_p and arbitrary coating thickness r_c, while $r_{\text{eff}} = r_c + r_p$, Eq. (1). The classical G-PSD $P(r)$ is absorbed by $r_{\text{eff}} = 0$.

B. Radical Voronoi applied to the benchmark

Consider a point $\mathbf{p}$ inside the void space of our naked benchmark (Fig. 8) with $r_\circ = 1$ and a probe particle radius $r_p \leqslant r_{\max}$. For all $\mathbf{p}$ residing inside the largest inscribed circle of radius $r_{\max}$, the radical N-Voronoi results in $\mathbf{c} = \mathbf{0}$ and $r(\mathbf{p};r_p) = r_{\max}$, which is correct. Now consider a point $\mathbf{p} = (p_x, p_y)$ with $p_x > r_{\max}$ located outside the central region, but for simplicity on the x axis, $p_y = 0$. Since $x(r) = 1 - \sqrt{(2+r)r}$ is the position of the center of the r circle that touches the naked material according to Eq. (22) with $r_s = r_\circ = 1$, and because this r circle must just touch the probe circle to make sure the probe is contained within the r circle, and the r circle maximally large, we know that the correct LPES radius $r(\mathbf{p};r_p)$ is given by the solution to

$x(r) + r = p_x + r_p$, hence

$$r(\mathbf{p}; r_p) = \frac{(p_x + r_p - 1)^2}{2(p_x + r_p)} \quad (p_x > r_{\max}, p_y = 0), \tag{28}$$

while $\mathbf{c} = (c_x, c_y)$ with $c_x = p_x + r_p - r(\mathbf{p}; r_p)$ and $c_y = 0$ is the center of the LPES sphere. If the radical Voronoi were applicable, this point $\mathbf{c}$ must be a vertex of either the N-Voronoi or the $(N+1)$-Voronoi networks. The N-Voronoi network has a single vertex at the origin in Fig. 8 and is therefore ruled out. The $(N+1)$-Voronoi network has four vertices, only one of them located on the positive x axis. It is sufficient to show that the above $\mathbf{c}$ does not fulfill Eq. (17) to disqualify the radical Voronoi method for calculating a proper G-PSD.

Inserting $\mathbf{c} = (c_x, 0)$, $\mathbf{p} = (p_x, 0)$ and $\mathbf{x}^i = (1, \pm 1)$ into Eq. (17) yields $(c_x - p_x)^2 - r_p^2 = (c_x - 1)^2$, further replacing c_x by $p_x + r_p - r(\mathbf{p}; r_p)$ shows that the equation defining the vertices of the $(N+1)$-Voronoi produces a vertex position at the proper $\mathbf{c}$ only for the single case of $p_x = 1 - r_p$. All other points with $p_x > r_{\max}$ on the x axis are not capturing the LPES radius correctly. Even worse, this single point $p_x = 1 - r_p$ is the irrelevant point for which the pore radius vanishes, and therefore does not exceed a finite r_p. The radical Voronoi does fail completely not only if $\mathbf{p}$ is located on the x axis, but whenever the probe particle covers a point outside the $r_{\max}$ circle. In general, the vertices of the $(N+1)$-Voronoi are located inside the material, except for the special case of $r_p = r_\circ$, that was most likely for this reason chosen in the original work [34].

C. Difference between G-PSD and T-PSD

Torquato's cumulative center-based T_c-PSD $\mathcal{P}^c_{\rm cum}(r) = 1 - E_V(r)$ is the probability of finding a region, which is a spherical cavity of radius r (centered at some arbitrary point), containing one or more particle centers [30]. Using our notation, we can write this as

$$\mathcal{P}^c_{\rm cum}(r) = 1 - \frac{V(0|r - r_\circ)}{V(0|-r_\circ)} \quad (r \geqslant 0), \tag{29}$$

noting that $V(0|r - r_\circ) = L^2 - \pi r^2$ if $r \leqslant r_\circ$ for our benchmark, i.e., if the particle is negatively coated. The situation was shown in Fig. 4(e). The often-employed surface-based version of the above center-based T_c-PSD is just the $r_\circ$-shifted variant of $\mathcal{P}^c_{\rm cum}(r)$,

$$\mathcal{P}^s_{\rm cum}(r) = 1 - \frac{V(0|r)}{V(0|0)} = \mathcal{P}^c_{\rm cum}(r + r_\circ) \quad (r \geqslant 0). \tag{30}$$

The T_c-PSD based on area T_c in Fig. 4(e), and the T_s-PSD based on area T_s in Fig. 4(f) are both qualitatively different from the G-PSD not only in general, but also for the benchmark, and the analytic expression is given by Eq. (11) or Eq. (30) with Eq. (26). This yields for our benchmark, recalling $r_s = r_\circ + r_{\rm eff}$, and using $\xi = r/r_\circ$,

$$\mathcal{P}^c_{\rm cum}(r) = \sqrt{\xi^2 - 1} - \frac{\pi \xi^2}{4} + \xi^2 \csc^{-1}(\xi) \quad (\xi \leqslant \sqrt{2}). \tag{31}$$

Note that Eq. (31) simplifies to $\mathcal{P}^c_{\rm cum}(r) = \pi \xi^2/4$ for $\xi \leqslant 1$; this expression is confirmed visually by Fig. 4(e). The largest pore radius is $r = \sqrt{2}\, r_\circ$ using $\mathcal{P}^c(r)$, exceeding the true largest radius $r_{\max}$ by 341% while the largest radius according to $\mathcal{P}^s(r)$ is the same as for the G-PSD, however contrary to the G-PSD, the volume accessible to the probe $V(0|r_{\max}) = 0$.

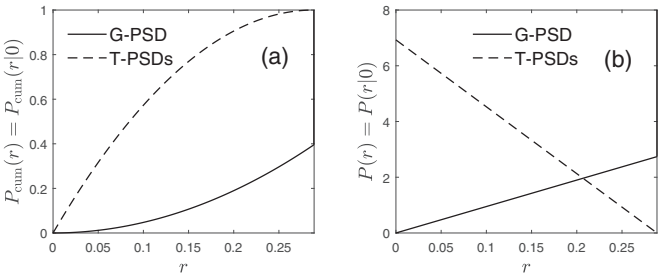

FIG. 12. Triangle. For the equilateral triangle with unit side length and a straight edges ($r_\circ = 0$), the volume $V(r|r_{\rm eff})$ can be calculated analytically, cf. Eq. (A4) and leads to the G-PSD quantities (a) $P_{\rm cum}(r) = [4(9 - \sqrt{3}\pi) r_1^2]/3$, (b) $P(r) = [8(9 - \sqrt{3}\pi) r_1]/3$, while both T-PSDs are identical in that case, (a) $\mathcal{P}^c_{\rm cum}(r) = \mathcal{P}^s_{\rm cum}(r) = 4(\sqrt{3} - 3r_1)r_1$, and (b) $\mathcal{P}(r) = 4(\sqrt{3} - 6r_1)$. This figure demonstrates the huge qualitative difference between G-PSD and T-PSDs. The same is true for our benchmark, cf. Eq. (31) versus Eq. (27) for $r_s = r_\circ$.

For systems made of point-like particles, $r_\circ = 0$, such as the equilateral triangle treated in Fig. 4(a), the two T-PSDs are identical by definition. The enormous differences between the G- and T-PSDs we quantify in Fig. 12 for this particularly simple example, for which analytic expressions are also available (Appendix 3).

D. Performance of the various G-PSD algorithms

Finally, we demonstrate how the newly proposed Voronoi-based algorithm performs in practice, and how it compares with the classical, grid-based approach, for both, the benchmark with $N = 1$, and a larger system with $N = 1000$. The larger system is chosen in addition, as one might be still in doubt about the relevance of the benchmark for the general case. We begin by inspecting the benchmark (Fig. 13). This figure quantitatively compares both the speed and quality of the two algorithms at either identical values for M, or identical computational cost. The Voronoi-based method outperforms the grid-based algorithm in all aspects. The deviation analyzed in panels (a) and (c) is measured against the analytical solution. Figure 13(b) highlights the CPU scaling behavior with M. While the Voronoi-based algorithm has a computational cost $\propto M$, the grid-based approach scales as $M^{5/3}$ for this example, and is more generally superlinear in M. To confirm that all observations for the benchmark overtake to real systems made of a large number of particles, we inspect such a configuration in Fig. 14. Indeed, the grid-based result converges only slowly against the exact result, which is (at least by eye) obtained using the Voronoi-based version at relatively small $M = 10N$ already. Still, a much larger M of the order of 10^7 is required to generate a smooth $P(r|0)$ for this system (result not shown). While such a value for M is easily within reach for the Voronoi-based method (several seconds), it poses both a memory problem for the grid-based version, as well as significantly enlarged computational cost (several hours).

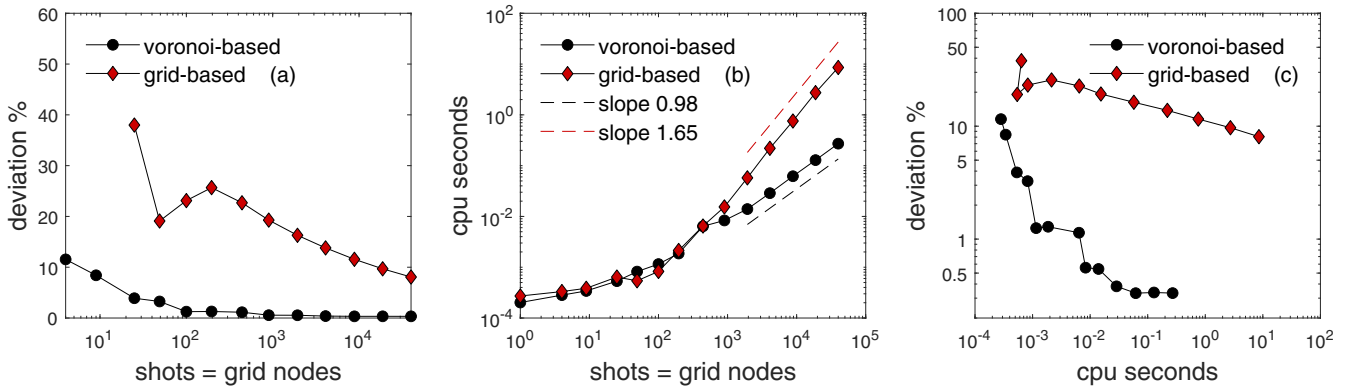

FIG. 13. Benchmark. Comparison (quality and speed) between methods to calculate the cumulative G-PSD $P_{\mathrm{cum}}(r|r_{\mathrm{eff}})$ for the benchmark setup with $r_{\mathrm{eff}} = 0$: (i) the Voronoi-based global analytical optimization developed in this work (Sec. III D) with M shots, (ii) regular grid-based brute force scan (Sec. III A) with M nodes. The deviation is a standard deviation calculated against the analytical solution (27). The computational effort is strictly linear in M for the Voronoi-based version, and generally independent of N, while the grid-based approach is $\propto M^{5/3}$ for this case, and also independent of N. The local constrained optimization approach is almost linear in M, but in addition linear in N except for systems with very small pores, where a neighbor list can remove the dependency on N (Sec. III B).

These aspects rule out the grid-based approach completely for systems with millions of particles.

V. CONCLUSIONS

We have presented a proper definition of the generalized G-PSD, following the notions introduced by Gelb and Gubbins [31]. Additionally we have described the relation of the G-PSD to other definitions in particular the different definitions based on the work of Torquato *et al.* [30], termed T-PSDs. Calculating the generalized G-PSD and classical G-PSD is a long-standing problem, as all known methods come with severe limitations. Therefore, to shed light on this problem, we set up a simple benchmark that is complex enough to test implementations, and calculated the generalized G-PSD and the T-PSDs analytically. With this at hand, we concluded that the method based on radical Voronoi tessellation is unfit for calculating the G-PSD for complex geometries. The grid-based approach and constrained optimization provide answers but are computationally expensive. The latter fails to locate the maxima at times especially in the case of complex geometries, while the former becomes highly time- and memory-consuming for high resolution.

We proposed a novel approach within the spirit of constrained optimization, which becomes a simpler problem to solve by considering an additional constraint. This reduces the search space of solutions for the optimization problem considerably. We were also able to calculate the solution to the optimization problem analytically for a given triangle located on a Voronoi face, which could prove crucial to a significant

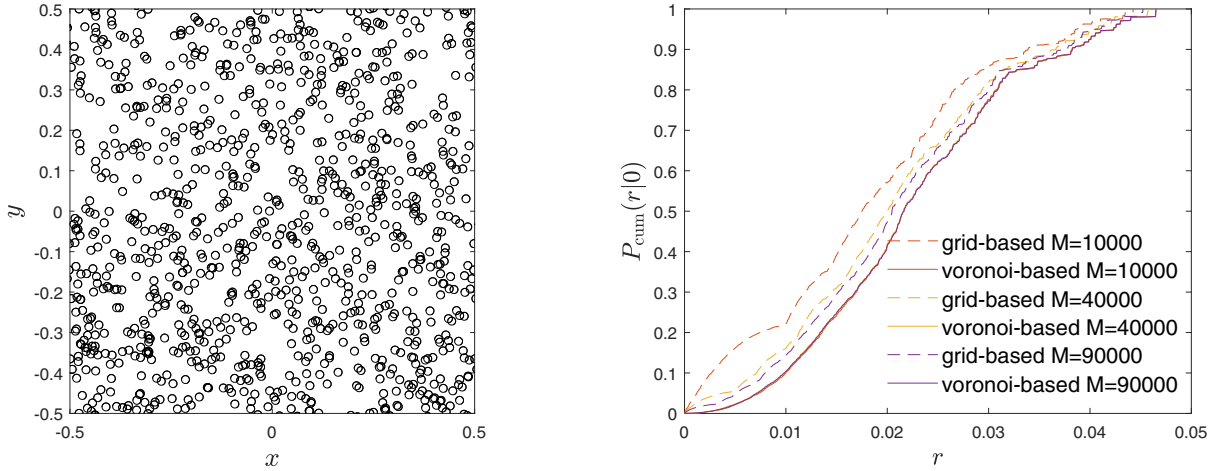

FIG. 14. Multidisk. Comparison of convergence behavior for Voronoi-based and grid-based methods during the calculation of the cumulative G-PSD $P_{\mathrm{cum}}(r|r_{\mathrm{eff}})$ for the multidisk $N = 1000$ setup shown on the left. The right panel displays $P_{\mathrm{cum}}(r|0)$ versus r for various choices of M, for both methods. While the grid-based result slowly converges against the Voronoi-based result, the latter remains basically unaffected by M for sufficiently large $M \gg N$. The underlying reason is the absence of a grid for the Voronoi-based method. At the same time it runs more efficiently and does not suffer from memory limitations. The present result just confirms the trends already available from our benchmark (Fig. 13).

increase in the efficiency and accuracy of determining generalized G-PSD for any given spherical monodisperse system. Instead of scanning the D-dimensional space in a random or deterministic fashion, which comes with the inherent danger of missing the global optimum, we can walk over a limited number of relevant triangles, those in the neighborhood of the test particle position. The size of the neighborhood is known a priori from the radius of the largest pore, and the global optimum is therefore found exactly (analytically) at relatively low computational cost.

While the 2D code is part of our Supplemental Material [42], we are planning to make the 3D code available as a LAMMPS [43] option, and as a stand-alone C++ application that makes use of the VORO++ library [44].

ACKNOWLEDGMENTS

This work was supported by Project IZCOZ0_189872 of the Swiss National Science Foundation (SNSF). M.K. would like to thank Yéléna Luap for creating Fig. 4.

APPENDIX: ANALYTICS RESULTS FOR THE GENERALIZED PSD

Analytical results for the benchmark configuration have been provided in Sec. IV A. Here we add additional results for simple configurations that serve to test the accuracy of the algorithms. As we have shown, the G-PSD, and also T-PSDs, are trivially calculated from an analytic expression for $V(r|r_{\text{eff}})$.

1. Circular or cylindrical void

The generalized G-PSD $P_{\bigcirc}(r|r_{\text{eff}})$ for a circular void with radius $r_{\bigcirc}$, or equivalently, an infinitely long cylindrical three-dimensional void with cylinder radius $r_{\bigcirc}$, is

$$P_{\bigcirc}(r|r_{\text{eff}}) = \delta(r - r_{\bigcirc} + r_{\text{eff}}) \quad (r \geqslant 0), \tag{A1}$$

because the full circle is available to circles of any radius r as long as $r \leqslant r_{\bigcirc} - r_{\text{eff}}$. The same result is obtained if we start from $V_{\bigcirc}(r|r_{\text{eff}})$, the area accessible to circles of radius r in the r_{eff}-coated substance. This is the area of a circle with radius $r_{\bigcirc} - r_{\text{eff}}$ as long as $r \leqslant r_{\bigcirc} - r_{\text{eff}}$, more formally,

$$V_{\bigcirc}(r|r_{\text{eff}}) = \pi(r_{\bigcirc} - r_{\text{eff}})^2 \Theta(r_{\bigcirc} - r_{\text{eff}} - r). \tag{A2}$$

Inserting this $V_{\bigcirc}(r|r_{\text{eff}})$ into Eq. (8) reproduces Eq. (A1). To complete this, Eqs. (4), (8) then imply $P_{\bigcirc}(r; r_p|r_c) = P_{\bigcirc}(r - r_p|r_c + r_p) = \delta(r - r_p - r_{\bigcirc} + r_{\text{eff}}) = \delta(r - r_{\bigcirc} + r_c)$ for $r \geqslant r_p$, just highlighting the fact that the only effect of probe radius r_p on the G-PSD for this simple example is its absence for $r \leqslant r_p$.

2. Square-shaped and cubic void

The pore radius distribution for a square-shaped two-dimensional void with linear size L, or equivalently, an infinitely long rectangular three-dimensional void with square cross-sectional area L^2, is most conveniently obtained from $V_{\square}(r|r_{\text{eff}})$, the area accessible to circles of radius r in the r_{eff}-coated square, given by

$$V_{\square}(r|r_{\text{eff}}) = [(L - 2r_{\text{eff}})^2 - (2r)^2 + \pi r^2]\Theta(L - 2r_{\text{eff}} - 2r). \tag{A3}$$

Here $(L - 2r_{\text{eff}})^2$ is the void area of the coated square, and $r^2 - \pi r^2/4$ is the area not reachable by an r circle in any of the four corners of the square. The area vanishes as soon as r exceeds $L/2 - r_{\text{eff}}$, giving rise to the Heaviside contribution. Inserting this $V_{\square}(r|r_p)$ into Eqs. (4) and (8) yields the G-PSD $P_{\square}(r; r_p|r_c)$ and all related quantities analytically.

3. Triangular void

For an equilateral triangle with unit side lengths, one finds

$$V_{\triangle}(r|r_{\text{eff}}) = V_{\triangle}(0|0) - (3\sqrt{3} - \pi)r^2 - 3r_{\text{eff}}(1 - \sqrt{3}r_{\text{eff}}), \tag{A4}$$

with $V(0|0) = \sqrt{3}/4$. This implies

$$P^{\text{cum}}_{\triangle}(r|r_{\text{eff}}) = \frac{4(9 - \sqrt{3}\pi)r_1^2}{3(1 - 2\sqrt{3}r_{\text{eff}})^2}, \tag{A5}$$

and

$$P_{\triangle}(r|r_{\text{eff}}) = \frac{dP^{\text{cum}}_{\triangle}(r|r_{\text{eff}})}{dr} = \frac{8(9 - \sqrt{3}\pi)r}{3(1 - 2\sqrt{3}r_{\text{eff}})^2}. \tag{A6}$$

The special case of $r_{\text{eff}} = 0$ is shown in Fig. 12.

[1] S. Sastry, D. S. Corti, P. G. Debenedetti, and F. H. Stillinger, Statistical geometry of particle packings. I. Algorithm for exact determination of connectivity, volume, and surface areas of void space in monodisperse and polydisperse sphere packings, Phys. Rev. E **56**, 5524 (1997).

[2] S. Sastry, P. G. Debenedetti, and F. H. Stillinger, Statistical geometry of particle packings. II. "Weak spots" in liquids, Phys. Rev. E **56**, 5533 (1997).

[3] N. V. Priezjev and M. A. Makeev, Evolution of the pore size distribution in sheared binary glasses, Phys. Rev. E **96**, 053004 (2017).

[4] A. Rege, S. Aney, and B. Milow, Influence of pore-size distributions and pore-wall mechanics on the mechanical behavior of cellular solids like aerogels, Phys. Rev. E **103**, 043001 (2021).

[5] M. A. Moret and G. F. Zebende, Amino acid hydrophobicity and accessible surface area, Phys. Rev. E **75**, 011920 (2007).

[6] J. Kim and B. J. Sung, Dynamics and spatial correlation of voids in dense two dimensional colloids, J. Chem. Phys. **141**, 014502 (2014).

[7] R. Kumar and B. Bhattacharjee, Porosity, pore size distribution and in situ strength of concrete, Cem. Concr. Res. **33**, 155 (2003).

[8] T. J. Richmond, Solvent accessible surface area and excluded volume in proteins: Analytical equations for overlapping

spheres and implications for the hydrophobic effect, J. Mol. Biol. **178**, 63 (1984).
[9] S. W. Cranford and M. J. Buehler, Packing efficiency and accessible surface area of crumpled graphene, Phys. Rev. B **84**, 205451 (2011).
[10] E. W. Washburn, The dynamics of capillary flow, Phys. Rev. **17**, 273 (1921).
[11] E. W. Washburn, Note on a method of determining the distribution of pore sizes in a porous material, Proc. Natl. Acad. Sci. U.S.A. **7**, 115 (1921).
[12] L. Drake, Pore-size distribution in porous materials, Ind. Eng. Chem. **41**, 780 (1949).
[13] C. M. Simon, J. Kim, D. A. Gomez-Gualdron, J. S. Camp, Y. G. Chung, R. L. Martin, R. Mercado, M. W. Deem, D. Gunter, M. Haranczyk *et al.*, The materials genome in action: Identifying the performance limits for methane storage, Energy Environ. Sci. **8**, 1190 (2015).
[14] S.-Y. Lee and S.-J. Park, Determination of the optimal pore size for improved CO_2 adsorption in activated carbon fibers, J. Colloid Interface Sci. **389**, 230 (2013).
[15] K. Kwac, J. H. Lee, J. W. Choi, and Y. Jung, Computational analysis of pressure-dependent optimal pore size for CO_2 capture with graphitic surfaces, J. Phys. Chem. C **120**, 3978 (2016).
[16] P. Nugent, Y. Belmabkhout, S. D. Burd, A. J. Cairns, R. Luebke, K. Forrest, T. Pham, S. Ma, B. Space, L. Wojtas *et al.*, Porous materials with optimal adsorption thermodynamics and kinetics for CO_2 separation, Nature (London) **495**, 80 (2013).
[17] E. P. Barrett, L. G. Joyner, and P. P. Halenda, The determination of pore volume and area distributions in porous substances. I. Computations from nitrogen isotherms, J. Am. Chem. Soc. **73**, 373 (1951).
[18] J. Villarroel-Rocha, D. Barrera, and K. Sapag, Introducing a self-consistent test and the corresponding modification in the Barrett, Joyner and Halenda method for pore-size determination, Microporous Mesoporous Mater. **200**, 68 (2014).
[19] D. Dollimore and G. Heal, An improved method for the calculation of pore size distribution from adsorption data, J. Appl. Chem. (London, U.K.) **14**, 109 (1964).
[20] P. I. Ravikovitch, G. L. Haller, and A. V. Neimark, Density functional theory model for calculating pore size distributions: Pore structure of nanoporous catalysts, Adv. Colloid Interface Sci. **76-77**, 203 (1998).
[21] P. I. Ravikovitch and A. V. Neimark, Density functional theory model of adsorption on amorphous and microporous silica materials, Langmuir **22**, 11171 (2006).
[22] G. Kupgan, T. P. Liyana-Arachchi, and C. M. Colina, NLDFT pore size distribution in amorphous microporous materials, Langmuir **33**, 11138 (2017).
[23] S. Hemes, G. Desbois, J. L. Urai, B. Schröppel, and J.-O. Schwarz, Multi-scale characterization of porosity in Boom Clay (HADES-level, Mol, Belgium) using a combination of X-ray μ-CT, 2D BIB-SEM and FIB-SEM tomography, Microporous Mesoporous Mater. **208**, 1 (2015).
[24] D. Wildenschild and A. P. Sheppard, X-ray imaging and analysis techniques for quantifying pore-scale structure and processes in subsurface porous medium systems, Adv. Water Resour. **51**, 217 (2013).
[25] A. Kaestner, E. Lehmann, and M. Stampanoni, Imaging and image processing in porous media research, Adv. Water Resour. **31**, 1174 (2008).
[26] L. D. Gelb and K. Gubbins, Characterization of porous glasses: Simulation models, adsorption isotherms, and the Brunauer-Emmett-Teller analysis method, Langmuir **14**, 2097 (1998).
[27] V. Sorichetti, V. Hugouvieux, and W. Kob, Structure and dynamics of a polymer-nanoparticle composite: Effect of nanoparticle size and volume fraction, Macromolecules (Washington, DC, U.S.) **51**, 5375 (2018).
[28] V. Sorichetti, V. Hugouvieux, and W. Kob, Determining the mesh size of polymer solutions via the pore size distribution, Macromolecules (Washington, DC, U.S.) **53**, 2568 (2020).
[29] K. Trepte and S. Schwalbe, porE: A code for deterministic and systematic analyses of porosities, J. Comput. Chem. **42**, 630 (2021).
[30] S. Torquato, B. Lu, and J. Rubinstein, Nearest-neighbor distribution functions in many-body systems, Phys. Rev. A **41**, 2059 (1990).
[31] L. D. Gelb and K. Gubbins, Pore size distributions in porous glasses: A computer simulation study, Langmuir **15**, 305 (1999).
[32] F.-X. Coudert and A. H. Fuchs, Computational characterization and prediction of metal–organic framework properties, Coord. Chem. Rev. **307**, 211 (2016).
[33] J. Rouquerol, P. Llewellyn, F. Rouquerol *et al.*, Is the bet equation applicable to microporous adsorbents, Stud. Surf. Sci. Catal. **160**, 49 (2007).
[34] M. Pinheiro, R. L. Martin, C. H. Rycroft, A. Jones, E. Iglesia, and M. Haranczyk, Characterization and comparison of pore landscapes in crystalline porous materials, J. Mol. Graphics Modell. **44**, 208 (2013).
[35] D. Vollath and D. V. Szabó, Coated nanoparticles: A new way to improved nanocomposites, J. Nanopart. Res. **1**, 235 (1999).
[36] Y. Li, M. Kröger, and W. K. Liu, Endocytosis of pegylated nanoparticles accompanied by structural changes of the grafted polyethylene glycol, Biomaterials **35**, 8467 (2014).
[37] A. Halperin, M. Kröger, and F. M. Winnik, Poly(n-isopropylacrylamide) phase diagrams: Fifty years of research, Angew. Chem., Int. Ed. **54**, 15342 (2015).
[38] B. Münch and L. Holzer, Contradicting geometrical concepts in pore size analysis attained with electron microscopy and mercury intrusion, J. Am. Ceram. Soc. **91**, 4059 (2008).
[39] S. Song, Q. Ding, and J. Wei, Improved algorithm for estimating pore size distribution from pore space images of porous media, Phys. Rev. E **100**, 053314 (2019).
[40] S. Bhattacharya and K. E. Gubbins, Fast method for computing pore size distributions of model materials, Langmuir **22**, 7726 (2006).
[41] T. Ng, J. Yeo, and Z. Liu, A molecular dynamics study of the thermal conductivity of nanoporous silica aerogel, obtained through negative pressure rupturing, J. Non-Cryst. Solids **358**, 1350 (2012).
[42] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.107.015307 for the codes for the calculation of the G-PSD and T_s-PSD using the Voronoi-based and

grid-based methods. We offer routines that allow us to redo the figures shown in this paper, as well as codes that can be used to investigate user-defined 2D structures.

[43] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton, LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, Comput. Phys. Commun. **271**, 108171 (2022).

[44] C. H. Rycroft, Voro++: A three-dimensional Voronoi cell library in C++, Chaos **19**, 041111 (2009).