Numerical studies of second- and fourth-order correlation functions in cluster-cluster aggregates in application to optical scattering

Vadim A. Markel,* Vladimir M. Shalaev,† and Evgeni Y. Poliakov
Department of Physics, New Mexico State University, Las Cruces, New Mexico 88003

Thomas F. George
Office of the Chancellor/Department of Chemistry and Department of Physics and Astronomy, University of Wisconsin-Stevens Point, Stevens Point, Wisconsin 54481-3897

(Received 13 November 1996)

Two- and four-point density correlation functions $p_2(r)$ and $p_4(r)$ are studied numerically and theoretically in computer-generated three-dimensional lattice cluster-cluster aggregates (CCA) with the number of particles N up to 20 000 in application to the light scattering problem. The ''pure'' aggregation algorithm is used, where subclusters of all possible sizes are allowed to collide. We find that large CCA clusters demonstrate pronounced multiscaling. In particular, the fractal dimension determined from the slope of $p_2(r)$ at small distances differs from that found from the dependence of the radius of gyration on the number of monomers (according to our data, 1.80 and 1.94, respectively). We also consider different functional forms for p_2 and their general properties and applicability. We find that the best fit to the numerical data is provided by the generalized exponential cutoff function with coefficients depending on N. The latter dependence is a manifestation of multiscaling. We propose some theoretical approaches for calculating $p_4(r)$, assuming $p_2(r)$ is known. In particular, we find the small-r asymptote for the $p_4(r)$ and verify it numerically. In addition, we find that $p_4(r)$ cannot be represented by a scaling dependence with a cutoff function, like $p_2(r)$. Instead, $p_4(r)$ is given by an expansion in terms of integer powers of r^{2D-3}, where D is the fractal dimension (≈ 1.8 for CCA clusters). [S1063-651X(97)14906-6]

PACS number(s): 61.43.Hv, 71.45.Gm, 42.25.Fx

I. INTRODUCTION

Computer algorithms play a very important role in understanding aggregation phenomena and physical properties of aggregates. In many instances, computer simulation is the only feasible theoretical approach to very complex stochastic aggregation processes, when traditional techniques of statistical mechanics cannot be applied due to strong fluctuations and the absence of equilibrium. These algorithms have attracted much attention since Witten and Sander [1], Meakin [2], and Kolb, Botet, and Jullien [3] proposed realistic algorithms which simulate natural aggregation processes to much detail.

One of the most extensively used applications of computer-generated clusters is the study of density correlation functions. A direct experimental measurement of these functions has also been carried out (see, for example, Ref. [4]), but such experiments face considerable difficulties. First, it is hard to calculate interparticle distances from two-dimensional electron micrograph images of three-dimensional clusters [4], and, second, the process of electron micrographing itself can damage the cluster structure.

Density correlation functions provide important geometrical characteristics of clusters, carrying valuable physical information. Correlation functions are especially useful in optics [4–8]. The two-point correlation function describes the average intensity of light scattered by an ensemble of clusters [5], while the four-point correlation function governs fluctuations in the scattered light [8].

In the present paper we study the two- and four-point correlation functions in aggregates obtained by the numerical algorithm known as the cluster-cluster aggregation (CCA) [2,3,9–11]. The CCA algorithm provides a very accurate simulation of aggregation processes which occur in nature, under the conditions that there is no spatially fixed center of aggregation and the concentration of aggregating material is sufficiently low. These conditions are well satisfied, for example, for formation of fractal carbone soot [5] and metallic colloids [12]. One of the possible applications of the results obtained in this paper is a theoretical description of light scattering from fractal soot clusters in the atmosphere.

In its most pure form, the CCA algorithm involves the following steps: First, a set of N pointlike particles are randomly placed on a simple cubic lattice of the size $L \times L \times L$. The size of the lattice is chosen so that the average density of particles, N/L^3, is much smaller than unity. Those particles which are separated by only one lattice unit are considered to be rigidly bound to each other and form a subcluster. Ideally, in the limit of zero density, there are initially N subclusters consisting of one particle each. Then a subcluster is picked randomly out of the set and moved one lattice unit in one of the six possible directions, chosen at random. If, after this move, the subcluster contacts other cluster(s) (via nearest-neighbor occupancy), these subclusters stick together to form a larger subcluster. The steps are re-

*Also with Institute of Automation and Electrometry, Russian Academy of Sciences, 630090 Novosikirsk, Russia.

†Also with the L. V. Kirenski Institute of Physics, Siberian Branch of the Russian Academy of Science, 660036 Krasnoyarsk, Russia.

peated until a single cluster of N particles is left. Periodic boundary conditions are usually assumed, which results in ''transparency'' of the boundaries of the lattice for moving subclusters and, on the other hand, in conservation of the number of particles.

A great number of modifications of the algorithm exists, among which are random rotations of subclusters, prescribing ''mobility'' to each subcluster (which is, typically, proportional to the inverse number of particles in this subcluster), hierarchical models, off-lattice models, and others [13]. Some of these modifications, such as the hierarchical model, are intended to simplify the numerical procedure, and to make it possible to generate large clusters in a feasible time; the price of these simplifications is sacrificing some essential features of real aggregation. Other modifications, such as random rotations of subclusters, are, on the contrary, intended to make the algorithm more realistic, while complicating the numerical procedure.

In the present paper we use only the pure form of the algorithm described above, and generate three-dimensional clusters with the number of particles N up to 20 000. For comparison, the correlation functions in Ref. [14] were studied using computer-generated CCA clusters with the maximum value of $N=900$. Clusters with a number of particles comparable to 20 000 have been generated earlier [10,11], but with the use of the hierarchical model [9], which allows only subclusters of the same size to collide, and considers aggregation of only two subclusters at a time. Though this model was shown to produce clusters very similar to those obtained by the direct calculation, it evidently lacks some important features of the real aggregation process. In particular, this model may be incapable of producing the multiscaling effect (which is established for the Witten-Sander clusters [15]).

Using the computer-generated CCA clusters with large N, we argue that the pure CCA model produces clusters possessing multiscaling. A manifestation of this effect is that the fractal dimension measured from the dependence of radius of gyration on the number of particles differs from that measured from the two-point correlation function at small distances. In this paper, we also consider different forms of cutoff functions for the two-point correlation function.

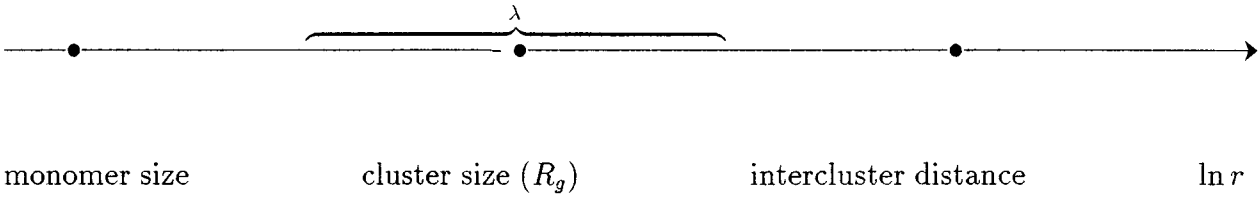
Concerning the four-point correlation function, we find theoretically its small-r asymptote, and confirm it numerically. However, it turns out that the four-point correlation function does not have a simple scaling form with a cutoff function. Instead, it can be expanded in a power series in r^{2D-3}, where r is the distance and D the fractal dimension (for CCA, $1.5<D<2$).

The paper is organized as follows: In Sec. II we review the use and definitions of correlation functions in application to the static light scattering problem. In Sec. III we discuss different models for the two-point correlation function and their properties and applicability. In Sec. IV we report some theoretical approaches for the calculation of the four-point correlation function, assuming that the two-point correlation function is known. In Sec. V we describe our numerical procedures for computer generation of CCA clusters and the computation of correlation functions. Section VI contains our numerical results for the two-point correlation function, and Sec. VII those for the four-point correlation function. Finally, Sec. VIII contains a summary and discussion.

II. CORRELATION FUNCTIONS IN OPTICS

The two- and four-point correlation functions are essential for the description of static light scattering from fractal clusters. The two-point correlation function describes the average scattered light intensity, while the four-point correlation function describes fluctuations of the scattered light. In this section we review briefly the use and definitions of these functions as applied to the scattering problem.

Here we adopt a basic assumption that the distance between different clusters is much larger than the wavelength λ, and the positions of clusters are random and uncorrelated, so that we add up intensities of light scattered by different clusters rather than amplitudes. (The effects of intercluster interference were considered in Refs. [6,7,16–18].) At the same time, we assume that λ is large compared to the size of a single monomer, but can be much smaller than the characteristic cluster size. We will use the ensemble-average radius of gyration, R_g, as the characteristic cluster size in this paper. The root-mean-square (rms) distance between monomers in a cluster, $R_{\rm rms}$, is equal to $\sqrt{2}R_g$. The hierarchy of sizes can be schematically expressed as



A. Monodisperse clusters

First, we introduce the correlation functions for the most simple case of monodisperse clusters. Consider scattering of a plain monochromatic wave from an ensemble of random fractal clusters containing N monomers each. If the electromagnetic interaction at the optical frequency between monomers in a cluster is weak (equivalently, if the frequency of incident radiation is far from internal electromagnetic resonance frequencies of a cluster), one can use the first Born approximation for the scattering problem. In this approximation, the intensity of light scattered by the mth cluster in the direction $\mathbf{k}'$ is proportional to the well known ''structural factor,'' $I_m(\mathbf{q})$ [5,8,16]:

$$I_m(\mathbf{q})=\left|\sum_{i=1}^{N}\exp(i\mathbf{q}\cdot\mathbf{r}_i^{(m)})\right|^2=\sum_{i=1}^{N}\sum_{j=1}^{N}\exp[i\mathbf{q}\cdot(\mathbf{r}_i^{(m)}-\mathbf{r}_j^{(m)})], \tag{1}$$

where $\mathbf{r}_i^{(m)}$ is the radius vector of the ith monomer belonging to the mth cluster, $\mathbf{q}=\mathbf{k}-\mathbf{k}'$, and $\mathbf{k}$ is the wave vector of the incident wave. The scattering angle θ (the angle between $\mathbf{k}$ and $\mathbf{k}'$) is related to the absolute value of $\mathbf{q}$ by the equation

$$q=k\sqrt{2(1-\cos\theta)}, \quad k=2\pi/\lambda. \tag{2}$$

The structural factor $I_m(\mathbf{q})$ depends only on the geometry of a cluster and the scattering angle. It does not include any factors depending on the polarization of the incident wave (if it is polarized) and scattering cross section of an individual monomer that the cluster is built from. Normally, these factors are the same for all clusters, independent of their geometry and number of particles [8], and, therefore, can be excluded from consideration.

The ensemble-average scattered intensity $\langle I(\mathbf{q})\rangle$ can be easily obtained from Eq. (1). Since here we consider CCA clusters which are spherically symmetrical on average, it is clear that $\langle I(\mathbf{q})\rangle$ can depend only on the scattering angle θ, but not on the azimuthal angle φ. This means that $\langle I(\mathbf{q})\rangle$ depends only on the absolute value of $\mathbf{q}$. Grouping together the terms with $i=j$ and $i\neq j$ in the double summation (1), we obtain

$$\langle I(\mathbf{q})\rangle=\langle I(q)\rangle=N+N(N-1)\phi_2(q), \tag{3}$$

$$\phi_2(q)=\langle\exp(i\mathbf{q}\cdot\mathbf{r}_{ij})\rangle, \tag{4}$$

where $\mathbf{r}_{ij}=\mathbf{r}_i-\mathbf{r}_j$, $i\neq j$, and the ensemble averaging $\langle\ \rangle$, on the right-hand side of Eq. (4) should be taken over pairs of distinct monomers belonging to the same cluster.

Next, we introduce the two-point correlation function $p_2(r)$ which is defined as a probability density to find a pair of distinct monomers belonging to the same cluster separated by a distance r. For clusters which are spherically symmetrical on average, this function does not depend on the direction of $\mathbf{r}$. We can also define the probability density to find a radius vector between two monomers in an element of volume $d^3\mathbf{r}$ as $P_2(\mathbf{r})=P_2(r)=p_2(r)/4\pi r^2$, so that the normalization rule reads $\int_0^\infty 4\pi P_2(r)r^2dr=\int_0^\infty p_2(r)dr=1$. To indicate the dependence of the correlation function on the number of particles N, we will also use the notation $p_2(r,N)$ where appropriate. Using the two-point correlation function $p_2(r)$, and performing an integration over the spatial angles, we obtain, for ϕ_2,

$$\phi_2(q)=\int_0^\infty p_2(r)\frac{\sin qr}{qr}dr. \tag{5}$$

In this paper, we are interested not only in the average scattered intensity, but also in a measure of its fluctuations, which originate due to the random nature of clusters. A conventional parameter characterizing such fluctuations is the relative dispersion, $\sigma_I/\langle I\rangle$, where the dispersion σ_I is defined as

$$\sigma_I^2=\langle I^2\rangle-\langle I\rangle^2. \tag{6}$$

Therefore, we need to calculate the average value $\langle I^2\rangle$. For an individual cluster (say, the mth), the squared scattered intensity is given by

$$I_m^2(\mathbf{q})=\left|\sum_{i=1}^N\exp(i\mathbf{q}\cdot\mathbf{r}_i^{(m)})\right|^4 =\sum_{i=1}^N\sum_{j=1}^N\sum_{k=1}^N\sum_{l=1}^N\exp[i\mathbf{q}\cdot(\mathbf{r}_i^{(m)}-\mathbf{r}_j^{(m)}+\mathbf{r}_k^{(m)}-\mathbf{r}_l^{(m)})]. \tag{7}$$

By grouping together the terms in Eq. (7) with different indices matching each other, and performing ensemble averaging, we find that

$$\langle I^2(q)\rangle=N(2N-1)+4N(N-1)^2\phi_2(q)+N(N-1)(N^2-3N+3)\phi_4(q). \tag{8}$$

Here $\phi_4(q)$ is defined, analogously to $\phi_2(q)$, as

$$\phi_4(q)=\langle\exp(i\mathbf{q}\cdot\mathbf{r}_{ijkl})\rangle, \tag{9}$$

where $\mathbf{r}_{ijkl}=\mathbf{r}_{ij}+\mathbf{r}_{kl}=(\mathbf{r}_i-\mathbf{r}_j)+(\mathbf{r}_k-\mathbf{r}_l)$, $i\neq j$, $k\neq l$, and any of the pair of indices (i,j) can coincide with any of the pair (k,l). The sum of all coefficients in Eq. (8) is equal to N^4, so that $\langle I^2(0)\rangle=N^4$.

Now we can introduce the four-point correlation function $p_4(r)$, which is defined as the probability density to find the absolute value of $\mathbf{r}_{ijkl}$ in a cluster (with the above limitations on indices) to be equal to r. Similar to Eq. (5), we can express ϕ_4 through p_4:

$$\phi_4(q)=\int_0^\infty p_4(r)\frac{\sin qr}{qr}dr. \tag{10}$$

Finally, the expression for the relative dispersion, $\sigma_I/\langle I\rangle$, follows from Eqs. (3), (6), and (8). In the limit of large N, we obtain

$$\frac{\sigma_I(q)}{\langle I(q)\rangle}=\frac{\sqrt{(1-1/N)[1+2(N-2)\phi_2(q)+(N^2-3N+3)\phi_4(q)-N(N-1)\phi_2^2(q)]}}{1+(N-1)\phi_2(q)} \approx\frac{\sqrt{1+2(N-2)\phi_2(q)+(N^2-3N+3)\phi_4(q)-N(N-1)\phi_2^2(q)}}{1+N\phi_2(q)}. \tag{11}$$

This limit should be understood as follows. We have set $1-1/N=1$ in the numerator of formula (11), and $(N-1)\phi_2(q)=N\phi_2(q)$ in its denominator. At the same time, we have left unchanged the prefactors of ϕ_2, ϕ_4, and ϕ_2^2 under the square-root sign. This is done because, for $q \ll 1/R_g$, ϕ_2 and ϕ_4 are close to unity, and the powers of N cancel out in the expression under the square root. When q is not small, Eq. (11) simplifies to

$$\frac{\sigma_I(q)}{\langle I(q)\rangle} \approx \frac{\sqrt{1+2N\phi_2(q)+N^2[\phi_4(q)-\phi_2^2(q)]}}{1+N\phi_2(q)}. \quad (12)$$

B. Polydisperse clusters

Now we turn to the case of a polydisperse ensemble of clusters. The ensemble-averaged correlation functions for polydisperse ensembles depend on the probability density $p(N)$ to find a cluster with a given N. Therefore, one encounters an infinite variety of correlation functions, which are difficult to classify. However, it was shown in [8] that, actually, we do not need to introduce correlation functions for polydisperse ensembles, if the fractal dimension of the clusters is less than 2, and q is much larger than $1/R_g$. Instead, one can express both the average scattered intensity and its relative dispersion in terms of correlation functions for some arbitrary (but large enough) $N=\text{const}$. We briefly summarize the arguments of [8] below. A very useful consideration of polydispersity effects can be also found in [19].

It is convenient to think of a polydisperse ensemble as of a set of many monodisperse subensembles, each with its own value of N (assume that in the ''large'' ensemble there are many clusters with each value of N). Then, for each subensemble, we can use monodisperse correlation functions. In particular, the two-point correlation function, $p_2(r)$, for a given value of N, has the well-known universal asymptote [13]

$$p_2(r)=ar^{D-1}/N, \quad \text{if } r \ll R_g, \quad (13)$$

where a is a constant and D the fractal dimension. If $D<2$, as it is the case for the CCA clusters, and $q \gg 1/R_g$, the integral in Eq. (5) converges while the asymptote (13) is still valid, and we have

$$\phi_2(q)=c/Nq^D, \quad c=a\Gamma(D-1)\sin[\pi(D-1)/2] \quad (14)$$

where $\Gamma(x)$ is the gamma function. Using Eqs. (3) and (14), and in the limit of large N, we find that

$$\langle I(q)\rangle_N = N(1+c/q^D), \quad (15)$$

where the subscript ''N'' denotes averaging over a subensemble with a given N. In order to find the average over the ''large'' ensemble, we simply perform an additional averaging of Eq. (15) over N:

$$\langle I(q)\rangle = \langle N\rangle(1+c/q^D). \quad (16)$$

The values of c and D can be found from a two-point correlation function $p_2(r)$ calculated for some fixed value of N, while Eq. (16) gives the average scattered intensity for an ensemble with an arbitrary distribution of N.

Now we turn to fluctuations. As suggested in Ref. [8], we can represent the intensity of light scattered by some individual random cluster with N_m particles as

$$I_m(\mathbf{q})=N_m J_m(\mathbf{q}), \quad (17)$$

where N_m and J_m are statistically independent random variables, and

$$\langle J(\mathbf{q})\rangle = \langle J(q)\rangle = 1+c/q^D. \quad (18)$$

Then, ensemble averaging in Eq. (17) results in Eq. (16). Further, the relative dispersion of the random variable J, $\sigma_J/\langle J\rangle$, coincides with the relative dispersion of the scattered intensity calculated in a monodisperse ensemble, which is given by formula Eq. (12). The factor $N\phi_2(q)$ on the right-hand side of Eq. (12) is independent of N, if $q \gg 1/R_g$ [because $\phi_2(q) \propto 1/N$ in this case; see Eq. (14)]. It is reasonable to assume that the factor $N^2\phi_4(q)$ is also N independent in the above limit of q. (It was numerically confirmed in Ref. [8] that $\sigma_J/\langle J\rangle$ is a universal N-independent value when $q \gg 1/R_g$.) Now we can use Eq. (17) together with the statistical independence of the variables N and J to calculate $\sigma_I/\langle I\rangle$ in terms of $\sigma_J/\langle J\rangle$ and the relative dispersion of N, $\sigma_N/\langle N\rangle$. Straightforward calculations yield

$$\frac{\sigma_I}{\langle I\rangle} = \frac{\sigma_J}{\langle J\rangle}\left[\frac{\sigma_N^2}{\langle N\rangle^2}\left(1+\frac{\langle J\rangle^2}{\sigma_J^2}\right)+1\right]^{1/2}. \quad (19)$$

This formula was numerically verified in [8].

Thus, if $D<2$ and $q \gg 1/R_g$, we can consider correlation functions for monodisperse ensembles only. However, if one of these inequalities does not hold, we cannot use the universal asymptote (13) to calculate $\phi_2(q)$. In this case one needs to know the exact form of $p_2(r)$, as well as the size distribution of clusters. It is still possible to use the idea of subensembles to calculate the average scattering intensity analogously to Eqs. (15) and (16). But if $\langle I(q)\rangle_N$ is a general function of N and q, it becomes impossible to make a factorization of random variables, as was done in formula (17).

C. Optical form factors and correlation functions

The optical form factors $\phi_2(q)$ and $\phi_4(q)$ can be always found by analytical or numerical integration according to Eqs. (5) and (10) if the correlation functions are known. In some instances, it is enough to know only a small-r asymptote for the correlation function, as in the case of calculating ϕ_2 when $D<2$ and $q \gg 1/R_g$. But in general, it is desirable to know the complete correlation functions. Such knowledge would allow us, for example, to calculate ϕ_2 in the ''intermediate'' region of q ($qR_g \sim 1$), or when $D>2$. We shall see below that the complete correlation function p_4 is required for a calculation of ϕ_4 for any q, since p_4 grows faster than r at small r, and it takes many oscillations of $\sin qr$ for integral (10) to converge.

Due to the very complex nonequilibrium nature of stochastic aggregation processes that lead to formation of fractal clusters, it seems impossible, in general, to find the correlation functions analytically [apart from the universal asymptote (13)]. This fact adds importance to approaches based on numerical simulations of aggregation and analytical approxi-

mations of numerically calculated correlation functions. Since it is impossible to approximate numerical functions in every detail, below we formulate some requirements for analytical functions which are intended to represent real correlation functions.

First, the correlation functions must have the correct small-r asymptote. As we saw above, this provides the correct average scattered intensity when $D<2$ and $q \gg 1/R_g$. If we assume scaling (power-law) behavior for correlation functions at small r, the correct asymptote includes two constants: the exponent and the structural coefficient (D and a in the case of p_2).

The second requirement is the correct normalization of correlation functions, which results in $\phi_2(0)=\phi_4(0)=1$. We also require that the correlation functions have correct second moments. This provides the correct asymptote of $\phi_2(q)$ and $\phi_4(q)$ at small q. For example, $\phi_2(q) \approx 1-(qR_{\rm rms})^2/6+\cdots$ when $r\to 0$, where $R_{\rm rms}^2=2R_g^2 = \int_0^\infty r^2 p_2(r)dr$.

A general requirement for any probability distribution function whose argument is distance in space is that the probability element, $p(r)dr$, is invariant under scale transformation of the form $\mathbf{r}\to b\mathbf{r}$. This requirement does not affect the functional form of $p(r)$ [provided it has the correct dimensionality $1/L$], but rather governs the rule for the scale transformation of dimensional coefficients, such as a in Eq. (13).

A specific requirement for $p_2(r)$ originates from the principle of irreversibility of aggregation (see Sec. III A 1 below). This principle requires that $N_1 p_2(r,N_1) \geqslant N_2 p_2(r,N_2)$, if $N_1>N_2$.

Finally, we adduce the first three terms of the Taylor expansion for the optical form factors $\phi_2(q),\phi_4(q)$ near $q=0$:

$$\phi_2(q)=1-\frac{q^2\langle r_{ij}^2\rangle}{6}+\frac{q^4\langle r_{ij}^4\rangle}{120}, \tag{20}$$

$$\phi_4(q)=1-\frac{q^2\langle r_{ij}^2\rangle}{3}+\frac{q^4}{60}\left[\langle r_{ij}^4\rangle+\frac{5}{3}\langle r_{ij}^2 r_{kl}^2\rangle\right], \tag{21}$$

$$\phi_4(q)-\phi_2^2(q)=\frac{q^4[\langle r_{ij}^2 r_{kl}^2\rangle-\langle r_{ij}^2\rangle^2]}{36}. \tag{22}$$

Here $\langle r_{ij}^2\rangle = R_{\rm rms}^2 = 2R_g^2$. From this example, we see that $\phi_4(q)$ differs from $\phi_2^2(q)$ in the fourth order of q, if there is a statistical dependence between r_{ij} and r_{kl}.

III. THEORETICAL MODELS FOR THE TWO-POINT CORRELATION FUNCTION

It is well known that the fractal dimension can be derived either from the two-point correlation function at small distances or from the dependence of the average radius of gyration on the number of monomers in a cluster. If there is no multiscaling, the two approaches must yield the same result. This is not necessarily the case if the clusters manifest multiscaling. Then the functional form of the two-point correlation function becomes more complex.

A. Models with no multiscaling

1. General nonmultiscaling dependence

If there is no multiscaling in the system, the two-point correlation function has the form

$$p_2(r,N)=aN^{-1}r^{D-1}f_2[r/R_g(N)], \tag{23}$$

where a is a constant and $R_g(N)$ is the average radius of gyration for a monodisperse ensemble of clusters, which scales with N as

$$R_g(N)=R_0 N^{1/D}, \quad N\gg 1, \tag{24}$$

and R_0 is a constant of the order of the minimum separation between monomers in a cluster. The function $f_2(x)$ describes the cutoff of the correlation function. It is close to unity if $R_0/R_g \ll x \ll 1$, so that Eq. (24) tends asymptotically to the universal scaling dependence (13). When $R_0 \ll R_g$, this can be expressed as $f_2(0)=1$. For large x, $f_2(x)$ must decrease quickly enough to allow normalization. From normalization of $p_2(r,N)$ and formula (24) for R_g, it follows that

$$aR_0^D\int_0^\infty x^{D-1}f_2(x)dx=1. \tag{25}$$

Another relation for f_2 can be obtained by calculating the rms distance between monomers, $R_{\rm rms}=\sqrt{2}R_g$, by using $R_{\rm rms}^2=\int_0^\infty r^2 p_2(r,N)dr$. This relation is

$$aR_0^D\int_0^\infty x^{D+1}f_2(x)dx=2. \tag{26}$$

Given rules (25) and (26), the possible choice for the cutoff function is still very wide. In order to satisfy Eqs. (25) and (26) with given constants a, D, and R_0, $f_2(x)$ must have, in general, at least two adjustable parameters. An alternative approach (see, for example, Refs. [4,5]) allows $f_2(x)$ to have only one adjustable parameter, and treats, instead, either a or R_0 as an adjustable parameter. We argue that a function $f_2(x)$ with two adjustable parameters is more appropriate, as it allows us to reproduce simultaneously such important features of the actual (experimental) distribution $p_2(r,N)$ as asymptote (13) and the gyration radius (24).

Another general property of $f_2(x)$ follows from the concept of irreversible aggregation [15]. This concept can be most directly applied to Witten-Sander clusters [15]. Consider the density of monomers measured at the distance r from the center of the Witten-Sander aggregation process (the ''seed''). This density can be written as $\rho^{(c)}(r,N) = Np_2^{(c)}(r,N)$, where the superscript ''(c)'' indicates that the distance r is measured from the center of aggregation. Due to irreversibility of the aggregation process, the density function can only increase as the aggregation process goes on and the number of monomers N increases. Therefore, we can state that $\rho^{(c)}(r,N_1)\geqslant\rho^{(c)}(r,N_2)$ if $N_1>N_2$. Obviously, this idea cannot be applied directly to the CCA clusters because the latter do not have centers of aggregation, and the two-point correlation function $p_2(r,N)$, which is important for optics, differs in its definition from $p_2^{(c)}(r,N)$. However, we will see below from numerical calculations that the inequal-

ity $N_1 p_2(r,N_1) \geq N_2 p_2(r,N_2)$ if $N_1 > N_2$ always holds for CCA clusters. By analogy, we can attribute this property to the irreversibility of the cluster-cluster aggregation process, which means that whenever a bond between two subclusters is made, it becomes stable. Application of the above inequality to Eq. (23) leads to the conclusion that $f_2(x)$ must be a monotonously decreasing function of x.

The value of aR_0^D is fixed; it does not depend on the choice of the unit of length, i.e., it is invariant with respect to any scale transformation. This fact follows from Eq. (25), and agrees with the general requirement of scale invariance of $p_2(r,N)dr$:

$$p_2(r,N)dr = aR_0^D x^{D-1} f_2(x)dx, \quad x = r/R_g . \tag{27}$$

In addition, $p_2(r,N)dr$ does not depend on N, if $r/R_g = \text{const}$. The latter property can be viewed as a manifestation of self-similarity, which is inherent only to fractal clusters (unlike the scale invariance of $p_2 dr$, which is not related to fractality).

If $p_2(r)$ has the form Eq. (23), the optical form factor (5) becomes a function of qR_g only: $\phi_2(q,N) = \phi_2(qR_g)$.

2. Generalized exponential model

One of the mostly common used forms for $f_2(x)$ [4,19] is the generalized exponential (GE)

$$f_2(x) = \exp(-\alpha x^\beta). \tag{28}$$

Note that the particular cases of simple exponential ($\beta = 1$) and Gaussian ($\beta = 2$) functions, which are widely used in the literature [4,19], contain only one adjustable parameter, and therefore can satisfy Eqs. (25) and (26) only by coincidence.

From the ratio of equations Eqs. (25) and (26), we find that the constants α and β must satisfy

$$\alpha = \left[\frac{\Gamma\left(\frac{D+2}{\beta}\right)}{2\Gamma\left(\frac{D}{\beta}\right)} \right]^{\beta/2} . \tag{29}$$

In addition, from Eq. (25) it follows that

$$aR_0^D = \frac{\beta\left[\Gamma\left(\frac{D+2}{\beta}\right)\right]^{D/2}}{2^{D/2}\left[\Gamma\left(\frac{D}{\beta}\right)\right]^{D/2+1}} \equiv F^{\mathrm{GE}}(\beta,D). \tag{30}$$

The function $F^{\mathrm{GE}}(\beta,D)$ on the right-hand side of Eq. (30) diverges as $\beta^{-1/2}(1+2/D)^{(D/\beta)(D/2+1)}$ in the vicinity of $\beta = 0$ and decreases monotonously with β to its lower bound,

$$F^{\mathrm{GE}}(\infty,D) = D^{D/2+1}(2D+4)^{-D/2}. \tag{31}$$

If the actual value of aR_0^D is smaller than the lower bound (31), the GE cutoff function cannot satisfy Eqs. (25) and (26).

Any reasonable cutoff function must have a finite first derivative at $x = 0$, to provide enough space for the small-r asymptote to manifest itself. This condition results in $\beta > 1$ and, consequently, $aR_0^D < F^{\mathrm{GE}}(1,D)$, where

$$F^{\mathrm{GE}}(1,D) = [D(D+1)/2]^{D/2}/\Gamma(D). \tag{32}$$

Thus, the GE cutoff model can be used when $F^{\mathrm{GE}}(\infty,D) < aR_0^D < F^{\mathrm{GE}}(1,D)$. Note that $F^{\mathrm{GE}}(x,D)$ is a monotonously decreasing function of x and, therefore, the above interval is not empty.

The GE cutoff satisfies the irreversibility principle, because Eq. (28) is a monotonously decreasing function. Below, we will see that it can be not always true in the case of multiscaling.

3. Overlapping sphere model

Another model for the cutoff function is the so-called model of overlapping spheres (OS's) [4]. This model originates from the two-point correlation function for nonfractal clusters of monomers distributed randomly and with no mutual correlation in a spherical volume of a radius R. This correlation function can be found analytically, and is given by the integral

$$p_2(r,R) = 4\pi r^2 \int \rho(\mathbf{r}-\mathbf{r}',R)\rho(\mathbf{r}',R)d^3\mathbf{r}', \tag{33}$$

where $\rho(\mathbf{r},R)$ is the density of particles (unity if $r < R$ and zero otherwise). The integral in Eq. (33) is equal to the volume of intersection of two spheres of radius R separated by the distance r between their centers. The result for p_2, that follows from Eq. (33), is

$$p_2(r,R) = \begin{cases} [(3r^2)/(2R^3)](r/2R-1)^2(r/2R+2) & \text{if } r \leq 2R \\ 0 \quad \text{if } r > 2R . \end{cases} \tag{34}$$

The idea of the OS model is to take the cutoff function for trivial clusters and to use it for fractal clusters. While in the trivial case the parameters of the cutoff function are defined by the radius R of the spherical volume occupied by the particles, for random fractal clusters such a parameter does not exist. Instead, we write the cutoff function in the following form:

$$f_2(x) = \begin{cases} (x-x_0)^2(x+2x_0)/(2x_0^3) & \text{if } x \leq x_0 \\ 0 \quad \text{if } x > x_0 . \end{cases} \tag{35}$$

Evidently, the OS model in its pure form contains only one adjustable parameter x_0. Its value can be found from the ratio of Eqs. (25) and (26):

$$x_0 = \left[2\frac{(D+2)(D+5)}{D(D+1)} \right]^{1/2} . \tag{36}$$

However, we must satisfy one of the Eqs. (25) and (26) independently. In order for this to be possible, the following relation between aR_0^D and D must hold:

$$aR_0^D = \frac{D(D+1)(D+3)}{3}\left[\frac{D(D+1)}{2(D+2)(D+5)} \right]^{D/2} . \tag{37}$$

If, by any chance, the experimental values of a, D, and R_0 satisfy Eq. (37), the OS cutoff with one adjustable parameter (35) can be used. If not, one can further generalize function (35) and write

$$f_2(x)=\begin{cases}(x-x_1)^2(x+2x_2)/(2x_1^2x_2) & \text{if } x\leqslant x_1\\ 0 \quad \text{if } x>x_1\,,\end{cases} \tag{38}$$

where x_1 and x_2 are independent adjustable parameters. Substitution of Eq. (38) into Eqs. (25) and (26) results in

$$x_2=\frac{D+2}{2}\,\frac{x_1(x_1^2-x_{1,0}^2)}{(D+3)x_{1,\max}^2-(D+5)x_1^2}, \tag{39}$$

$$aR_0^D=\frac{D(D+1)(D+2)^2(D+3)}{12}\,\frac{x_1^2-x_{1,0}^2}{x_1^{D+2}}\equiv F^{\mathrm{OS}}(x_1,D), \tag{40}$$

$$x_{1,0}=\left[\frac{2(D+4)(D+5)}{(D+1)(D+2)}\right]^{1/2},$$
$$x_{1,\max}=\left[\frac{2(D+4)(D+5)}{D(D+1)}\right]^{1/2}, \tag{41}$$

where x_1 must be found from Eq. (40). The maximum value of $F^{\mathrm{OS}}(x_1,D)$ as a function of x_1 is reached when $x_1=x_{1,\max}$, and is equal to

$$F^{\mathrm{OS}}(x_{1,\max},D)=\frac{D(D+1)(D+2)(D+3)}{6}\left[\frac{D(D+1)}{2(D+4)(D+5)}\right]^{D/2}. \tag{42}$$

If it happens that the actual value of aR_0^D is greater than the value of $F^{\mathrm{OS}}(x_{1,\max},D)$ Eq. (42), the generalized OS model cannot satisfy Eqs. (25) and (26). If aR_0^D is smaller than $F(x_{1,\max})$, Eq. (40) has two solutions for x_1, both exceeding the value of $x_{1,0}$.

Theoretically, any of the two possible solutions to Eqs. (39) and (40) can be chosen, but in practice we need also to satisfy the irreversibility principle which states that $f_2(x)$ must be a positive monotonously decreasing function. Applied to Eq. (38), this requirement results in

$$x_2>x_1/4 \quad \text{or} \quad x_2<-x_1/2, \tag{43}$$

where x_1 is positive by definition [otherwise, the interval where cutoff (38) is defined becomes empty]. Using Eqs. (39) and (43), we find the allowed interval for x_1 as

$$x_{1,0}<x_{1,\min}<x_1<x_{1,\max}, \tag{44}$$

where

$$x_{1,\min}=\left[\frac{2(D+4)(D+5)}{D(D+3)}\right]^{1/2}=x_{1,\max}\left(\frac{D+1}{D+3}\right)^{1/2}. \tag{45}$$

Thus one solution (with $x_{1,\min}x_1<x_{1,\max}$) to Eq. (40) is acceptable; the second solution (with $x_1>x_{1,\max}$) either violates the irreversibility principle or leads to negative values of p_2.

The value of $F^{\mathrm{OS}}(x_{1,\min},D)$ is given by

$$F^{\mathrm{OS}}(x_{1,\min},D)=\frac{1}{D+1}\left(\frac{D+3}{D+1}\right)^{D/2}F^{\mathrm{OS}}(x_{1,\max},D). \tag{46}$$

The values of aR_0^D that can be compatible with the OS model lie in the interval $F^{\mathrm{OS}}(x_{1,\min},D)<aR_0^D<F^{\mathrm{OS}}(x_{1,\max},D)$. We note that the ratio $F^{\mathrm{OS}}(x_{1,\max},D)/F^{\mathrm{OS}}(x_{1,\min},D)$ increases with D. For $D=2$, this ratio is equal to 9/5.

B. Multiscaling

The multiscaling phenomenon is well established for Witten-Sander clusters [15]. It also was recently observed in computer-generated CCA clusters [21]. The presence of multiscaling results in a more general scaling behavior of the two-point correlation function than the one specified by formulas (23) and (24). In particular, the value of D defined from the asymptotic behavior (13) of $p_2(r,N)$ at small distances can differ from D found from the dependence (24) of R_g on N. However, we emphasize that it does not follow from $D_1=D_2=D$ that the two-point correlation function must necessarily have the form Eq. (23). In this paper, we refer to multiscaling as *any* deviation from the most simple functional form Eq. (23) including, but not limited, to the case of $D_1\neq D_2$.

In this subsection we will assume that multiscaling takes place only in the two-point correlation function, while the radius of gyration obeys the simple scaling form (24) (probably with a different exponent D_2). The case where the correlation function and the gyration radius both demonstrate multiscaling is considered in Sec. III B 1.

1. General multiscaling dependence

If two different exponents D_1 and D_2 exist, the most general functional form for $p_2(r,N)$, which has the universal asymptote Eq. (13), is

$$p_2(r,N)=aN^{-1}r^{D_1-1}g_2[r/R_g(N),N], \tag{47}$$

$$R_g(N)=R_0N^{1/D_2}, \quad N\gg 1, \tag{48}$$

where $g_2(0,N)=1$ for any N (assuming that $R_g\gg R_0$). Note that since $g_2(x,N)\to 1$ when $x\to 0$ for any N, it is impossible to factorize $g_2(x,N)$ as $g_2(x,N)=g_{2a}(x)g_{2b}(N)$.

The two rules for $g_2(x,N)$, analogous to Eqs. (25) and (26), are

$$aR_0^{D_1}\int_0^\infty x^{D_1-1}g_2(x,N)\,dx=N^{1-D_1/D_2}, \tag{49}$$

$$aR_0^{D_1}\int_0^\infty x^{D_1+1}g_2(x,N)\,dx=2N^{1-D_1/D_2}. \tag{50}$$

As above, $aR_0^{D_1}$ is invariant with respect to any scale transformation; in addition, it does not depend on N by the defi-

nitions of the constants a and R_0. Therefore, the only source of the dependence on N on the left-hand sides of Eqs. (49) and (50) is the function $g_2(x,N)$.

The value $p_2(r,N)dr$ now is defined by

$$p_2(r,N)dr = aR_0^{D_1} N^{D_1/D_2-1} x^{D_1-1} g_2(x,N)dx. \quad (51)$$

The explicit dependence of $p_2(r,N)dr$ on N (apart from the factor $x=r/R_g$) may be viewed as a violation of the self-similarity. However, the factor $N^{D_1/D_2-1}g_2(x,N)$ can, in principle, be independent of N in some region of x, which would mean self-similarity in the intermediate asymptote. In other words, a cluster can be statistically not self-similar as a whole, but its smaller parts of different sizes lying in the region $R_0 \ll r \ll R_g$ can still be self-similar.

The optical form factor ϕ_2, Eq. (5), now depends on both qR_g and N. It becomes a general function of two variables q and N. From the scattering perspective, this means that in the case of multiscaling the differential scattering cross section (normalized to unity for $q=0$) depends not only on the dimensionless size parameter qR_g, but also on the number of particles in a cluster.

2. Generalized exponential cutoff in the case of multiscaling

Let us consider how the model cutoff functions can be used in the case of multiscaling. For the GE cutoff function (28), we can assume that α and β are functions of N:

$$g_2(x,N) = \exp[-\alpha(N)x^{\beta(N)}]. \quad (52)$$

The equations for α and β, which follow from Eqs. (49) and (50), are quite similar to Eqs. (29) and (30):

$$\alpha(N) = \left[\frac{\Gamma\left(\frac{D_1+2}{\beta(N)}\right)}{2\Gamma\left(\frac{D_1}{\beta(N)}\right)} \right]^{\beta(N)/2}. \quad (53)$$

$$aR_0^{D_1} N^{D_1/D_2-1} = \frac{\beta(N)\left[\Gamma\left(\frac{D_1+2}{\beta(N)}\right)\right]^{D_1/2}}{2^{D_1/2}\left[\Gamma\left(\frac{D_1}{\beta(N)}\right)\right]^{D_1/2+1}} \equiv F^{GE}[\beta(N),D_1]. \quad (54)$$

The last equation specifies the dependence $\beta(N)$ implicitly.

If $D_1<D_2$, the left-hand side of Eq. (54) decreases with N, and, when N is large enough, becomes smaller than the lower bound of the right-hand side, given by Eq. (31) (with D replaced by D_1). This means that the multiscaling GE cutoff cannot be used for arbitrary large N when $D_1<D_2$. However, in practice we may not need to deal with extremely large N. If the ratio $|D_2-D_1|/D_2$ is much smaller than unity, Eq. (54) can still have a solution for some practical values of N, even if $D_1<D_2$.

Consider the case of $D_1<D_2$ in more detail. As mentioned above, the expression on the right-hand side of Eq. (54) is a monotonously decreasing function of β. Therefore, the solution to this equation, $\beta(N)$, is a monotonously increasing function of N if $D_1<D_2$, provided this solution exists. That is, $\beta(N)$ grows toward infinity when N approaches some critical value, and there is no solution for larger values of N. The critical value of N is equal to $[aR_0^{D_1}/F^{\mathrm{GE}}(\infty,D_1)]^{D_2/(D_2-D_1)}$, where $F^{\mathrm{GE}}(\infty,D_1)$ is defined by Eq. (31). If $D_2-D_1 \ll D_2$, the critical value of N can be large. In practice, we can use the GE cutoff with $D_1<D_2$, if the number of monomers is much smaller than the above critical value. (According to our calculations, the critical value of N is $\sim 10^7$ monomers for three-dimensional CCA clusters; see Sec. VI B.)

If $D_1>D_2$, $\beta(N)$ is a monotonously decreasing function of N, and a solution to Eq. (54) always exists. However, for some value of N, β will become smaller than unity, and function (52) cannot serve as a cutoff function. The characteristic value of N in this case is $[F^{\mathrm{GE}}(1,D_1)/aR_0^{D_1}]^{D_2/(D_1-D_2)}$.

We now consider the limitations which are set upon the dependence $\beta(N)$ by the principle of irreversible aggregation. For the GE cutoff function, it states that

$$\alpha(N_1)\left[\frac{r}{R_g(N_1)}\right]^{\beta(N_1)} < \alpha(N_2)\left[\frac{r}{R_g(N_2)}\right]^{\beta(N_2)} \quad \text{if } N_1>N_2. \quad (55)$$

Using formula (48), we can rewrite this inequality as

$$\left(\frac{r}{R_0}\right)^{\beta(N_1)-\beta(N_2)} < \frac{\alpha(N_2)N_1^{\beta(N_1)/D_2}}{\alpha(N_1)N_2^{\beta(N_2)/D_2}} \quad \text{if } N_1>N_2. \quad (56)$$

By looking at the limit of $r \gg R_0$, we find that $\beta(N_1)-\beta(N_2)$ must be a negative number or zero; otherwise the expression on the left-hand side of Eq. (56) can exceed any given number when r/R_0 is large enough. Therefore, we conclude that $\beta(N)$ must be a monotonously decreasing function of N in order to satisfy the irreversibility principle.

As we saw above, $D_1<D_2$ results in an increasing function $\beta(N)$, which is incompatible with the irreversibility principle. However, we should note that in real clusters the violation of this principle can occur at very large values of $x=r/R_g$, where the theoretical cutoff is no longer valid. Indeed, the real correlation functions become exactly equal to zero for large values of x, while the GE cutoff is always positive. This means that the real correlation function must decrease faster for large x than any GE model can provide. On the other hand, it should pose no difficulty for a theoretical description of the correlation function, since the physically important integrals of the type of Eq. (5) converge much sooner than the violation takes place.

Lastly, we consider the particular case of $D_1=D_2=D$, where α and β become N-independent constants. This means that the GE cutoff function (52) becomes N independent too. However, a general function $g_2(x,N)$ can depend on N even if $D_1=D_2$ (see Appendix A). As a result, multiscaling can exist even if $D_1=D_2=D$, but the GE cutoff is incapable of describing this kind of multiscaling.

3. Overlapping spheres cutoff in the case of multiscaling

Analogously to Sec. III B 2, we consider the OS model (38), with x_1 and x_2 being functions of N:

$$g_2(x,N)=\begin{cases}[x-x_1(N)]^2[x+2x_2(N)]/[2x_1^2(N)x_2(N)] & \text{if } x\leqslant x_1(N)\\ 0 \quad \text{if } x>x_1(N)\ .\end{cases} \tag{57}$$

The relation between $x_1(N)$ and $x_2(N)$, Eq. (39), still holds in this case (with D being replaced by D_1), but instead of Eq. (40) we have

$$aR_0^{D_1}N^{D_1/D_2-1}=F^{OS}(x_1,D_1). \tag{58}$$

Comparing this equation to its counterpart, Eq. (54), we can notice the main differences between OS and GE cutoffs.

First, the function $F^{\mathrm{GE}}(\beta,D_1)$ has a finite lower bound, but is not limited from above. $F^{\mathrm{OS}}(x_1,D_1)$ has, on the other hand, a finite upper bound but can approach zero. This means that in the OS model the value of $aR_0^{D_1}N^{D_1/D_2-1}$ must go to zero for large N. Subsequently, the case of $D_1<D_2$ is more compatible with the OS cutoff than $D_1>D_2$ (contrary to the GE case). Exactly as in the GE case, we can still use the OS cutoff with $D_1>D_2$ if the value of N is not too large. The critical value of N is $[F^{\mathrm{OS}}(x_{1,\max},D_1)/aR_0^{D_1}]^{D_2/(D_1-D_2)}$, where $F^{\mathrm{OS}}(x_{1,\max},D_1)$ is given by Eq. (42).

Second, in the OS model there are, generally, two possible solutions to Eq. (58) which, in the case of multiscaling, result in two possible branches of the function $x_1(N)$. As we saw above, application of the irreversibility principle leaves only one allowed solution, namely the one with $x_1<x_{1,\max}$, if there is no multiscaling in the system. But if the multiscaling is manifested, consideration of the irreversibility principle becomes more complicated. It is shown in Appendix B that, in the limit of very large N and $D_1<D_2$, the upper branch of the solution ($x_1<x_{1,\max}, x_2>0$) violates the irreversibility principle, while the lower branch ($x_1>x_{1,\max}, x_2<0$), though satisfying the irreversibility principle, leads to negative values of p_2 for large r.

Finally, the OS model is incapable of describing multiscaling with $D_1=D_2$ due to the same reasons that were discussed in Sec. III B 2 for the GE model.

4. Model of continuous fractal dimension

The model of continuous fractal dimension (CFD) was proposed in Refs. [15,20] for description of multiscaling. The CFD was calculated numerically in Ref. [15] for Witten-Sander aggregates.

The two-point correlation function in the CFD model is given by the expression

$$p_2(r,N)=aN^{-1}\frac{r^{D[r/R_g(N)]-1}}{R_0^{D[r/R_g(N)]-D_1}}h_2[r/R_g(N)]. \tag{59}$$

Here $D(x)$ is the CFD, and R_g depends on N according to Eq. (48). Since Eq. (59) must coincide with Eq. (13) when $r\ll R_g$, we require that $D(0)=D_1$ and $h_2(0)=1$. Note that there is no general requirement that $D(\infty)=D_2$; at least it cannot be deducted from the definitions of D_2 [Eq. (48)] and h_2 [Eq. (59)]. But, as can be seen from Eqs. (51) and (60) (see below), a cluster becomes self-similar in the limit of $x\gg 1$ if $D(\infty)=D_2$.

The factor $R_0^{D[r/R_g(N)]-D_1}$ in Eq. (59) provides the correct dimensionality of p_2. In order to satisfy the dimensionality requirement, we can use R_g instead of R_0 in the denominator of Eq. (59). But it easy to see that, in this case, Eq. (59) would degenerate into a nonmultiscaling dependence with $f_2(x)=x^{D(x)-D_1}h_2(x)$.

The functional dependence (59) is less general than Eq. (47). The latter is, evidently, the most general form of a function of two variables, r and N. In order for Eqs. (47) and (59) to represent the same function, the following relation must hold:

$$g_2(x,N)=x^{D(x)-D_1}N^{[D(x)-D_1]/D_2}h_2(x). \tag{60}$$

Substituting this expression for $g_2(x,N)$ into Eqs. (49) and (50), we find that

$$aR_0^{D_1}\int_0^\infty x^{D(x)-1}N^{D(x)/D_2-1}h_2(x)dx=1, \tag{61}$$

$$aR_0^{D_1}\int_0^\infty x^{D(x)+1}N^{D(x)/D_2-1}h_2(x)dx=2. \tag{62}$$

It was shown in [15] from the irreversibility principle that $D(x)$ must be a monotonously decreasing function. On the other hand, it follows from Eq. (61) that the function $D(x)/D_2-1$ must change its sign [to prove this, consider the derivative $\partial/\partial N$ of Eq. (61)]. At last, we know that $D(0)=D_1$. These three facts are compatible with each other only if $D_1>D_2$.

Note that in order to calculate numerically $D(x)$, we do not need to know the cutoff function $h(x)$. Indeed, we can take clusters with different N and different R_g, and choose some value of x common for all clusters. Then the value of r would be different for clusters with different N, which makes it possible to calculate $D(x)$ from the slope of $p_2(r,x=\mathrm{const})$ as a function of r.

C. Double multiscaling

In Secs. III B 1–III B 4, we considered multiscaling only in the correlation function, but assumed that the radius of gyration obeys the simple scaling dependence (48). Now we assume that it is not so, and $R_g(N)$ is an arbitrary monotonously increasing function. In this case, the crucial role for all theoretical models plays the function $R_g^{D_1}(N)/N$. We emphasize that the constant D_1 should be found from the small-r asymptote of p_2 rather than the radius of gyration.

We can still use the GE and OS models for cutoff functions. Relations (53) and (39) do not change, but Eqs. (54) and (58) transform to

$$aR_g^{D_1}(N)/N = F^{(*)}(\text{adjustable parameter}), \qquad (63)$$

where $(*)$ denotes either GE or OS, and the adjustable parameter is β or x_1 for the GE or OS models, respectively. To select an appropriate model, one should analyze how the function $R_g^{D_1}(N)/N$ behaves when $N \to \infty$.

It is a little more difficult to use the CFD model in the case of double multiscaling. For this model, we need a parameter of dimensionality of length. As mentioned in Sec. III B 4, R_g cannot be used as such a parameter, because it leads to degeneration of CFD into a nonmultiscaling model. But the parameter R_0 is not precisely defined in the case of double multiscaling. In fact, there might be several parameters of the dimensionality of length, which govern the dependence of R_g on N. However, we can choose the smallest of them and call it R_0. It can be a lattice period or a minimum separation between particles, or the average distance from a monomer to its nearest neighbor (the latter characteristic always exists). Regardless of the definition of R_0, we can always write $R_g(N) = R_0 \psi(N)$, where $\psi(N)$ is a dimensionless function. Then, instead of Eq. (49), we obtain

$$aR_0^{D_1} \int_0^\infty N^{-1} \psi^{D(x)}(N) x^{D(x)-1} h(x)\, dx = 1. \qquad (64)$$

Applying the derivative $\partial/\partial N$ to Eq. (64), we find that that $D(x)N[\partial\psi/\partial N]/\psi - 1$ must change its sign as a function of x for any N. This, in particular, must be true for $N \to \infty$. If we make the reasonable assumption that $0 < c_1 < D(x) < c_2 < \infty$, we come to the conclusion that there must exist a finite limit of $N[\partial\psi/\partial N]/\psi$ when $N \to \infty$. The condition for such a limit to exist is that $\psi(N)$ is a power function for large N. In other words, R_g must depend on N according to Eq. (48), at least for large values of N.

To summarize, the CFD model can be used in the case of double multiscaling only when $\lim_{N\to\infty} R_g(N) = R_0 N^{1/D_2}$. Note that this requirement follows only from normalization of $p_2(r,N)$

IV. THEORETICAL MODELS FOR THE FOUR-POINT CORRELATION FUNCTION

In this section we propose a method to deduce the small-r asymptote of p_4 theoretically, considering that p_2 is known. The problem of the cutoff function for p_4 will be considered numerically in Sec. VII.

From a mathematical point of view, the four-point correlation function $p_4(r)$ is just some probability distribution function. It has the same normalization as $p_2(r)$ but a different second moment. That is, the second moment of p_2 is $\langle \mathbf{r}_{ij}^2 \rangle = R_{\rm rms}^2 = 2R_g^2$, while the second moment of p_4 is $\langle \mathbf{r}_{ijkl}^2 \rangle = \langle (\mathbf{r}_{ij} + \mathbf{r}_{kl})^2 \rangle = 2R_{\rm rms}^2 = 4R_g^2$.

A. Approximation of statistical independence

It is well known that the positions of monomers in a cluster are correlated. In other words, the information that there is a monomer at the point $\mathbf{r}_i$ influences the probability to find another monomer somewhere else, say, at the point $\mathbf{r}_j$. The smaller the distance $|\mathbf{r}_i - \mathbf{r}_j|$, the stronger the correlation. The source of this statistical dependence is evident: monomers tend to group close to each other, forming ''blobs'' and empty ''holes'' of all sizes, up to the maximum cluster size.

But if we consider $\mathbf{r}_{ij}$ and $\mathbf{r}_{kl}$ as random variables, the reason for their statistical dependence is not as clear. Imagine that we found a pair of monomers in a cluster separated by $\mathbf{r}_{ij}$. The question that we have to answer is: does this fact change the probability of finding another pair of monomers separated by $\mathbf{r}_{kl}$ (as compared to the *a priori* probability)? If the answer is negative, then $\mathbf{r}_{ij}$ and $\mathbf{r}_{kl}$ are statistically independent random variables, which constitutes the main assumption of the approximation of statistical independence (ASI).

Strictly speaking, if the ASI was correct, we would not need to know the correlation function p_4 at all. Indeed, the optical form factor $\phi_4(q)$ could be calculated simply as $\langle \exp(i\mathbf{q}\cdot\mathbf{r}_{ijkl}) \rangle = \langle \exp(i\mathbf{q}\cdot\mathbf{r}_{ij}) \rangle \langle \exp(i\mathbf{q}\cdot\mathbf{r}_{kl}) \rangle = \phi_2^2(q)$. This equality always holds for small q [up to second order in q, see Eqs. (20)–(22)]. But for larger q, it is not necessarily so.

To illustrate the above statement, consider the relative dispersion of the scattered light given by formula (12). As follows from Eqs. (12) and (14), if $\phi_4(q) = \phi_2^2(q)$, we have

$$\frac{\sigma_I(q)}{\langle I(q) \rangle} = \sqrt{\frac{1 + 2c/q^D}{1 + c/q^D}}, \qquad (65)$$

where the constant c is given in Eq. (14). On the other hand, it was shown numerically in Ref. [8] that the relative dispersion is very close to unity when $1/R_g \ll q \ll 1/R_0$, and from formula (65) we see that this would be true only if $2c/q^{D_1} \ll 1$. With the use of Eqs. (2) and (14), this inequality translates into

$$\frac{2aR_0^{D_1} \sin[(D_1 - 1)\pi/2]}{[2\pi\sqrt{2(1-\cos\theta)}]^{D_1}} \left(\frac{\lambda}{R_0} \right)^{D_1} \ll 1. \qquad (66)$$

Since the prefactor for $(\lambda/R_0)^{D_1}$ is the order of unity or larger, the wavelength must be much smaller than the characteristic separation between monomers, R_0, in order to provide the above inequality. But this contradicts the numerical results of Ref. [8], where the value of λ was chosen to be significantly larger than R_0, but the relative dispersion was still very close to unity for large scattering angles.

The above example demonstrates that the hypothesis of statistical independence is not confirmed numerically [8]. Still, it is convenient to use this hypothesis as a starting point. Below, we adopt the following theoretical approach: First, we use the ASI to find the small-r asymptote of p_4. Then we show that the same small-r asymptote follows from the requirement of correct small- and large-q asymptotes of $\phi_4(q)$. In Sec. VII we find the higher corrections to p_4 numerically.

Now we turn to the determination of the functional form of p_4 in the ASI. The function $p_4(r)$ was defined as a probability density to find the absolute value of $\mathbf{r}_{ijkl} = \mathbf{r}_{ij} + \mathbf{r}_{kl}$ to be equal to r. Since in the ASI $\mathbf{r}_{ijkl}$ is a sum of two statistically independent random variables, we can apply the general formalism of the theory of functions of random variables, which allows us to express the probability distribution $p(x+y=r)$ through a convolution of $p(x=r)$ and $p(y=r)$. Here x and y are statistically independent random

variables, and r is the argument of the probability distribution functions. It is more convenient to apply this formalism to the functions $P_m(\mathbf{r})=P_m(r)=p_m(r)/4\pi r^2$, where $m=2$ and 4, and $P_m(\mathbf{r})$ is the corresponding probability density to find the value of $\mathbf{r}_{ij}$ (for $m=2$) or $\mathbf{r}_{ijkl}$ (for $m=4$) in the three-dimensional element of volume $d^3\mathbf{r}$:

$$P_4(\mathbf{r})=\int P_2(\mathbf{r}-\mathbf{r}')P_2(\mathbf{r}')d^3\mathbf{r}'. \tag{67}$$

The normalization of P_4 follows immediately from Eq. (67) if P_2 is normalized correctly. Returning to our usual notation, p_2 and p_4, and performing integration over spatial angles, we find that

$$p_4(r)=\frac{r}{2}\int_0^\infty \frac{p_2(r')}{r'}dr' \int_{|r-r'|}^{r+r'} \frac{p_2(r'')}{r''}dr''. \tag{68}$$

First, we analyze the form of p_4 in the case of no multiscaling. Substituting the function $p_2(r,N)$ in the form Eq. (23) into Eq. (68), we find that

$$p_4(r,N)=\frac{a^2r^{2D-1}}{2N^2}Y(r/R_g), \tag{69}$$

$$Y(x)=\int_0^\infty (t')^{D-2}f_2(xt')dt' \int_{|t'-1|}^{t'+1}(t'')^{D-2}f_2(xt'')dt''. \tag{70}$$

Here $f_2(x)$ is the nonmultiscaling cutoff for the two-point correlation function.

Consider the asymptotical behavior of the integral $Y(x)$ near $x=0$. If $D<1.5$, one can always choose x to be small enough for integral (70) to converge while $f_2(xt')\approx f_2(xt'')\approx 1$. This means that $\lim_{x\to 0}Y(x)=\text{const}$ when $D<1.5$. Though, from the mathematical point of view, this constant is a leading term the expansion of $Y(x)$ near $x=0$, in practice we are not always able to achieve really small values of $x=r/R_g$. Therefore, it is desirable to obtain the second term in the expansion of $Y(x)$, especially when D is close to 1.5 and the convergence of integral (70) with $f_2=1$ is slow. Strictly speaking, the second term of the expansion of $Y(x)$ depends on the form of f_2. This is clear already from the fact that the integral of the first correction to the integrand in Eq. (70) [which can be obtained with the use of $f(x)=1+f'(0)x$] diverges when $D>1$, i.e., for any reasonable fractal dimension. However, we can assume that the function $f_2(x)$ decreases with x faster than any power function of x and has a ''characteristic length'' $x_c\sim 1$. Then we can replace the infinite upper limit of integration in the first integral in Eq. (70) by x_c/x while keeping $f_2=\text{const}$, and the expansion of $Y(x)$ acquires the form

$$Y(x)=\frac{c_2}{3-2D}-\frac{c_1}{3-2D}x^{3-2D} \quad \text{if } x\to 0, \tag{71}$$

where c_1 and c_2 are positive constants. While the above derivation of the second term of the expansion is not mathematically rigorous, it turns out to be correct for two seemingly different forms of f_2: a steplike function [$f_2(x)=1$ if $x<x_0$ and $f_2(x)=0$ if $x>x_0$] and a linear cutoff [$f_2(x)=1-x/x_0$ if $x<x_0$ and $f_2(x)=0$ if $x>x_0$].

Returning to the small-r asymptote of $p_4(r,N)$, we find that

$$p_4(r,N)=\frac{c_1r^2}{(2D-3)N^{3/D}}-\frac{c_2r^{2D-1}}{(2D-3)N^2} \quad \text{when } r\ll R_g. \tag{72}$$

Here c_1 and c_2 are positive constants, different from c_1 and c_2 used in Eq. (71).

A similar analysis can be applied for the case of $D>1.5$, but it is easier to start directly from formula (68) in this case. Indeed, the two-dimensional area of integration in Eq. (68) is a semi-infinite rectangular strip that is symmetrical with respect to the line $r'=r''$ and with two corners at the points $(0,r)$ and $(r,0)$. When $r\to 0$, the integral over this area can be replaced by $\sqrt{2}r$ times a one-dimensional integral along the line $r'=r''$ from r to infinity. When $D>1.5$, this integral converges and returns the result of Eq. (72). But now the first term in this expression is leading, and the second term is the first correction, contrary to the case of $D<1.5$. It should be noted that in both cases the first correction is negative.

Since we are mostly interested in the situation where $D>1.5$ (CCA clusters), we can rewrite Eq. (72) as

$$p_4(r,N)=\frac{br^2}{N^{3/D}}\left[1-c_1\left(\frac{r}{R_g}\right)^{2D-3}\right]. \tag{73}$$

Again, c_1 is a positive coefficient. The expression in the square brackets can be viewed as the first two terms of an expansion of p_4 in terms of integer powers of $(r/R_g)^{2D-3}$. In principle, we can think of a sequence of coefficients c_k (with $c_0=1$) which define this expansion completely. Such an expansion does not coincide with the Taylor expansion. This can be easily seen from the example of a GE function of the form $f_4(x)=\exp(-\alpha x^{2D-3})$. The first derivative of this function at $x=0$ is equal to $-\infty$ if $1.5<D<2$. A function with an infinite derivative at $x=0$, though formally satisfying the condition $f_4(x)\approx 1$ if $x\ll 1$, does not leave any space for the universal asymptote to manifest itself, as pointed out in Secs. III A 2 and III B 2.

Thus we have to conclude that, in the framework of the ASI, the functional form of $p_4(r)$ is more complicated than the simple scaling behavior of the kind of Eq. (23). It cannot be described by a scaling behavior with a cutoff that has a finite first derivative at $r=0$. Note that one of the alternative possibilities is to consider separate cutoff functions for each term in Eq. (72).

Finally, the expression for p_4, Eq. (72), is modified in the case of multiscaling as

$$p_4(r,N)=\frac{c_1r^2}{(2D_1-3)N^{[2(1-D_1/D_2)+3/D_2]}} -\frac{c_2r^{2D_1-1}}{(2D_1-3)N^2} \quad \text{when } r\ll R_g, \tag{74}$$

where both coefficients c_1 and c_2, are different from those in formula (72), and can, in principle, depend on N.

B. Method of Fourier transformation

As discussed in Sec. IV A, the ASI correctly describes optical form factors for small values of q (up to second order in q) but fails when $q \gg 1/R_g$. In this section we show that it is possible to modify the cutoff functions in Eqs. (69) and (74) in such a way that the resultant correlation function would produce correct results for large values of q as well.

We start from the asymptotic behavior of $\phi_4(q)$. As follows from Eqs. (20) and (21), for $q \ll 1/R_g$, we have $\phi_4(q) = \phi_2^2(q)$. For $q \gg 1/R_g$, we know from numerical calculations that the relative dispersion of scattered light, given by Eq. (12), is very close to unity [8]. As follows from Eq. (12), this means that $\phi_4(q) = 2\phi_2^2(q)$ for $q \gg 1/R_g$. We can write ϕ_4 in the form

$$\phi_4(q,N) = \xi(qR_q,N)\phi_2^2(q,N), \tag{75}$$

where $\xi(0,N) = 1$, $\xi(\infty,N) = 2$, and $\xi(x,N)$ is a monotonously increasing function of x with a finite first derivative. It follows from formula (22) that the first two terms of the Taylor expansion of $\xi(x,N)$ near $x = 0$ have the form $\xi(x,N) \approx 1 + c(N)x^4$. The basic assumption of this section is that there exists a region of x, $0 < x < x_c$, where $\xi(x,N)$ is close to unity (say, $|\xi(x) - 1| < \delta$, where δ is a predetermined small constant) for any N, where the right bound of this region, x_c, does not depend on N. This is always true if there is no multiscaling, since both ϕ_2 and ϕ_4 are functions of qR_g only in this case, and $\xi(x,N)$ does not depend on N.

Now we can use the definition of ϕ_4, Eq. (10), and make the inverse Fourier transformation to obtain an expression for $p_4(r)$:

$$p_4(r) = \frac{2r}{\pi}\int_0^\infty q\xi(qR_g,N)\phi_2^2(q)\sin(qr)dq. \tag{76}$$

Next, we substitute expression (5) for ϕ_2 to Eq. (76) to obtain a closed expression for $p_4(r)$ in terms of $p_2(r)$ and $\xi(x,N)$:

$$p_4(r) = \frac{2r}{\pi}\int_0^\infty \int_0^\infty \frac{p_2(r')p_2(r'')}{r'r''}dr'dr'' \int_0^\infty \times \xi(qR_g,N)\frac{\sin(qr)\sin(qr')\sin(qr'')}{q}dq. \tag{77}$$

If $\xi(x,N) = 1$, the integral over q can be taken, and we return to expression (68). If $\xi(x,N)$ is a general function with the properties specified above, integrals (77) and (68) are not identical. However, it is easy to show that the small-r asymptote of the right-hand part of Eq. (77) coincides with Eq. (72). The mathematical arguments here are quite similar to those used in Sec. IV A. Since the asymptotic behavior of p_4 at small r is determined by large values of r' and r'', the integral over q in Eq. (77) would converge while $\xi \approx 1$. According to the assumption made above, a finite region where $\xi \approx 1$ exists for any N. In this case the first two terms of expansion of p_4 near $r = 0$ do not depend on the form of $\xi(x,N)$, and coincide with those in the ASI. But the higher corrections to $p_4(r)$ differ, of course, from those in the ASI.

V. NUMERICAL PROCEDURES

A. Simulation of fractal cluster aggregation

We have implemented the CCA algorithm on a simple cubic lattice with periodic boundary conditions. We have built 40 random clusters for each value of the number of particles in a cluster N, except for $N = 20\,000$, when we have built only 20 clusters. The size of the lattice L varied depending on N. The following values of L were selected: $L = 200$ for $N = 5000$, $L = 260$ for $N = 7500$, $L = 300$ for $N = 10\,000$, $L = 310$ for $N = 12\,500$, $L = 340$ for $N = 15\,000$, and $L = 350$ for $N = 20\,000$. This ensured that the density of monomers was sufficiently low (3.7×10^{-4} for $N = 10\,000$), and that the cluster size was smaller than the size of the lattice. For example, the rms distance between two particles in clusters with $N = 10\,000$ was 99.4 (lattice units), which is substantially less than the corresponding lattice size. This was true also for the largest clusters with $N = 20\,000$, with the rms distance between particles equal to 141.7. However, we should note that due to computational limitations, L was not proportionally large for $N = 20\,000$ in comparison with other N. Comparing to $N = 10\,000$ (with $L = 300$), we could have expected that $L \approx 300\sqrt{2} \approx 424$ for $N = 20\,000$.

During the aggregation, each subcluster was moved with equal probability (independently of its size), and no rotations were allowed. Our simulations showed that clusters of all sizes collided during the aggregation. At approximately half of the full aggregation time, a main subcluster was formed, which accounted for about half of all aggregating mass, while the rest of the particles were aggregated in subclusters of widely varying sizes, including single nonaggregated monomers. Closer to the end of aggregation, one large subcluster and a number of small subclusters were left. This aggregation pattern suggests that the hierarchical model [9,13], which allows only clusters of the same size to collide, is, in principle, different from the ''pure'' aggregation algorithm. A two-dimensional projection of a typical cluster with $N = 15\,000$ is shown in Fig. 1.

B. Numerical calculation of correlation functions

For finite lattice clusters, both p_2 and p_4 are, strictly speaking, highly singular functions. However, if the correlation function is used for calculation of some average values $\langle F \rangle$ according to

$$\langle F \rangle = \int_0^\infty F(r)p_m(r)dr, \tag{78}$$

and the function $F(r)$ changes slowly enough, we can replace the exact function p_m by some ''smoothed'' function according to

$$p_m(r) \to \frac{1}{\delta}\int_{r-\delta/2}^{r+\delta/2} p_m(r')dr'. \tag{79}$$

Below, we will use the notation p_m for the smoothed function.

The natural choice for the constant δ in Eq. (79) is the lattice unit. If we choose δ to be less than the lattice unit, the

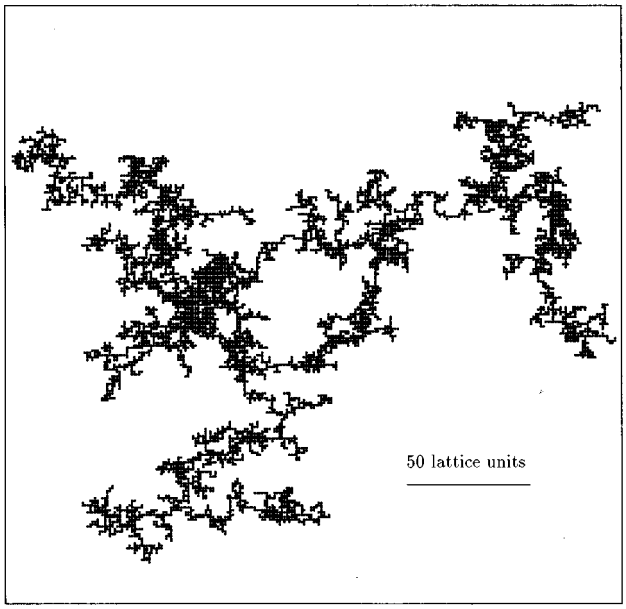


FIG. 1. Two-dimensional projection of a typical cluster with $N=15\,000$.

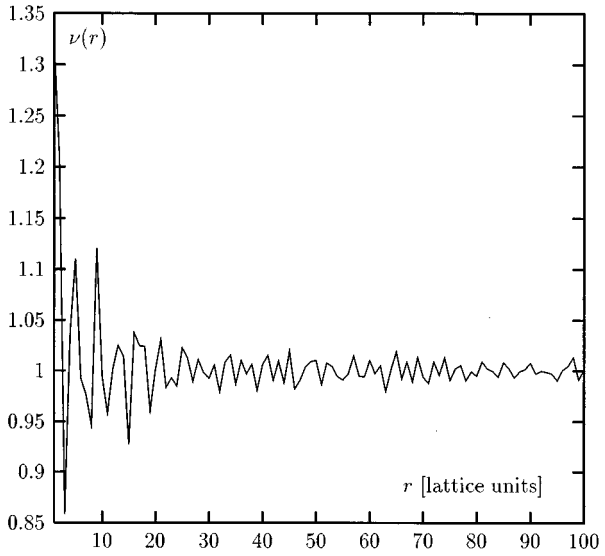


FIG. 2. Lattice density function $\nu(r)$.

resultant function p_m fluctuates strongly at small distances, and if we choose δ to be larger than the lattice unit, we lose precision at large distances.

We have set the lattice unit and the value of δ to be equal to unity. We calculated the function $p_2(r)$ for integer values of r beginning with $r=1$ [$p_2(0)=0$ by definition] by enumeration of all possible pairs of particles in each cluster. For calculation of $p_4(r)$, it was computationally impossible to enumerate all possible combinations of indices i, j, k, and l, since the number of such combinations is proportional to N^4. Instead, we picked a fixed number of index combinations at random. To ensure statistical reliability of calculations, the number of random combinations was set to 2×10^8 for each cluster, irrespective of N. However, this number is still much smaller than the number of all possible combinations, even in the smallest clusters with $N=5000$.

Even after smoothing according to Eq. (79), the correlation functions of a lattice cluster possess some random irregularities, which are more pronounced at small distances and may seem to be random but are, in essence, artifacts of the lattice on which the cluster was built. The origin of these irregularities is that the density of sites of the lattice itself (measured at a certain distance from the origin and averaged over angles) has certain fluctuations which disappear at large distances, where the discrete structure of the lattice is no longer of importance. To eliminate these irregularities in the correlation function, we suggest using the following procedure: First, we define the density of lattice sites $\nu(r)$ as

$$\nu(r)=\Delta N_{\mathrm{latt}}(r)/V(r), \tag{80}$$

$$V(r)=\frac{4\pi}{3}[(r+1/2)^3-(r-1/2)^3], \tag{81}$$

where the variable r takes integer values beginning from $r=1$, and $\Delta N_{\mathrm{latt}}(r)$ is the number of lattice sites which lie in the spherical shell $r-1/2<r'<r+1/2$ with the center at the origin. Evidently, this function becomes close to unity as r grows, but for small r it can significantly differ from unity (see Fig. 2). Then we define the corrected correlation function as

$$\varrho_m(r)=p_m(r)/\nu(r). \tag{82}$$

Since $\nu(r)$ looks much like a statistical noise and is unity on average (see Fig. 2), we can assume that $\int_0^\infty \varrho_m dr = \int_0^\infty p_m(r)dr$, which means that normalization is conserved by procedure (82). In Fig. 3 we compare the original function $p_2(r)$ and the modified function $\varrho_2(r)$ for an ensemble of 40 clusters with $N=10\,000$.

VI. NUMERICAL CALCULATIONS: TWO-POINT CORRELATION FUNCTION

A. Small-r asymptote and fractal dimension

The small-r asymptote of $p_2(r,N)$ is illustrated in Fig. 4, where the lattice-corrected function $\varrho_2(r,N)$ is plotted for different values of N. We also show the theoretical asymptote (13) in this figure. The numerical values of constants a and D_1 were found from the linear regression. Table I shows the results for a and D_1 for different values of N. In the third column of the table, we specify the range of r where the linear regression was used. This range was determined for each N from the condition that a double-log plot of the function $\varrho_2(r,N)$ demonstrates no visually apparent deviation from a straight line, apart from random fluctuations. For this purpose, a plot was made in a square frame for each N, starting from the minimum value of r and ending with the maximum value, and the vertical range of the plot was chosen from the condition that the line started in the lower left corner of the frame and ended in the upper right corner (these graphs are not shown). This method of selecting a range for r is, of course, not precise, since it relies on visual interpretation. For example, the ranges for $N=10\,000$ and

TABLE I. Constants a and D_1, determined by linear regression.

$N\times10^{-3}$	R_g	Interval of r	Using p_2		Using ϱ_2	
			a	D_1	a	D_1
5	49.7±1.1	[3,15]	4.3±0.5	1.74±0.04	4.07±0.06	1.771±0.006
10	70.3±1.4	[3,30]	4.2±0.2	1.77±0.01	4.16±0.05	1.776±0.004
15	84.1±1.6	[3,30]	3.9±0.2	1.80±0.01	3.89±0.01	1.804±0.001
20	100.2±2.3	[3,40]	3.8±0.1	1.82±0.01	3.75±0.02	1.822±0.002

15 000 appeared to be the same. However, for larger values of N it was possible to choose a range for r with stronger confidence.

The statistical errors shown in Table I are at the level of one standard deviation. These errors represent only the uncertainties associated with the linear regression procedure; they do not reflect any errors associated with the fact that the number of clusters was finite. However, we believe that the latter factor is insignificant, since the small-r asymptote was reproduced with a very high precision for different ensembles of clusters.

Now we return to the discussion of numerical results demonstrated in Fig. 4 and Table I. First, we see that the universal asymptote (13) describes the behavior of $p_2(r,N)$ with high precision when $r\ll R_g$. This precision is even more apparent if we consider a figure analogous to Fig. 4, but showing $Np_2(r,N)$ instead of $N\varrho_2(r,N)$ and with a strong magnification of the region $r<20$ (not shown). In such a figure, we would see that the points $Np_2(r,N)$ for $r=\text{const}$, and different values of N coincide with a precision which by far exceeds the random-lattice-related fluctuations.

On the other hand, the data of Table I suggest that there is a weak but systematic dependence of the constants a and D_1 on N. This fact contradicts, in principle, the idea of a universal small-r asymptote. However, it has a simple explanation: If N is not large enough, the interval of r, in which the universal asymptote is valid shrinks. Though we tried to use values of r which are much smaller than R_g in our calculations, we did not know how strong this inequality should be in practice. Due to computational limitations, we were not able to make this inequality as strong, for example, as two orders of magnitude. The radius of gyration was only ~50 for $N=5000$ and ~100 for $N=20\,000$. On the other hand, it is possible that if N is not large enough, and the strong inequality $1\ll r\ll R_g$ is not achievable, the interval where the universal asymptote is valid shrinks to a point. This would mean that the function $N\varrho_2(r,N)$ plotted in a double-log scale as a function of r would always deviate from a straight line, curving toward the lower right corner of the frame, though this tendency might not be visible to the eye.

We believe that the best values of a and D_1 are those for $N=15\,000$. The data for $N=20\,000$ cannot be so precise, because, due to computational limitations, we were not able to make the lattice size L appropriately large (it should have been larger than 400, but we were able to use $L=350$). In addition, the data for $N=15\,000$ have the best precision. The values of a and D_1 found from the linear regression of $\varrho(r,N=15\,000)$ were used for the theoretical line shown in Fig. 4. Below, we will adopt the numerical values $a=3.9$ and $D_1=1.8$. Another result which can be seen from Table I is that the use of the lattice-modified function $\varrho_2(r,N)$ allows us to increase the precision of linear regression coefficients, especially for smaller values of N.

Now we turn to the dependence $R_g(N)$ and to determining the constants R_0 and D_2 (see Fig. 5). We have found that $R_0=0.61\pm0.08$ and $D_2=1.94\pm0.06$. Again, the errors are shown at the level of one standard deviation. We see that the difference between D_1 and D_2 is more than two standard deviations of D_2 (the uncertainty of D_1 is too small to be considered). This is strong evidence in favor of multiscaling.

Now we can calculate the value of the important constant $aR_0^{D_1}$:

$$aR_0^{D_1}\approx1.6. \tag{83}$$

B. Cutoff functions for p_2

First, we consider the GE and OS cutoff functions. We start the discussion from the nonmultiscaling perspective. As

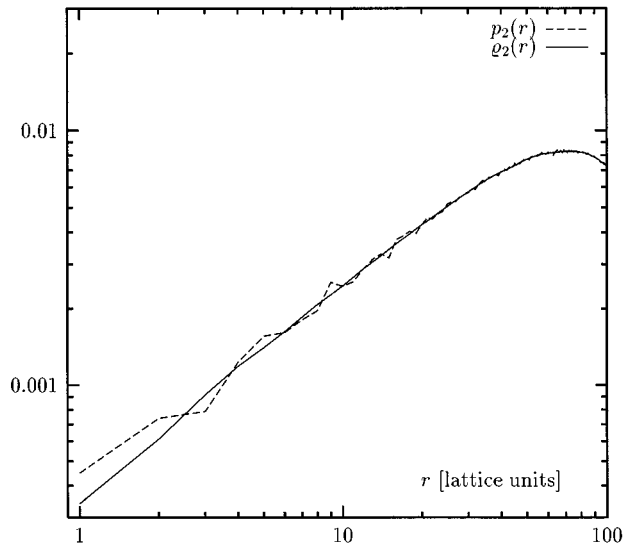


FIG. 3. Comparison of the original $p_2(r)$ and corrected $\varrho_2(r)$ correlation functions for $N=10\,000$.

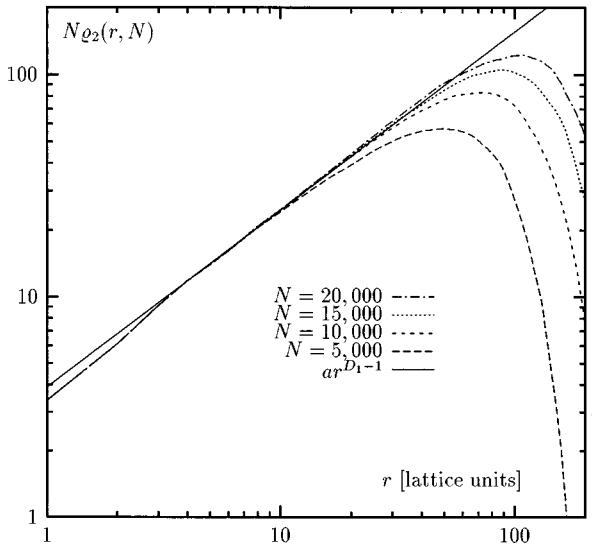


FIG. 4. Numerically calculated function $N\varrho_2(R,N)$ for different N and the theoretical asymptote (13) with $a=3.9$ and $D=1.8$.

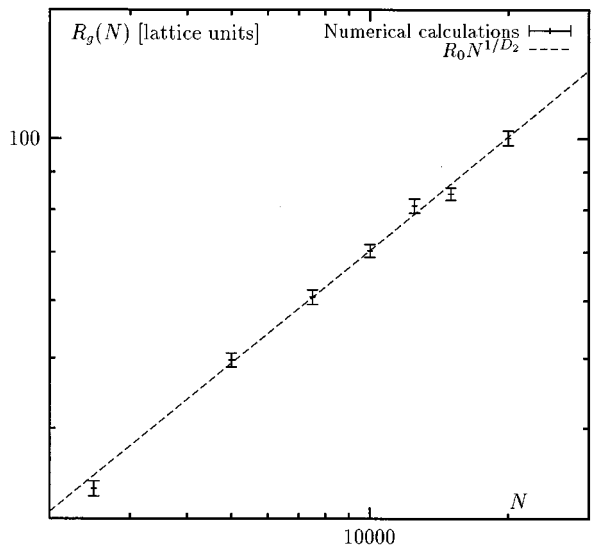


FIG. 5. Radius of gyration as a function of N and theoretical formula (48) with $R_0=0.61$ and $D_2=1.94$.

shown in Sec. III A 2, the condition of applicability of the GE model is $F^{\mathrm{GE}}(\infty,D_1)<aR_0^{D_1}<F^{\mathrm{GE}}(1,D_1)$, where $F^{\mathrm{GE}}(\infty,D_1)$ and $F^{\mathrm{GE}}(1,D_1)$ are given by Eqs. (31) and (32). Using the value of D_1 found in Sec. VI A, we find that $F^{\mathrm{GE}}(\infty,D_1)\approx 0.49$ and $F^{\mathrm{GE}}(1,D_1)\approx 2.47$. We see that the above inequality is fulfilled for the GE cutoff.

For the OS cutoff, the condition is $F^{\mathrm{OS}}(x_{1,\min},D_1)<aR_0^{D_1}<F^{\mathrm{OS}}(x_{1,\max},D_1)$ (see Sec. III A 3). Using Eqs. (42) and (46), we find that $F^{\mathrm{OS}}(x_{1,\min},D_1)=0.75$ and $F^{\mathrm{OS}}(x_{1,\max})=1.29$. Therefore, the condition for the OS cutoff is not fulfilled. This means that we cannot find a set of parameters for the OS cutoff function that would correctly reproduce the gyration radius and the small-r asymptote simultaneously.

The above consideration was based on the assumption that $D_2=D_1$. If this is not the case, we must replace $aR_0^{D_1}$ by $aR_0^{D_1}N^{D_1/D_2-1}$ in the corresponding inequalities. Since $D_1<D_2$, the inequality for GE cutoff is fulfilled only for $N<N_{\max}$, where $N_{\max}=[aR_0^{D_1}/F^{\mathrm{GE}}(\infty,D_1)]^{D_2/(D_2-D_1)}]\approx 1.3\times 10^7$. Thus we see that the GE cutoff can be used in the case of multiscaling if $N\ll N_{\max}$. For the values of N used here, this inequality is well fulfilled.

For the OS model we have a contrary situation. The condition $F^{\mathrm{OS}}(x_{1,\min},D_1)<aR_0^{D_1}N^{D_1/D_2-1}<F^{\mathrm{OS}}(x_{1,\max},D_1)$ is not fulfilled for $N\sim 1$, but becomes fulfilled for $N_{\min}<N<N_{\max}$, where $N_{\min}=[aR_0^{D_1}/F^{\mathrm{OS}}(x_{1,\max},D_1)]^{D_2/(D_2-D_1)}\approx 20$ and $N_{\max}=[aR_0^{D_1}/F^{\mathrm{OS}}(x_{1,\min},D_1)]^{D_2/(D_2-D_1)}\approx 36\,000$. As a result, in the case of multiscaling, the conditions for the OS and GE models are fulfilled for the values of N studied in this paper (from 5000 to 20 000).

The GE and OS cutoff functions for $N=15\,000$ are shown in Fig. 6. First, we discuss the GE curve. The values of α and β found from the numerical solution to Eqs. (53) and (54) are equal to 0.273 and 2.489, respectively. It is interesting to note that the values of α and β obtained from the best fit to the experimental curve (minimizing χ^2), which is depicted in Fig. 6 by circles, came out very close to the above numbers: 0.276 ± 0.002 and 2.49 ± 0.01, respectively. In principle, the conditions of minimizing χ^2 and preserving the second moment of the distribution [Eqs. (53) and (54)) are different in their nature, and should not lead necessarily to the same results. In our case, it turned out that the solution to Eqs. (53) and (54) provides at the same time the best fit to the experimental cutoff curve within the framework of the two-parameter GE cutoff model.

The GE cutoff functions for two different values of N are shown in Fig. 7. The corresponding values of α and β for $N=5000$ are 0.344 and 2.238 [from numerical solutions to Eqs. (53) and (54)] and 0.356 ± 0.003 and 2.18 ± 0.02 (from the best fit). The analogous data for $N=15\,000$ are given above. Note that $\beta(N)$ is, indeed, an increasing function of N, as shown theoretically in Sec. III B 2, for the case of $D_1<D_2$. A systematic dependence of $g_2(x,N)$ on N which is

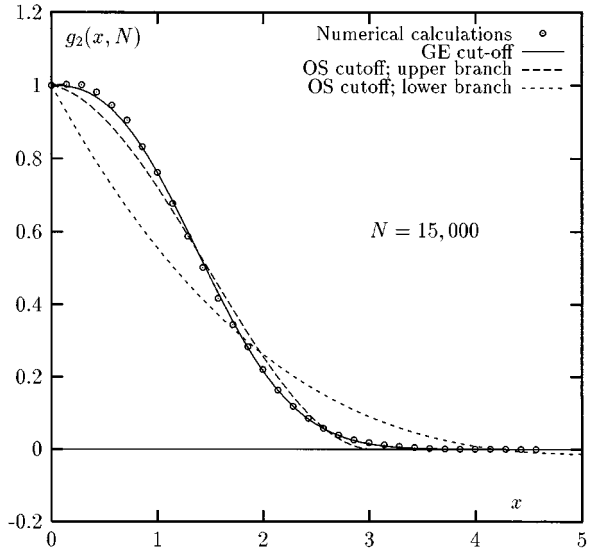


FIG. 6. Comparison of GE and OS cutoff functions for $N=15\,000$.

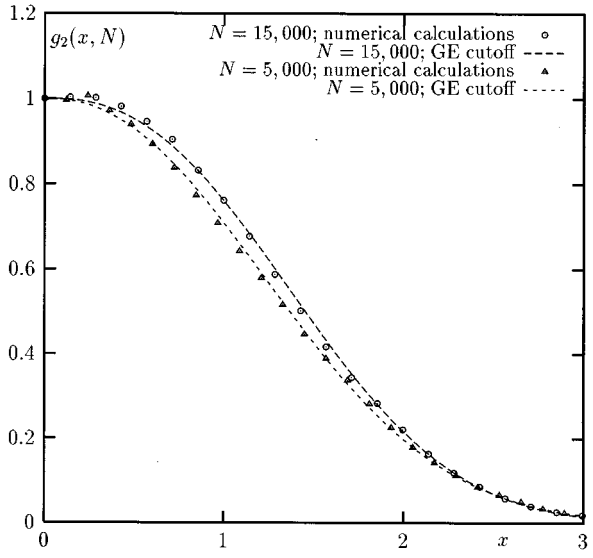


FIG. 7. GE cutoff functions for different N.

apparent from Fig. 7 confirms the multiscaling nature of the two-point correlation function.

Theoretically, the GE cutoff functions with $D_1<D_2$ violate the irreversibility principle. More specifically, inequality (55) is violated for all values of r larger than some critical value. But in practice, such a violation happens for extremely large values of r. This is illustrated in Fig. 8, where we plot both sides of inequality (55) as functions of r, with $N_1=15\,000$ and $N_2=5000$. The intersection of the curves occurs at $r=[\alpha(N_2)R_g^{\beta(N_1)}(N_1)/\alpha(N_1)R_g^{\beta(N_2)}(N_2)]^{1/[\beta(N_1)-\beta(N_2)]}$. Given the above numerical values of N_1, N_2, α, and β, the critical value of r becomes equal to $\sim 2.3\times 10^4$, which is more than two orders of magnitude larger than the gyration radius for $N_1=15\,000$. Evidently, the experimental function $p_2(r)$ is exactly equal to zero for such a large r, and the GE cutoff function is

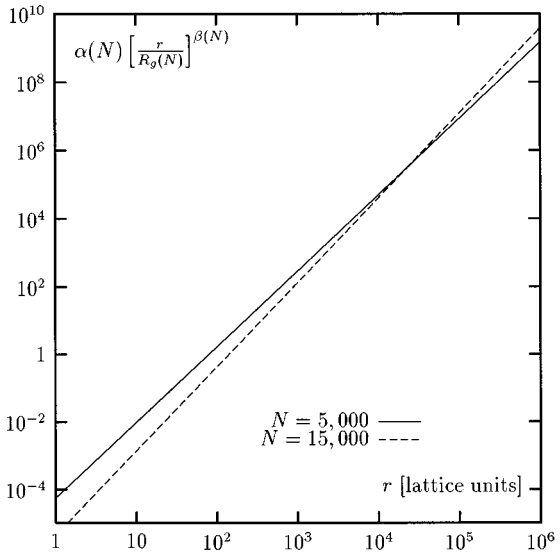


FIG. 8. Illustration of violation of the irreversibility principle by the GE cutoff functions [see inequality (55)].

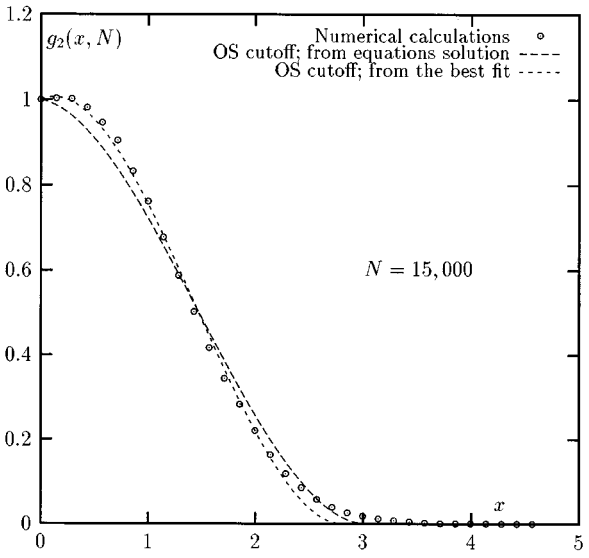


FIG. 9. OS cutoff functions with coefficients determined from Eqs. (49) and (50) and from the best fit, compared to the numerical data.

negligibly small. As a result, the irreversibility principle is satisfied by the GE cutoff for all physically reasonable values of r.

Now we turn to the OS cutoff. The values of parameters x_1 and x_2 found from the numerical solution to Eqs. (39) and (40) are equal to 3.032 and 0.825, respectively, for the upper branch ($x_2>0$) and 6.831 and -2.103 for the lower branch ($x_2<0$). First, as seen in Fig. 6, the lower branch (with $x_2<0$) provides a very poor fit to the experimental curve. Therefore, we will not discuss the lower branch here. The upper branch provides a fair fit, but not as good as the GE model. The constants x_1 and x_2 found from the best fit to the experimental curve are 2.81 ± 0.02 and 0.61 ± 0.02, respectively, which are close to the corresponding values shown above (but not as close as we had it for the parameters α and β in the GE model). The difference between the two OS cutoff functions [one with constants found from the solution to Eqs. (39) and (40), and the other with constants found from the best fit] is shown in Fig. 9. We see that the second set of constants provides a better fit to the experimental curve for $x<2$; however, using these constants would result in a wrong value of R_g.

If one is not interested in the small-q asymptote of the optical form factor $\phi_2(q)$, Eq. (5), it can be appropriate to use the OS cutoff with the constants x_1 and x_2 found from the best fit. The advantage of using the OS cutoff function is that it allows us to take integral (5) in terms of elementary functions.

As shown in Sec. III B 3, selection of the upper branch of the solution for $x_1(N)$ can lead to violation of the irreversibility principle for sufficiently large N, if $D_2-1<D_1<D_2$. Though D_1 actually lies in this interval, the values of N used in this paper are not large enough for the violation to take place. To illustrate this, we plotted the density functions $Np_2(r,N)$ for different values of N, using the OS cutoff with the parameters x_1 and x_2 found both from Eqs. (39) and (40)

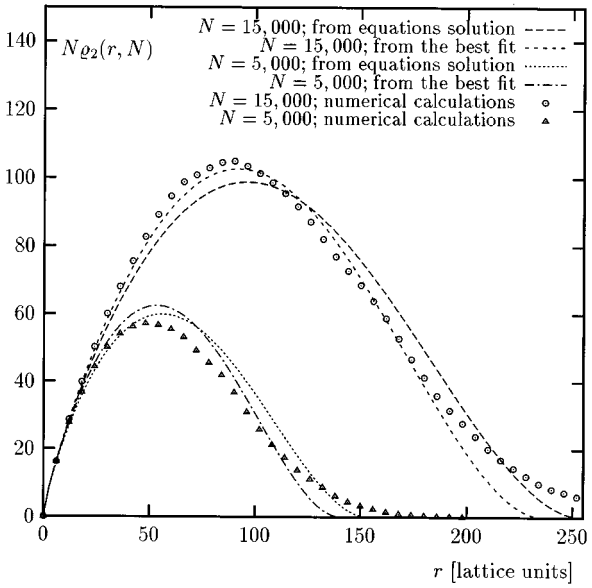


FIG. 10. Numerical two-point correlation functions for different N compared to the OS model functions with coefficients determined from Eqs. (49) and (50) and from the best fit.

and from the best fit. We see that the irreversibility principle is not violated in both cases. We also show corresponding numerical curves in Fig. 10, which allows one to visualize the quality of fit of the actual correlation functions (rather than the cutoff functions) in the OS model. From comparing the curves in Fig. 10, a conclusion can be made that the quality of fit increases with N.

Analogous data for the GE model are shown in Fig. 11. In this figure, we do not show the curves with parameters determined from the best fit, since they are practically indistinguishable from the curves with parameters found from solution to Eqs. (53) and (54). In general, we can state that the quality of fit is much better in the GE model. This fact supports the experimental data of [4,19], where the cutoff function was found to be close to Gaussian ($\beta=2$). For example, the value of β found in this paper was ≈ 2.2 for $N=5000$.

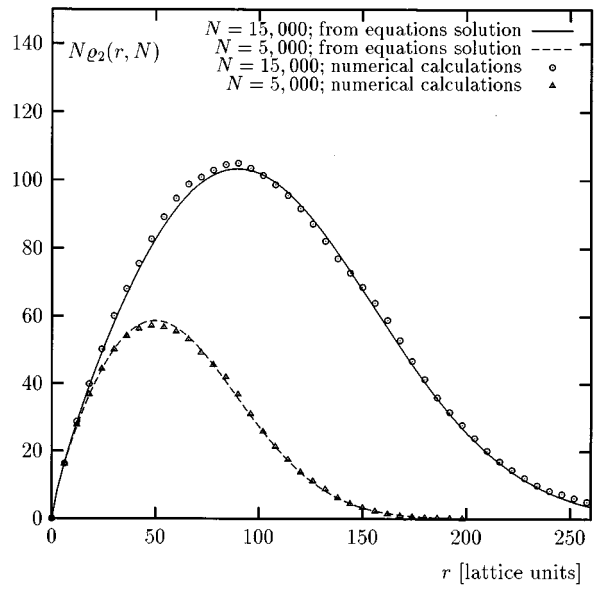


FIG. 11. The same as Fig. 10, but for the GE model function. The coefficients determined from Eqs. (49) and (50) and from the best fit practically coincide.

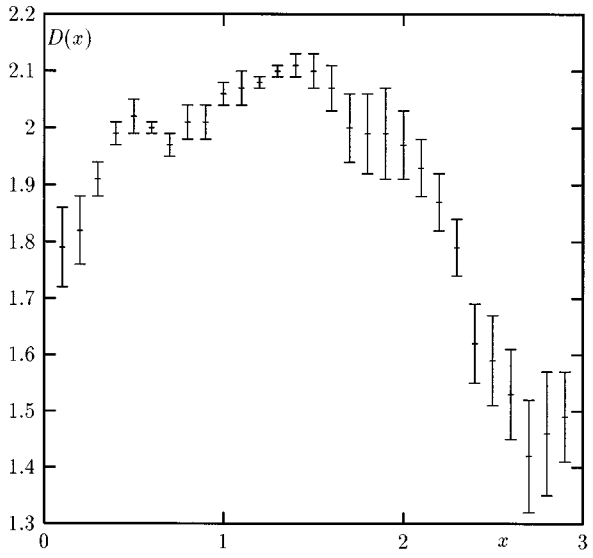


FIG. 12. Function $D(x)$ calculated numerically with uncertainties of linear regression at the level of one standard deviation.

Finally, we briefly discuss the CFD model. It was shown in Sec. III B 4 that this model is incompatible with the irreversibility principle if $D_1<D_2$. However, the question of how large the value of r/R_g should be for the violation of the ireversibility principle to take place is open. As we saw for the example of the GE cutoff, this violation can happen for very large values of r/R_g, which is not an essential drawback of the model.

Technically, we can calculate the function $D(x)$ using ensembles of clusters with different N by considering the function $p_2[r,N(r)]$, where $N(r)$ is determined from $r/R_g(N)=x=\text{const}$. If this function, plotted on a double-log scale, looks like a straight line, we can conclude that the CFD model is correct. A mathematical measure of deviation of a curve from a straight line is the uncertainty of linear regression coefficients. If this uncertainty is comparable to the coefficients themselves, the curve cannot be considered as a straight line. In Fig. 12 we show the function $D(x)$, calculated as a coefficient of linear regression, together with its uncertainty. Though it is difficult to make a definitive conclusion, the uncertainties of $D(x)$ are fairly high, at the level of 10% on average. For comparison, we would obtain this level of uncertainty, if we try to approximate the function $y(x)=x^2$ by a linear function in the interval $x\in[0,10]$.

VII. NUMERICAL CALCULATIONS: FOUR-POINT CORRELATION FUNCTION

The small-r behavior of $p_4(r,N)$ is illustrated in Fig. 13. In this figure we plotted numerical curves $\varrho_4(r,N)$ for different N along with the theoretical asymptotes (72). The coefficients c_1 and c_2, defined by Eqs. (72), were found from the best fit and are shown in Table II.

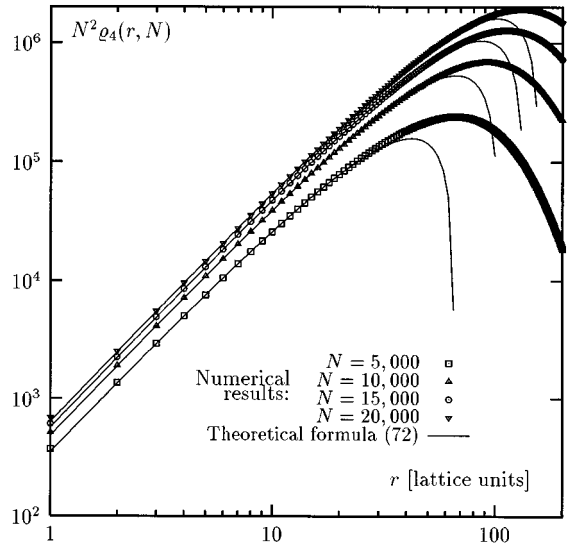


FIG. 13. Small-r asymptote of $p_4(r,N)$. Numerical data are compared to the theoretical formula (72).

To obtain the data of Table II, the interval of r was selected as follows: First, an arbitrary interval was taken, starting at $r=1$ and ending approximately near the maximum of $p_4(r)$. Then the coefficients c_1 and c_2 were calculated by minimizing χ^2 in this interval. Next, the right bound of the interval was decreased gradually, with the values of c_1 and c_2 calculated at each step, until they became constants (independent of the right bound) within the precision of calculation. It can be seen from Fig. 13 that the theoretical asymptote (72) is reproduced with great precision. We emphasize that both terms in (72) are important for a good fit. However, the relative significance of the second term decreases with N. We can also conclude that $N^2p_4(r,N)$ is not a universal function (i.e., independent of N) when $r \ll R_g$, as is $Np_2(r,N)$. Instead, there exists a universal limit: $\lim_{r/R_g\to 0}\lim_{N\to\infty}N^2p_4(r,N)=c_1r^2/(2D_1-3)$. The condition $N\to\infty$ is rather strong, as seen from the numerical results. We can see from Fig. 13 that the curves for $N=15\,000$ and $20\,000$ are still distinguishable. The weak dependence of the coefficients c_1 and c_2 on N can be viewed as a result of multiscaling. However, it is important to emphasize that the dependence of c_1 and c_2 on N is very weak in Table II. We expect this dependence to be much stronger for any expansion in terms of r/R_g other than Eq. (72).

TABLE II. Coefficients c_1 and c_2 defined in Eq. (72) found from the best fit.

$N\times 10^{-3}$	Interval of r	c_1	c_2
5	[1,15]	13.45±0.02	18.72±0.06
10	[1,30]	14.40±0.01	19.18±0.02
15	[1,40]	14.61±0.01	19.04±0.02
20	[1,50]	14.756±0.004	19.21±0.01

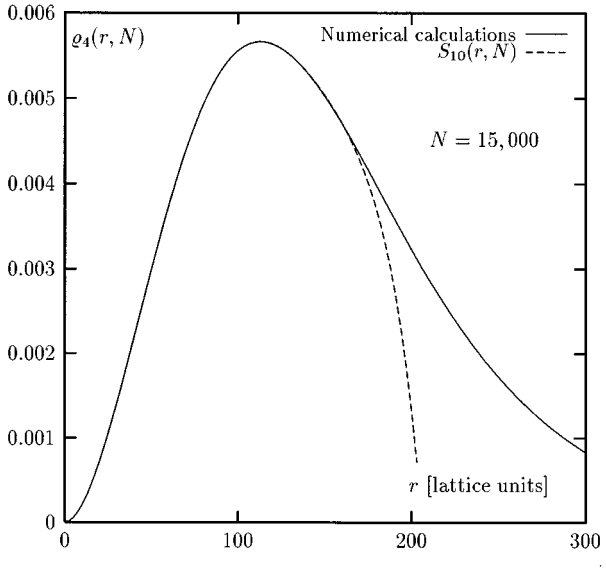


FIG. 14. Comparison of the numerically calculated function $p_4(r,N)$ and expansion (84) for $N=15\,000$ and $n=10$. The coefficients $c_1,\ldots,c_n$ are found from the best fit (minimization of χ^2).

In order to obtain a better fit of p_4 for larger values of r and, in particular, in the area where $r\sim R_g$, one can generalize Eq. (73) and approximate $p_4(r,N)$ by sums of the form

$$S_n(r,N)=\frac{br^2}{N^{3/D}}\sum_{k=0}^{n}(-1)^k c_k\left(\frac{r}{R_g}\right)^{k(2D-3)}, \qquad (84)$$

where $c_0=1$ and c_k are coefficients different from the coefficients c_1 and c_2 used in Table II above. We emphasize that this particular expansion of p_4 is anticipated to have universal (independent of N) coefficients c_k, apart from a possible weak multiscaling dependence, as demonstrated in Table I for c_1 and c_2. Any expansion of p_4 in terms of other powers of r/R_g would have coefficients which depend strongly on N and, therefore, are not useful in practice. Also, expansion (84) is, in principle, different from the Taylor expansion, since the first derivative of the ratio of its first two terms is infinite. As a result, an integer power polynomial in r/R_g will not provide a good fit to p_4. For the same reason, we cannot represent p_4 as a simple power function with a cutoff.

An example of function S_n, Eq. (84) with $n=10$ is shown in Fig. 14. The coefficients $c_1,\ldots,c_{10}$ were found from the best fit. Practically, the best-fit algorithm was performed as follows: We calculated four coefficients at a time. The first two of them were known from Table II. The next four coefficients were calculated in the interval of r [0,120] by minimizing χ^2. Then the values of coefficients $c_1,\ldots,c_6$ were considered to be fixed, and we calculated the remaining coefficients in the same interval of r. Figure 14 demonstrates an excellent fit up to $r=160$. Higher n can be used if it is desirable to fit p_4 at larger values of r.

Though it is tempting to use a GE function for p_4, e.g.,

$$p_4(r,N)=\frac{br^2}{N^{3/D_1}}\exp[-\alpha(r/R_g)^{2D_1-3}]; \quad (85)$$

the coefficients c_k calculated using the best fit do not support this choice. The above function has only one adjustable parameter. Unlike the GE functions studied in Secs. III A 2 and III B 2, function (85) has a fixed power of r/R_g in the exponent.

VIII. SUMMARY AND DISCUSSION

The correlation functions p_2 and p_4 studied in this paper are important for an interpretation of scattering experiments when the geometry of scattering objects is similar to that of cluster-cluster aggregates, and the first Born approximation can be used. In particular, the two-point correlation function p_2 defines the angular dependence of the average scattering intensity according to Eqs. (3)–(5). In order to describe fluctuations of the scattered intensity, one needs to employ, in addition, the four-point correlation function p_4 [see formulas (8), (10), and (11)]. The optical form factors ϕ_2 and ϕ_4 defined by Eqs. (4) and (10) are experimentally measurable. It is important that the correlation functions, which carry information about the geometry of scattering objects, can always be calculated by the Fourier transformation if the form factors are known (this constitutes the solution to the so-called inverse problem). On the other hand, if the correlation functions are known from some other type of experiments, or from numerical simulations, one can predict the results of scattering experiments (the direct problem) by evaluating integrals (5) and (10).

We start the summary of our results from the two-point correlation function. The following results were obtained in our numerical calculations:

(i) The cluster-cluster aggregates (CCA's) demonstrate a pronounced multiscaling in the two-point correlation function with $D_1<D_2$.

(ii) The generalized exponential (GE) cutoff function describes the numerical two-point correlation function better than the other models considered in this paper.

(iii) Nevertheless, the overlapping sphere (OS) cutoff can be used in some instances, if it is desirable to obtain a closed expression for the optical form factor $\phi_2(q)$ in terms of elementary functions.

Though the GE cutoff provides an excellent fit for $N\sim10^4$, it may not work for larger N. We estimated that the critical value of N is 10^7. This number is probably much larger than any practical value for clusters of nanoparticles, but cannot be so large for atomic and molecular clusters. Another drawback of the GE cutoff is that it does not allow one to calculate the optical form factor $\phi_2(q)$, Eq. (5), in terms of elementary, or, at least, commonly used special functions. From the practical point of view, this is not really important, since numerical integration is always available. But in some instances, and especially if we want to extract some expression for $\phi_2(q)$ in the intermediate region of q, an analytical expression may be desirable. In this case the OS cutoff can be used.

Considering the four-point correlation function, we can conclude the following:

(i) The theoretical small-r asymptote (72), though obtained from a wrong assumption, is confirmed by the numerical data with a high precision.

(ii) The four-point correlation function is not described by a simple scaling behavior with a cutoff function, similar to Eqs. (23) and (47). Instead, it can be approximated by a sum of noninteger powers of r/R_g of the form Eq. (84). As a result, p_4, unlike p_2, does not have a universal (N independent) small-r asymptote (compare Figs. 4 for p_2 and 13 for p_4).

(iii) The four-point correlation function can be approximated with high precision by sum (84) up to some maximum value of r; this value of r grows with the number of terms in Eq. (84), n; for sufficiently large n, the complete function $p_4(r,N)$ can be approximated. The coefficients of this expansion are expected to be universal (N independent), apart from a possible weak multiscaling dependence. This was confirmed numerically for the first two coefficients (see Table II).

It should be noted that for the practical purposes [calculation of the form factor $\phi_4(r)$] integration according to Eq. (10) should be carried out only up to the maximum point of r for which the approximation works. It is clear that an integral, Eq. (10), of each individual term in Eq. (84) diverges at the upper limit.

Though expansion (84) proved to be useful, we cannot rule out other possible functional forms for p_4. In particular, we can consider a model where p_4 is given by a formula similar to Eq. (72) [or Eq. (74) in the case of multiscaling] where each term is multiplied by its own cutoff function, which is expandable in the Taylors series, while the sum of the two terms is not. Another possibility is to assume that the first two terms of the expansion of p_4 coincide with Eq. (72) or (74), but the higher terms contain only integer powers of r. At last, our numerical data were insufficient for quantitative verification of multiscaling behavior of the coefficients c_k in expansion (84).

ACKNOWLEDGMENTS

This research was supported in part by the U.S. Environmental Protection Agency under Grant No. R822658-01-0, and by the National Science Foundation under Grant No. DMR-9500258.

APPENDIX A: MULTISCALING WITH $D_1=D_2$

To illustrate that multiscaling can exist even if $D_1=D_2$, consider Eqs. (49) and (50) in the case of $D_1=D_2=D$:

$$aR_0^D\int_0^\infty x^{D-1}g_2(x,N)dx=1, \quad (A1)$$

$$aR_0^D\int_0^\infty x^{D+1}g_2(x,N)dx=2. \quad (A2)$$

An evident example of a function $g_2(x,N)$ which depends explicitly on N, but satisfies at the same time Eqs. (A1) and (A2) for any N, is a function with three adjustable parameters. Let one of the adjustable parameters be an arbitrary function of N. Then the other two parameters can be found

from Eqs. (A1) and (A2) for any value of N (provided all the functions involved are mathematically ''good'' and the solution exists).

APPENDIX B: IRREVERSIBILITY PRINCIPLE IN THE OS MODEL FOR LARGE VALUES OF N

Mathematically, the irreversibility principle can be formulated as $\partial g_2[r/R_g(N),N]/\partial N>0$ for any r. We should add to this a condition for positive definition of p_2 and, consequently, g_2: $g_2[r/R_g(N),N]>0$. The solution of these two inequalities requires an explicit expression for $x_1(N)$, $x_2(N)$, which, in general, cannot be obtained analytically from Eq. (58). However, we can find approximate solutions to Eq. (58) with F^{OS} defined by Eq. (40), in the limiting case of large N and $D_1<D_2$.

First, we reformulate the irreversibility principle in a more suitable way. From the general form of polynomial (57), we notice that the inequality $\partial g_2/\partial N>0$ (for any r) can be equivalently replaced by the following two inequalities: $\partial(R_g x_1)/\partial N>0$ and $\partial(\partial g_2/\partial r|_{r=0})/\partial N>0$. If x_1 and x_2 do not depend on N, the above inequalities lead to the condition on x_1, Eq. (44), obtained in Sec. III A 3.

Now we analyze the function $g_2(r/R_g,N)$ in the limit of large N and $D_1<D_2$. As mentioned above, we have two possible branches for the solutions to Eq. (58): $x_1(N)$ and $x_2(N)$. The first (lower) branch is

$$\lim_{N\to\infty} x_1(N)=x_{1,0}+\delta(N), \tag{B1}$$

$$\lim_{N\to\infty} x_2(N)=\frac{(D_1+2)x_{1,0}^2\delta(N)}{(D_1+3)x_{1,\max}^2-(D_1+5)x_{1,0}^2}, \tag{B2}$$

$$\delta(N)=\frac{6aR_0^{D_1}N^{D_1/D_2-1}x_{1,0}^{D_1+2}}{D_1(D_1+1)(D_1+2)^2(D_1+3)}\ll 1. \tag{B3}$$

Substituting these expressions into Eq. (57), we find that

$$\lim_{N\to\infty}\left.\frac{\partial g_2}{\partial r}\right|_{r=0}=\frac{1}{R_g}\left[\frac{1}{2x_2}-\frac{2}{x_1}\right]\propto N^{1-(D_1+1)/D_2}. \tag{B4}$$

This is an increasing function of N only if $D_1<D_2-1$. The latter condition guarantees that the irreversibility principle is satisfied for the lower branch of $x_1(N)$. If $D_2-1<D_1<D_2$, the curves $p_2(r,N)$ intersect for sufficiently large N. In the latter case, if we consider the *exact* OS solution for $p_2(r,N)$ and gradually increase N, we find that for a certain value of N the derivative $\partial(\partial g_2/\partial r|_{r=0})/\partial N$ becomes equal to zero. A curve with a slightly larger N intersects the previous curve at some value of r close to zero. As N grows, the point of intersection moves further to the right from $r=0$, and the violation of the irreversibility principle becomes more apparent.

Now we turn to the upper branch of $x_1(N)$:

$$\lim_{N\to\infty} x_1(N)=\left[\frac{D_1(D_1+1)(D_1+2)^2(D_1+3)}{12aR_0^{D_1}N^{D_1/D_2-1}}\right]^{1/D_1}, \tag{B5}$$

$$\lim_{N\to\infty} x_2(N)=-\frac{D_1+2}{2(D_1+5)}x_1(N). \tag{B6}$$

For the derivative of g_2, we have

$$\lim_{N\to\infty}\left.\frac{\partial g_2}{\partial r}\right|_{r=0}\propto -N^{-1/D_1}. \tag{B7}$$

It follows from Eq. (B7) that the condition $\partial(\partial g_2/\partial r|_{r=0})/\partial N>0$ is met. But, since $x_2<0$, we must also verify that g_2 is a positively defined function. Evidently, this is not so. This function with the constants x_1 and x_2 defined by Eqs. (B5) and (B6) crosses the horizontal axis at $r=R_g x_1(D_1+2)/(D_1+5)<R_g x_1$. However, the minimum value of g_2 is $-2/[(D_1+2)(D_1+5)^2]$, which can be very close to zero (~ -0.01 for $D_1=2$, for example). In addition, $R_g x_1$ grows faster with N than R_g and, for sufficiently large N, the correlation function p_2 acquires a long tail, which extends to values of r far exceeding R_g. The violation of positive definition of p_2 occurs in this tail where $r\gg R_g$, and for some practical applications can be ignored. If, however, negative values of p_2 are not acceptable, one can redefine the OS cutoff function and truncate it at the point of its intersection with the horizontal axis.

[1] T. A. Witten and L. M. Sander, Phys. Rev. Lett. **47**, 1400 (1981).
[2] P. Meakin, Phys. Rev. Lett. **51**, 1119 (1983).
[3] M. Kolb, R. Botet, and R. Jullien, Phys. Rev. Lett. **51**, 1123 (1983).
[4] J. Cai, N. Lu, and C. M. Sorensen, J. Colloid Interf. Sci. **171**, 470 (1995).
[5] M. V. Berry and I. C. Percival, Opt. Acta **33**, 577 (1986).
[6] M. Carpineti, M. Giglio, and V. Degiorgio, Nuovo Cimento **16D**, 1243 (1994).
[7] D. Asnaghi, M. Carpineti, M. Giglio, and A. Vailati, Physica A **213**, 148 (1995).
[8] V. A. Markel, V. M. Shalaev, E. Y. Poliakov, and T. F. George, J. Opt. Soc. Am. A **14**, 60 (1997).
[9] R. Jullien, M. Kolb, and R. Botet, J. Phys. Lett. (Paris) **45**, L211 (1983).
[10] P. Meakin, Phys. Rev. A **29**, 997 (1984).
[11] P. Meakin, Phys. Lett. **107A**, 269 (1985).
[12] Yu. E. Danilova, A. I. Plekhanov, and V. P. Safonov, Physica A **1-4**, 61 (1992).
[13] R. Botet and R. Jullien, Phase Transitions **24-26**, 691 (1990).
[14] M. Y. Lin, R. Klein, H. M. Lindsay, D. A. Weitz, R. C. Ball, and P. Meakin, J. Colloid. Interf. Sci. **137**, 263 (1990).
[15] C. Amitrano, A. Coniglio, P. Meakin, and M. Zannetti, Phys. Rev. B **44**, 4974 (1991).

[16] M. Carpineti, M. Giglio, and V. Degiorgio, Phys. Rev. B **51**, 590 (1995).
[17] J. C. Earnshow and D. J. Robinson, Nuovo Cimento **16D**, 1141 (1994).
[18] J. C. Earnshow and D. J. Robinson, Physica A **214**, 23 (1995).
[19] N. G. Khlebtsov, Colloid J. **58**, 100 (1996).
[20] A. Coniglio and M. Zanetti, Physica A **163**, 325 (1990).
[21] V. A. Markel, V. M. Shalaev, E. Y. Poliakov, and T. F. George, in *Fractal Frontiers*, edited by M. M. Novak and T. G. Dewey (World Scientific, Singapore, 1997).