Weibull function to describe the cumulative size distribution of clumps formed by two-dimensional grains randomly arranged on a plane

M. Fanfoni, B. Bonanni, R. Martini, S. Addessi, C. Goletti, and A. Sgarlata
Dipartimento di Fisica, Università di Roma Tor Vergata, Via della Ricerca Scientifica 1, 00133 Roma, Italy

(Received 11 September 2023; accepted 18 March 2024; published 12 April 2024)

Many manifestations of natural processes give rise to interesting morphologies; it is all too easy to cite the corrugation of the Earth's surface or of planets in general. However, limiting ourselves to 2D cases, the morphology to which crystal growth gives rise is also intriguing. In particular, it is interesting to study some characteristics of the cluster projection in 2D, namely the study of the shapes of the speckles (fractal dimension of their rims) or the distribution of their areas. Recently, for instance, it has been shown that the size cumulative distribution function (*cdf*) of "voids" in a corrole film on Au(111) is well described by the well known Weibull distribution. The present article focuses on the *cdf* of cluster areas generated by numerical simulations: the clumps (clusters) are generated by overlapping grains (disks) whose germs (disk centers) are chosen randomly in a 2000 × 2000 square lattice. The obtained *cdf* of their areas is excellently fitted to the Weibull function in a given range of surface coverage. The same type of analysis is also performed for a fixed-time clump distribution in the case of Kolmogorov-Johnson-Mehl-Avrami (KJMA) kinetics. Again, a very good agreement with the Weibull function is obtained.

DOI: 10.1103/PhysRevE.109.044131

I. INTRODUCTION

The Weibull cumulative distribution function (*cdf*), or simply the Weibull function, [1–3]

$$W(x) = 1 - e^{-k\,x^n}, \tag{1}$$

i.e., a stretched exponential, which in probability theory texts is written $W(x) = 1 - \exp[-(\frac{x}{\lambda})^n]$, has a very broad application in many fields of human knowledge [4–11], including many aspects of the materials science. In particular it well describes the distribution of the particle size after a fragmentation process, for example in milling, graining, or crushing operations [12–17]. Precisely in this context, the scientists who first employed Eq. (1) were Rosin and Rammler [18]. In fact it is often referred to as the RRSB distribution after Rosin, Rammler, Sperling, and Bennet [18–21]; the last two scientists were involved in the study of the problem of determining the size distribution of the coal dust. For further information about the history of Eq. (1), see, for example, [1,12,21] and references therein. We have recently measured and empirically demonstrated that also the areas of the 2D holes (here clumps) created in a film of corrole molecules deposited on an Au(111) surface are distributed according to a Weibull function [22]. In the same paper we also presented a simulation in order to understand how universal the distribution of Weibull was. On that example, we extend and deepen our numerical study to different cluster (here clump) formation modalities that we wish to discuss in this article. We begin by emphasizing that, from the point of view of what we mean to prove, the difference between holes, islands, or clusters is completely immaterial: they are simple speckles in the images and all are referred to as clumps in a 2D space.

The simulation requires to generate clumps with irregular edges, consistently with the experimental data [22]. Among the various possibilities we have chosen the simplest stochastic process [23,24], that is, the random overlapping of disks. Therefore, in our process the distribution of clump sizes is obtained by the random accumulation of disks while the distributions studied in Refs. [12–17] are the result of a fragmentation process. The two processes differ from each other not only because the first takes place in 2D space whereas the second in 3D space, but also in the way the distributions are formed, one may say: bottom-up for the first, and top-down for the second.

The nature of this article is purely geometrical, even if the conclusions can certainly be useful also for interpreting some experimental results.

II. RESULTS AND DISCUSSION

A. First stochastic model

We have carried out several simulations, with a code written in MATLAB12, where 2D clumps were created with different modalities to then determine the cumulative distribution of their areas. The results were fitted to the Weibull function Eq. (1).

In this article there are three entities, which, depending on the scientific environment where they are used, are called by different names. We have chosen to use the nomenclature adopted in the book by Chiu *et al.* [23]. The entities are: clump instead of cluster, grain instead of disk, and germ instead of disk center. The reason for this lies basically in the fact that the subject matter is also of interest in stochastic geometry.

In the simulations, the germs $\{\mathbf{x}_i, i = 1, ..., N\}$, and the corresponding grains of radius r, were generated at random [25] on a $L \times L$ square lattice, with $L = 2000$ px, until an assigned coverage Θ was reached. If c_i is the ith grain set, its normalized area is $\mathscr{A}(c_i) = \frac{\pi r^2}{A}\ \forall i$, where A is the area of the lattice, it follows that $\Theta = \mathscr{A}(\bigcup_{i=1}^{N} c_i)$. The clump is

TABLE I. Simulation results for coverage $\Theta \in [0.35, 0.60]$ corresponding to clumps created on a 2000×2000 square lattice from grains with random germs and fixed radius r that range from 30 to 50 in steps of ten. The cumulative distributions of areas of the resulted clumps were fitted to the Weibull function Eq. (1); best fit shape parameter n and adjusted R squared are reported.

∞ overlaps			One overlap			Two overlaps		
Θ	n	adj.R squared	Θ	n	adj.R squared	Θ	n	adj.R squared
Grain radius 30			Grain radius 30			Grain radius 30		
0.35	0.739	1.00000	0.35	0.842	0.99989	0.35	0.744	0.99999
0.40	0.709	0.99998	0.40	0.741	1.00000	0.40	0.717	0.99999
0.45	0.640	0.99999	0.45	0.704	0.99998	0.45	0.659	0.99996
0.50	0.576	0.99999	0.50	0.624	0.99994	0.50	0.576	0.99999
0.55	0.491	0.99991	0.55	0.516	0.99995	0.55	0.489	0.99997
0.60	0.385	0.99974	0.60	0.401	0.99983	0.60	0.379	0.99983
Grain radius 40			Grain radius 40			Grain radius 40		
0.35	0.771	0.99998	0.35	1.055	0.99912	0.35	0.714	0.99999
0.40	0.676	0.99995	0.40	0.730	0.99997	0.40	0.710	0.99997
0.45	0.655	0.99997	0.45	0.679	0.99992	0.45	0.672	0.99989
0.50	0.590	0.99997	0.50	0.621	0.99998	0.50	0.573	0.99997
0.55	0.502	0.99998	0.55	0.530	0.99988	0.55	0.511	0.99995
0.60	0.406	0.99992	0.60	0.379	0.99988	0.60	0.394	0.99986
Grain radius 50			Grain radius 50			Grain radius 50		
0.35	0.862	0.99995	0.35	0.978	0.99958	0.35	0.783	0.99999
0.40	0.648	0.99985	0.40	0.693	0.99997	0.40	0.652	0.99995
0.45	0.623	0.99990	0.45	0.684	0.99999	0.45	0.642	0.99993
0.50	0.578	0.99994	0.50	0.615	0.99991	0.50	0.577	0.99995
0.55	0.512	0.99996	0.55	0.553	0.99994	0.55	0.525	0.99996
0.60	0.362	0.99986	0.60	0.455	0.99986	0.60	0.395	0.99972

generated by overlapping grains (see Chapter 3 of Ref. [23]) and its area is simply determined by counting the pixels that make it up without assigning any kind of weight to the different clumps.

The cumulative distributions of the areas of the resulted clumps were determined as a function of the parameters Θ and r. The final distribution (histogram) was obtained as the average over 200 iterations; the number of intervals (bins) was chosen maximizing the adjusted R squared in the fitting procedure to Eq. (1); it is worth noting that the bin must be greater than or equal to πr^2, the latter corresponding to the lower limit of the simulated clump area. A summary of the results is reported in Table I, organized in three different columns: ∞ overlaps, one overlap, and two overlaps, corresponding to three different stochastic processes. In the first column (referred to as "∞ overlaps") the code accepts any number of grain overlaps (Boolean model) [23,24]. In the second and third columns the generated grain is accepted only if it does not have more than one or two overlaps with those already accepted, respectively. Although the code allows to choose any number of overlaps, for four overlaps the resulting morphology is in fact indistinguishable from that obtained without any limit in the number of overlaps (∞ overlaps). Indeed, the contribution of the overlapped areas to the entire coverage Θ decreases according to the relation $\Theta_m = \Theta/m!$, where m is the number of overlaps [26] and Θ_m the corresponding fraction of coverage.

Data in Table I are calculated for three different values of grain radius (namely $r = 30, 40, 50$, corresponding to about 1.5%, 2.0%, 2.5% of the side L, respectively). For each of them coverage ranges from $\Theta = 0.35$ to $\Theta = 0.60$ in steps of 0.05; therefore, each datum of the table can be identified by the 3-tuple, $\{\Theta, r, n\_overlaps\}$, where $n\_overlaps$ indicates the maximum number of accepted overlaps. In each column we report, for every coverage Θ, the shape parameter n of the Weibull function Eq. (1) and the corresponding adjusted R squared which evaluates the goodness of the fit and consequently how close the chosen function is to the exact one. Figure 1 depicts some examples of the resulting morphologies along with the cumulative distributions and the corresponding best fitting Weibull functions. The fit procedure obviously excludes everything below the first bin that has a width greater or equal to the area of the minimum clump. For this very reason the curve within the first bin has been drawn dashed [27].

The adjusted R squared values reported in Table I evidence that the clump areas distribution is surprisingly well described by the Weibull function. As a matter of fact, for the majority of them the difference to unity is of the order of $10^{-5} - 10^{-4}$ and only for a few exceptions in the range of 10^{-3}. The worst is for $\{0.35, 40, 1\}$, for which $\bar{R}^2 = 0.99912$; looking in detail at the point of the maximum deviation between the fit data and the simulation data, the difference, in this case, is slightly lower than 2%. Additionally, for $\bar{R}^2$ in the range 0.99958–0.99985, the maximum deviation ranges between 1% and 0.16%, whereas it ranges between 0.10% and 0.05% if $\bar{R}^2$ is in the interval 0.99995–0.99999. Based on the above discussion, it can be summarized that the Weibull function

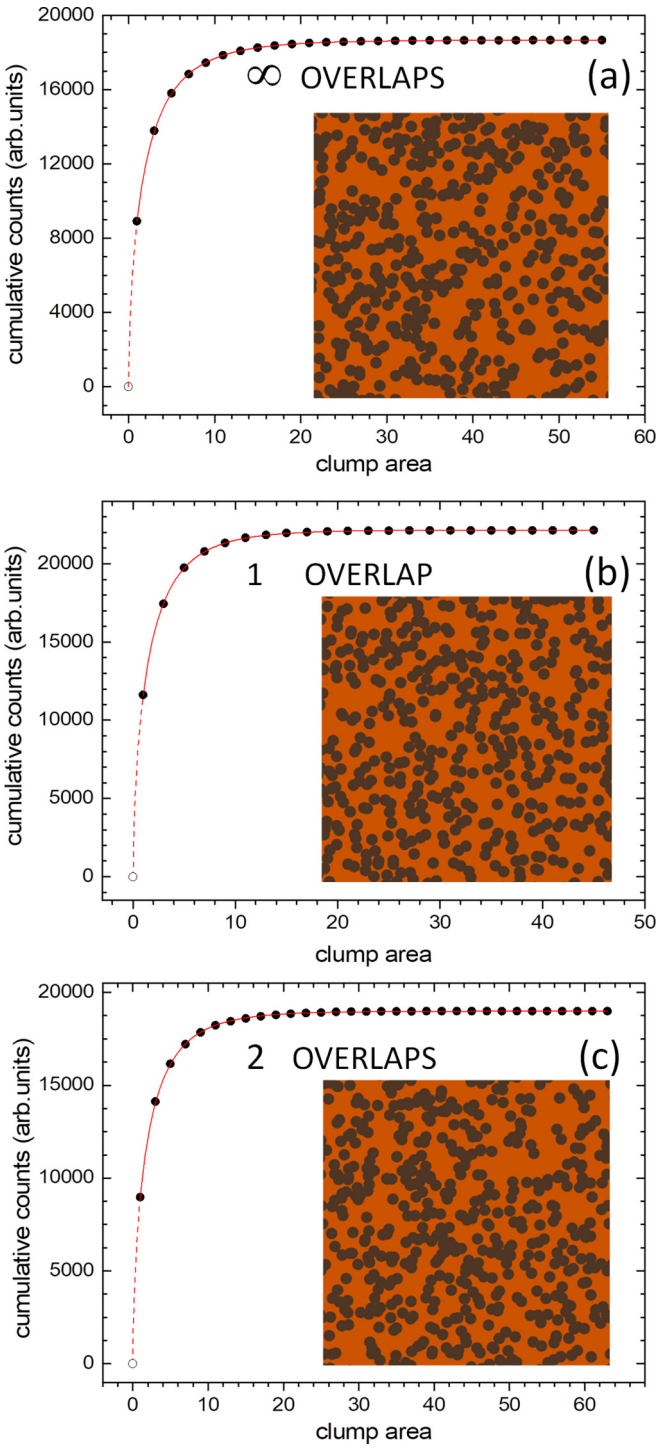

FIG. 1. An example of morphology and cumulative counts (full symbols) of clump areas (grain area units) for grains generated at random for $\Theta = 0.45$, $r = 40$; and (a) ∞ overlaps, (b) one overlap, and (c) two overlaps. Continuous line represents the best fit of simulation results to the Weibull function (see Table I). First part of the curve is dashed since the theoretical point (0,0) is not included in the fitting process, (see the text for details).

describes very well the distribution of clump areas just beyond the area of a grain (numerically: beyond the first bin); actually, the clump cumulative size distribution is discontinuous at πr^2 [29].

TABLE II. Comparison among adjusted R squared values resulting from fitting the cumulative distributions of areas to the Weibull and to the log-normal function for three different 3-tuples. More details can be found in the Supplemental Material [28].

$\{\Theta, r, n\_overlaps\}$	$\bar{R}^2_{\text{Weibull}}$	$\bar{R}^2_{\text{Log-norm}}$
{0.40, 30, 2}	0.99999	0.99965
{0.45, 50, 1}	0.99999	0.99965
{0.55, 30, ∞}	0.99991	0.99953

In Table I, results obtained for low coverage, $\Theta < 0.35$, are not reported, neither are those for $\Theta > 0.60$. As a matter of fact, it happens that outside the range $\Theta \in [0.35, 0.60]$ the Weibull function no longer works, or at least not very well anymore. For instance in tests performed for $\Theta = 0.20$ and $\Theta = 0.25$ the agreement between the numerical simulation and the Weibull function becomes poorer, providing (for instance for $r = 40$ and one overlap) $\bar{R}^2 = 0.99804$ and $\bar{R}^2 = 0.99677$, respectively; incidentally this is accompanied by the transition of the shape parameter n from less to greater than unity; the latter event in itself is not particularly significant, but indicates that the derivative of the Weibull function $W(x)$ at $x = 0$ is zero and the probability distribution function (*pdf*) displays a maximum. Incidentally, in Table I we find $n \gtrsim 1$ also for {0.35, 40, 1} (for details see [30]). The reason of the worsening of the adjusted R squared for $\Theta < 0.35$ is that below this coverage the number of single grains becomes too high and the cumulative distribution gets closer and closer to the step function. As far as the coverage values $\Theta > 0.60$ are concerned, tests performed for $\Theta = 0.70$ and $\Theta = 0.75$ show that the distributions are simply meaningless. As a matter of fact, for $\Theta > 0.60$, the film coverage is close to the percolation threshold $\Theta_p = 0.676$ [31–33] and the cumulative distribution of clump areas begins to be dominated by the clumps of increasingly larger dimensions until, for $\Theta = \Theta_p$, the infinite clump appears. However, it should be noted that an accurate reading of Table I shows that already at $\Theta = 0.60$ we begin to notice a worsening of the $\bar{R}^2$ value.

Since in the literature there are a plethora of papers which compare the Weibull function with the log-normal function [14,34–38], we could not avoid trying to fit the results of our simulations to the aforementioned function: the fit quality with the log-normal worsens with respect to the Weibull function as evidenced, for example, in the comparison among the $\bar{R}^2$ values reported in Table II for three different 3-tuples. (See the Supplemental Material [28] for more details).

B. Second stochastic model: Kolmogorov-Johnson-Mehl-Avrami (KJMA) kinetics

A different process was also addressed where after the random generation of N points on a 2000×2000 square lattice, they become the germs of grains whose radii start growing simultaneously and linearly in time. In this case, the aim of determining the kinetics of the transformation $\Theta(t)$ is a problem whose solution has been found in 1937–1941, independently, by Kolmogorov, Johnson, Mehl, and Avrami hence the acronym KJMA [39–44]. In particular, the aforementioned process is known as Dirac-δ nucleation [45] and the clump

TABLE III. Simulation results for KJMA Dirac-δ nucleation kinetics corresponding to germs randomly distributed throughout a 2000×2000 square lattice and selected coverage $\Theta \in [0.36, 0.55]$. The cumulative distributions of areas of the resulted clumps were fitted to the Weibull function Eq. (1); best fit shape parameter n and adjusted R squared are reported.

KJMA Dirac δ		
Θ	n	adj.R squared
0.36	0.834	0.99986
0.41	0.819	0.99993
0.46	0.682	0.99999
0.51	0.598	0.99998
0.55	0.524	0.99993

TABLE IV. Simulation results in the case of KJMA linear nucleation kinetics corresponding to nuclei randomly distributed throughout a 2000×2000 square lattice and selected coverage $\Theta \in [0.35, 0.54]$. The cumulative distributions of areas of the resulted clumps were fitted to the Weibull function Eq. (1); best fit shape parameter n and adjusted R squared are reported.

KJMA linear nucleation		
Θ	n	adj.R squared
0.35	0.672	0.99999
0.40	0.621	0.99999
0.45	0.566	0.99998
0.50	0.510	0.99998
0.54	0.449	0.99995

size distribution has rarely been studied; this is one of the reasons that pushed us to study the topic.

At any value of t (measured in iterations) there exists a configuration of clumps; in Table III we report the summary of the fits for selected coverage values. Besides, the morphology at time $t = 28$, which corresponds to $\Theta = 0.46$, is shown in Fig. 2(a) along with the analysis of cumulative distribution of the areas. Also in this case, the cumulative distribution of clump areas fits to Eq. (1) in an absolutely satisfactory way. The reader can easily realize that the KJMA kinetics for a Dirac-δ nucleation at time t and grain radius $r(t)$ is, in fact, a Boolean model, which is virtually the same model as for ∞ overlaps previously described (except for a different distribution of the number of grains).

Always within the ambit of the KJMA model, we wrote a code in which the N points on the 2000×2000 square lattice do not start growing simultaneously, but rather "turn on" according to a linear law (linear nucleation). In detail, 100 new points out of a total of 1000 (which are the germs) start growing at the end of each growth cycle. In other words, the growth law that, as in the Dirac-δ case is linear, is of the type $r_j(t) = a(t - t_j)$, where t is the running time, r_j the radius of the jth grain, and t_j the germ birth time. The results are reported in Table IV and, as in the KJMA Dirac-δ kinetics, the numerical data fit Eq. (1) very satisfactorily, as shown in Fig. 2(b) where also the morphology at time t = 36 and $\Theta = 0.45$ is reported. It is worth emphasizing once again that fitting the data from the above simulations to the log-normal function provides poorer results. Incidentally, we recall that the KJMA model is often employed to describe the kinetics of some kinds of phase transition.

C. Effects of hard correlation

In all the stochastic processes previously described, the simulated cumulative distribution agrees really well with Eq. (1). Besides, they share the fact that the choice of the germs is Poissonian. So as to check whether this is a prerequisite to obtain a good agreement between the simulation and the Weibull function, we considered a certain degree of correlation among the germs; for this purpose we introduced the constraint that does not allow the germs $\mathbf{x}_k$ to lay closer

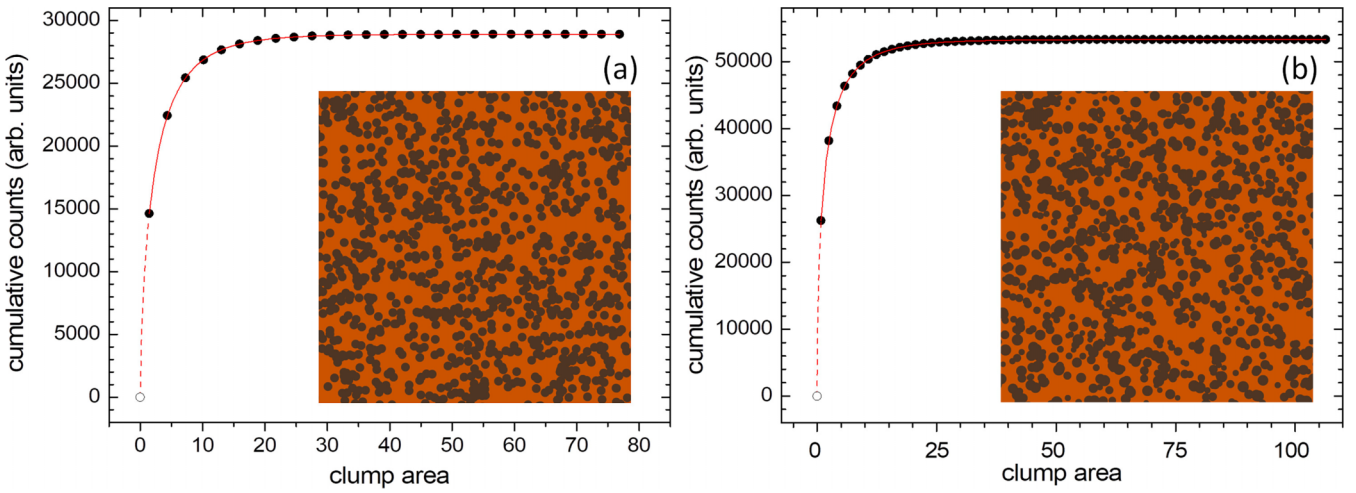

FIG. 2. KJMA kinetics morphology for (a) instantaneous (Dirac δ) nucleation at time t = 28 and $\Theta = 0.46$; (b) linear nucleation at time t = 36 and $\Theta = 0.45$. Full symbols are the corresponding cumulative counts (arbitrary units) of clump areas fitted tothe Weibull function (continuous line), (see Table III and Table IV). First part of the curve is dashed since the theoretical point (0,0) is not included in the fitting process, (see the text for details).

than a given distance R_c (correlation distance) from each other [46].

In other words, with each germ $\mathbf{x}_k$ is associated a circle of radius R_c; if $\bigcup_{i=1}^{k} C_i$ is the set of the first k circles of radius R_c, then the $(k+1)$th germ will be accepted if it satisfies the condition

$$\mathbf{x}_{k+1} \notin \bigcup_{i=1}^{k} C_i. \tag{2}$$

It is the case to underline that R_c can vary within the interval $R_c \in [0, 2r)$, but it cannot overcome $2r$ because in such case there would be no overlap between grains and the *pdf* would simply be a Dirac δ. A certain number of simulations for different values of coverage ($\Theta = 0.35, 0.45, 0.50, 0.60$) were performed for constant grain radius $r = 40$ and R_c from 50 to 70 in steps of five (so that $R_c < 2r$). The cumulative distribution is still well described by the Weibull function, with $\bar{R}^2$ varying in the range $0.99972 - 0.99999$ for all the tested coverages and the different R_c (see [47]).

D. Effects of inhomogeneous spatial distribution

The introduction of the correlation does not change the germ distribution homogeneity; however, it is precisely homogeneity that plays a key role for the distribution of grain areas to fit the Weibull function. In order to corroborate the latter statement, we introduced three functions $f(x, y)$, which allow the point (x_i, y_i) to be accepted if, after having randomly chosen a number $p \in [0, 1]$, the relation $p < f(x_i, y_i)$ is satisfied.

The functions are as follows:

a Gaussian-like function

$$f^{(G)}(x, y) = \exp\left[-\left(\frac{x^2 + y^2}{2\,\xi^2}\right)\right] \tag{3}$$

and two periodic functions

$$f^{(ss)}(x, y) = \left[\sin\left(\frac{x}{\xi}\right)\sin\left(\frac{y}{\xi}\right)\right]^2 \tag{4}$$

$$f^{(s^2)}(x, y) = \sin^2\left[\left(\frac{x}{\xi}\right)\left(\frac{y}{\xi}\right)\right], \tag{5}$$

represented in Figs. 3(a)–3(c), respectively, where $\xi = L/c$ with $c \in \mathbb{Q}$. In particular, in our simulations $c = 1/2, 2/3, 1, 2, 3$ for $f^{(G)}$, while $c = 5$ and $c = 2$ for $f^{(ss)}$ and $f^{(s^2)}$, respectively.

As far as the $f^{(G)}$ is concerned, the simulation results do not fit Eq. (1) satisfactorily for $\xi \leqslant L/3$; specifically, for $\xi = L/3$, $\bar{R}^2 = 0.995$ [see Fig. 3(d)]. Such effect is even more evident when the Gaussian-like maximum coincides with a lattice corner, and in this case the fit is meaningless (more details can be found in the Supplemental Material [28]).

The effect of the functions $f^{(ss)}(x, y)$ and/or $f^{(s^2)}(x, y)$ over the grain spatial arrangement is quite strong and, in both cases, the cumulative distribution of clump areas utterly differs from Eq. (1). The effect is more drastic for $f^{(ss)}(x, y)$, to the point that the fit does not make sense [see Fig. 3(e)]. The reason depends on the fact that the grains are concentrated in very compact clumps, while for $f^{(s^2)}(x, y)$ the grains are more scattered throughout the whole space [Fig. 3(f)]; nevertheless, in this case, the adjusted R squared is 0.99881 instead of typical values $\bar{R}^2 \in [0.99989, 0.99997]$ for the random case with the same grain radius and coverage ($r = 40$, $\Theta = 0.45$). Although we have made use of only three functions, the message is rather apparent: the Weibull function describes the cumulative distribution of clump areas only if the spatial distribution of their initial germs is sufficiently spread.

III. FINAL REMARKS

Before summarizing the main results in the conclusion section, a couple of comments are mandatory.

The first comment concerns the boundary conditions. The simulations are performed in a square lattice of side $L = 2000$ and periodic boundary conditions (PBCs) are taken into account, i.e., grains are generated onto a torus. The PBCs guarantee the correct value of the coverage Θ. By referring to the paradigmatic example of the left panel of Fig. 4, a correct evaluation of the cumulative cluster distribution should consider (thanks to the PBCs) the two clumps inside the $L \times L$ square lattice as if they were one. Our code, on the contrary, consider them as two separated clumps whose area is given, as for the clump on the north side, by that of the blue clump, and as for that on the south side by the sum of the red and the blue clump areas (first method). From what has just been said, a certain degree of error is introduced in the cumulative area distribution by the clumps overflowing from the edge of the lattice; however, their number is proportional to the side L of the lattice, while the contribution of the clumps inside the lattice is proportional to L^2, so that the error decreases as $1/L$. Nevertheless, when increasing L from 2000 to 6000, no signs of significant improvement in the $\bar{R}^2$ values were observed against an increase in computation time of 150% and more.

A possible method to overcome the problems of the edge effects, is to introduce a ΔL wide guard region ($\Delta L = r$) around the original lattice, making the PBCs conditions unnecessary. Indeed, since the germs lie within the $L \times L$ lattice, all clumps lie within the $(L + r) \times (L + r)$ lattice, so that all pixels of each clump are counted (see for instance the clump at the north edge in the right panel of Fig. 4) [48]. However, since the germs are generated at random on the $L \times L$ region, in the $(L + r) \times (L + r)$ region the process is no longer Poissonian. Precisely for this reason, between the two methods, we preferred the former.

The second comment has to do with interesting papers in which the problem of determining the distribution function of grains or crystallites in the growth or crystallization processes has been tackled analytically [49–52]. Based on the titles, these articles may be associated with our analysis but this is not the case. In our case the clumps are separated, while in those articles, particles in question touch each other or share a portion of their rim like the tiles in the Voronoi tessellation.

IV. CONCLUSIONS

The numerical simulations reported here are inspired by some experimental results that show that the areas of voids (here, clumps) observed in a molecular film distribute according to the Weibull function [22]. With the aim of verifying

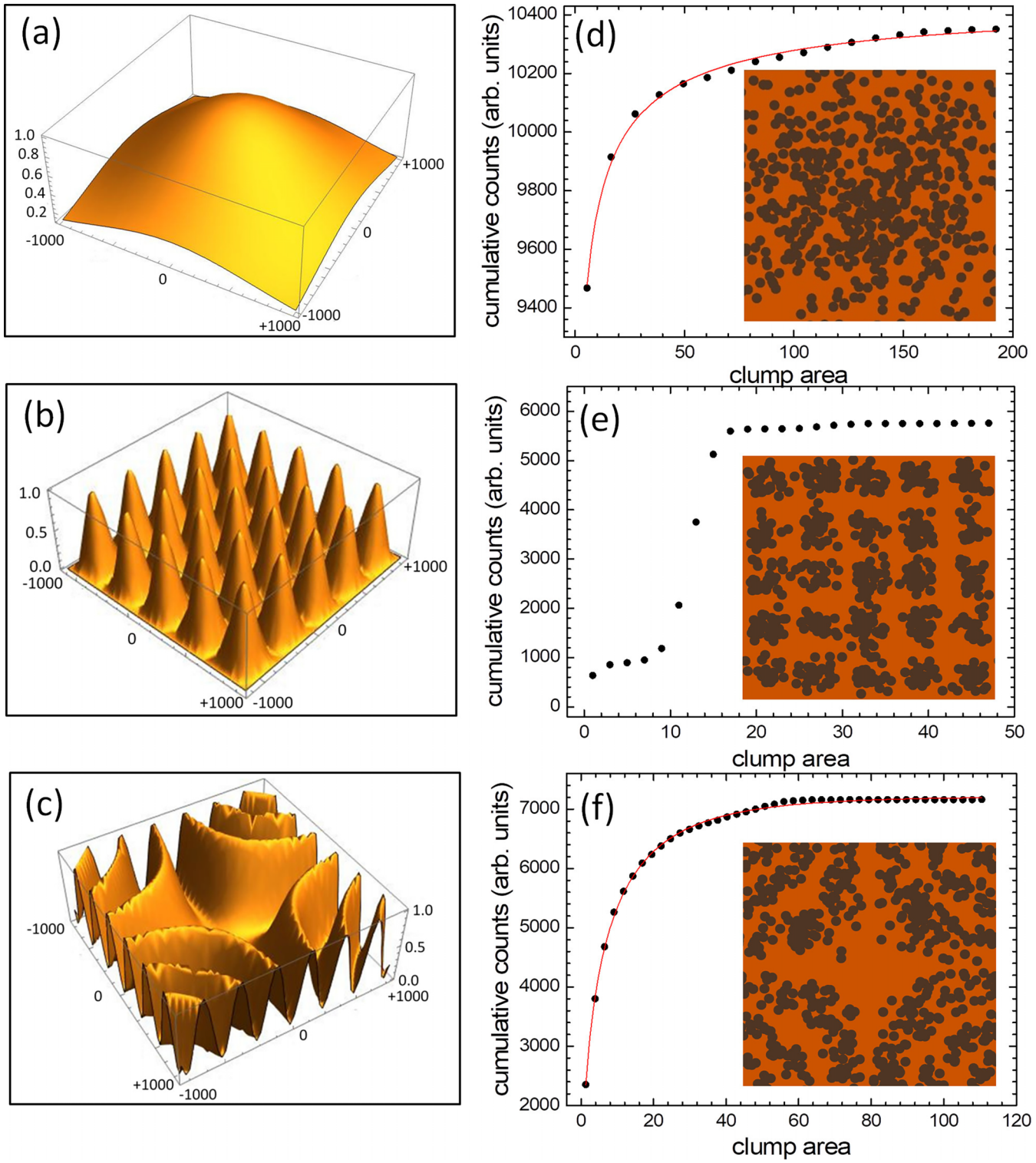

FIG. 3. Panels (a)–(c) show a graphical representation of the functions $f^{(G)}(x, y)$, $f^{(ss)}(x, y)$, and $f^{(s^2)}(x, y)$. Panels (d)–(f) report an example of morphology and the cumulative counts (arbitrary units) of clump areas when respective functions are introduced (see text for details); $\Theta = 0.45$, $r = 40$.

how general this last evidence is, we approached the problem numerically.

The results of our simulations demonstrate that, within a certain range of coverage, there is a close connection between the Weibull function and the cumulative clump area distribution if clumps are generated by the overlap of randomly distributed grains of radius r. The same analysis was applied to the distribution of clumps obtained from two typical KJMA kinetics. Also in this case, the agreement between the clump size distribution and Weibull function is excellent.

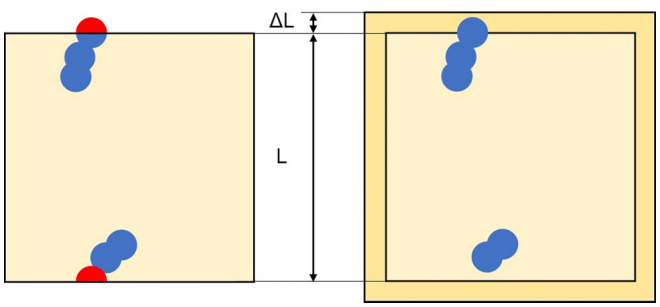

FIG. 4. Schematic graphical representation to show how the periodic boundary conditions are taken into account in the simulations (see text for details).

We also show that hard-correlated spatially distributed grains give rise to a clump configuration whose area distribution still fits quite well to the Weibull function. On the other hand, by significantly changing the grain arrangement by making space inhomogeneous, the numerical data do not fit to the Weibull function anymore.

Eventually, the Weibull function is not only an excellent descriptor of the clump size distribution on a surface, but it would be, in principle, preferable to the log-normal function. It goes without saying that, from the experimental point of view, the possibility of obtaining an adjusted R squared with values differing from the unity by about $10^{-5} - 10^{-4}$ should be excluded. In other words, fitting the data to one of the two functions is, in fact, equivalent. However, our study seems to go beyond: in fact, given the exceptional goodness of the $\bar{R}^2$ values, it suggests the possibility that, under certain conditions, the Weibull function could be, in the interval $[\pi r^2, \infty)$, the exact clump size distribution function for the 2D processes we have described.

Often, simulations such as those reported here have been employed in stochastic geometry [23,24] in order to describe and explain several properties of heterogeneous materials in many areas of science and technology. Naturally, the results obtained here are far from being a mathematical proof, but being well supported by numerical analysis, we hope they will give rise to a new challenge in the field of mathematical morphology [53,54] and stochastic geometry.

ACKNOWLEDGMENTS

B.B. and A.S. thank the University of Rome Tor Vergata for funding (Project "GREEN, Growth of coRrole layErs: from propErty tuning to applicatioNs" Progetti di Ateneo Università degli Studi di Roma "Tor Vergata", CUP E83C22000350005). A.S. thanks the Regione Lazio POR FESR Lazio 2014-2020 for funding (Project id 1076780300-0327), Proposal code (POR A0375E0102) "High electron mobility 2D polymers (P2DAME, HEM-2DP)".

[1] H. Rinne, *The Weibull Distribution: A Handbook* (CRC Press, Taylor and Francis Group, Boca Raton, NW, USA 2009).
[2] W. Weibull, A statistical distribution function of wide applicability, J. Appl. Mech. **18**, 293 (1951).
[3] D. N. Prabhakar Murthy, M. Xie, and R. Jiang, *Weibull Models* (John Wiley & Sons, New York, 2003).
[4] G. Zachariev, A Statistical theory of the damage of materials, Mod. Mech. Eng. **06**, 129 (2016).
[5] M. N. Sharif and M. N. Islam, The Weibull distribution as a general model for forecasting technological change, Forecast. Soc. **18**, 247 (1980).
[6] T. Tsuboi, J. Takami, S. Okabe, K. Inam, and K. Aono, Weibull parameter of oil-immersed transformer to evaluate insulation reliability on temporary overvoltage, IEEE Trans. Dielectr. Electr. Insul. **17**, 1863 (2010).
[7] P. Wais, A review of Weibull functions in wind sector, Renewable Sustainable Energy Rev. **70**, 1099 (2017).
[8] S. Bhardwaj, N. Bhardwaj, and V. Kumar, The study of reliability of diesel locomotive engine using Weibull distribution, Int. J. Agric. Stat. Sci. **15**, 549 (2019).
[9] B. Zhang, L. Tang, and M. Roemer, Probabilistic weather forecasting analysis for unmanned aerial vehicle path planning, J. Guid. Control. Dyn. **37**, 309 (2014).
[10] V. P. Singh, On application of the Weibull distribution in hydrology, Water Resources Management **1**, 33 (1987).
[11] A. Putkonen, A. Nioche, M. Laine, C. Kuuramo, and A. Oulasvirta, Fragmented visual attention in web browsing: Weibull analysis of item visit times, in *Advances in Information Retrieval: 45th European Conference on Information Retrieval, Dublin, Ireland*, edited by J. Kamps *et al.*, Lecture Notes in Computer Science, Vol. 13981 (Springer, 2023), pp. 62–78.
[12] M. Alderliesten, Mean particle diameters. Part VII. The Rosin-Rammler size distribution: Physical and mathematical properties and relationships to moment-ratio defined mean particle diameters, Part. Part. Syst. Charact. **30**, 244 (2013).
[13] Z. Fang Jr., B. R. Patterson, and M. E. Turner, Modeling particle size distributions by the Weibull distribution function, Mater. Charact. **31**, 177 (1993).
[14] H. J. H. Brouwers, Packing fraction of particles with a Weibull size distribution, Phys. Rev. E **94**, 012905 (2016).
[15] A. Paluszny, X. H. Tang, M. Nejati, and R. W. Zimmerman, A direct fragmentation method with Weibull function distribution of sizes based on finite- and discrete element simulations, Int. J. Solids Struct. **80**, 38 (2016).
[16] W. K. Brown and K. H. Wohletz, Derivation of the Weibull distribution based on physical principles and its connection to the Rosin–Rammler and lognormal distributions, J. Appl. Phys. **78**, 2758 (1995).
[17] T. M. Zobeck, T. E. Gill, and T. W. Popham, A two-parameter Weibull function to describe airborne dust particle size distributions, Earth Surf. Process. Landforms **24**, 943 (1999).
[18] P. Rosin and E. Rammler, The laws governing the fineness of powdered coal, J. Inst. Fuel. **7**, 29 (1933).
[19] P. Rosin, E. Rammler, and K. Sperling, Korngrßenprobleme des Kohlenstaubes unde ihre Bedeutung für die Vermahlung. Report C52 of the National Coal Council (VDI-Verlag Berlin) 1 (1933).
[20] J. G. Bennett, Broken coal, J. Inst. Fuel. **10**, 22 (1936).

[21] D. Stoyan, Weibull, RRSB or extreme-value theorists? Metrika **76**, 153 (2013).
[22] M. Fanfoni, B. Bonanni, R. Martini, A. Sgarlata, F. Caroleo, R. Paolesse, and C. Goletti, Morphological study of a self-organized corrole film on a gold crystal, both in solution and after solvent evaporation, J. Phys. Chem. C **127**, 12702 (2023).
[23] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and Its Applications*, 3rd ed. (Wiley, New York 2013).
[24] S. Torquato, *Random Heterogeneous Materials* (Springer, New York 2002).
[25] The random number generator (rng), which is applied to place the germs in the lattice, uses the seed "shuffle" so that the rng function initializes the generator seed based on the current time, resulting in a different sequence of random numbers after each call to rng.
[26] V. Sessa, M. Fanfoni, and M. Tomellini, Validity of Avrami's kinetics for random and nonrandom distributions of germs, Phys. Rev. B **54**, 836 (1996).
[27] In fact, also in experimental data there exists a clump area greater than zero, below which it is impossible to go. The reason lies in instrumental limitations, as the smallest detectable island depends on the resolving power, electronic and thermal noise, instrument sensitivity, etc.
[28] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.109.044131 for (a) results obtained by fitting the cumulative distribution of clump areas to the log-normal function, (b) graphical representation of the Gaussian-like function with maximum at the left upper corner of the lattice, corresponding morphology and cumulative distribution of clump areas.
[29] The discontinuity is particularly evident when the radius of the grain tends to zero: in this case the *cdf* tends to the step function. In fact, as r approaches zero the probability of having an isolated disk diverges, as shown in Ref. [23], Eq. (3.104).
[30] It is worth noting that $n = 1.055$, so that the cumulative distribution is very close to an exponential. To eliminate any doubts about the goodness of the result we repeated the simulation three times, obtaining the same value within 0.004%.
[31] D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, London 1992).
[32] M. Li, H. Chen, J. Lin, R. Zhang, and L. Liu, Effects of the pore shape polydispersity on the percolation threshold and diffusivity of porous composites: Theoretical and numerical studies, Powder Technol. **386**, 382 (2021).
[33] Z. Koza, P. Brzeski, and G. Kondrat, Percolation of fully penetrable disks using the three-leg cluster method, J. Phys. A: Math. Theor. **56**, 165001 (2023).
[34] J. Laherrère and D. Sornette, Stretched exponential distributions in nature and economy: "fat tails" with characteristic scales, Eur. Phys. J. B **2**, 525 (1998).
[35] D. T. Kinsella and K. Persson, On the applicability of the Weibull distribution to model annealed glass strength and future research needs, in *Challenging Glass 5: Conference on Architectural and Structural Applications of Glass*, edited by Belis, Bos, and Louter (Ghent University, Belgium, 2016), p. 593.
[36] W. J. Szajnowski, *Discrimination between Log-Normal and Weibull Clutter* (IEEE, 1977), Vol. AES-13, pp. 480–485.
[37] D. Kundu and A. Manglick, Discriminating between the Weibull and log-normal distributions, Nav. Res. Logist. **51**, 893 (2004).
[38] J. F. Ortiz-Yañez and M. R. Piña-Monarrez, Discrimination between the lognormal and Weibull distributions by using multiple linear regression, DYNA **85**, 9 (2018).
[39] A. N. Kolmogorov, On the statistical theory of the crystallization of metals, Bull. Acad. Sci. USSR (Math. Series) **3**, 355 (1937).
[40] W. A. Johnson and R. F. Mehl, Reaction kinetics in processes of nucleation and growth, Trans. AIME **135**, 416 (1939).
[41] M. Avrami, Kinetics of phase change, J. Chem. Phys. **7**, 1103 (1939).
[42] M. Avrami, Kinetics of phase change. II transformation-time relations for random distribution of nuclei, J. Chem. Phys. **8**, 212 (1940).
[43] M. Avrami, Granulation, phase change and microstructure kinetics of phase change. III, J. Chem. Phys. **9**, 177 (1941).
[44] M. Fanfoni and M. Tomellini, The Johnson-Mehl-Avrami-Kolmogorov model: A brief review, Il Nuovo Cimento D **20**, 1171 (1998).
[45] R. Polini, M. Tomellini, M. Fanfoni, and F. Le Normand, Dirac-δ nucleation in the framework of Avrami's model: the case of diamond growth on deformed Si(100), Surf. Sci. **373**, 230 (1997).
[46] M. Tomellini, M. Fanfoni, and M. Volpe, Spatially correlated nuclei: How the Johnson-Mehl-Avrami-Kolmogorov formula is modified in the case of simultaneous nucleation, Phys. Rev. B **62**, 11300 (2000).
[47] Unfortunately, in the previous paper [22], because of a not optimal choice of the histogram's interval, the reported value for $\bar{R}^2$ was erroneous. However, that error is of secondary importance and the only thing that needs to be clarified is that, on the basis of the new analysis reported here, the random origin of the experimental holes in [22] cannot be asserted.
[48] We tested this approach for several 3-tuples$\{\Theta, 40, 1\}$ finding a clump area distribution that fits very well to the Weibull function. Incidentally, the returned shape parameter, n, differs by a few percent from that of the former method and only for $\Theta = 0.35$ the percentage difference reaches the value of about 15%.
[49] A. V. Teran, A. Bill, and R. B. Bergamm, The grain size distribution in crystallization processes with anisotropic growth rate, Phys. Rev. B **81**, 075319 (2010).
[50] E. Pineda, P. Bruna, and D. Crespo, Cell size distribution in random tessellations of space, Phys. Rev. E **70**, 066119 (2004).
[51] J. Farjas and P. Roura, Cell size distribution in a random tessellation of space governed by the Kolmogorov-Johnson-Mehl-Avrami model: Grain size distribution in crystallization, Phys. Rev. B **78**, 144101 (2008).
[52] M. Tomellini, Fokker–Planck equation for the particle size distribution function in KJMA transformations, Physica A **615**, 128515 (2023).
[53] K. Michielsen and H. De Raedt, Integral-geometry morphological image analysis, Phys. Rep. **347**, 461 (2001).
[54] K. Michelsen, H. De Raedt, and J. T. H. M. De Hosson, Aspects of mathematical morphology, Adv. Imaging Electron Phys. **125**, 119 (2003).