Volume 385, issue 2 15 November 2007 ISSN 0378-4371

Available online at
ScienceDirect
www.sciencedirect.com
http://www.elsevier.com/locate/physa

Available online at www.sciencedirect.com
ScienceDirect
Physica A 385 (2007) 518–526
PHYSICA A
www.elsevier.com/locate/physa

On the size distribution of Poisson Voronoi cells

Járai-Szabó Ferenc[a,b], Zoltán Néda[a,*]

[a]Department of Theoretical and Computational Physics, Babeş-Bolyai University, str. Kogălniceanu 1, RO-400084 Cluj-Napoca, Romania
[b]Interdisciplinary Computer Simulation Group, KMEI, str. Tipografiei 12, RO-400101, Cluj-Napoca, Romania

Received 18 April 2007
Available online 6 August 2007

Abstract

Poisson Voronoi diagrams are useful for modeling and describing various natural patterns and for generating random lattices. Although this particular space tessellation is intensively studied by mathematicians, in two- and three-dimensional (3D) spaces there is no exact result known for the size distribution of Voronoi cells. Motivated by the simple form of the distribution function in the 1D case, a simple and compact analytical formula is proposed for approximating the Voronoi cell's size-distribution function in the practically important 2D and 3D cases as well. Denoting the dimensionality of the space by d ($d = 1, 2, 3$) the $f(y) = Const * y^{(3d-1)/2} \exp(-(3d+1)y/2)$ compact form is suggested for the normalized cell-size distribution function. By using large-scale computer simulations the viability of the proposed distribution function is studied and critically discussed.

PACS: 89.75.Kd; 02.50.Ng; 02.10.−v

Keywords: Voronoi diagrams; Monte Carlo methods; Cell-size distribution

1. Introduction

Voronoi diagrams [1] are a particular case of space tessellation where, given a set of centers, the space is divided according to their "spheres of influence". Each Voronoi cell contains those points of the space that are closest to the same center. A Voronoi tessellation in two dimensions would look like the polygons sketched in Fig. 1 or Fig. 2d.

Given a set of centers there are two relatively easy ways to generate the corresponding Voronoi diagram. We sketch this methods for the two-dimensional (2D) case, and the generalization to any other dimension is immediate. In the *first method* (the perpendicular bisectors method [1,2]) one starts from a given center (P_0) and detects the nearest P_1 center to it. A part of the perpendicular bisector on the P_0P_1 line will form the first edge of the Voronoi polygon corresponding to P_0. Then the second nearest center (P_2) is detected and the perpendicular bisector on P_0P_2 is constructed again. This algorithm is continued with the third (P_3), fourth (P_4), fifth (P_5),... nearest center, until the perpendicular bisectors on P_0P_3, P_0P_4, P_0P_5... will close a stable

*Corresponding author. Tel.: +40 745 310531; fax: +40 264 591906.
E-mail address: zneda@phys.ubbcluj.ro (Z. Néda).

0378-4371/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2007.07.063

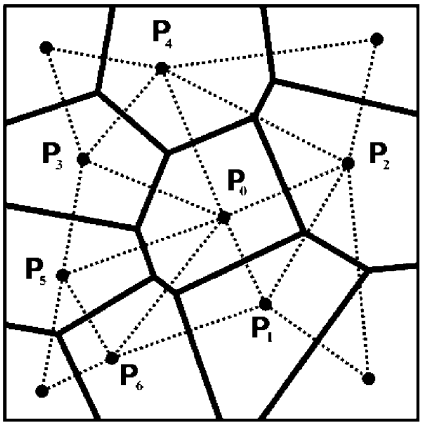

Fig. 1. The "perpendicular bisector method" for constructing Voronoi diagrams in 2D.

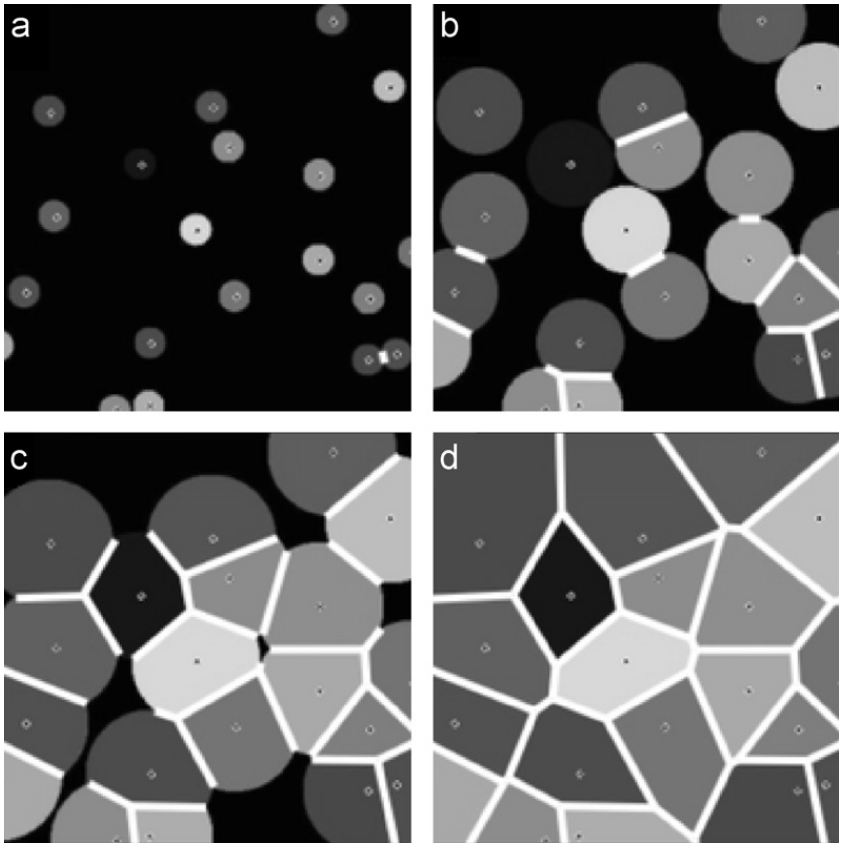

Fig. 2. The Avrami–Johnson–Mehl method for constructing Voronoi diagrams in 2D: (a)–(d) presents snapshots from a small graphical simulation.

polygon which does not change after considering any more distant points. Repeating the above algorithm for all centers the Voronoi tessellation of the whole space (Fig. 1) can be obtained.

The *second method* (called the Avrami–Johnson–Mehl method [3]) is especially useful for computer simulations. In this algorithm each center is identified as a nucleation point from where a virtual disc with uniform radial velocity is growing (Fig. 2). When two discs touch each other the growth in the contact direction is stopped for both of them and the contact point becomes a point on the corresponding Voronoi diagram. The growth in all other directions is continued until a nearby disc is reached. In this way the same space tessellation as in the perpendicular bisector algorithm is achieved. In computer simulations it is handy to

implement this algorithm not on continuous space but on large lattices since the contact points are easier to identify.

In the 2D case the Voronoi diagram can be obtained also from Delaunay triangulation. The Delaunay triangulation of a point set is a collection of edges satisfying an empty-circle criteria, which means that for each edge we can find a circle containing the edge's endpoints but not containing any other point from the initial set. In two dimensions the Delaunay triangulation is the dual structure of the Voronoi diagram [2].

Particular cases of Voronoi diagrams, where the centers are randomly and uncorrelated distributed, are called Poisson Voronoi diagrams. Poisson Voronoi diagrams (PVDs) are especially important for modeling and describing a wide variety of natural and social phenomena. PVD has been used to construct random lattices in quantum field theory [4] or in the studies of conductivity and percolation in granular composites [5]. PVD was also used in modeling growth of metal clusters on amorphous substrates [6], in studying conduction and percolation in continuous media [7], in modeling microemulsions [8], in interpreting small angle X-ray scattering for heterogeneous catalyst [9], in evaluating the actual galaxy distribution [10], in describing sections through various geological materials [11], in biology [12], in animal ecology [13], in sociology [14], etc.... The above list is far from being complete and suggests just a few possible applications for this particular space tessellation. For a more complete discussion of the use of Voronoi diagrams many good review works are available [15–17].

Despite their importance in science, our knowledge on the geometrical and statistical properties of PVD is far from being complete [1,2]. One of the most debated and less clarified aspect is the $g(S)$ size-distribution function of Voronoi cells $g(S) = P(S, S + \mathrm{d}S)/\mathrm{d}S$, where $P(S, S + \mathrm{d}S)$ is the probability that the size of a Voronoi cell is between S and $S + \mathrm{d}S$. Instead of $g(S)$ it is more convenient to use the more general $f(y)$ distribution function for the $y = S/\langle S\rangle$ normalized cell sizes, which is independent of the center's density and it is universal for all PVD in a given dimension. Alternatively, one could determine the $F(y)$ cumulative distribution function defined as $F(y) = \int_0^y f(x)\,\mathrm{d}x$.

Apart of the simple one-dimensional case, presently there is no exact result or handy analytical approximation for the form of $f(y)$. Since $f(y)$ in 2D and 3D is of primary interest in many practical applications, there is a growing need for a simple and analytically usable formula. This would help in characterizing and classifying several experimental patterns, and would give an important starting point also for modeling these structures. In contrast with mathematicians experimental scientists need a simple expression that could give a first hint about the nature of the measured cell-size distribution, which is usually determined with poor statistics.

There are many conjectures on the analytical form of $f(y)$ and many computer simulations were done to prove the suggested forms. Up to our knowledge in 2D the largest computer simulations were done by Tanemura [18,19] with 10^7 Voronoi cells and Hinde and Miles [20] with 2×10^6 cells. In 3D the largest ensembles were studied again by Tanemura [18,19] (3×10^6 cells) and Kumar et al. [21] (3.6×10^6 cells).

As a generally accepted result emerges a three parameter (a, b and c) generalized gamma function fit

$$f(y) = c\frac{b^{a/c}}{\Gamma(a/c)} y^{a-1}\exp(-by^c) \tag{1}$$

describes the computer simulation data reasonable well. Some authors [6,22] suggested, however, that a simpler two-parameter (a and b) gamma function fit

$$f(y) = \frac{b^a}{\Gamma(a)} y^{a-1}\exp(-by), \tag{2}$$

works also well.

In 2D for the three parameter fit (1) Tanemura [18,19] found $a = 3.315$, $b = 3.04011$ and $c = 1.078$, in good agreement with the results of Hinde and Miles [20] $a = 3.3095$, $b = 3.0328$ and $c = 1.0787$. For the two parameter fit (2) the values $a = b = 3.61$ [22] or $a = 3.61$ and $b = 3.57$ [6] were reported.

In the 3D case Tanemura found [19] $a = 4.8065$, $b = 4.06342$ and $c = 1.16391$ for (1), while Kiang [23] suggested a fit of the form (2) with $a = b = 6$. We have to mention, however, that the simulations of Tanemura [18,19] did not support Kiang's results [23] at all.

For the sake of completeness it also has to be mentioned that there is an exact analytical result for the second moment ($\langle y^2 \rangle$) of the PVD's both in the 2D and 3D cases [24]. According to this $\langle y_{2D}^2 \rangle = 1.280$ and $\langle y_{3D}^2 \rangle = 1.180$ [15], offering an excellent possibility for testing the computer simulation results and the correctness of the proposed fit. It was the enormous discrepancy between Gilbert's and Kiang's results for this second moment that condemned Kiang's simulation results.

The aim of the present work is not to give a better and more complicated fit for $f(y)$. We would rather intend to prove that a simple two parameter fit of form (2) used by Kiang can be still a fair approximation for all practical applications. Experimental scientists instead of focusing on a more accurate but difficult fit for the presumed PVD type patterns can use with confidence a simple approximation of form (2). In the present work large-scale computer simulations are also considered for the problem, generating more Voronoi cells than in all previous studies we are aware of. The statistics of 3×10^7 and 1.8×10^7 cells are studied in 2D and 3D, respectively. Using this improved statistics Kiang's conjecture will be followed and a first approximation for $f(y)$ in form (2) will be given with simple and handy values of $a = b$. Solving the problem exactly in 1D will give us further motivation for this simpler form of $f(y)$.

2. The 1D case

Let us first study theoretically the simple problem in 1D and prove the validity of (2) with $a = b$. One has to mention however that several other methods are known to obtain the exact form of the $f_{1D}(y)$ distribution function in this simple case [3,15].

A line with length L is considered on which N centers are randomly and independently distributed. The density of centers is given thus as $n = N/L = 1/\langle d \rangle$, where $\langle d \rangle$ stands for the average distance between centers. We will study the limit $L \to \infty$, $N \to \infty$, but n is finite. It is necessary to construct the Voronoi diagrams for these centers (Fig. 3). If a center P is considered, first its neighbor in the left (P_l) and right (P_r) directions will be detected. Then the PP_l and PP_r lines are divided into two equal parts, by the D_l and D_r points, respectively. The segment $D_l D_r$ is then the Voronoi cell corresponding to the center P. It is obvious that for the considered limit the average length of Voronoi cells is $\langle d \rangle = 1/n$.

In order to get the distribution function $g(d)$ of the Voronoi cell's length, first the distribution function $h(s)$ for the lengths between centers will be determined. Let us start from the well-known Poisson distribution $P(N, t)$, giving the probability that inside a length t there are N centers:

$$P(N, t) = \frac{1}{N!} \langle N \rangle_t^N \exp(\langle N \rangle_t). \tag{3}$$

In the above equation $\langle N \rangle_t = nt$ stands for the expected (average) number of centers on a length t. The probability that in an interval of length t situated on the immediate right of P there are no other centers is

$$P(0, t) = \exp(-nt). \tag{4}$$

The cumulative distribution $P_r(d_r > t)$ that the first neighbor at the right is at a distance d_r bigger than t is $P_r(d_r > t) = P(0, t)$. The distribution function $g_r(d_r)$ for the lengths d_r can be thus calculated as

$$g_r(d) = -\frac{\partial P_r(d_r > d)}{\partial d} = n \exp(-nd). \tag{5}$$

Due to symmetry constraints, the same distribution function should apply for the d_l lengths relative to the first neighbor in the left direction. The distribution function for the half of these intervals ($z = d_r/2$ or $z = d_l/2$) is given as

$$w(z) = 2n \exp(-2nz). \tag{6}$$

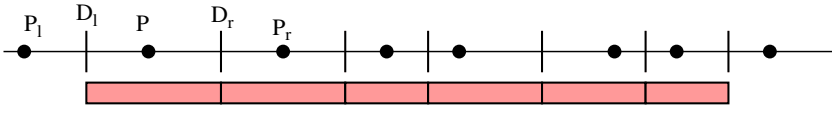

Fig. 3. Construction of Voronoi cells in 1D.

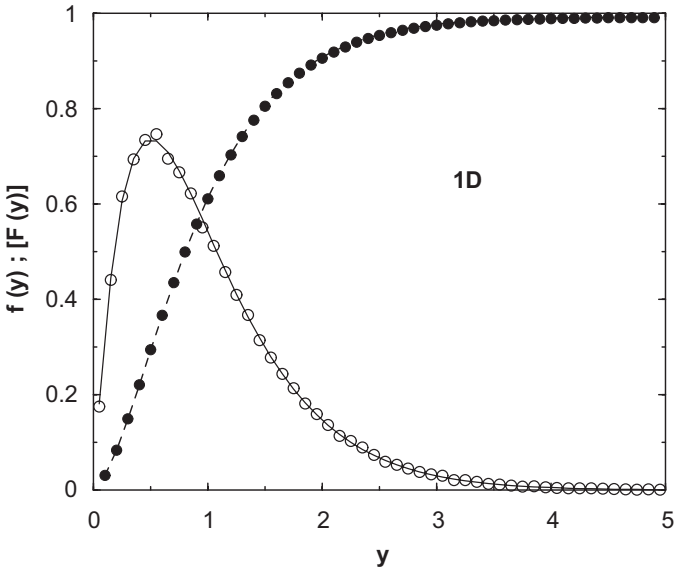

Fig. 4. Simulation results (empty circles) in 1D in comparison with the (8) exact result (solid line). Results for the cumulative distribution function are also plotted. Filled circles are simulation data and the dashed line is given by Eq. (9).

The length d of the Voronoi cell is $d = d_l/2 + d_r/2$, and its distribution $g(d)$ can be calculated as the convolution of two distributions of form (6):

$$g(d) = \int_0^d w(z)w(d-z)\,\mathrm{d}z = 4n\,\exp(-2nd). \tag{7}$$

It is necessary to realize that this distribution function is normalized for $L \to \infty$. The distribution function for the adimensional quantity $y = d/\langle d\rangle$ is given then as

$$f_{1D}(y) = 4y\,\exp(-2y), \tag{8}$$

which has the general form (2) with $a = b = 2$. The cumulative distribution function $F_{1D}(y)$ is given by

$$F_{1D}(y) = 1 - (2y+1)\mathrm{e}^{-2y}, \tag{9}$$

and the moments of $f_{1D}(y)$ are immediately calculable: $\langle y\rangle_{1D} = 1$; $\langle y^2\rangle_{1D} = \frac{3}{2}$; $\langle y^3\rangle_{1D} = 3$. The most probable normalized length obtained from (8) is $y_{1D} = \frac{1}{2}$.

A simple computer simulation exercise can easily convince us about the validity of our calculations. Results in this sense are presented on Fig. 4. As an interesting observation one can realize that the distribution function for the lengths between randomly displaced centers (given by (5)) is qualitatively different from the (7) distribution function for the length of Voronoi cells (see also Ref. [25]).

3. The 2D case

Theoretical attempts to get analytical result for $f_{2D}(y)$ ($y = S/\langle S\rangle$, with S the area of Voronoi cells) in 2D failed. We thus considered Monte Carlo-type computer simulations and fitted our simulation data in different forms. In particular, we focused on a three-parameter fit in the generally accepted (1) form and tried also a simple two-parameter approximation (2) with handy $a = b$ values. It was found that the simple choice $a = b = \frac{7}{2}$ gives a visually good fit. For the normalized distribution function of Voronoi cell areas in 2D we thus proposed the

$$f_{2D}(y) = \frac{343}{15}\sqrt{\frac{7}{2\pi}}y^{5/2}\,\exp\left(-\frac{7}{2}y\right) \tag{10}$$

simple approximation. In Fig. 5 we plotted with a continuous line the curve (10) in comparison with simulation data obtained on 29.889×10^6 Voronoi cells (almost three times more than the number of cells used

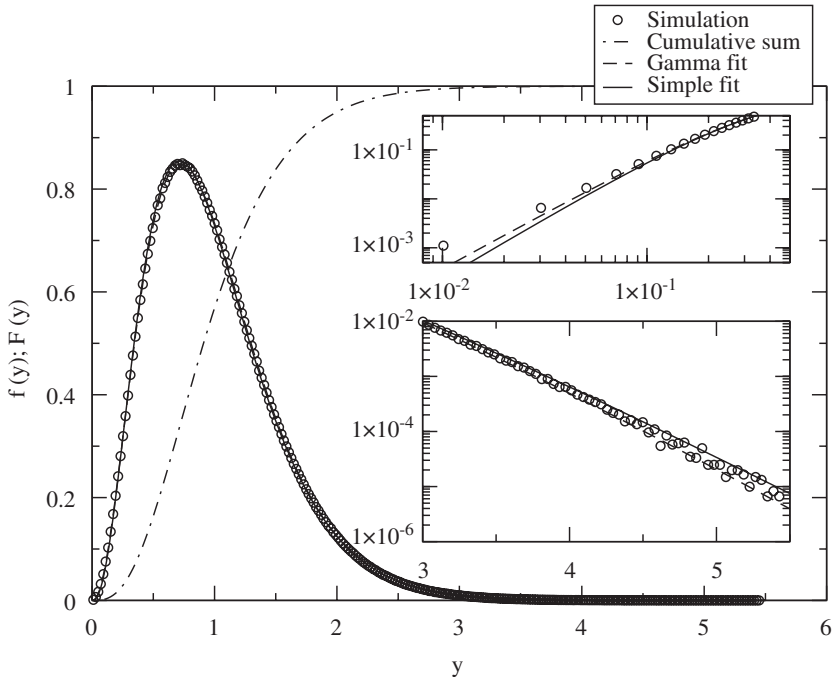

Fig. 5. The Voronoi cell's normalized area-distribution function in 2D. Empty circles are simulation results and the solid line is the (10) formula suggested in this study. With point-dashed line the cumulative distribution function, and with dashed line the (1) best gamma function fit is plotted. The two inset figures are magnification of the small and large y limits plotted on log–log and log-normal scales, respectively. On the scale of the figure there is no detectable difference between the cumulative distribution function calculated from simulation and the analytical expression given by (1) and (10).

by Tanemura). On the same graph the best gamma function (1) fit is drawn with a dashed line. The $F_{2D}(y)$ cumulative distribution function is also plotted with a dash-dotted line.

At a first glance there is no detectable difference between computer simulation results, the curve suggested by (10) and the gamma-function fit. Magnifying however the initial part and tail of the distribution function and plotting it on log–log and log-normal scales (insets in Fig. 5), respectively, one can observe slight differences. As expected, the three-parameter gamma-function fit is better, but the improvement relative to (10) is not spectacular. The best-fit parameters obtained by us for (1) are $a = 2.2975, b = 3.01116$ and $c = 1.0825$, in comparison with the values $a = 3.315$, $b = 3.04011$ and $c = 1.078$ obtained by Tanemura [18,19]. For the analytically known second moment of the distribution ($\langle y^2_{theor2D}\rangle = 1.280$) our simulation data gives $\langle y^2_{sim2D}\rangle = 1.28231$, and the three-parameter gamma-fit yields $\langle y^2_{gamma2D}\rangle = 1.27947$. The error relative to the exactly known result is of the same order (0.04%) as in the case of the fit given by Tanemura to his own computer simulation results.

Using (10) all the important moments can be analytically calculated: $\langle y\rangle_{2D} = 1$; $\langle y^2\rangle_{2D} = \frac{9}{7}$; $\langle y^3\rangle_{2D} = \frac{99}{49}$. The second moment has of course a much bigger relative error (0.4%) respective to the exactly known value than the one obtained with the more sophisticated three-parameter gamma-function fit. This relative error is, however, still quite small and usual experimental data on Poisson Voronoi type patterns give deviations of the order of a few percents. The most probable normalized area is $y_{prob2D} = \frac{5}{7}$.

4. The 3D case

Due to the complex geometry involved, the possibility to analytically calculate $f_{3D}(y)$ ($y = V/\langle V\rangle$, V the volume of Voronoi cells) in 3D is even more gloomy. Thus we again performed large-scale computer simulations, studying the statistics of 18.27×10^6 Voronoi cells (six times more than the statistics considered by Tanemura). The three-parameter gamma-function gives a good fit for the simulation data, but again as in

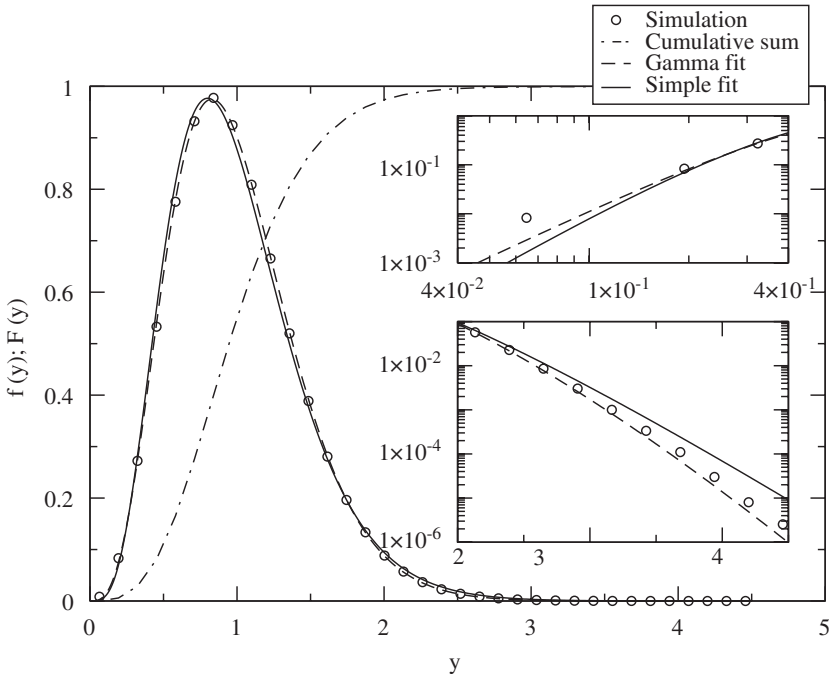

Fig. 6. The Voronoi cell's normalized volume-distribution function in 3D. Empty circles are simulation results and the solid line is the (11) formula suggested in this study. With point-dashed line the cumulative distribution function, and with dashed line the (1) best gamma function fit is plotted. The two inset figures are magnification of the small and large y limits plotted on log–log and log-normal scales, respectively. On the scale of the figure there is no detectable difference between the cumulative distribution function calculated from simulation and the analytical expression given by (1) and (11).

the 2D case a simple fit of form (2) works also reasonably well and handy $a = b = 5$ values can be considered. We suggest thus that in 3D the Voronoi cell's normalized volume distribution can be approximated as

$$f_{3D}(y) = \frac{3125}{24} y^4 \exp(-5y). \tag{11}$$

In Fig. 6 simulation results (empty circles) are compared with the (11) approximation (continuous line) and the three-parameter gamma-function fit (dashed line). In a first visual approximation one will realize that both curves describe the simulation data well. Magnifying, however, the initial part and tail of the distribution function and plotting it on log–log and log-normal scales (insets in Fig. 6), respectively, one can observe the differences. As expected, the (1) gamma-function fit is better, and follows the trend of the simulation results more. The best-fit parameters obtained in this study are $a = 3.24174$, $b = 3.24269$ and $c = 1.26861$ (in contrast with $a = 4.8065$, $b = 4.06342$ and $c = 1.16391$ found by Tanemura [18,19]). The improvement relative to the simple (11) approximation is, however, again not spectacular, and is relevant only in the limit of very large or very small Voronoi cells. These limits do not appear usually in real experimental data, due to the fact that much weaker statistics are achieved (patterns with less than 10^4 cells are studied). In the figure the form of the cumulative distribution function $F_{3D}(y)$ is also plotted (point-dashed line). In the scale of the image there is no detectable difference in the cumulative distribution function determined from simulation and forms (11) or (1).

By using (11) the important moments are analytically calculable: $\langle y\rangle_{3D} = 1$; $\langle y^2\rangle_{3D} = \frac{6}{5}$; $\langle y^3\rangle_{3D} = \frac{42}{25}$. The most probable normalized volume is $y_{prob3D} = \frac{4}{5}$. For the second moment the relative error respective to the analytical exact results ($\langle y^2_{theor3D} = 1.18$) is $\varepsilon = 1.7\%$. The gamma fit for the simulation data yields $\langle y^2_{3Dgamma}\rangle = 1.18683$ ($\varepsilon = 0.57\%$) while Tanemura's fit seems better yielding 1.17830 ($\varepsilon = 0.14\%$). The second moment computed directly from simulation data is $\langle y^2_{sim3D}\rangle = 1.19$, giving the $\varepsilon = 0.85\%$ relative error.

In agreement with the simulations of Tanemura [18,19] we have also found that the values $a = b = 6$ suggested by Kiang [23] are not appropriate and give no good fit to our simulation data.

5. Conclusions

Motivated by the simple form of the exact result (8) for the size-distribution function of Poisson Voronoi cells in 1D we proposed simple expressions for approximating the distribution in 2D and 3D where no exact results are available. Exceeding the statistics considered in all previous studies computer simulations were used to investigate numerically the distribution function. It was shown that a simple form (2) with $a = b$ is appropriate for all practical applications to approximate the size distribution of the Poisson Voronoi cells. In 1D the exact results gives $a = 2$. In 2D and 3D we found that $a = 7/2$ and $a = 5$, respectively, gives fair approximation. The simple values suggested for $a = b$ allow also to write approximations (10) and (11) in a compact form. If we denote by d the dimensionality of the problem ($d = 1, 2, 3$), the value of a can be given as $a = (3d + 1)/2$. Eqs. (8),(10) and (11) can be written then in a compact form as

$$f_d(y) = \frac{((3d+1)/2)^{(3d+1)/2}}{\Gamma((3d+1)/2)} y^{(3d-1)/2} \exp\left(-\frac{3d+1}{2} y\right). \tag{12}$$

This distribution function is not an exact one and it is less accurate than a more complicated three-parameter fit given by the generalized gamma function. Mathematicians will probably not appreciate it... but due to its simplicity it will definitely be of importance for experimental scientists studying and characterizing complex Voronoi diagram-like patterns.

Acknowledgments

This work was supported by the Romanian CNCSIS 41/183 research grant. The research of F. Járai-Szabó has also been supported by the Fellowship Program for Transborder Hungarian Scientific Research, Hungarian Academy of Science.

References

[1] A. Okabe, B. Boots, K. Sugihara, S.N. Chiu, Spatial Tessellations: Concepts and Applications of Voronoi Diagrams, Wiley, Chichester, 2000.
[2] Voronoi diagrams, in: M. de Berg, M. van Kreveld, M. Overmars (Eds.), Computational Geometry: Algorithms and Applications, second ed., Utrecht, The Netherlands, 2000, Otfried Schwarzkopf, Hong Kong, China (Chapter 7).
[3] J.L. Meijering, Philips Res. Rep. 8 (1953) 270.
[4] J.M. Drouffe, C. Itzykson, Nucl. Phys. B 235 (1984) 45.
[5] G.R. Jerauld, J.C. Hatfield, H.T. Davis, J. Phys. C 17 (9) (1984) 1519;
G.R. Jerauld, L.E. Scriven, H.T. Davis, J. Phys. C 17 (19) (1984) 3429.
[6] S.B. DiCenzo, G.K. Wertheim, Phys. Rev. B 39 (10) (1989) 6792.
[7] P.H. Winterfield, L.E. Scriven, M.T. Davis, J. Phys. C 14 (1981) 2361.
[8] Y. Talmon, S. Prager, Nature 267 (1977) 333;
Y. Talmon, S. Prager, J. Chem. Phys. 69 (7) (1978) 2984.
[9] H. Brumberger, J. Goodisman, J. Appl. Cryst. 16 (1983) 83.
[10] S. Yoshioka, S. Ikeuchi, Astrophys. J. 341 (1989) 16.
[11] I.K. Crain, Random Processes in Geology, Springer, New York, 1976.
[12] A. Okabe, A. Suzuki, Eur. J. Oper. Res. 98 (1997) 445;
S.C. Müller, T. Mair, O. Steinbock, Biophys. Chem. 72 (1998) 37;
M. Tanemura, H. Honda, A. Yoshida, J. Theor. Biol. 153 (1991) 287.
[13] W.D. Hamilton, J. Theor. Biol. 31 (1971) 295;
M. Tanemura, M. Hasegawa, J. Theor. Biol. 82 (1980) 477.
[14] B.N. Boots, Geografiska Ann. 55B (1973) 34;
B.N. Boots, Can. Geogr. 19 (1975) 107.
[15] N. Rivier, in: D. Bideau, A. Hansen (Eds.), Disorder and Granular Media, Elsevier Science Publishers B.V., New York, 1993.
[16] G. Le Caer, R. Delannay, J. Phys. I France. 3 (1993) 1777.
[17] D. Weaire, J.P. Kermode, J. Wejchert, Philos. Mag. B 53 (5) (1986) L101;
D. Weaire, N. Rivier, Contemp. Phys. 25 (1984) 59.
[18] M. Tanemura, Forma 18 (2003) 221.
[19] M. Tanemura, in: G. Lugosi, D. Nagy (Eds.), Proceedings of Intersections of Art and Science, July 8–14, 2001, Sydney ⟨http://www.mi.sanu.ac.yu/vismath/proceedings/⟩.

[20] A.L. Hinde, R.E. Miles, J. Stat. Comput. Simul. 10 (1980) 205.
[21] S. Kumar, S.K. Kurtz, J.R. Banavar, M.G. Sharma, J. Stat. Phys. 67 (1992) 523.
[22] D. Weaire, J.P. Kermode, J. Wejchert, Philas. Mag. B 53 (5) (1986) L101.
[23] T. Kiang, Z Astrophys. 64 (1966) 433.
[24] E.N. Gilbert, Ann. Math. Stat. 33 (1962) 958.
[25] M.A. Fortes, P. Pina, Philas. Mag. B 67 (1993) 263.