New Monte Carlo method for planar Poisson–Voronoi cells

2007 J. Phys. A: Math. Theor. 40 2615

(http://iopscience.iop.org/1751-8121/40/11/002)

IOP Publishing Journal of Physics A: Mathematical and Theoretical

J. Phys. A: Math. Theor. **40** (2007) 2615–2638 doi:10.1088/1751-8113/40/11/002

New Monte Carlo method for planar Poisson–Voronoi cells

H J Hilhorst

Laboratoire de Physique Théorique[1], Bâtiment 210, Université de Paris-Sud, 91405 Orsay Cedex, France

Received 16 December 2006, in final form 27 January 2007
Published 28 February 2007
Online at stacks.iop.org/JPhysA/40/2615

Abstract

By a new Monte Carlo algorithm, we evaluate the sidedness probability p_n of a planar Poisson–Voronoi cell in the range $3 \leqslant n \leqslant 1600$. The algorithm is developed on the basis of earlier theoretical work; it exploits, in particular, the known asymptotic behaviour of p_n as $n \to \infty$. Our p_n values all have between four and six significant digits. Accurate n dependent averages, second moments and variances are obtained for the cell area and the cell perimeter. The numerical large-n behaviour of these quantities is analysed in terms of an asymptotic power series in n^{-1}. Snapshots are shown of typical occurrences of extremely rare events, implicating cells of up to $n = 1600$ sides embedded in an ordinary Poisson–Voronoi diagram. We reveal and discuss the characteristic features of such many-sided cells and their immediate environment. Their relevance for observable properties is stressed.

PACS numbers: 02.50.Ey, 02.50.Ng, 02.70.Uu, 87.18.Hf, 87.18.Bb

(Some figures in this article are in colour only in the electronic version)

1. Introduction

A Voronoi diagram partitions space into convex cells constructed around a set of point-like 'seeds' or 'particles', in such a way that each point of space is in the cell of the particle to which it is closest. When the particles are distributed randomly and uniformly, the partitioning is called a Poisson–Voronoi diagram or a random Voronoi froth.

Voronoi cells play a role in science and engineering and are also of interest to mathematicians. Applications include cellular structures that either arise spontaneously in nature (e.g. in biological cellular structures, in soap froths or in granular materials) or are employed as a tool of analysis (e.g. to identify lattice defects in simulations of melting crystals). Many references are given in [1] and in the encyclopaedic monograph on tessellations by Okabe *et al* [2].

[1] Research unit UMR 8627 of the Centre National de la Recherche Scientifique.

The simplest Voronoi diagrams are of the Poisson type. It is important, therefore, that the properties of Poisson–Voronoi diagrams be understood as well as possible. Here we pursue, by means of a new Monte Carlo method, earlier investigations [1, 3, 4] on such diagrams in the Euclidean plane $\mathbb{R}^2$.

The most prominent statistical property of the planar Poisson–Voronoi cell is its 'sidedness'. We denote by p_n the probability that a cell is n-sided, for arbitrary integer $n \geqslant 3$. Other properties of interest include the average area of an n-sided cell and the average length of its perimeter, the statistics of the angles at the vertices and correlations between neighbouring cells. All these properties may be expressed as multiple integrals on the particle positions [2, 5], but only a few of them can be calculated explicitly. In particular, no simple closed form expression for p_n is known. An exact relation derived from Euler's theorem ensures that the average sidedness $\overline{n} \equiv \sum_{n=3}^{\infty} n p_n$ is equal to $\overline{n} = 6$.

It is known numerically that p_n peaks at $n = 6$ and falls off rapidly for large n. Hayen and Quine [6] have numerically evaluated the integral for p_3 with high accuracy. For $n = 4, 5, \ldots$ the values of p_n stem only from Monte Carlo work. The most accurate reported values of p_n are due to Calka [7] for $4 \leqslant n \leqslant 7$ and to Brakke [8] for $8 \leqslant n \leqslant 16$. One has $p_{16} \approx 10^{-8}$, and the largest sidedness observed in simulations by conventional algorithms is around $n = 16$. Drouffe and Itzykson [9, 10], as part of an effort to construct field theories on random lattices, developed an improved algorithm by which they estimated p_n for n up to 50. Their results, however, have error bars that for $n \gtrsim 30$ become of the same order as p_n themselves. Hence, simulating many-sided Voronoi cells has remained a challenge.

The interest of investigating Voronoi cells for asymptotically large n was stressed by Le Caër and Delannay [11]. Analytic knowledge of the large-n behaviour of p_n, apart from the insight that it provides, also constrains the laws that describe the finite n behaviour as observed in experiments and simulations. An example of this interplay between the regimes of finite and of asymptotic n is the theoretical explanation given in [4] of the failure of Aboav's law [12] for Poisson–Voronoi diagrams.

The analytic study of p_n in the limit $n \to \infty$ was taken up in [1, 3]. It was shown there, among many other things, that asymptotically

$$p_n \simeq C p_n^{(0)}, \qquad n \to \infty, \tag{1.1}$$

with $C = 0.344\,347\ldots$ [2] and

$$p_n^{(0)} = \frac{1}{4\pi^2} \frac{(8\pi^2)^n}{(2n)!}. \tag{1.2}$$

In the present work, we exploit this asymptotic knowledge. Going beyond equation (1.1) we write an equality that is exact for all n rather than merely asymptotic, namely

$$p_n = C_n p_n^{(0)}, \tag{1.3}$$

whence necessarily $\lim_{n\to\infty} C_n = C$. We focus on C_n and show that it can be expressed as an average

$$C_n = \langle \Theta\, \mathrm{e}^{-\mathbb{V}} \rangle, \tag{1.4}$$

where $\mathbb{V}$ is a known expression in the angular variables that describe the n-sided cell, and Θ is an indicator (i.e. equal to 0 or to 1) imposing a geometric constraint on the set of angles. We will determine the prefactor C_n in (1.3) by Monte Carlo evaluation of the right-hand side of equation (1.4) for finite $n = 3, 4, \ldots$. The Monte Carlo algorithm is new for this problem. Whereas all previously used methods become rapidly inefficient with increasing n,

[2] Its analytic expression is $C = \prod_{q=1}^{\infty} (1 - q^{-2} + 4q^{-4})^{-1}$.

the performance of the algorithm presented here is, very roughly, independent of n. This makes it possible, in particular, to explore the structure of Voronoi cells in the hitherto inaccessible large-n regime.

The remaining sections of this paper are as follows. In section 2 the algorithm is described. In section 3 results are presented and discussed for the sidedness probability p_n as well as for the averages and second moments of the cell perimeter and cell area. The asymptotic large-n behaviour of these quantities is analysed numerically. In section 4 we present and discuss characteristic pictures of many-sided Voronoi cells in an environment of ordinary cells. In section 5 we summarize our results.

The algorithm requires the explicit expressions for $\mathbb{V}$ and Θ in equation (1.4). Finding these is a matter of considerable technical complexity; it is based on results of [1] and is the subject of appendices A and B.

2. Monte Carlo algorithm

2.1. Context

Monte Carlo methods for generating Voronoi cells of Poisson distributed particles are discussed by Okabe *et al* [2]. One class of methods simply determines p_n as the relative frequency of occurrence of n-sided cells. But since p_n decreases to zero faster than exponentially for $n \gtrsim 12$, the statistical precision goes down accordingly. With such methods, it is hardly possible to accumulate sufficient statistics for even single-digit precision as soon as $n \approx 16$.

Another class of methods generates cells for a value of n fixed in advance. The first to have done so seem to have been Drouffe and Itzykson [9]. The method employed by Calka [7] is also in this class. These methods face the problem of attrition: a Monte Carlo-generated geometrical object, in order to represent a valid n-sided cell, must satisfy certain geometrical constraints. The probability that an attempted generation satisfies the constraints again decreases rapidly with growing n.

The present algorithm, which also fixes n in advance, completely solves the problem of attrition: the geometric constraints are satisfied with a probability that tends to unity when $n \to \infty$. In order to arrange things this way, a certain amount of rather technical rewriting of the initial problem is necessary. We have confined this rewriting to the appendices. If one accepts its results, the method is easy to apply.

2.2. Angular variables

An n-sided Voronoi cell around a particle in the origin, as shown in figure 1, is specified completely by its n vertex vectors $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_n$. It may be specified alternatively by its n mid-point vectors, i.e. the projections $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_n$ of the origin onto the sides. The explicit expression [1, 7, 9, 13] for p_n as a multiple integral on $\mathbf{R}_m$ is given in appendix A. It has not, however, been possible to evaluate this integral analytically. By choosing other sets of variables of integration one may recast the original integral in numerous different forms, none of which is simple. For our purpose, it is essential to use the angular variables that we will define now.

Let Φ_m and Ψ_m be the polar angles of $\mathbf{R}_m$ and $\mathbf{S}_m$, respectively. Other angles relevant for this problem are defined in figure 1. The quantities $\eta_m = \Psi_{m+1} - \Psi_m$ are the angles between two consecutive vertex vectors and the quantities $\xi_m = \Phi_m - \Phi_{m-1}$ are those between two consecutive mid-point vectors; n-periodicity in the index m is understood. For fixed sets $\xi = \{\xi_m\}$ and $\eta = \{\eta_m\}$, one may still jointly rotate the set of vertex vectors with respect to

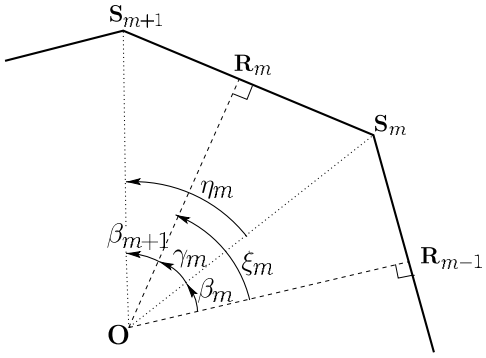

Figure 1. Heavy line: the perimeter of the Voronoi cell around a particle in the origin **O**. The dashed and dotted lines connect the origin to the midpoints $\mathbf{R}_m$ and vertices $\mathbf{S}_m$, respectively. The particles of the neighbouring cells are located at $2\mathbf{R}_1, \ldots, 2\mathbf{R}_n$. The right angles have been marked. The figure defines the sets of angles ξ_m, η_m, β_m and γ_m.

the set of mid-point vectors: this modifies only the relative angles β_m and γ_m between the two sets. We may select any one of these relative angles and call it 'the' angle of rotation, since it will determine all others; we will select β_1. When for a generic β_1 we draw the cell boundary by clockwise constructing its successive segments, then after a turn of 2π it appears not close onto itself but to spiral. A 'no-spiral condition' must therefore determine the appropriate value of the angle of rotation $\beta_1 = \beta_*(\xi, \eta)$ for which the cell boundary closes. This condition reads [1]

$$G(\xi, \eta; \beta_*) = 0, \tag{2.1}$$

where

$$\mathrm{e}^{2\pi G} = \prod_{m=1}^{n} \frac{\cos\gamma_m}{\cos\beta_m}. \tag{2.2}$$

One may note that equation (2.2) involves β_m and γ_m that are themselves determined by the solution $\beta_1 = \beta_*$ of (2.1). For an arbitrary pair (ξ, η), there need not exist a solution to equation (2.1). In appendix B we show that it has a solution, which moreover is unique, if and only if

$$\max_{1\leqslant m\leqslant n}\left[\sum_{\ell=1}^{m-1}(\xi_\ell - \eta_\ell) + \xi_m\right] - \min_{1\leqslant m\leqslant n}\left[\sum_{\ell=1}^{m-1}(\xi_\ell - \eta_\ell)\right] < \pi, \tag{2.3}$$

which is a criterion expressed entirely in terms of the supposedly given sets ξ and η.

After these preliminaries we return to (1.4). The symbol Θ in that expression denotes the indicator function of the domain in (ξ, η) space where (2.3) is satisfied. Finally, the 'interaction' $\mathbb{V}$ in (1.4) is given explicitly in terms of the angular variables in appendix A through a sequence of definitions, equations (A.10) and (A.5)–(A.8), that we will not display here.

2.3. Algorithm for determining p_n

The sidedness probability p_n is given by equations (1.2)–(1.4). We determine it numerically by evaluating $\langle\Theta\,\mathrm{e}^{-\mathbb{V}}\rangle$ as follows. We fix the sidedness n, after which the simulation proceeds according to the five steps given below.

(i) We draw $n-1$ random numbers uniformly distributed on $[0, 1]$ and order them. After multiplication by 2π, this gives[3] $0 < \bar{\Psi}_1 < \bar{\Psi}_2 < \cdots < \bar{\Psi}_{n-1} < 2\pi$. We set $\bar{\Psi}_n = 2\pi$ and choose

$$\begin{aligned} \eta_m &= \bar{\Psi}_{m+1} - \bar{\Psi}_m, \qquad m = 1, \ldots, n-1, \\ \eta_n &= \bar{\Psi}_1. \end{aligned} \tag{2.4}$$

We next draw $2n-1$ random numbers, order them and discard those of odd rank so that only $n-1$ are left. After multiplication by 2π this gives $0 < \bar{\Phi}_1 < \bar{\Phi}_2 < \cdots < \bar{\Phi}_{n-1} < 2\pi$. We set $\bar{\Phi}_0 = 0$ and choose

$$\begin{aligned} \xi_m &= \bar{\Phi}_m - \bar{\Phi}_{m-1}, \qquad m = 1, \ldots, n-1, \\ \xi_n &= 2\pi - \bar{\Phi}_{n-1}. \end{aligned} \tag{2.5}$$

(ii) We check if the pair of sets (ξ, η) thus obtained satisfies equation (2.3). If so, then we know that there exists $\beta_*(\xi, \eta)$ which may be determined from equation (2.1); hence $\Theta = 1$ and we proceed with (iii). If not, then it is impossible to satisfy equation (2.1), we have $\Theta = 0$, and the attempt to generate an n-sided cell fails. We increase the attempt counter by one unit and return to (i).

(iii) We solve $\beta_*(\xi, \eta)$ from equation (2.1) by a numerical iteration procedure which also yields the derivative $G'(\xi, \eta; \beta_*)$ needed in the next step.

(iv) We calculate the weight $\exp(-\mathbb{V})$ according to equations (A.10) and (A.5)–(A.8) of appendix A and add it to the accumulated weight. We increase the attempt counter by one unit and return to (i).

(v) In the end the total accumulated weight is divided by the total number of attempts, including those that failed. The result is an estimate for p_n.

We remark that the successive cells generated by this procedure are all statistically independent.

2.4. Algorithm for n dependent averages

The simulation method described above allows us to study arbitrary cell properties $F(\mathbf{R}_1, \ldots, \mathbf{R}_n)$. Writing $\langle F \rangle_n$ for the average of F subject to a given sidedness n, we have

$$\langle F \rangle_n = \frac{\langle I_F \Theta \, \mathrm{e}^{-\mathbb{V}} \rangle}{\langle \Theta \, \mathrm{e}^{-\mathbb{V}} \rangle}. \tag{2.6}$$

Here the numerator, which generalizes (1.4), has an insertion I_F that derives from F by a radial integration. We recall that the average $\langle \cdots \rangle$, defined in (A.15), applies to quantities that depend exclusively on the angular variables. To find I_F from F, we set $R_{\mathrm{av}} = n^{-1} \sum_{m=1}^{n} R_m$. We may then express the ratios $\rho_m = R_m / R_{\mathrm{av}}$ entirely in terms of the angular variables (see appendix A). Then, if F is of dimension d_F, it may be factorized into a radial and an angular part according to

$$F(\mathbf{R}_1, \ldots, \mathbf{R}_n) = \left(R_{\mathrm{av}}^2 / 4\lambda \right)^{d_F/2} \hat{F}(\xi, \eta), \tag{2.7}$$

where we show explicitly the areal particle density λ which had heretofore been scaled away[4]. When (2.7) is integrated over the radial scale R_{av}, an extra factor appears as compared to the same operation for p_n and we find

$$I_F = \frac{\Gamma\left(n + \frac{1}{2} d_F\right)}{\Gamma(n)} W^{-d_F/2} \hat{F}, \tag{2.8}$$

[3] The angles $\bar{\Psi}_m$ are equal to the angles Ψ_m of figure 1 up to a common additive constant which drops out of equation (2.4). An analogous remark applies to $\bar{\Phi}_m$.

[4] A factor 4λ instead of λ appears because the particle positions were defined as $2\mathbf{R}_m$ instead of $\mathbf{R}_m$.

where Γ denotes the gamma function and where we abbreviated

$$W = 4\lambda\pi(1 + n^{-1}V) \tag{2.9}$$

with V given by (A.8).

We will limit ourselves to considering the first and second moments of two quantities that are frequently encountered in applications and that have therefore been the subject of much earlier work, namely the cell perimeter P and the cell area A. These are explicitly given by

$$P = R_{\mathrm{av}}(4\lambda)^{-1/2}\hat{F}_1, \qquad A = R_{\mathrm{av}}^2(4\lambda)^{-1}\hat{F}_2, \tag{2.10}$$

with the angular factors

$$\hat{F}_k = \frac{1}{k}\sum_{m=1}^{n} \rho_m^k(\tan\gamma_m + \tan\beta_{m+1}), \qquad k = 1, 2. \tag{2.11}$$

Setting successively $F = P, P^2, A, A^2$ we find that the corresponding insertions in the numerator of equation (2.6) are

$$\begin{aligned} I_P &= \left[\Gamma\left(n + \tfrac{1}{2}\right)\Big/\Gamma(n)\right]W^{-1/2}\hat{F}_1, \\ I_{P^2} &= nW^{-1}\hat{F}_1^2, \\ I_A &= \tfrac{1}{2}nW^{-1}\hat{F}_2, \\ I_{A^2} &= \tfrac{1}{4}n(n+1)W^{-2}\hat{F}_2^2. \end{aligned} \tag{2.12}$$

The simulation steps for finding the numerator of equation (2.6) are the same as for p_n except that (iv) and (v) are respectively replaced with (iv′) and (v′) given below.

(iv′) We multiply the insertion I_F of the quantity F of interest by the weight $\exp(-\mathbb{V})$ and accumulate the value thus obtained.

(v′) In the end the total accumulated value is divided by the total number of attempts and by the estimate obtained for p_n. This provides an estimate for $\langle F\rangle_n$. The numerical data shown will all be for $\lambda = 1$.

3. Results and discussion

3.1. The distribution of $\mathbb{V}$ *and the indicator* Θ

Before discussing our results for the sidedness probability p_n, we briefly consider the quantities $\mathbb{V}$ and Θ that via (1.3) and (1.4) enter into their definition. Let $P(\mathbb{V})$ denote the probability distribution of $\mathbb{V}$ and $\phi_n \equiv \langle\Theta\rangle$ the probability for an attempted cell generation to be successful. In terms of these, we may rewrite (1.4) as

$$C_n = \phi_n \int \mathrm{d}\mathbb{V}\, P(\mathbb{V})\, \mathrm{e}^{-\mathbb{V}}, \tag{3.1}$$

which exhibits the important intermediate role of $P(\mathbb{V})$ and ϕ_n.

In order to show what $P(\mathbb{V})$ looks like, we have plotted its logarithm in figure 2 for $n = 50, 100, 200$ and 400. The curves clearly demonstrate that for $n \to \infty$ there is convergence to a limit. For $\mathbb{V} \to \pm\infty$ the limit distribution appears to decay exponentially, $P(\mathbb{V}) \sim \exp(-\kappa_\pm|\mathbb{V}|)$, but with very different decay constants: we obtain $\kappa_+ = 0.185 \pm 0.005$ from a fit in the range $3 \leqslant \mathbb{V} \leqslant 30$ followed by extrapolation to $n = \infty$, and $\kappa_- = 2.47 \pm 0.02$ from a fit in the range $-3 \leqslant \mathbb{V} \leqslant -1.5$. This large $\mathbb{V}$ behaviour has not yet been explained theoretically.

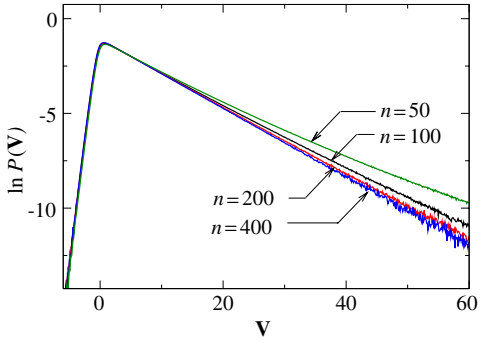

Figure 2. Logarithm of the probability distribution $P(\mathbb{V})$ of $\mathbb{V}$ (see equation (3.1)) for four different values of n, showing convergence to a limit distribution for $n = \infty$.

The Θ function in (1.4) imposes constraint (2.3) and is at the origin of the failed generation attempts. Whereas these do not contribute to p_n in step (iv) of the algorithm above, equation (3.1) shows that via $\langle\Theta\rangle = \phi_n$ they do enter into the determination of its normalization. In the last column of table 1 we list the fractions ϕ_n of successful attempts as determined from the simulation. Although ϕ_n is equal only to $\phi_3 = 0.058$ for $n = 3$, it turns out to rise rapidly with n, is already as high as $\phi_{10} = 0.8$ for $n = 10$, and tends to unity for $n \to \infty$. That is, attrition disappears in the large-n limit.

This brings out the two key steps that are responsible for the success of the present algorithm: (i) the limit distribution of $P(\mathbb{V})$ has become n independent since we extracted from the initial expression for p_n the appropriate n dependent prefactor $p_n^{(0)}$ given in (1.2) and (ii) attrition disappears for large n because of our choice of the angles (ξ, η) as the variables of integration.

3.2. Sidedness probability p_n

In table 1 we present the results for the sidedness probability p_n for n in the range between $n = 3$ and $n = 1600$. They are based on a number N_n of generation attempts given in the fourth column of that table.

The second column of table 1 shows the best results for p_n found in the literature for each value of n. The p_3 value was obtained by numerical integration; the other p_n are Monte Carlo results. For $p_4, \ldots, p_7$ the statistical error is in the last decimal; for higher n the standard deviations are indicated.

Our own results for p_n, given in the third column of table 1, are accurate up to absolute errors of order less than 10^{-5}. Standard deviations were calculated by subdividing the data into 20 or more groups and considering the dispersion of the group averages. We will now discuss these results as a function of n.

Case $n = 3$. The probability p_3 for a cell to be three-sided is the only one that has been evaluated by numerical integration. This was done by Hayen and Quine [6], who reduced the original integral to a four-dimensional one. They present a 12-digit result of which the second column of table 1 shows only the first seven significant decimals. For $n = 3$ we performed an especially long run with the purpose of testing our Monte Carlo method and checking the result of [6]. As shown in table 1, our method reproduces six significant digits of Hayen and Quine's result and leaves their value within our error bars.

Table 1. The sidedness probability p_n. Second column: literature data taken from Hayen and Quine [6] for p_3; from Calka [7] for $p_4, \ldots, p_7$; from Brakke [8] for $p_8, \ldots, p_{15}$ and from Drouffe and Itzykson [9] for p_n with $n \geqslant 16$. Third column: p_n and its standard deviation calculated by the Monte Carlo method of this work. Fourth column: number N_n of cell generation attempts. Fifth column: fraction ϕ_n of successful attempts.

	Literature [6–9]	This work		
n	p_n	p_n	N_n	ϕ_n
3	$1.124\,001\ldots \times 10^{-2}$	$(1.124\,000 \pm 0.000\,021) \times 10^{-2}$	1.2×10^{10}	0.0580
4	$1.068\,38 \times 10^{-1}$	$(1.068\,454 \pm 0.000\,025) \times 10^{-1}$	2×10^{9}	0.1730
5	2.5946×10^{-1}	$(2.594\,44 \pm 0.000\,07) \times 10^{-1}$	1.6×10^{9}	0.3077
6	2.9473×10^{-1}	$(2.947\,23 \pm 0.000\,09) \times 10^{-1}$	2×10^{9}	0.4391
7	1.9877×10^{-1}	$(1.987\,68 \pm 0.000\,07) \times 10^{-1}$	4×10^{8}	0.5564
8	$(9.0116 \pm 0.0020) \times 10^{-2}$	$(9.0131 \pm 0.0006) \times 10^{-2}$	10^{8}	0.6554
9	$(2.9644 \pm 0.0012) \times 10^{-2}$	$(2.9652 \pm 0.0002) \times 10^{-2}$	8×10^{7}	0.7361
10	$(7.4471 \pm 0.0059) \times 10^{-3}$	$(7.4487 \pm 0.0006) \times 10^{-3}$	8×10^{7}	0.8002
11	$(1.4796 \pm 0.0026) \times 10^{-3}$	$(1.4818 \pm 0.0002) \times 10^{-3}$	6×10^{7}	0.8501
12	$(2.409 \pm 0.011) \times 10^{-4}$	$(2.4000 \pm 0.0002) \times 10^{-4}$	6×10^{7}	0.8884
13	$(3.18 \pm 0.04) \times 10^{-5}$	$(3.2324 \pm 0.0003) \times 10^{-5}$	6×10^{7}	0.9175
14	$(3.60 \pm 0.13) \times 10^{-6}$	$(3.6835 \pm 0.0004) \times 10^{-6}$	4×10^{7}	0.9393
15	$(3.7 \pm 0.4) \times 10^{-7}$	$(3.6017 \pm 0.0004) \times 10^{-7}$	4×10^{7}	0.9556
16	$(2.3 \pm 0.3) \times 10^{-8}$	$(3.0574 \pm 0.0004) \times 10^{-8}$	4×10^{7}	0.9677
17		$(2.2762 \pm 0.0002) \times 10^{-9}$	4×10^{7}	0.9765
18	$(1.3 \pm 0.5) \times 10^{-10}$	$(1.4989 \pm 0.0002) \times 10^{-10}$	4×10^{7}	0.9830
19		$(8.7983 \pm 0.0013) \times 10^{-12}$	4×10^{7}	0.9878
20	$(1.5 \pm 0.8) \times 10^{-13}$	$(4.6314 \pm 0.0004) \times 10^{-13}$	8×10^{7}	0.9912
21		$(2.1994 \pm 0.0004) \times 10^{-14}$	2×10^{7}	0.9937
22		$(9.4835 \pm 0.0017) \times 10^{-16}$	2×10^{7}	0.9955
23		$(3.7227 \pm 0.0005) \times 10^{-17}$	2×10^{7}	0.9968
24		$(1.3379 \pm 0.0003) \times 10^{-18}$	2×10^{7}	0.9977
25	$(9.6 \pm 5.9) \times 10^{-21}$	$(4.4184 \pm 0.0004) \times 10^{-20}$	4×10^{7}	0.9984
30	$(1.3 \pm 1.1) \times 10^{-29}$	$(5.4595 \pm 0.0005) \times 10^{-28}$	4×10^{7}	0.9997
40	2.4×10^{-50}	$(6.7349 \pm 0.0006) \times 10^{-46}$	8×10^{7}	1.0000
50	1.5×10^{-75}	$(5.223 \pm 0.001) \times 10^{-66}$	1.6×10^{7}	1.0000
60		$(7.192 \pm 0.002) \times 10^{-88}$	1.2×10^{7}	1.0000
70		$(3.4805 \pm 0.0004) \times 10^{-111}$	3×10^{7}	1.0000
80		$(9.598 \pm 0.002) \times 10^{-136}$	10^{7}	1.0000
90		$(2.1616 \pm 0.0005) \times 10^{-161}$	0.8×10^{7}	1.0000
100		$(5.2691 \pm 0.0006) \times 10^{-188}$	1.6×10^{7}	1.0000
150		$(1.0535 \pm 0.0002) \times 10^{-332}$	4×10^{6}	1.0000
200		$(3.818 \pm 0.001) \times 10^{-492}$	4×10^{6}	1.0000
300		$(1.084 \pm 0.001) \times 10^{-841}$	2×10^{6}	1.0000
400		$(9.863 \pm 0.003) \times 10^{-1221}$	4×10^{6}	1.0000
600		$(3.645 \pm 0.002) \times 10^{-2040}$	10^{6}	1.0000
800		$(1.326 \pm 0.001) \times 10^{-2918}$	2×10^{6}	1.0000
1000		$(6.365 \pm 0.003) \times 10^{-3841}$	1.6×10^{6}	1.0000
1200		$(3.262 \pm 0.002) \times 10^{-4798}$	1.2×10^{6}	1.0000
1400		$(1.385 \pm 0.001) \times 10^{-5784}$	0.8×10^{6}	1.0000
1600		$(7.4306 \pm 0.0020) \times 10^{-6796}$	4×10^{6}	1.0000

For all $n > 3$ the literature results are based on Monte Carlo evaluation.

Cases $n = 4$ through $n = 7$. The most accurate literature results in this intermediate regime are due to Calka [7], whose algorithm like ours fixed n in advance. Our results are fully compatible with those of [7].
Cases $n = 8$ through $n = 15$. The Monte Carlo results obtained in the 1980s by Brakke [8] for $3 \leqslant n \leqslant 16$ have long remained unsurpassed. Our simulations confirm all of Brakke's results. Beyond $n \approx 10$ the accuracy of the Monte Carlo algorithm of [8] rapidly goes down with increasing n, and for $n = 16$ its error bars are as large as the result itself. This effect is just due to the low relative frequency of cells of so many sides, the number n not being fixed in this method. By contrast, the accuracy of our method, for a fixed amount of computer time invested per value of n, stays roughly constant.

Case $n \geqslant 16$. Drouffe and Itzykson [9] developed a more powerful simulation method aimed at simulating cells of larger sidedness. In their method, n is again fixed in advance. Their accuracy amounts to roughly a single significant digit in the regime $16 \leqslant n \lesssim 25$; for $n \gtrsim 25$ the error becomes again of the order of p_n itself. This error increase is due to attrition, i.e. an increasing rejection rate of configurations that are generated but do not satisfy the required geometrical constraints. From our data it appears that for $n \gtrsim 25$ the authors [9] miss the true values by an ever larger factor and that only their logarithmic order of magnitude is right[5]. Again, the method of the present work maintains an error in the fifth digit, i.e. a relative error not larger than 10^{-4}.

Case of extremely large n. The range of n from 50 to 1600 had so far been an unexplored territory. In this very large-n regime attrition is negligible and, for a constant calculational effort per value of n, the method keeps producing results with an error only in the fourth significant digit. The probabilities p_n are extremely small. Numerically, we could easily handle such small numbers by first factorizing out $p_n^{(0)}$ of which we computed and stored only the logarithm. As discussed in subsection 3.1, the remaining factor $C_n = \langle \Theta \, \mathrm{e}^{-\mathbb{V}} \rangle$ has a finite distribution and hence causes no underflow problems.

Generating the values of such 'unphysically' small probabilities is much more than a mere technical achievement. First, it provides another check that the programme works correctly; indeed we find that for $n \to \infty$ the ratio $p_n / C p_n^{(0)} = C_n / C$ tends to unity as it should. Second, it gives access to the large-n expansion of p_n to be discussed in subsection 3.3. Thirdly and most importantly, values of n this large are required to see the separation of length scales that occurs in the many-sided cell; this is the subject of section 4.

Sum rules. The probabilities p_n should obey certain sum rules. Upon summing p_n of table 1 and writing $\overline{X_n} = \sum_{n=3}^{\infty} X_n p_n$, we find

$$\begin{aligned} &\sum_{n=3}^{\infty} p_n = 1.000\,010(15), \\ &\overline{n} = 6.0001(1), \\ &\overline{n^2} = 37.7816(7), \\ &\mu \equiv \overline{n^2} - \overline{n}^2 = 1.7804, \end{aligned} \tag{3.2}$$

with an error in μ at most equal to ± 0.0015 but probably smaller due to partial cancellation of the errors in $\overline{n}^2$ and $\overline{n^2}$. The first and second relations of (3.2) may be compared to the exactly known values 1 and 6, respectively. The second moment $\overline{n^2}$ has an exact expression as

[5] For a figure comparing their data to our asymptotic result, we refer to [3].

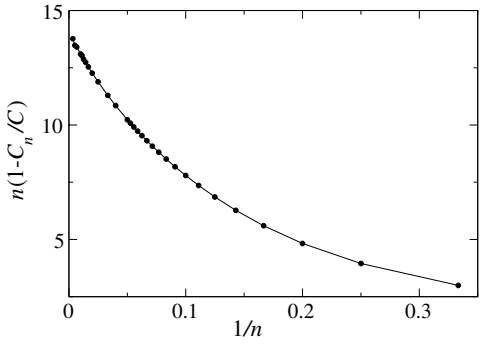

Figure 3. To study the asymptotic large-n behaviour of the sidedness probability $p_n = C_n p_n^{(0)}$, we plot the quantity $n(1 - C_n/C)$ where $C = \lim_{n\to\infty} C_n$ (see equations (1.2) and (1.3)). The solid line connects the data points. Data shown are in the range $3 \leqslant n \leqslant 300$. The largest error bars occur for small $1/n$ and are of the order of the data symbols. The intercept of the curve with the vertical axis is the coefficient e_1 of the leading correction term in the expansion of equation (3.3).

a double integral [14], which when evaluated numerically gives $\overline{n^2} = 37.780\,811\ldots$. Hence $\mu = 1.780\,811\ldots$ numerically exactly. We therefore see that, when their error bars are taken into account, our Monte Carlo data are in excellent agreement with these sum rules.

Conclusion. The general conclusion of this subsection is that for low n (say $n \lesssim 8$), our method is probably as good as several of the existing ones. If we did slightly outperform them in that small n regime, that was only because of the length of our runs. However, for larger N (say $n \gtrsim 8$) our method has a decisive advantage over the existing ones.

3.3. Asymptotic behaviour of p_n

On the basis of the numerical data, we will now discuss the asymptotic behaviour of p_n for $n \to \infty$. Analytically it is known that $p_n = C_n p_n^{(0)}$ with $p_n^{(0)}$ given by equation (1.2) and where the correction factor C_n may be obtained by a series expansion that classifies contributions according to their power in $n^{-1/2}$. On that basis, Hilhorst [3] fitted the limited p_n data available at that time (essentially $n \lesssim 30$) by a correction term proportional to $n^{-1/2}$. It remained possible, however, that the coefficient of the $n^{-1/2}$ term would cancel, and indeed Drouffe and Itzykson [9] had hypothesized earlier that the leading correction was of the order n^{-1}.

The numerical data of this work now indicate unambiguously that the series is actually one in powers of n^{-1},

$$p_n = \frac{C}{4\pi^2}\frac{(8\pi^2)^n}{(2n)!}\left[1 - \frac{e_1}{n} + \frac{e_2}{n^2} - \frac{e_3}{n^3} + \cdots\right], \tag{3.3}$$

where $e_1, e_2, \ldots,$ are numerical coefficients.

The factor in square brackets in equation (3.3) is equal to C_n/C, which for $n \to \infty$ is known to tend to unity. In figure 3, in order to find the corrections to the leading order term in (3.3), we have plotted $n(1 - C_n/C) = e_1 - e_2 n^{-1} + \cdots$ against n^{-1}. This figure shows that the intercept with the vertical axis is located at $e_1 = 14.00 \pm 0.05$. We may now proceed by subtracting this estimated value of e_1 from the curve of figure 3, multiply it again by n and look for a new intercept with the vertical axis which, if it is well defined, is equal to $-e_2$.

Upon iterating until the statistical errors obscure a well-defined intercept, we obtained in this way estimates for the first few e_i. The uncertainties increase with the index i. We found

$$e_1 = 14.00 \pm 0.05, \qquad e_2 = 94 \pm 4, \qquad e_3 = 375 \pm 80, \tag{3.4}$$

in which the errors are correlated: the values deviate together upward or downward. The important conclusion is that p_n has a series expansion in powers of n^{-1}. The cancellation of the half-integer powers in the expansion of [1] is no doubt due to a symmetry in the theory that still remains to be identified.

In a final remark we wish to stress that finding this asymptotic expansion is different from finding a 'best fit', which we do not attempt here. The curve of figure 3 is close to the sum of a constant and an exponential in n^{-1}, but we have no reason to believe that there exists a simple analytical expression that fits all data within their error bars.

3.4. Perimeter and area

We have simulated the two cell properties that have received the greatest attention in the literature, namely the cell perimeter P and the cell area A. We determined the average, the second moment and the variance of both of these quantities as a function of n. The perimeter results are summarized in table 2 and the area results in table 3.

Similar tables extracted from the literature were compiled by Okabe *et al* [2]. However, by far the most accurate ones appear in an unpublished work by Brakke [8] and concern the regime $3 \leqslant n \leqslant 16$. All our area and perimeter data are compatible with those of [8], but our error bars are strongly reduced. A further check on the numerical data is provided by two more sum rules,

$$\overline{P_n} = 4.000\,035(65), \qquad \overline{A_n} = 1.000\,02(2), \tag{3.5}$$

for which the exact values are 4 and 1, respectively.

We now turn to the large-n behaviour. Our data indicate the expansions

$$\begin{aligned} \langle P\rangle_n &= (\pi n)^{\frac{1}{2}} + a_{\frac{1}{2}} n^{-\frac{1}{2}} + a_{\frac{3}{2}} n^{-\frac{3}{2}} + \cdots, \\ \langle P^2\rangle_n &= \pi n + b_0 + b_1 n^{-1} + \cdots, \\ \langle A\rangle_n &= \tfrac{1}{4} n + c_0 + c_1 n^{-1} + \cdots, \\ \langle A^2\rangle_n &= \left(\tfrac{1}{4} n\right)^2 + d_{-1} n + d_0 + \cdots, \end{aligned} \tag{3.6}$$

which again go down by integer powers of n. They imply that the variances have the series

$$\begin{aligned} \langle P^2\rangle_n - \langle P\rangle_n^2 &= b_0 - 2\pi^{\frac{1}{2}} a_{\frac{1}{2}} + \cdots \\ \langle A^2\rangle_n - \langle A\rangle_n^2 &= \left(d_{-1} - \tfrac{1}{2} c_0\right) n + \cdots. \end{aligned} \tag{3.7}$$

The leading terms in each of the four series of equation (3.6) are known from theoretical analysis [1, 3]. Heuristically they follow from the sole observation that in the large-n limit the Voronoi cell becomes a circle of radius $R_c = (n/4\pi)^{1/2}$. Theoretical analysis can in principle also produce the higher order terms in (3.6), but this has not been attempted yet. Consequently, the leading coefficients of the series in (3.7) are not known analytically. Here, our simulation results provide answers.

In figure 4 we have plotted $[\langle \delta P^2\rangle_n/\pi]^{1/2}$ and $2[\langle \delta A^2\rangle_n/n]^{1/2}$. The numerical data strongly point to limit values equal to $\frac{1}{2}$ for both quantities when $n \to \infty$. Conjecturing that these limits are exact, we then conclude that

$$b_0 - 2\pi^{\frac{1}{2}} a_{\frac{1}{2}} = \tfrac{1}{4}\pi, \qquad d_{-1} - \tfrac{1}{2} c_0 = \tfrac{1}{16}. \tag{3.8}$$

Table 2. Estimates of the average $\langle P\rangle_n$, the second moment $\langle P^2\rangle_n$ and the root-mean-square fluctuation $\langle \delta P^2\rangle_n^{1/2}$ of the cell perimeter P. The numbers in parentheses represent the standard deviation in the last digit. The entries of the third column have an error of at most one unit in their last digit. The limit value $\frac{1}{2}\pi^{1/2} = 0.886\,226\ldots$ for $n = \infty$ has the status of a conjecture.

n	$\langle P\rangle_n$		$\langle P^2\rangle_n$		$\langle \delta P^2\rangle_n^{1/2}$
3	2.740 296	(2)	8.171 30	(2)	0.813 68
4	3.219 524	(3)	11.048 19	(2)	0.826 34
5	3.642 658	(3)	13.966 26	(3)	0.835 04
6	4.026 307	(4)	16.919 58	(4)	0.841 69
7	4.380 000	(6)	19.902 72	(6)	0.847 02
8	4.710 196	(8)	22.910 84	(8)	0.851 40
9	5.020 869	(12)	25.940 26	(12)	0.855 06
10	5.315 211	(11)	28.987 90	(12)	0.858 16
11	5.595 488	(10)	32.050 43	(12)	0.860 78
12	5.863 536	(11)	35.125 88	(13)	0.863 04
13	6.120 72	(2)	38.2114	(2)	0.864 97
14	6.368 24	(1)	41.3055	(2)	0.866 64
15	6.607 05	(2)	44.4066	(2)	0.868 09
16	6.837 97	(2)	47.5136	(3)	0.869 36
17	7.061 73	(2)	50.6258	(2)	0.870 47
18	7.278 84	(2)	53.7410	(3)	0.871 45
19	7.489 92	(2)	56.8598	(3)	0.872 33
20	7.695 44	(1)	59.9820	(3)	0.873 10
21	7.895 76	(3)	63.1066	(4)	0.873 80
22	8.091 18	(3)	66.2318	(5)	0.874 42
23	8.282 15	(2)	69.3596	(3)	0.874 99
24	8.468 92	(3)	72.4890	(4)	0.875 51
25	8.651 71	(2)	75.6198	(4)	0.875 98
30	9.513 79	(2)	91.2825	(5)	0.877 83
40	11.039 71	(1)	122.6501	(4)	0.880 05
50	12.379 83	(2)	154.0370	(5)	0.881 35
60	13.588 87	(3)	185.4355	(7)	0.882 19
70	14.698 96	(2)	216.8384	(6)	0.882 79
80	15.731 05	(2)	248.2460	(7)	0.883 23
90	16.699 51	(3)	279.6545	(8)	0.883 57
100	17.614 87	(2)	311.0645	(8)	0.883 84
150	21.618 17	(3)	468.1280	(12)	0.884 65
200	24.988 33	(3)	625.1998	(11)	0.885 05
300	30.636 07	(2)	939.3527	(12)	0.885 44
400	35.393 84	(2)	1253.5092	(13)	0.885 64
600	43.370 90	(4)	1881.820	(3)	0.885 84
800	50.093 46	(2)	2510.143	(2)	0.885 93
1000	56.014 90	(1)	3138.456	(2)	0.885 99
1200	61.367 64	(1)	3766.773	(2)	0.886 03
1400	66.289 53	(2)	4395.087	(4)	0.886 06
1600	70.870 47	(2)	5023.408	(2)	0.886 08
∞					0.886 226 ...

Analysis of the $\langle P^2\rangle_n$ data to next order suggests that $\langle P^2\rangle_n/(\pi n) - 1$ tends to -1 as $n \to \infty$. Conjecturing that this, too, is exact and combining it with the first one of equations (3.8) we arrive at

$$a_{\frac{1}{2}} = -\tfrac{5}{8}\pi^{\frac{1}{2}}, \qquad b_0 = -\pi. \tag{3.9}$$

Table 3. Estimates of the average $\langle A\rangle_n$, the second moment $\langle A^2\rangle_n$ and the normalized root-mean-square fluctuation $n^{-1/2}\langle \delta A^2\rangle_n^{1/2}$ of the cell area A. The numbers in parentheses represent the standard deviation in the last digit. The entries of the third column have an error of at most one unit in their last digit. The limit value $\frac{1}{4}$ for $n = \infty$ has the status of a conjecture.

n	$\langle A\rangle_n$		$\langle A^2\rangle_n$		$n^{-1/2}\langle \delta A^2\rangle_n^{1/2}$
3	0.343 087	(3)	0.161 573	(3)	0.120 92
4	0.558 052	(4)	0.401 285	(5)	0.149 89
5	0.774 080	(4)	0.736 75	(1)	0.165 86
6	0.995 789	(5)	1.179 53	(2)	0.176 98
7	1.222 51	(1)	1.735 16	(3)	0.185 41
8	1.453 28	(1)	2.407 24	(3)	0.192 00
9	1.687 36	(1)	3.198 47	(4)	0.197 56
10	1.924 08	(2)	4.110 64	(7)	0.202 13
11	2.162 95	(2)	5.1451	(1)	0.205 99
12	2.403 66	(2)	6.3033	(1)	0.209 29
13	2.645 78	(2)	7.5854	(1)	0.212 17
14	2.889 06	(3)	8.9920	(1)	0.214 69
15	3.133 31	(3)	10.5234	(1)	0.216 89
16	3.378 35	(3)	12.1797	(2)	0.218 85
17	3.624 16	(2)	13.9619	(2)	0.220 59
18	3.870 34	(3)	15.8680	(2)	0.222 15
19	4.117 03	(3)	17.8996	(3)	0.223 57
20	4.364 15	(3)	20.0570	(3)	0.224 84
21	4.611 58	(4)	22.3394	(3)	0.226 01
22	4.859 23	(4)	24.7464	(4)	0.227 06
23	5.107 15	(4)	27.2790	(4)	0.228 03
24	5.355 31	(5)	29.9371	(5)	0.228 91
25	5.603 58	(6)	32.7198	(6)	0.229 74
30	6.846 86	(6)	48.5090	(6)	0.233 06
40	9.339 13	(4)	89.470	(1)	0.237 23
50	11.834 58	(7)	142.932	(2)	0.239 77
60	14.331 83	(6)	208.900	(2)	0.241 46
70	16.829 79	(6)	287.363	(3)	0.242 67
80	19.3283	(1)	378.331	(4)	0.243 58
90	21.827 26	(8)	481.800	(4)	0.244 29
100	24.326 27	(6)	597.763	(3)	0.244 85
150	36.8236	(2)	1365.10	(1)	0.246 58
200	49.3224	(1)	2444.94	(1)	0.247 42
300	74.3214	(2)	5542.16	(3)	0.248 26
400	99.3206	(1)	9889.33	(3)	0.248 71
600	149.3196	(3)	22 333.58	(1)	0.249 14
800	199.3201	(2)	39 778.23	(6)	0.249 35
1000	249.3192	(2)	62 222.3	(1)	0.249 48
1200	299.3193	(2)	89 666.7	(1)	0.249 57
1400	349.3187	(3)	122 110.8	(2)	0.24963
1600	399.3190	(2)	159 555.4	(2)	0.249 68
∞					0.250 000 . . .

We do not attempt, however, a similar analytical conjecture for the second pair of coefficients, c_0 and d_{-1}, nor will we pursue estimates for the higher order coefficients in the series (3.6) and (3.7), except below in connection with Lewis' law.

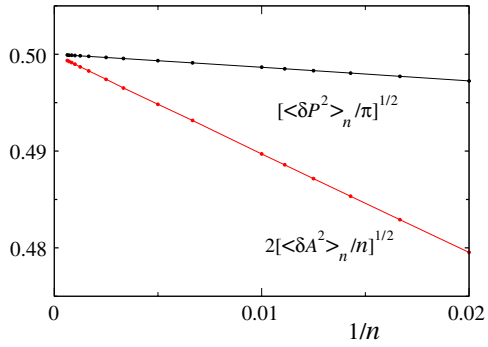

Figure 4. The asymptotic large-n behaviour of the root-mean-square fluctuations $\langle \delta P^2 \rangle_n^{1/2}$ and $\langle \delta A^2 \rangle_n^{1/2}$ of the perimeter and the area, respectively, of the n-sided cell. Data shown are in the range $50 \leqslant n \leqslant 1600$. The error bars are smaller than the data points. The two curves that connect the data points are asymptotically straight lines which for $n \to \infty$ both appear to converge $\frac{1}{2}$.

Lewis' celebrated law [15] is an empirical statement about one of the cell's most conspicuous properties, namely the relation between its area and its number of sides. The law states that the average area $\langle A \rangle_n$ of an n-sided cell increases with n as

$$\langle A \rangle_n = \frac{a_0}{\lambda}(n - n_0), \tag{3.10}$$

where a_0 and n_0 are constants and where we have displayed again the dependence on the areal particle density λ. Sometimes (see the discussion in [2]) this law is written in a more restricted one-parameter form

$$\langle A \rangle_n = \frac{b(n-6)-1}{\lambda}. \tag{3.11}$$

It is found, however, that $\langle A \rangle_n$ deviates from linearity with n in simulations of Poisson–Voronoi diagrams as well as in the data from most experimental systems. We now look at what the asymptotic analysis has to say.

In [1, 3] we proved that asymptotically

$$\langle A \rangle_n \simeq \pi R_c^2 = \frac{n}{4\lambda}, \qquad n \to \infty, \tag{3.12}$$

and this result has been incorporated in the series for $\langle A \rangle_n$ in (3.6). A coefficient $a_0 \approx \frac{1}{4}$ had since long been suspected by various authors [9, 16, 17]. Going now beyond (3.12) and proceeding in the same way as for p_n, we can determine the coefficients of the series of (3.6) for $\langle A \rangle_n$ on the basis of our simulation results of table 2. This yields

$$c_0 = -0.6815(5), \qquad c_1 = 0.750(5), \qquad c_2 = 3.15(10). \tag{3.13}$$

We now consider the laws (3.10) and (3.11). The fact that we found $c_1, c_2, \ldots$ to be nonvanishing confirms once more that $\langle A \rangle_n$ is not strictly linear in n. From the above it follows that in (3.10) one should choose

$$n_0 = -4c_0 = 2.7260(4) \tag{3.14}$$

if one wishes it to correctly represent the asymptotic behaviour of $\langle A \rangle_n$ for Poisson–Voronoi cells. This is of course different from finding a best fit to a limited set of $\langle A \rangle_n$ data in a restricted n interval. If that is the purpose, the values of a_0 and n_0 will depend on the available data and on the way the fit is carried out. The one-parameter law (3.11) postulates a relation

between a_0 and n_0 that is violated in the asymptotic expansion. Hence (3.11) cannot be used to describe the large-n behaviour of $\langle A\rangle_n$ and merely has the status of an empirical fit to the data, the value of b again depending on the data set and on how the fit is done.

4. Characteristic cell shapes

It has been established [1, 3] that in the large-n limit, the Voronoi cell tends to a disc of radius $R_c = (n/4\pi)^{1/2}$.[6] In [1] it was furthermore shown that the cell perimeter undergoes 'elastic' deformations from circularity, the elasticity being, of course, of an entropic origin. The probability law of these deformations was given in the large-n limit. In this section we show how our Monte Carlo method allows us what was hitherto impossible, namely to simulate for any finite n the detailed statistics of the cell shape.

We Monte Carlo generated cells of prescribed sidedness n in a 'natural', that is, an unbiased, environment. This was done as follows. For a given n the cell angles (ξ, η) were drawn randomly, and β_* was found according to the rules of section 2.3. The cell radius was taken equal to its most probable value $R_c = (n/4\pi)^{1/2}$ and the cell boundary was constructed. This boundary, together with the position of the central particle, fixes the positions of the n first-neighbour particles. We then determined the cell's fundamental domain $\mathcal{F}$, that is, the union of the n discs of radius S_m centred at the vertices $\mathbf{S}_m$. The complement of $\mathcal{F}$ in a large rectangle of suitable size was subsequently filled randomly with particles of a uniform density equal to 1. The particles added by this procedure are all necessarily second or higher order neighbours of the central one. The Voronoi construction was finally applied to the full collection of particles to complete the Voronoi cell diagram.

4.1. Cells of $n = 24, 48$ *and* 96 *sides*

We have generated typical cell shapes for a sequence of values of n, starting with $n = 3$ and doubling n each time. Figures 5–7, in which the dots represent the particles, show the results for cells of $n = 24, 48$ and 96 neighbours. The three pictures are at different scales, but all have unit particle density. This picture sequence illustrates the tendencies that characterize many-sided cells. One tendency is for the first neighbour cells to be elongated. This feature is apparent already for $n = 24$ and becomes very pronounced for $n = 48$, whereas the $n = 96$ cell has *only* very elongated neighbours. The same phenomenon was observed by Lauritsen *et al* [17], but in a different system. These authors consider Poisson–Voronoi diagrams to which they assign an 'energy' that favours many-sided cells. Snapshots of their configurations show a dense structure of many-sided cells (of sidednesses higher than $n = 60$) separated by mostly four-sided elongated cells. Their procedure does not, however, provide estimates for p_n in an unbiased Poisson–Voronoi diagram.

Another tendency, similarly appearing in [17], is for the first-neighbour particles to align on what tends towards a continuous curve. Whereas for $n = 24$ some imagination is still needed to see this curve, it becomes clearly distinguishable for $n = 48$ and is immediately obvious to the eye for $n = 96$. The typical distance between nearest neighbour particles along this curve decreases as $2\pi(2R_c)/n \sim n^{-1/2}$. We note that whereas Voronoi cells are always convex, the 'curve' connecting the first neighbours need not enclose a convex area; in each of the figures 5–7, there are small but clearly distinguishable deviations from convexity.

[6] The approach of a Voronoi cell to a disc has been rigorously proved [18–20] under the condition of the cell becoming large (e.g., by letting its area tend to infinity), rather than many-sided ($n \to \infty$). The two conditions lead to distinctly different statistical ensembles.

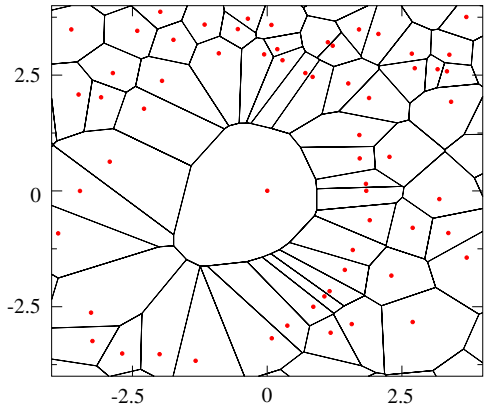

Figure 5. A typical Voronoi cell with $n = 24$ neighbours. The dots represent the particles.

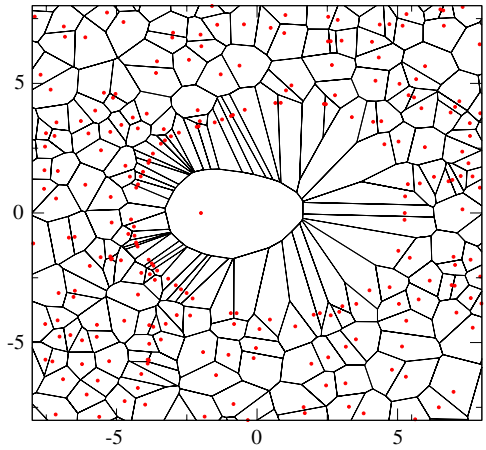

Figure 6. A typical Voronoi cell with $n = 48$ neighbours.

Figure 8 enlarges a detail of figure 7 and shows a collection of first-neighbour cells. All first neighbours fully visible in the figure are four-sided except those marked A, B, D, E, which are five-sided, and the cell C, which is either five- or six-sided (this depends on how the two almost coinciding three-vertices are arranged at the point marked '2V'; a higher resolution is needed to decide this question). Figure 8 illustrates that in the large-n limit four-sided first neighbours become dominant. In [4] it was argued that five-sided cells constitute a fraction only of order $n^{-1/2}$ of all first-neighbour cells, and that the probability of six- and higher-sided first neighbours is still of higher order in $n^{-1/2}$. In figure 8 the cell marked P is a second neighbour to the central cell. The boundary separating it from the first neighbours has been shown as a heavy solid line on which we will further comment shortly.

4.2. Very large cells

Focusing now on very large n, we show in figure 9 a central particle located in the origin and having 1536 neighbours. As before, the dots represent the particles. The inner contour, which

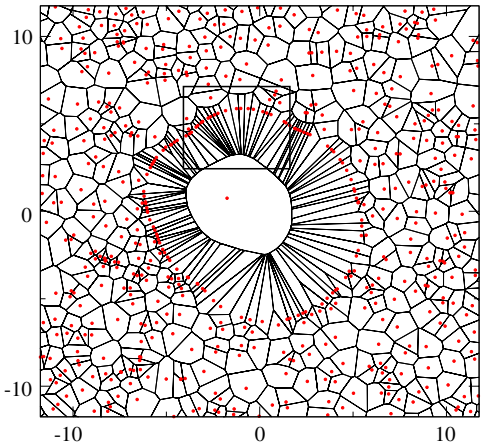

Figure 7. A typical Voronoi cell with $n = 96$ neighbours. The region inside the box is shown enlarged in figure 8.

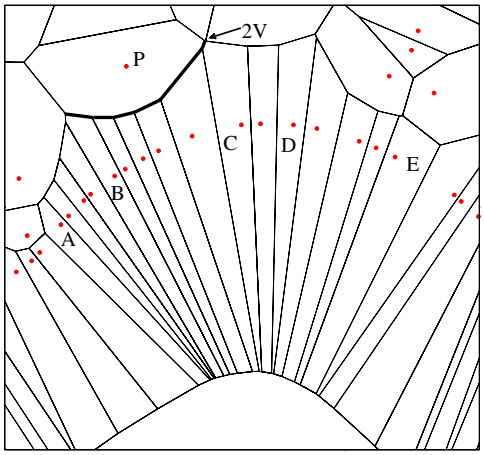

Figure 8. Enlargement of the box in figure 7, showing some of the strongly elongated first neighbours of the central cell. Among the first neighbours fully visible, cells A, B, C, D and E have more than four sides. The arrow marked '2V' points to two three-vertices that coincide within the resolution of the figure. Cell P is a second neighbour whose boundary with the first neighbours (heavy lines) is an example of an 'incipient parabola segment' (see the text).

is nearly indistinguishable from a circle of radius R_c, represents the boundary of the Voronoi cell of the central particle. The outer 'curve', which is also very close to circular but of radius $2R_c$, represents the alignment of the 1536 first-neighbour particles. Their high line density gives the impression of a continuous curve. Cell boundaries other than those of the central cell have not been drawn; they would totally blacken the empty annular region between the boundary of the central cell and its first-neighbour particles.

The boxed region in figure 9 is shown enlarged in figure 10, where we did draw all Voronoi cell boundaries. The extreme elongation of the first-neighbour cells is what first strikes the eye. The discrete structure of the 'curve' of first neighbours is also apparent now. The

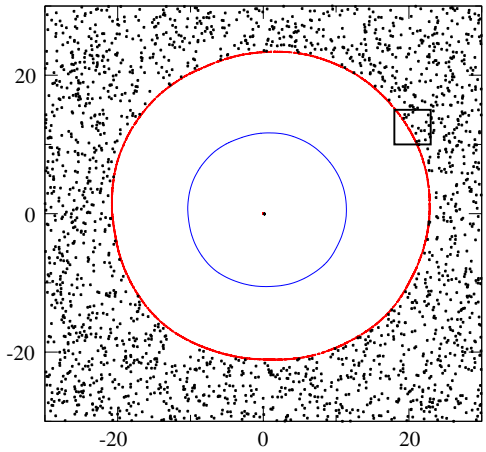

Figure 9. Approach to the infinite n limit. The origin is occupied by a particle whose Voronoi cell has $n = 1536$ sides. The almost circular inner curve is the cell boundary of the central cell. The other cell boundaries have not been drawn. The almost circular outer curve is made up of 1536 first-neighbour particles. The region inside the box is shown enlarged in figure 10.

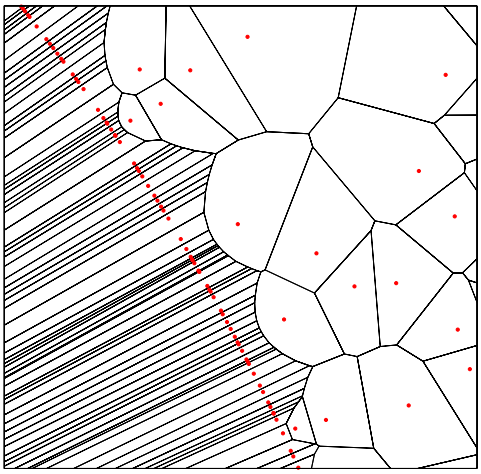

Figure 10. Enlargement of the box in figure 9, where now all cell boundaries have been drawn. The discrete structure of the outer 'curve' of figure 9 has become visible here.

distances ℓ_m between successive first-neighbour particles along this curve are of order $n^{-1/2}$. More precisely, if we set $\ell_m = \lambda_m(4\pi/n)^{1/2}$, then the theory [1] implies that for $n \to \infty$ the qualities λ_m are independent identically distributed random variables of probability law $\lambda_m \exp(-\lambda_m)$. Random deviations from a local straight line are too small to be discernible to the eye; they may be argued [21] to decrease as $n^{-3/2}$, which is also the order of magnitude of the systematic deviations due to the radius of curvature $2R_c$. The large cell and its environment are characterized, therefore, by four different length scales, each varying with its own power of n. They have been summarized in table 4. One has to go to n values as high as we did in order for the separation of scales to become clearly visible.

Table 4. Four length scales characterizing the n-sided Voronoi cell in the large-n limit; see figures 9 and 10.

Scale	Length
$n^{1/2}$	Cell radius
n^{0}	Typical interparticle distance outside the first-neighbour circle
$n^{-1/2}$	Typical distance between successive first-neighbour particles
$n^{-3/2}$	Random deviations of first neighbours from full alignment

Very large n is also still required for another feature to become apparent. In figure 10 the boundary between a second-neighbour cell and its adjacent first-neighbour cells is, by construction, composed of points equidistant to the second-neighbour particle and the almost continuous straight line of first-neighbour particles. But such a boundary is a parabola. Hence, in the limit $n \to \infty$ the boundary separating the set of first from the set of second-neighbour cells is *piecewise parabolic*. Indeed, with this observation in mind one now recognizes the boundary segment of cell P in figure 8 (heavy solid line), and others in that same figure, as 'incipient parabolic'. Such knowledge was at the basis of the theory of two-cell correlations exposed in [4]. There, laws discovered in the $n \to \infty$ limit were extrapolated backward and shown to be relevant for finite n. It was shown, in particular, that Aboav's linear relationship [2, 12] between n and the total average sidedness nm_n of the neighbours of an n-sided cell cannot hold in Poisson–Voronoi diagrams. We expect that in the future the study of large cells will shed further light also on various issues relevant for the finite n behaviour.

5. Summary and conclusion

In this paper we have developed a new Monte Carlo method for evaluating the sidedness probability p_n for arbitrary n. The method, which is constructed on the basis of an extension of earlier theory [1, 3], is not difficult to implement once the rather complicated analytic expressions that intervene are available.

We have determined p_n as well as the first and second moments of the n dependent cell perimeter and cell area. A full agreement is obtained with earlier results for p_n due to Hayen and Quine [6], Calka [7] and Brakke [8], whose data extend up to $n = 16$. For $n \gtrsim 10$ we have reduced the error bars on p_n very considerably. In the range up to $n = 50$, we improved and corrected the p_n data due to Drouffe and Itzykson [9]. For $50 < n \leqslant 1600$, we obtained data in a range that had so far remained inaccessible. This enabled us to investigate the asymptotic large-n behaviour of p_n and of the perimeter and area moments. On the basis of our numerical results we conclude that they all have asymptotic series in entire powers of n^{-1}, possibly up to an overall prefactor $n^{\frac{1}{2}}$.

Exploiting our full control of the cell statistics, we have exhibited occurrences of extremely rare many-sided cells in a typical environment of ordinary cells. Their embedding involves four distinct length scales, varying with four different powers of n. This has confirmed, among several other things, the very elongated shape of the first-neighbour cells.

No particular effort was made to optimize the code. Our total investment of computer time on a recent PC model was limited to a few hundred hours and allowed us to obtain p_n to at least four or five significant decimals for the set of n values listed in the tables. We have also not attempted to provide any 'best fits' to the numerical curves, as we have no reason to believe that there exist simple analytic expressions that fit the data within our error bars over their full range.

The Monte Carlo work of this paper became possible only after initial analytic progress [1, 3]. We believe that it will in return spur further analytic investigation. One branch of study may concern the nature of the asymptotic expansions uncovered here. Another one may deal with correlations between a cell and its second, third and higher topological neighbours, which are a recurrent theme in the theory and applications of Voronoi cells.

Acknowledgments

The author thanks Daniela de Oliveira Maionchi for providing him with the computer program that draws the Voronoi cells of a given set of particles. He thanks Pierre Calka for discussions and for referring [6] to him.

Appendix A. Theory

We present here the extension of an earlier work that opens the way to the numerical simulations of this work. We consider uniformly and independently distributed particles in the plane. Let a particle be placed in the origin and let n other particles occupy the positions $2\mathbf{R}_1, 2\mathbf{R}_2, \ldots, 2\mathbf{R}_n$. The sidedness probability p_n of the cell containing the origin may then be written as a $2n$-fold integral on the midpoint coordinates [7, 9, 10, 13],

$$p_n = \frac{1}{n!}\int \mathrm{d}\mathbf{R}_1 \ldots \mathrm{d}\mathbf{R}_n\, \chi(\mathbf{R}_1, \ldots, \mathbf{R}_n)\, \mathrm{e}^{-\mathcal{A}(\mathbf{R}_1,\ldots,\mathbf{R}_n)}. \tag{A.1}$$

Here the indicator function χ is equal to unity (or to zero) on the domain of phase space where the perpendicular bisectors of the vectors $2\mathbf{R}_m$, for $m = 1, 2, \ldots, n$, define an n-sided (or a fewer-sided) cell around the origin, and $\mathcal{A}$ is the two-dimensional volume of the area that should be void of particles if this cell is not to be intersected by any of the bisectors of the position vectors of the remaining particles. Explicit expressions for $\mathcal{A}$ and χ are given in [1, 7].

A.1. Starting point

After one integrates over a common radial scale, expression (A.1) takes the form [1] of an integral on the angle β_1 and on the sets of angles $\xi = \{\xi_m\}$ and $\eta = \{\eta_m\}$,

$$p_n = \frac{(n-1)!}{2n}\int_{-\pi/2}^{\pi/2}\mathrm{d}\beta_1 \int \mathrm{d}\xi\, \mathrm{d}\eta\, \delta\left(\sum_{m=1}^{n}\xi_m - 2\pi\right)\delta\left(\sum_{m=1}^{n}\eta_m - 2\pi\right)$$
$$\times \frac{\delta(\beta_1 - \beta_*)}{G'(\xi, \eta; \beta_*)}\left[\prod_{m=1}^{n}\rho_m^2 T_m\right][\pi(1 + n^{-1}V)]^{-n}, \tag{A.2}$$

where G and β_* are as defined in the main text (equations (2.2) and (2.1)) and the derivative $G' = \mathrm{d}G/\mathrm{d}\beta_1$ is given explicitly by (B.8) and (B.4) of appendix B. The definitions of the new symbols occurring in (A.2) follow below.

First of all, the symbol $\int \mathrm{d}\xi\, \mathrm{d}\eta$ in (A.2) is shorthand for the nested integrations

$$\int \mathrm{d}\xi\, \mathrm{d}\eta = \int_0^{\pi/2+\beta_1}\mathrm{d}\xi_1 \int_0^{\pi/2+\gamma_1}\mathrm{d}\eta_1 \int_0^{\pi/2+\beta_2}\mathrm{d}\xi_2$$
$$\ldots \int_0^{\pi/2+\gamma_{n-1}}\mathrm{d}\eta_{n-1}\int_0^{\pi/2+\beta_n}\mathrm{d}\xi_n \int_0^{2\pi}\mathrm{d}\eta_n. \tag{A.3}$$

The notation is hybrid; the variables $\gamma_1, \beta_2, \gamma_3, \ldots, \beta_n$ occurring here should be viewed as functions of ξ, η and the 'angle of rotation' β_1. They are given by

$$\begin{aligned} \beta_m &= \beta_*(\xi,\eta) - \sum_{\ell=1}^{m-1}(\xi_\ell - \eta_\ell), \\ \gamma_m &= -\beta_*(\xi,\eta) + \sum_{\ell=1}^{m-1}(\xi_\ell - \eta_\ell) + \xi_m, \qquad m = 1,\ldots,n, \end{aligned} \tag{A.4}$$

where $\sum_{\ell=1}^{0}$ denotes an empty sum. Next, T_m and ρ_m are functions of γ_m and β_m given by

$$T_m = \frac{\sin(\beta_m+\gamma_m)}{\cos^2\beta_m}, \qquad m = 1,\ldots,n, \tag{A.5}$$

$$\rho_m = \frac{\cos\gamma_m \cos\gamma_{m-1}\ldots\cos\gamma_1}{\cos\beta_m\cos\beta_{m-1}\ldots\cos\beta_1}\rho_n, \qquad m = 1,\ldots,n-1, \tag{A.6}$$

and the condition

$$n^{-1}\sum_{m=1}^{n}\rho_m = 1. \tag{A.7}$$

Finally, V is given by

$$\begin{aligned} V = {} & \frac{n}{4\pi}\sum_{m=1}^{n}\rho_m^2[\tan\gamma_m - \gamma_m + \tan\beta_{m+1} - \beta_{m+1} + \gamma_m\tan^2\gamma_m + \beta_{m+1}\tan^2\beta_{m+1}] \\ & + \frac{n}{2\pi}\sum_{m=1}^{n}\left(\rho_m^2 - 1\right)(\gamma_m + \beta_{m+1}). \end{aligned} \tag{A.8}$$

The factor n included in its definition makes that, typically, V is of order n^0 as $n \to \infty$. This completes the definition of the multiple integral (A.2) for p_n.

A.2. Transformations

We now depart from the development of [1] and transform expression (A.2) as follows. We integrate over β_1 and henceforth when writing β_1 it will be understood that it takes the value $\beta_1 = \beta_*(\xi,\eta)$. The integration requires that equation (2.1) has a solution. In [1] a unique solution was shown to exist perturbatively for large n; in appendix B of the present work we provide the demonstration for general n. We furthermore replace the upper integration limits of the integrals over ξ_m and η_m by ∞ at the expense of introducing Heaviside theta functions. Using that $\xi_m - \beta_m = \gamma_m$ and $\eta_m - \gamma_m = \beta_{m+1}$ and introducing a factor $\theta\left(\frac{\pi}{2} - \beta_1\right)$, which may be done for free, we find that equation (A.2) may be converted into

$$\begin{aligned} p_n = {} & \frac{(n-1)!}{2n\pi^n}\int_0^\infty \mathrm{d}\xi_1\,\xi_1\ldots\mathrm{d}\xi_n\,\xi_n\int_0^\infty \mathrm{d}\eta_1\ldots\mathrm{d}\eta_n \\ & \times\delta\left(\sum_{m=1}^{n}\xi_m - 2\pi\right)\delta\left(\sum_{m=1}^{n}\eta_m - 2\pi\right)\Theta\,\mathrm{e}^{-\mathbb{V}}, \end{aligned} \tag{A.9}$$

in which

$$\mathrm{e}^{-\mathbb{V}} = G'(\xi,\eta;\beta_*)^{-1}\left[\prod_{m=1}^{n}\rho_m^2 T_m\xi_m^{-1}\right](1+n^{-1}V)^{-n} \tag{A.10}$$

and

$$\Theta = \prod_{m=1}^{n} \theta\left(\frac{\pi}{2} - \beta_m\right) \prod_{m=1}^{n} \theta\left(\frac{\pi}{2} - \gamma_m\right). \tag{A.11}$$

Expression (A.9) is more symmetric than (A.2)–(A.3). Its integrand is a function exclusively of ξ_m and η_m. We have purposefully included extra weights ξ_m in the integrations in (A.9) and compensated for these by factors ξ_m^{-1} in the product on m in (A.10). In this way we obtain the property that $T_m \xi_m$ remains finite when $\xi_m \to 0$, which was desirable analytically [1, 3] and is also necessary numerically.

The same quantity $\mathbb{V}$ as defined in (A.10) was studied analytically in [1, 3], where it was shown that for $n \to \infty$ it remains, typically, of order n^0.

One further rewriting is useful. We set

$$\xi_m = \alpha_{2m-1} + \alpha_{2m}, \qquad m = 1, 2, \ldots, n, \tag{A.12}$$

and use that

$$\int_0^\infty \mathrm{d}\alpha_{2m-1}\, \mathrm{d}\alpha_{2m}\, f(\alpha_{2m-1} + \alpha_{2m}) = \int_0^\infty \mathrm{d}\xi_m\, \xi_m f(\xi_m) \tag{A.13}$$

for any function $f(\xi_m)$. This converts (A.9) into the final result

$$p_n = p_n^{(0)} \langle \Theta\, \mathrm{e}^{-\mathbb{V}} \rangle, \tag{A.14}$$

where for any function X of the angular variables the average $\langle X \rangle$ is defined by

$$\langle X \rangle = \mathcal{N} \int_0^\infty \mathrm{d}\alpha_1 \ldots \mathrm{d}\alpha_{2n} \int_0^\infty \mathrm{d}\eta_1 \ldots \mathrm{d}\eta_n \delta\left(\sum_{m=1}^{2n} \alpha_m - 2\pi\right) \delta\left(\sum_{m=1}^{n} \eta_m - 2\pi\right) X, \tag{A.15}$$

where $\mathcal{N} = (n-1)! \big/ \left[2n\pi^n p_n^{(0)}\right]$. The factor $p_n^{(0)}$ extracted from the right-hand side of (A.14) is easily calculated as

$$\begin{aligned} p_n^{(0)} &= \frac{(n-1)!}{2n\pi^n} \left[\int_0^\infty \mathrm{d}\alpha_1 \ldots \mathrm{d}\alpha_{2n} \delta\left(\sum_{m=1}^{2n} \alpha_m - 2\pi\right)\right] \left[\int_0^\infty \mathrm{d}\eta_1 \ldots \mathrm{d}\eta_n \delta\left(\sum_{m=1}^{n} \eta_m - 2\pi\right)\right] \\ &= \frac{(n-1)!}{2n\pi^n} \times \frac{(2\pi)^{2n-1}}{(2n-1)!} \times \frac{(2\pi)^{n-1}}{(n-1)!} \\ &= \frac{(8\pi^2)^n}{4\pi^2 (2n)!}, \end{aligned} \tag{A.16}$$

which is (1.2). This way of arriving at $p_n^{(0)}$ is slightly simpler than the original calculation of [1]. Expressions (A.14)–(A.15) are new and are at the basis of the Monte Carlo simulation of this work. The integrals in (A.15) directly suggest step (i) of the algorithm of subsection (2.3).

Appendix B. The equation $G = 0$

We discuss here the function G defined by

$$\mathrm{e}^{2\pi G} = \frac{\cos\gamma_1 \cos\gamma_2 \ldots \cos\gamma_n}{\cos\beta_1 \cos\beta_2 \ldots \cos\beta_n}. \tag{B.1}$$

The transformation to angular variables in appendix A led to equation (A.11), i.e. to the upper limits of integration $\beta_m, \gamma_m < \frac{\pi}{2}$. Since $\beta_m + \gamma_m = \xi_m$ and since $\xi_m > 0$, we have in fact that in the integral for p_n the angles β_m and γ_m are restricted by

$$-\frac{\pi}{2} < \beta_m, \gamma_m < \frac{\pi}{2}. \tag{B.2}$$

Hence $G \to -\infty$ whenever any angle γ_m tends to $\pm\frac{\pi}{2}$, and $G \to \infty$ whenever any β_m tends to $\pm\frac{\pi}{2}$. We now set

$$\begin{aligned} \beta_m &= \tilde{\beta}_m + \beta_1, \\ \gamma_m &= \tilde{\gamma}_m - \beta_1, \qquad m = 1, \ldots, n, \end{aligned} \tag{B.3}$$

where $\tilde{\beta}_m$ and $\tilde{\gamma}_m$ are functions of ξ_m and η_m that may be read off by a comparison of equations (B.3) and (A.4), respectively,

$$\begin{aligned} \tilde{\beta}_m &= -\sum_{\ell=1}^{m-1}(\xi_\ell - \eta_\ell), \\ \tilde{\gamma}_m &= \sum_{\ell=1}^{m-1}(\xi_\ell - \eta_\ell) + \xi_m, \qquad m = 1, \ldots, n, \end{aligned} \tag{B.4}$$

where again $\sum_{\ell=1}^{0}$ denotes the empty sum. Making all β_1 dependence explicit, we get

$$\mathrm{e}^{2\pi G} = \frac{\cos(\tilde{\gamma}_1 - \beta_1)\cos(\tilde{\gamma}_2 - \beta_1)\ldots\cos(\tilde{\gamma}_n - \beta_1)}{\cos(\tilde{\beta}_1 + \beta_1)\cos(\tilde{\beta}_2 + \beta_1)\ldots\cos(\tilde{\beta}_n + \beta_1)}, \tag{B.5}$$

which we wish to study as a function of the single variable β_1, at fixed (ξ, η). Expression (B.5) shows that $\exp(2\pi G)$ is positive on the interval

$$-\frac{\pi}{2} + \max_{1\leqslant m\leqslant n} \tilde{\gamma}_m < \beta_1 < \frac{\pi}{2} - \max_{1\leqslant m\leqslant n} \tilde{\beta}_m, \tag{B.6}$$

provided this interval is not empty, that is, provided

$$\max_{1\leqslant m\leqslant n} \tilde{\gamma}_m + \max_{1\leqslant m\leqslant n} \tilde{\beta}_m < \pi. \tag{B.7}$$

Because of the preceding discussion, G approaches $-\infty$ and ∞ as β_1 approaches the left and right hand end points of this interval, respectively. To show that G is actually monotonous in β_1 on the interval (B.6), it suffices to analyse the derivative

$$\frac{\mathrm{d}G}{\mathrm{d}\beta_1} = \frac{1}{2\pi}\sum_{m=1}^{n}[\tan(\tilde{\gamma}_m - \beta_1) + \tan(\tilde{\beta}_m + \beta_1)]. \tag{B.8}$$

Since $\tilde{\beta}_m + \tilde{\gamma}_m = \xi_m > 0$, it follows that there are three cases, namely (i) $\tilde{\beta}_m, \tilde{\gamma}_m > 0$; (ii) $\tilde{\beta}_m > 0$, $\tilde{\gamma}_m < 0$ and (iii) $\tilde{\beta}_m < 0$, $\tilde{\gamma}_m > 0$. By considering each of them separately, one deduces that the summand in equation (B.8) is always positive. It follows that $\mathrm{d}G/\mathrm{d}\beta_1 > 0$ and hence that $G = 0$ has a unique solution $\beta_1 = \beta_*(\xi, \eta)$ in the interval (B.6).

Hence we have shown that the conditions $\xi_m, \eta_m > 0$ and $\beta_m, \gamma_m < \frac{\pi}{2}$ suffice for the equation $G = 0$ to have a unique solution β_* in the physical interval (B.6). This condition involves, however, the angles β_m and γ_m which are determined by the solution β_*. We would like to have a criterion for the existence of a solution in terms of the sole sets (ξ, η) given at the outset. By retracking the solution method, we see that it is valid for all (ξ, η) as long as the 'physical' interval (B.6) is not empty, that is, as long as equation (B.7) is satisfied. When made explicit with the aid of (B.4), equation (B.7) becomes condition (2.3) of the main text.

References

[1] Hilhorst H J 2005 *J. Stat. Mech.* P09005
[2] Okabe A, Boots B, Sugihara K and Chiu S N 2000 *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams* 2nd edn (Cichester: Wiley)
[3] Hilhorst H J 2005 *J. Stat. Mech.* L02003

[4] Hilhorst H J 2006 *J. Phys. A: Math. Gen.* **39** 7227
[5] Meijering J L 1953 *Philips Res. Rep.* **8** 270
[6] Hayen A and Quine M P 2000 *J. Stat. Comput. Simul.* **67** 351
[7] Calka P 2003 *Adv. Appl. Probab.* **35** 863
[8] Brakke K A 1986 unpublished. Available on http://www.susqu.edu/brakke/aux/downloads/200.pdf
[9] Drouffe J M and Itzykson C 1984 *Nucl. Phys.* B **235 [FS11]** 45
[10] Itzykson C and Drouffe J M 1989 *Statistical Field Theory* vol 2 (Cambridge: Cambridge University Press) chapter 11
[11] Le Caër G and Delannay R 1993 *J. Phys. A: Math. Gen.* **26** 3931
[12] Aboav D A 1970 *Metallography* **3** 383
[13] Calka P 2003 *Adv. in Appl. Probab.* **35** 551
[14] Finch S R 2003 unpublished. Available on http://algo.inria.fr/csolve/vi.pdf. Addendum to *Mathematical Constants* Encyclopedia of Mathematics and its Applications 94 (Cambridge: Cambridge University Press)
[15] Lewis F T 1928 *Anat. Records* **38** 341
Lewis F T 1930 *Anat. Records* **47** 59
Lewis F T 1931 *Anat. Records* **50** 235
[16] Miles R E and Maillardet R J 1982 *J. Appl. Probab.* A **19** 97
[17] Lauritsen K B, Moukarzel C and Herrmann H J 1993 *J. Phys. I France* **3** 1941
[18] Calka P 2002 *Adv. Appl. Probab.* **34** 702
[19] Hug D, Reitzner M and Schneider R 2004 *Adv. Appl. Probab.* **36** 667
[20] Calka P and Schreiber T 2005 *Ann. Probab.* **33** 1625
[21] Hilhorst H J unpublished