Planar line processes for void and density statistics in thin stochastic fibre networks

C.T.J. Dodson School of Mathematics University of Manchester Manchester, M60 1QD, UK. W.W. Sampson School of Materials University of Manchester Manchester, M60 1QD, UK.

Abstract

Using results for the distribution of perimeters of random polygons arising from random lines in a plane, we obtain new analytic approximations to the distributions of areas and local line densities for random polygons and compute various limiting properties of random polygons. Using simulation, we show that the lengths of adjacent sides of polygons generated by random line processes in the plane are correlated with $\rho = 0.616 \pm 0.001$.

Keywords: Poisson line process, Random polygons, line density, polygon area distribution, limiting distributions, approximations. **2000 MSC**: 60D05, 60G55

1 Introduction

Random lines with uniformly distributed orientation and each passing through a point distributed according to a Poisson point process in a plane, partition the space with random polygons and much is known about such processes. There are analytic results of Miles [1, 2] and Tanner [3] (cf. also Stoyan et al. [4]) for infinite random lines in a plane, for example:

- Expected number of sides per polygon $\bar{n} = 4$
- Variance of the number of sides per polygon $var(n) = \frac{\pi^2 + 24}{2}$
- Perimeter P of polygons with n sides has a χ^2 distribution with 2(n-2) degrees of freedom and probability density function q given by

$$q(P,n) = \frac{P^{n-3} e^{-P/2}}{2^{n-2} \Gamma(n-2)}, \quad n = 3, 4, \dots$$
 (1.1)



Figure 1: Probability density functions for perimeter P of random polygons with n=3,4,5,6 sides; P plotted in units of mean polygon side length.

where P is given as a multiple of the mean polygon side length and the n = 3 case for perimeter of triangles coincides with an exponential distribution. See Figure 1 for some cases of small n.

• Probability of triangles $p_3 = (2 - \frac{\pi^2}{6}) \approx 0.355$

• Probability of quadrilaterals
$$p_4 = \frac{1}{3} - \frac{7\pi^2}{36} + 4 \int_0^{\pi/2} x^2 \cot x \, dx \approx 0.381$$

Stoyan et al. [4] p325 collect further results from Monte Carlo methods:

 $p_5 \approx 0.192, \ p_6 \approx 0.059, \ p_7 \approx 0.013, \ p_8 \approx 0.002$ (1.2)

and mention the empirical approximation obtained by Crain and Miles for the distribution of the number of sides per random polygon

$$p_n \approx \frac{e^{-1}}{(n-3)!}.$$
 (1.3)

We can obtain estimates for the mean and variance of side lengths of the random polygons with n sides:

$$\bar{x}(n) = \int_0^\infty \frac{P}{n} q(P, n) \, dP = \frac{2(n-2)}{n} \quad \text{similarly} \quad var(x)_n = \frac{4(n-2)}{n^2}.$$
(1.4)

though this estimate of $var(x)_n$ will, of course, underestimate the true variance.

The limiting arithmetic average (unweighted by n) over all n is given by

$$\bar{x} = \lim_{k \to \infty} \sum_{n=3}^{k} \frac{\bar{x}(n)}{k-2} = 2 = \lim_{n \to \infty} \bar{x}(n).$$
 (1.5)

Another overall average for mean polygonal side length that may be of more practical value is that obtained via weighting by the number fractions p_n of polygons with n sides, given approximately by the expression (1.3). We estimate this by obtaining first an approximate probability density function for the perimeter of all polygons:

$$q(P) = \sum_{n=3}^{\infty} p_n q(P, n) = \frac{1}{2} e^{-\left(1 + \frac{P}{2}\right)} I_0(\sqrt{2P})$$
(1.6)

where $I_0(\zeta)$ is the zeroth order modified Bessel function of the first kind. The probability density function given by Equation (1.6) has mean, $\overline{P} = 4$ and variance var(P) = 12 and is plotted in Figure 2. Recall that the probability density functions we have used give P as a multiple of the mean polygon side length and the expected number of sides per polygon is four. It follows that weighted mean polygon side length is $\frac{\overline{P}}{4} = 1$, as expected. We may state therefore that the overall arithmetic mean polygon side length is twice the weighted mean polygon side length.

Simulations of random lines by Piekaar and Clarenburg [5] found that around 39 % of the polygons were triangles, around 35 % had four sides and the remaining 25 % or so had 5 or more sides in the random isotropic case. Such practical realizations of isotropic random lines tend to yield random polygons that appear 'roundish', nearly regular, rather than irregular in shape and this is reported also by Corte and Lloyd [6] for fibre networks made using laboratory and commercial filtration processes. See Miles [7] and Kovalenko [8] for proofs that this regularity is in fact a limiting property for random polygons as A, P or n become large.

Note that, for a planar process of lines resulting in polygon side lengths following a *Gaussian* distribution, the perimeter distributions for *n*-sided polygons would be χ^2 with n - 1 degrees of freedom—so for triangles it gives the same as Miles [1], but the Gaussian case differs increasingly for larger n.

The analytic distribution function for the area A of the random polygons remains an unsolved analytic problem; however, approximating analyses have been given by Corte and Lloyd [6] for the random line case. This was extended by Dodson and Sampson [9, 10] to non-random stochastic line



Figure 2: Probability density function approximation for perimeter P of all polygons averaged over $n \ge 3$; P plotted in units of mean polygon side length.

and fibre processes using for the polygon edge distribution a gamma distribution and then that the radii of circles with equivalent area to polygons was closely approximated by a gamma distribution also. Similar findings have been reported by Castro and Ostoja-Starjewski [11] from Monte-Carlo analyses. Dodson and Sampson [12, 13] discussed also porosity statistics and transfer properties in stochastic fibre networks. In section 2.1 below we provide a new analytic approximation to the distribution of areas of random polygons.

The spatial covariance function for local fibre density in random processes of straight finite fibres was derived and used by Dodson [14] to obtain analytically the variance of local density as a function of the zone size for arbitrary schemes of complete sampling with rectangular zones.

These geometric features of random line and fibre processes are of importance to physicists studying stochastic fibrous materials at a range of scales. For example, Berhan *et al.* [15, 16, 17] have recently considered the influences of the structure of random networks of nanotubes on their mechanical properties. Similarly, the pore size of electrospun nano-fibrous networks for application as nano-filters has been recently shown to depend strongly on the diameter of the electrospun fibres [18] as predicted by theory [19]. Note also that the fracture of random fibre networks is of interest

to the study of paper, glass fibre mats and randomly oriented short fibre composites [20, 21]. There is a significant body of work discussing fracture of such materials; see, *e.g.* [22] for a review summarising results in the context of paper, and [23] for more general discussion of fracture processes including the scaling of fracture path roughness in random fibre networks. A specific outstanding problem is that it is not known whether the distribution of polygons through which a fracture path passes in a random fibre network is representative of the network as a whole or is drawn selectively from larger or smaller polygons in the global distribution. Here we provide approximations for the distribution of local line density which we expect to be relevant to such studies of fracture paths. We provide also numerical estimates of the correlation between adjacent polygon sides which may be useful in analytic approaches to estimate the global and local deformation of networks under an imposed load.

2 Approximating distributions for local density and void areas

Here we make a simplifying approximation in our analysis, in order to make approximating estimates for probability density functions of local line density and area for random polygons. The assumption is that the polygons are all sufficiently close to being regular for us to be able to use the standard formulae for areas of regular polygons. As we have mentioned, this is a limiting property for large n and we will see in Section 3 that significant correlation exists between the lengths of adjacent sides of the polygons arising from random line processes in the plane.

In order to test the validity of the assumption of regularity of polygons, it is helpful to consider the probability density of radii of circles inscribed within these polygons. Without this assumption, Miles [1] provides the analytic result that this probability density is exponential for infinite lines of arbitrary width and arbitrary distributions of width. For a regular *n*sided polygon, the radius of the inscribed circle is,

$$r = \frac{P}{2n} \cot\left(\frac{\pi}{n}\right) \quad . \tag{2.7}$$

From Equation (1.1), the probability density of inscribed circle radii for an n-sided regular polygon is,

$$s(r,n) = \frac{(n r \tan(\pi/n))^{n-2} e^{-n r \tan(\pi/n)}}{r \Gamma(n-2)} \quad . \tag{2.8}$$

It follows that the probability density of inscribed radii of all regular poly-

gons is approximated by

$$s(r) = \sum_{n=3}^{\infty} p_n \, s(r,n)$$
 . (2.9)

We have been unable to obtain a closed form of the sum given in Equation (2.9), though, numerically, we find that the sum to n = 8 accounts for more than 99.9 % of the distribution. This distribution has mean, $\bar{r} = 0.506$ and coefficient of variation, cv(r) = 0.950, close to the unit coefficient of variation of the exponential distribution provided by Miles. The probability density for pore radii as given by Equation (2.9) summed to n = 8 is plotted in Figure 3 along with that for the exponential distribution with the same mean. The similarity between the two probability densities is immediately evident. It is worth bearing in mind that inscribed circles will contact every side of a regular polygon but for the irregular polygons considered by Miles [1] these will, in general, be tangential to 3 sides only. Nevertheless, the analysis suggests that the assumption of regularity provides a reasonable basis for our subsequent analysis.

2.1 Void area statistics

For an n-sided regular polygon with perimeter P, we can compute the area

$$A = \frac{P^2}{4n} \cot \frac{\pi}{n}, \quad n = 3, 4, 5....$$
 (2.10)

Then using (1.1) the probability density function for area A of our random n-sided polygons is approximated by

$$p(A,n) = \frac{e^{-\sqrt{A n \tan\left(\frac{\pi}{n}\right)}} \left(A n \tan\left(\frac{\pi}{n}\right)\right)^{\frac{n-2}{2}}}{2 A \Gamma(n-2)}, \quad n = 3, 4, \dots$$
(2.11)

See Figure 4 for some cases of small n and note that we have the following limiting properties:

$$\lim_{n \to \infty} p(A, n) = 0 \text{ but } \lim_{n \to \infty} \int_0^\infty p(A, n) \, dA = 1 \text{ and } \lim_{n \to \infty} \bar{A}(n) = \infty.(2.12)$$

The mean area of an n-sided polygon is

$$\bar{A}(n) = \frac{(n-2)(n-1)}{n} \cot\left(\frac{\pi}{n}\right).$$
 (2.13)

Again, we have been unable to obtain an analytic form of the weighted probability density function for polygon areas, as given by

$$q(A) = \sum_{n=3}^{\infty} p_n \, p(A, n) \tag{2.14}$$



Figure 3: Approximate probability density function for radii of circles inscribed in regular polygons as given by Equation (2.9) compared with probability density for exponential distribution with the same mean. r is plotted in units of mean polygon side length.

7



Figure 4: Probability density function approximations for area A of random polygons with n=3,4,5,6 sides. A is plotted in units of the mean polygon side length squared.



Figure 5: Probability density function approximation for area A of all polygons averaged over $n \ge 3$. A is plotted in units of the mean polygon side length squared.

though the weighted probability density function determined for terms up to n = 20 should be sufficient for practical applications and is shown in Figure 5; numerical analysis of this probability density function yields $\bar{A} =$ 1.84 and var(A) = 12.42.

2.2 Local density statistics

The local density of line length per unit area for a regular *n*-sided polygon is b = P/A and so using (1.1), *b* in our random polygons has probability density function *f* approximated by

$$f(b,n) = \frac{\left(\frac{2n\,\tan(\pi/n)}{b}\right)^{n-2} \,e^{-\frac{2\,n\,\tan(\pi/n)}{b}}}{b\,\Gamma(n-2)}, \quad n = 3, 4, \dots$$
(2.15)

See Figure 6 for some cases of small n.

We note that the mean line density $\bar{b}(3)$ for regular triangles is undefined because the integral is divergent, but the means exist for polygons of $n \ge 4$



Figure 6: Probability density function approximations for local line density b of random polygons with n=3,4,5,6 sides. b is plotted in units of reciprocal mean polygon side length.

sides:

$$\bar{b}(n) = \frac{2n \tan(\pi/n)}{(n-3)}$$
 (2.16)

and
$$\lim_{n \to \infty} \bar{b}(n) = 0$$
. (2.17)

The divergence of the integral for the case when n = 3 is a consequence of the unboundedness of the expectation of the reciprocal of an exponentially distributed variable.

The weighted probability density function for local line density is given by

$$g(b) = \sum_{n=3}^{\infty} p_n f(b, n) \quad .$$
 (2.18)

Again, this requires numerical evaluation and we consider that terms up to n = 20 suffice for practical purposes; the probability density function given by such an evaluation is shown in Figure 7. Just as the mean line density $\bar{b}(3)$ for regular triangles is undefined because the integral is divergent, so is the weighted average line density. We can estimate this however for



Figure 7: Probability density function approximation for local line density b of all polygons averaged over $n \ge 3$. b is plotted in units of reciprocal mean polygon side length.

polygons of four or more sides using

$$\bar{b}_{n\geq4} = \frac{\sum_{n=4}^{\infty} p_n \,\bar{b}(n)}{\sum_{n=4}^{\infty} p_n} \tag{2.19}$$

which yields $\bar{b}_{n\geq 4} \approx 6$.

3 Product models, crossing clustering and correlation

Practical applications have exploited the fact that the expected number of sides per random polygon is four. They work by substituting for the polygon area distribution a direct product of two exponential [6] or gamma [9] distributions, or by using a bivariate gamma [10]. These models have been and continue to be successful in predicting properties of real and simulated networks [6, 10, 12, 13, 25, 24].

An interesting question [10] concerns the evident fact that, even for random isotropic homogeneous line processes, denser regions tend to have shorter polygon sides than less dense regions. A product model for polygon

areas might represent this feature through correlation of the constituent variables, such as a bivariate exponential or bivariate gamma distribution.

Consider first the case of a random isotropic homogeneous line process in a plane. We represent the distribution of random polygon areas A by a rectangular grid with rectangle sides the random variables x, y drawn from an exponential distribution with unit mean with the ordering $x \leq y$; the area of each rectangular polygon is A = x y. Thus, since the underlying process is isotropic and homogeneous, we have the random variables related by

$$\bar{x} = (1-\epsilon) \text{ and } \bar{y} = (1+\epsilon)$$
 (3.20)

$$\bar{x} + \bar{y} = 2$$
 so $\bar{x}\bar{y} = (1 - \epsilon^2).$ (3.21)

For random isotropy we have correlation coefficient

$$\rho = \frac{cov(x,y)}{\sqrt{var(x) var(y)}}$$
(3.22)

$$= \frac{A - \bar{x}\bar{y}}{\sqrt{var(x) var(y)}}$$
(3.23)

$$= \frac{A - (1 - \epsilon^2)}{\sqrt{\operatorname{var}(x) \operatorname{var}(y)}}.$$
(3.24)

It remains to estimate ρ , on the assumption that x, y come from a common exponential distribution with unit mean. We start with pairs of randomly chosen numbers from the exponential distribution; hence the mean value of the product of these pairs is 1. Then convert each pair $\{x_i, y_i\}$ into an ordered pair (x_i, y_i) such that $x_i \leq y_i$ and create now two distributions, one for the first member x and one for the second member y. Intuitively, we take any $y_i < x_i$ from the source distribution of y and add them to source distribution of x; also we take any $x_i > y_i$ from the source distribution of x and add these to the source distribution of y. Note that the mean product of pairs $\overline{xy} = 1$ unaltered; however, the ordered pairs are no longer independent. This yields the probability density function for x

$$g_{\leq}(x) = \frac{1}{2}e^{-2x}$$
 so $\bar{x} = \frac{1}{2}$ and $var(x) = \frac{1}{4}$. (3.25)

Also, the probability density function for y is

$$g_{\geq}(y) = 2 e^{-2y} (1 - e^{-y})$$
 so $\bar{y} = \frac{3}{2}$ and $var(y) = \frac{5}{4}$. (3.26)

It follows that

$$\rho = \frac{\overline{x}\,\overline{y} - \overline{x}\,\overline{y}}{\sqrt{var(x)\,var(y)}} = \frac{1}{\sqrt{5}}.$$
(3.27)



Figure 8: Graphical representation of random line processes in the plane.

The estimate that we have obtained for the correlation between ordered pairs of x and y takes no account of clustering of crossings and arises only from sorting pairs such that $x_i < y_i$. We expect clustering to increase the correlation between adjacent pairs of inter-crossing distances.

We have written *Mathematica* code to extract pairs of x and y representing the lengths of the adjacent sides of polygons arising from a Poisson line process in a unit square. The code works by solving the equations of lines drawn at random within the unit square for the coordinates of all crossings that occur between them. Each of the coordinates is identified by the lines that generate it, allowing the coordinates of the adjacent crossings on these lines to be extracted; from these the lengths of adjacent pairs of polygon sides are calculated. Graphical representations of these random line networks are shown in Figure 8. Note that we consider only pairs of polygon sides bounded entirely by the unit square. Where either of a pair of adjacent polygon sides cross the sides of the unit square, these are discounted from the analysis. In Figure 8 these polygon sides are represented by broken lines. Importantly, discarding these polygon sides from our analvsis had no significant influence on the distribution of polygon sides, which was exponential, as expected. Given this, we can be confident that any difference between the correlation computed from our simulation and that calculated for independent polygon sides is an intrinsic feature of the network structure and not an artefact arising from the way the problem has been encoded within the software.

In Figure 9, the correlation between adjacent polygon side lengths is plotted against the intensity of the line process generating them. The process intensity is calculated as the total length of lines in the unit square. The error bars on ρ represent 95 % confidence intervals calculated from 10 different random seed numbers. Networks with an increasing number of lines per unit area were generated each random seed, permitting the correlation to be tracked as a function of process intensity; this process is illustrated in the inset figure. For processes of 1000 lines in the unit square



Figure 9: Correlation plotted against process intensity. Data represent the means of 10 simulations; error bars represent 95 % confidence intervals. Inset diagram shows data for the 10 simulations.

we calculate the correlation between more than a million pairs of adjacent polygon sides and observe a correlation of $\rho = 0.616 \pm 0.001$; this is represented by the broken horizontal line in Figure 9. We observe the same correlation in networks of 500 lines with a confidence interval varying only in the fourth decimal place. Note also that whereas the correlation for individual line processes may exceed this value at low process intensities, the mean correlation observed over our 10 cases was always less than 0.616 for process intensities less than 500. It is interesting that for process of 20 or more lines per unit area, the correlation is always greater than that calculated for independent and ordered pairs, *i.e.* $\rho = \frac{1}{\sqrt{5}} \approx 0.447$ and increases rapidly towards its stable value with increasing intensity.

Examples of scatter plots showing the distribution of sorted x and y pairs are shown in Figure 10. The figures on the left show pairs (x, y) drawn independently from exponential distributions with unit mean and sorted such that $x_i \leq y_i$; those on the right show pairs (x, y) obtained from our analysis of random polygons arising from a planar line process of 500 lines in the unit square. A three-dimensional histogram approximating the joint probability density of x, y is shown beneath each scatter plot wherein the heights of the bars are scaled such that the volume under the histogram sums to 1. For the independent sorted pairs we have $\bar{x} = 0.501$, $\bar{y} = 1.503$, var(x) = 0.251, var(y) = 1.250, $\rho = 0.445$ in agreement with Equations (3.25) to (3.27). For the sides of random polygons we have $\bar{x} = 0.608$, $\bar{y} = 1.392$, var(x) = 0.413, var(y) = 1.399, $\rho = 0.615$.

References

- R.E. Miles. Random polygons determined by random lines in a plane. Proc. Nat. Acad. Sci. USA 52 901-907,1157-1160 (1964).
- [2] R.E. Miles. The various aggregates of random polygons determined by random lines in a plane. Advances in Math. 10, 256-290 (1973).
- [3] J.C. Tanner. The proportion of quadrilaterals formed by random lines in a plane. J. Appl. Probab. 20(2), 400-404 (1983).
- [4] D. Stoyan, W.S. Kendall and J. Mecke. Stochastic geometry and its applications 2nd edition, John Wiley, Chichester (1995).
- [5] H.W. Piekaar and L.A. Clarenburg. Aerosol filters—Pore size distribution in fibrous filters. *Chem. Eng. Sci.* 22, 1399-1408 (1967).
- [6] H. Corte and E.H. Lloyd. Fluid flow through paper and sheet structure. In Consolidation of the Paper Web Trans. IIIrd Fund. Res. Symp. (F. Bolam, ed.), pp981-1009, BPBMA, London, (1966).
 - 15

- [7] R.E. Miles. A heuristic proof of a long-standing conjecture of D.G. Kendall concerning the shapes of certain large random polygons. Adv. in Appl. Probab. 27(2), 397-417 (1995).
- [8] I. Kovalenko. A simplified proof of a conjecture of D.G. Kendall concerning shapes of random polygons. J. Appl. Math. Stochastic Anal. 12(4), 301-310 (1999).
- [9] C.T.J. Dodson and W.W. Sampson. Modeling a class of stochastic porous media. App. Math. Lett. 10(2), 87-89 (1997).
- [10] C.T.J. Dodson and W.W. Sampson. Effect of correlated free fibre lengths on pore size distribution in fibrous mats. In Advances in Paper Science and Engineering Trans. XIIIth Fund. Res. Symp.(S.J. I'Anson, ed.) pp943-960, FRC, Manchester (2005).
- [11] J. Castro and M. Ostoja-Starzewski. Particle sieving in a random fiber network. Appl. Math. Modelling 24(8-9):523-534, 2000.
- [12] C.T.J. Dodson and W.W. Sampson. Spatial statistics of stochastic fibre networks. J. Statist. Phys. 96(1/2), 447-458 (1999).
- [13] C.T.J. Dodson, Y. Oba and W.W. Sampson. Bivariate normal thickness-density structure in real near-planar stochastic fibre networks. J. Statist. Phys. 102(1/2), 345-353 (2001).
- [14] C.T.J. Dodson. Spatial variability and the theory of sampling in random fibrous networks. J. Roy. Statist. Soc. B 33(1), 88-94 (1971).
- [15] L. Berhan, Y.B. Yi, A.M. Sastry, E. Munoz, M. Selvidge and R. Baughman. Mechanical properties of nanotube sheets: Alterations in joint morphology and achievable moduli in manufacturable materials. J. App. Phys. 95(8):4335-4345, 2004.
- [16] L. Berhan, Y.B. Yi and A.M. Sastry. Effect of nanorope waviness on the effective moduli of nanotube sheets. J. App. Phys. 95(9):5027-5034, 2004.
- [17] Y.B. Yi, L. Berhan and A.M. Sastry. Statistical geometry of random fibrous networks, revisited: Waviness, dimensionality, and percolation. J. App. Phys. 96(3):1318-1327, 2004.
- [18] Y.C. Ahn, S.K. Park, G.T. Kim, Y.J. Hwang, C.G. Lee, H.S. Shin and J.K. Lee. Development of high efficiency nanofilters made of nanofibers. *Current App. Phys.* 6(6):1030-1035, 2006.

- [19] S.J. Eichhorn and W.W. Sampson. Statistical geometry of pores and statistics of porous nanofibrous assemblies. J. Roy. Soc. Interface 2(4):309-318, 2005.
- [20] J. Åström, S. Saarinen, K. Niskanen and J. Kurkijärvi. Microscopic mechanics of fiber networks. J. App. Phys. 75(5):2383-2392, 1994.
- [21] V.I. Räisänen, M.J. Alava and R.M. Nieminen. Failure of planar fiber networks. J. App. Phys. 82(8):3747-3753, 1997.
- [22] M. Alava and K. Niskanen. The physics of paper. *Rep. Prog. Phys.* 69(3):669-723, 2006.
- [23] M.J. Alava, P.K.V.V. Nukala and S. Zapperi. Statistical models of fracture. Adv. Phys. 55(3-4):349-476, 2006.
- [24] M. Deng and C.T.J. Dodson. Paper: An Engineered Stochastic Structure. Tappi Press, Atlanta (1994).
- [25] W.W. Sampson. Comments on the pore radius distribution in nearplanar stochastic fibre networks. J. Mater. Sci. 36(21), 5131-5135 (2001).



Figure 10: Scatter plots and 3D histograms for sorted x, y pairs. Left: independent and sorted x, y; right: sorted x, y representing adjacent sides of random polygons.