Image Anal Stereol 2003;22:147-152
Original Research Paper

ESTIMATING THE INTENSITY OF GERM-GRAIN MODELS WITH OVERLAPPING GRAINS

HAMID GHORBANI AND DIETRICH STOYAN

Institut für Stochastik, TU Bergakademie Freiberg, 09596 Freiberg, Germany
e-mail: ghorbani@math.tu-freiberg.de, stoyan@orion.hrz.tu-freiberg.de
(Accepted October 29, 2003)

ABSTRACT

Formulas are derived for the spherical contact distribution of a planar germ-grain model Z with circular grains where the germs form either a 'segment cluster' process or a 'line-based' Poisson point process. They are used in order to estimate the intensity λ of the germ process by means of the spherical contact distribution function. As an application the number of dislocations on a silicon wafer is estimated.

Keywords: dislocations, germ-grain model, intensity, Poison cluster process, Poison line process, silicon wafer, spherical contact distribution function.

INTRODUCTION

This paper describes methods for estimating the intensity λ of a germ-grain model (Stoyan *et al.*, 1995) with deterministic circular grains. While the aim is estimating the intensity of a point process X, the observable is a random set, the union Z of overlapping discs with fixed radius R, $Z = X \oplus b(o,R)$, where $\oplus$ denotes Minkowski-addition and o is the origin. Because of overlappings, intensity estimation in this setting is a complicated problem.

In the classical case of a Boolean model (X is a stationary Poisson process) one of the successful statistical methods uses the spherical contact distribution function, see Serra (1988) and Stoyan *et al.* (1995). This approach is used also in the cases considered here, where X is modelled as a segment cluster process, i.e. a Poisson cluster process where the cluster points are scattered on segments, or a 'line-based' Poisson point process, where the points are scattered randomly on the lines of a Poisson line process. As an application, the number of dislocations on a silicon wafer is estimated. Here the point pattern X of centers of dislocations shows a high degree of clustering and the observable is a system of overlapping etch pits, which grow around the centers of dislocations with circular shape after etching process.

FUNDAMENTALS

SPHERICAL CONTACT DISTRIBUTION FUNCTION

In the description and statistical analysis of a stationary point process X in the plane the spherical contact distribution function $H_s^X(r)$ plays an important role. It is defined as

$$H_s^X(r) = P(X \cap b(o,r) \neq \emptyset) \quad \text{for} \quad r \geq 0\,, \qquad (1)$$

where $b(o,r)$ is the disc with radius r centered at the origin o.

By definition $1 - H_s^X(r)$ is the probability that in the disc $b(o,r)$ there is no point of the point process. Thus $H_s^X(r)$ is the distribution function of the distance from the origin to the point of the point process closest to o.

The spherical contact distribution function of a stationary random closed set Z is defined as

$$\begin{aligned} H_s^Z(r) &= P(Z \cap b(o,r) \neq \emptyset \mid o \notin Z) \\ &= 1 - \frac{P(Z \cap b(o,r) = \emptyset)}{1-p} \quad \text{for} \quad r \geq 0\,, \quad (2) \end{aligned}$$

where $p(< 1)$ is the probability of covering the origin by Z or the area fraction of Z,

$$p = P(o \in Z)\,.$$

For the particular case of a germ-grain model with circular grains, $Z = X \oplus b(o,R)$, the spherical contact distribution of Z can be easily expressed in terms of that of the germ process X:

$$\begin{aligned} H_s^Z(r) &= 1 - \frac{1 - P(o \in Z \oplus b(o,r))}{1-p} \\ &= 1 - \frac{1 - P(o \in X \oplus b(o,r+R))}{1-p} \quad \text{for} \quad r \geq 0 \end{aligned}$$

where

$$p = P(o \in Z) = P(o \in X \oplus b(o,R))\,.$$

Consequently,

$$H_s^Z(r) = 1 - \frac{1 - H_s^X(r+R)}{1 - H_s^X(R)} \quad \text{for} \quad r \geq 0 \,. \tag{3}$$

If X is a Poisson point process with intensity κ then its spherical contact distribution function is

$$H_s^X(r) = 1 - e^{-\kappa\pi r^2} \quad \text{for} \quad r \geq 0 \,. \tag{4}$$

see Stoyan *et al.* (1995).

POINT PROCESS MODELS

POISSON CLUSTER PROCESSES

The basic ingredient of a Poisson cluster process is a stationary Poisson process of intensity λ_p; its points are called 'parent points'. Independent identically distributed clusters of daughter points are scattered around the parent points and the union of all daughter points forms the Poisson cluster process, see Stoyan *et al.* (1995).

The spherical contact distribution of a Poisson cluster processes can be obtained theoretically, by means of the formulas for the Boolean model, where the 'germs' are the parent points and the 'grains' the clusters. Of course, here the grains are not convex and the general methods for Boolean models with non-convex grains developed in Hug *et al.* (2002) and Last and Holtmann (1999) can be applied. A direct approach starts from the formula

$$H_s^X(r) = 1 - \exp\left[-\lambda_p EA\left(C \oplus b(o,r)\right)\right] \quad \text{for} \quad r \geq 0 \,. \tag{5}$$

Here $A(\cdot)$ denotes the area and C denotes the random set consisting of the points of a typical cluster centered at o.

The difficulty of calculating $H_s^X(r)$ in formula (5) is determining $EA(C \oplus b(o,r))$ because the discs may overlap. Various Poisson cluster processes differing in the arrangement of points forming the cluster and their spherical contact distribution functions have been considered in Saxl (1993) and Saxl and Rataj (1996).

SEGMENT CLUSTER PROCESS

The segment cluster process is a special Poisson cluster process in which the typical cluster C is a set of points that are distributed on a segment of fixed length l and uniform random orientation according to a homogeneous one-dimensional Poisson process with intensity λ_l. The parent points are located in the centers of the segments.

The intensity λ of this model is given by

$$\lambda = \lambda_p \cdot \lambda_l \cdot l \,.$$

The aim of the following is to derive a formula for the term $EA(C \oplus b(o,r))$, the mean area of $C \oplus b(o,r)$, which appears in formula (5). Fig. 1 shows the set the area of which has to be determined, the dilation of the typical cluster by $b(o,r)$.

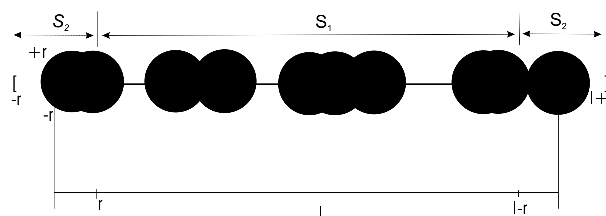

Fig. 1. *A typical dilated cluster of a segment cluster process, which is divided in two parts S_1 and S_2.*

The problem of calculating the exact area is not easy, since the discs may overlap. A possible approximation is as follows.

The set $C \oplus b(o,r)$ is divided in two parts, namely S_1 and S_2. The part of this set which lies inside the segment $[r, l-r]$ is called S_1 and the rest outside the segment $[r, l-r]$ is called S_2. The mean area of S_1 can be calculated exactly using the Cavalieri principle and the theory of the Boolean model, while the area of S_2 is calculated approximately using simulation results.

For the set S_1, the Cavalieri principle yields

$$EA(S_1) = \int_{-r}^{r} El(t)\mathrm{d}t \,,$$

where $l(t)$ denotes the sum of lengths of the chords produced by intersection of S_1 with the segment $\{(x,y) : r \leq x \leq (l-r), y = t\}$ with $t \in [-r, r]$.

In order to calculate $El(t)$, a one-dimensional stationary Boolean model is considered with intensity λ_l and a typical grain which is a segment of deterministic length $2\sqrt{r^2 - t^2}$ (this is the length of the chord obtained by intersection of $b(o,r)$ and the line $y = t$). The length fraction $L_L(t)$ (the proportional of line occupied with discs) of this model satisfies

$$L_L(t) = \frac{El(t)}{l - 2r} \quad \text{for} \quad -r \leq t \leq r \,.$$

The known formulas for the Boolean model yield

$$L_L(t) = 1 - e^{-2\lambda_l\sqrt{r^2 - t^2}} \quad \text{for} \quad -r \leq t \leq r \,.$$

Therefore,

$$EA(S_1) = (l-2r)\int_{-r}^{r}(1-e^{-2\lambda_l\sqrt{r^2-t^2}})\,\mathrm{d}t$$
$$= r(l-2r)\int_{-1}^{1}(1-e^{-2\lambda_l r\sqrt{1-t^2}})\,\mathrm{d}t\,.$$

With

$$f(\mu) = \int_{-1}^{1}(1-e^{-2\mu\sqrt{1-t^2}})\,\mathrm{d}t \qquad (6)$$

one obtains

$$EA(S_1) = r(l-2r)f(\lambda_l r)\,. \qquad (7)$$

Because there is no analytical solution for the integral in (6), it was numerically calculated for different values of μ. By a least square technique an exponential function was fitted to the results, yielding the approximation

$$\hat{f}(\mu) \approx 2-2e^{-1.5\mu}\,, \qquad (8)$$

and

$$EA(S_1) \approx 2r(l-2r)(1-e^{-1.5\lambda_l r}) \quad \text{for} \quad r \le \frac{l}{2}\,. \qquad (9)$$

In order to calculate $EA(S_2)$ a one-dimensional homogeneous Poisson process on the segment $[0,2r]$ with intensity λ_l is considered. Let C_0 denotes the set of points of the process which are distributed on this segment. Define

$$a(\lambda_l,r) = EA\left(C_0 \oplus b(o,r) \cap ([-r,r]\times[-r,r])\right)\,. \qquad (10)$$

Then clearly,

$$EA(S_2) = 2a(\lambda_l,r)\,.$$

Moreover,

$$a(\lambda_l,r) = r^2 a(\lambda_l r,1)\,, \qquad (11)$$

which is a consequence of the definition of $a(\lambda_l,r)$ in Eq. (10).

The values of $a(\lambda_l,1)$ were calculated by simulation for different λ_l, what yielded by numerical investigations the following approximation for $a(\lambda_l,1)$:

$$\hat{a}(\lambda_l,1) = 0.92\frac{e^{-\lambda_l}}{\lambda_l} - 2.50\,e^{-\lambda_l} - \frac{0.92}{\lambda_l} + 3.41 \quad \text{for} \quad \lambda_l > 0\,, \qquad (12)$$

see for more details Ghorbani (2003).

The formulas (11) and (12) yield

$$\hat{a}(\lambda_l,r) = 0.92\frac{r}{\lambda_l}e^{-\lambda_l r} - 2.50\,r^2 e^{-\lambda_l r} - 0.92\frac{r}{\lambda_l} + 3.41\,r^2 \quad \text{for} \quad \lambda_l > 0\,.$$

Consequently,

$$\hat{E}A(S_2) = 1.84\frac{r}{\lambda_l}e^{-\lambda_l r} - 5r^2 e^{-\lambda_l r} - 1.84\frac{r}{\lambda_l} + 6.82\,r^2 \quad \text{for} \quad \lambda_l > 0\,. \qquad (13)$$

Finally, substituting the summation of Eqs. (9) and (13) in (5) yield

$$-\log(1-H_s^X(r)) \approx \lambda_p\Big\{(4r^2-2rl)e^{-1.5\lambda_l r} + \left(\frac{1.84r}{\lambda_l} - 5r^2\right)e^{-\lambda_l r} + 2.82r^2 + 2rl - \frac{1.84r}{\lambda_l}\Big\} \quad \text{for} \quad r \le \frac{l}{2}\,. \qquad (14)$$

LINE-BASED POISSON PROCESS

Consider a randomly distributed system of lines which are distributed as a stationary-symmetric Poisson line process Φ with intensity L_A, see Stoyan *et al.* (1995). On each line, points are distributed according to a homogeneous one-dimensional Poisson process of intensity λ_l. The point process obtained by the union of all points is called line-based Poisson process. This process is a special case of Cox process.

The intensity λ of the line-based Poisson process is given by

$$\lambda = L_A \cdot \lambda_l\,,$$

where L_A is the mean length of lines per area unit.

The number of lines of Φ hitting $b(o,r)$ has a Poisson distribution of mean $2rL_A$. Consequently, the spherical contact distribution function of Φ is

$$H_s^{\Phi}(r) = P(o \in \Phi \oplus b(o,r)) = 1-e^{-2rL_A} \quad \text{for} \quad r \ge 0\,.$$

An approximation for the spherical contact distribution function of a line-based Poisson process X can be obtained by taking the limit of the spherical contact distribution function of the segment cluster process in (14) for $l \to \infty$ and $\lambda_p \to 0$ with $\lambda_p \cdot l \to L_A$. This yields

$$H_s^X(r) \approx 1-e^{-2rL_A(1-e^{-1.5r\lambda_l})} \quad \text{for} \quad r \ge 0\,. \qquad (15)$$

If small values of r are of interest, the approximation (15) should give good results also for points scattered on *random curves*.

SUPERPOSITION

In order to come to more realistic models superposition is a useful operation. Let X_1 and X_2 be two independent stationary point processes. Consider the union

$$X = X_1 \cup X_2 .$$

With probability one the sets X_1 and X_2 do not overlap, i.e. never points of X_1 and X_2 coincide. The spherical contact distribution function of the union point process $X, H_s^X(\cdot)$, can be easily obtained in terms of $H_s^{X_1}(\cdot)$ and $H_s^{X_2}(\cdot)$ as follows.

Clearly, for any set K

$$\begin{aligned} P\big((X_1 \cup X_2) \cap K = \emptyset\big) \\ &= P(X_1 \cap K = \emptyset \quad \text{and} \quad X_2 \cap K = \emptyset) \\ &= P(X_1 \cap K = \emptyset) \cdot P(X_2 \cap K = \emptyset) . \end{aligned}$$

This yields for $K = b(o,r)$

$$H_s^X(r) = 1 - \big(1 - H_s^{X_1}(r)\big)\big(1 - H_s^{X_2}(r)\big) \quad \text{for} \quad r \geq 0 . \quad (16)$$

If the point process X is a superposition of a segment cluster process X_1 and a stationary Poisson process X_2 with intensity κ then its intensity λ is given by

$$\lambda = \lambda_p \cdot \lambda_l \cdot l + \kappa ,$$

and the formulas (1), (14) and (16) yield the spherical contact distribution function

$$\begin{aligned} -\log(1 - H_s^X(r)) \approx \lambda_p \Big\{ (4r^2 - 2rl)e^{-1.5\lambda_l r} \\ + \left(\frac{1.84\,r}{\lambda_l} - 5r^2\right) e^{-\lambda_l r} + 2.82\,r^2 + 2rl - \frac{1.84\,r}{\lambda_l} \Big\} \\ + \kappa\pi r^2 \quad \text{for} \quad r \leq \frac{l}{2} . \quad (17) \end{aligned}$$

If the point process X is a superposition of a line-based Poisson process X_1 and a Poisson process X_2 with intensity κ then its intensity λ is given by

$$\lambda = L_A \cdot \lambda_l + \kappa ,$$

and the formulas (1), (15) and (16) yield the spherical contact distribution function

$$-\log(1 - H_s^X(r)) \approx 2rL_A(1 - e^{-1.5r\lambda_l}) + \kappa\pi r^2 \quad \text{for} \quad r \geq 0 . \quad (18)$$

APPLICATION

Multicrystalline silicon (mc-Si) is a cost-effective, solar graded raw material for industrial solar cell production. The mc-Si materials contain a high density of extended crystal defects such as dislocations, grain boundaries and micropercipitate. These defects largely control the electrical properties of the material and often impact the solar cell characteristics adversely. In this paper only dislocations are considered. Dislocations are line defects in crystal, which either begin and end on the crystal surface or form a closed curve inside the crystal. A volume of 1 mm^3 may contain dislocations of total length of 1 km.

An important problem is estimating the number of dislocations which appear on silicon wafers, which are planar sections trough the original crystal.

This problem has been investigated by Rinio (2003). After a one minute etching process, microscopic images are made from etched samples. These pits grow around of dislocation objects and have a circular shape. Fig. 2 shows a typical etched dislocation pattern which consists of single etch pits and groups of overlappings etch pits. For such patterns Rinio (2003) estimates the number of dislocations as

$$\begin{aligned} \text{number of dislocations} &= \text{number of single etch pits} \\ &+ \frac{\text{total area of overlapping etch pits}}{\text{mean area of a single etch pit}} . \quad (19) \end{aligned}$$

Because of heavy overlappings of etch pits this estimator is suspected to underestimate the number of dislocations.

For improving the method of estimating the number of dislocations the etched dislocation pattern Z is considered as a germ-grain model. There the germ point pattern X is the system of etch pits centers or dislocation line points and the grains are discs with fixed radius R. Therefore,

$$Z = X \oplus b(o,R) . \quad (20)$$

Experimental observations show that $R = 1\mu m$ is reasonable. The following models for the point process X are considered;

1. Poisson process,
2. segment cluster process,
3. line-based Poisson process,
4. superposition of a segment cluster process and a Poisson process,
5. superposition of a line-based Poisson process and a Poisson process.

By help of the ImageC software of Imtronic the empirical spherical contact distribution function $\hat{H}_s^Z(\cdot)$ of a typical dislocation pattern, which is shown in Fig. 2, was estimated.

The theoretical spherical contact distributions of the germ-grain models Z in (20) for the point process models X mentioned above were derived by means of formula (3). The corresponding parameters were estimated using non-linear regression. Finally, the number of dislocations was estimated by

$$\hat{n} = A(W) \cdot \hat{\lambda} \ ,$$

where $A(W)$ is the area of the window of observation and λ is the intensity of the fitted model. The results are shown in Tab. 1. Figs. 3 and 4 show the empirical spherical contact distribution functions and the fitted curves for the Poisson process and the superposition of a line-based Poisson process and Poisson process.

Obviously, the Poisson process is by no means a suitable model; it was considered only to show the reader that a more general model is needed for describing the dislocation pattern. Moreover, none of the other models is a realistic model for dislocation patterns. A possible criterion for comparison between the models would be the error mean square MS_E which is defined as

$$MS_E = \frac{1}{m-k+1} \sum_{i=1}^{m} \left(H_s^Z(r_i) - \hat{H}_s{}^Z(r_i) \right)^2 \ ,$$

in which m is the number of empirical observations and k is the number of estimated parameters. The values of MS_E for different models are given in Tab. 1. These values suggest that the empirical spherical contact distribution function are very well fitted by the spherical contact distribution functions of the models 2, 4 and 5. While all these models are geometrically unrealistic, they fit the empirical contact distribution function well and the corresponding estimates of n are similar and larger than the Rinio estimate.

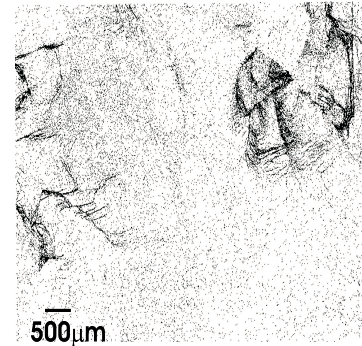

Fig. 2. *A typical sample of dislocation pattern on mc-Si wafer. The darker regions indicate a higher degree of dislocation clusters. By courtesy of Markus Rinio.*

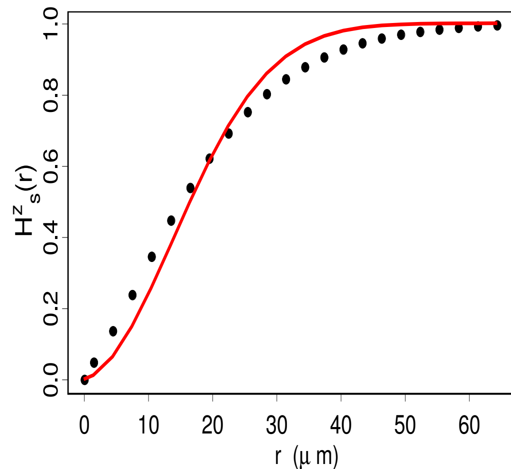

Fig. 3. *The empirical spherical contact distribution function (•) for the dislocation pattern of Fig. 2 and the fitted theoretical function for a Poisson process (—).*

Table 1. *Estimated parameters of different models for the dislocation pattern in Fig. 2. The estimate n by Rinio's method is* 5.7×10^4. *The models are 1: Poisson process, 2: Segment cluster process, 3: Line-based Poisson process, 4: Superposition of a segment cluster process and a Poisson process, 5: Superposition of a line-based Poisson process and a Poisson process.*

Model	$\hat{\lambda}_p$ $[\mu m^{-2}]$	$\hat{\lambda}_l$ $[\mu m^{-1}]$	$\hat{l}$ $[\mu m]$	$\hat{L}_A$ $[\mu m^{-1}]$	$\hat{\kappa}$ $[\mu m^{-2}]$	$\hat{n}$ $[10^4]$	$\sqrt{MS_E}$
1	-	-	-	-	0.0007	1.9	0.0500
2	0.0003	0.161	69.055	-	-	7.7	0.0045
3	-	0.048	-	0.032	-	4.1	0.0140
4	0.0003	0.161	66.055	-	0.00	7.7	0.0045
5	-	0.152	-	0.016	0.0002	7.1	0.0050

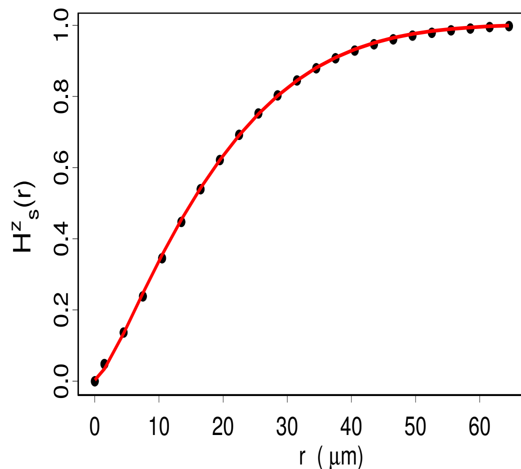

Fig. 4. *The empirical spherical contact distribution function (•) for the dislocation pattern of Fig. 2 and the fitted theoretical function for a superposition of a line-base Poisson process and a Poisson process (—).*

ACKNOWLEDGMENTS

The authors are very grateful to Professor H. J. Moeller and Markus Rinio of Experimental Physics Institute of TU Bergakademie Freiberg for valuable discussions on problems related to dislocations and kind permission for using their dislocation data.

REFERENCES

Ghorbani H (2003). Methods of spatial statistics for the characterization of dislocation systems. PhD thesis, Technical University Bergakademie Freiberg. In preparation.

Hug D, Last G, Weil W (2002). A survey on contact distributions. *Morphology of Condensed Matter. Physics and Geometry of Spatially Complex Systems.* Lecture Notes in Physics (Mecke K, Stoyan D, eds.), 600:317-57. Berlin: Springer.

Last G, Hotmann M (1999). On the empty space function of some germ-grain models. Patt Recogn 32:1587-600.

Rinio M (2003). PhD thesis, Technical University Bergakademie Freiberg. In preparation.

Rinio M, Peters S, Werner M, Möller HJ (2002). Measurement of the normalized recombination strength of dislocations in multicrystalline silicon solar cells. Solid State Phenomena 82-84:701-6.

Saxl I. (1993). Spherical contact distances in Neyman-Scott process of regular clusters. Acta Stereol 12:155-122.

Saxl I, Rataj J (1996). Spherical contact and nearest neighbour distances in Boolean cluster fields. Acta Stereol 15:91-6.

Serra J (1988). Image Analysis and Mathematical Morphology. New York: Academic Press.

Stoyan D, Kendall W, Mecke J (1995). Stochastic Geometry and its Applications. Chichester: J. Wiley and Sons.

Stoyan D, Stoyan H (1994). Fractals, Random Shapes and Point Fields. Chichester: J. Wiley and Sons.