

# TECHNICAL RESEARCH REPORT

## Bayesian Hypothesis Testing for Boolean Random Sets with Radial Convex Primary Grains Using Morphological Skeleton Transforms

*by N. Sidiropoulos, J. Baras,  
C. Berenstein*

**T.R. 91-40**



*Sponsored by  
the National Science Foundation  
Engineering Research Center Program,  
the University of Maryland,  
Harvard University,  
and Industry*

# Bayesian Hypothesis Testing for Boolean Random Sets with Radial Convex Primary Grains using Morphological Skeleton Transforms \*

N. Sidiropoulos<sup>†</sup>      J. Baras<sup>‡</sup>      C. Berenstein<sup>§</sup>

Systems Research Center  
University of Maryland  
College Park, MD 20742  
March 1991

## Abstract

We consider the problem of binary hypothesis testing for planar Boolean random sets with radial convex primary grains. We show that this problem is equivalent to the problem of binary hypothesis testing for Poisson points on a subset of  $\mathcal{R}^3$ . The log-likelihood ratio for Poisson points can therefore be applied to observation points on this subset of  $\mathcal{R}^3$ . Several interesting results pertaining to the asymptotic performance of the log-likelihood ratio for Poisson points are known. A major difficulty with this approach is that the test is based on observation points on a subset of  $\mathcal{R}^3$ , and is not directly given in terms of the observation of a realization of a Boolean random set. An efficient means of mapping realizations of planar Boolean random sets to corresponding realizations of Poisson point processes on this subset of  $\mathcal{R}^3$  is needed in order to implement the test. We show that this can be achieved via a class of morphological transformations known as morphological skeleton transforms. These transforms are flexible shape-size analysis tools based on elementary morphological and set-theoretic operations. This is the principal contribution of this paper.

---

\*Research partially supported by NSF grant NSFD CDR 8803012, through the Engineering Research Centers Program

<sup>†</sup>Also with the Department of Electrical Engineering

<sup>‡</sup>Also with the Department of Electrical Engineering

<sup>§</sup>Also with the Department of Mathematics

# 1 Introduction

The *Boolean random set* is an important and relatively simple example of a random set. Its importance stems from two principal considerations: its analytical tractability and its power in modeling many real-life applications. Despite its simplicity, it has many interesting properties, and, in fact, there are many unanswered questions [15, p65]. A Boolean model is a basic model in stereology and stochastic geometry [15, 7, 13, 1]. Typical applications include: random clumping of dust, or powder particles, or blood cells; modeling of geological structures, patterns in photographic emulsion, colloids in gel form, and structural inhomogeneities in amorphous matter [15, p68, and references therein].

Statistical inference techniques similar to maximum likelihood, or maximum a posteriori are practically non-existent in random set theory [2]. In this paper, we consider a restricted version of the Boolean random set, the *Boolean random set with radial convex primary grains*. This version is still powerful enough to model many real-life applications, yet restricted enough to allow for the analysis and design of efficient algorithms for the statistical inference of various model parameters. Recently, this version has been successfully used to model the degradation process in an attempt to derive the “optimal” reconstruction filters for communication of morphologically coded images [12]. The tools needed for successfully undertaking such a task come from the fields of signal estimation and detection, stochastic geometry, image analysis, and mathematical morphology. The main tool needed is a class of morphological transformations collectively known as *morphological skeleton transforms*. This research is concerned with the application of morphological methods to the problem of hypothesis testing for the case of Boolean random sets with radial convex primary grains.

The rest of this paper is organized as follows. Section (2) is a brief introduction to Poisson point processes and their statistical inference, as it pertains to our case. The scope of this introduction is limited, and it merely serves to state some concrete probabilistic definitions and a few relevant results. Section (3) develops the Boolean random set with radial convex primary grains model. Close to the end of section (3), a major result is presented in the form of a theorem. It states that a planar Boolean random set with radial convex primary grains, under certain assumptions, can be viewed as a Poisson point process on a subset of  $\mathcal{R}^3$ . Section (3.1) formulates the log-likelihood ratio for the case of Boolean random sets with radial convex primary grains, in terms of the corresponding Poisson point process.

Section (4) presents two morphological set representation schemes, and proposes their use to estimate the hidden realization of the Poisson point process from the observation of a realization of a Boolean RS with radial convex grains. Section (5) considers the equivalent discrete-case problem. Due to the existence of certain asymmetries between the continuous and discrete case in the morphological formalization of the notion of size, and the inherently continuous nature of the Poisson point process, it is necessary to consider the discrete case separately. The results are shown to be very similar with the continuous case. An interesting subtlety is thoroughly discussed towards the end of section (5). It has to do with the non-uniqueness in representing a closed set  $X$  as a union of overlapping “primitive” sets. A statistically plausible solution is proposed. Finally, section (6) contains some concluding remarks, and guidelines for future research.

## 2 Point Processes and related results

We begin with some background about point processes and a brief survey of related results.

**Definition 1** [15, p96] *A Point Process,  $\Phi$ , on  $\mathcal{R}^n$ , is a measurable mapping of a probability space  $(\Omega, \Sigma(\Omega), P)$  into a measurable space  $(N, \Sigma(N))$ , where  $N$  is the family of all subsets  $\phi$ , of  $\mathcal{R}^n$ , satisfying the following two regularity conditions*

(i)  $\phi$  is locally finite (each bounded subset of  $\mathcal{R}^n$  must contain only a finite number of points in  $\phi$ ).

(ii)  $\phi$  is simple (no two points in  $\phi$  coincide).

Here  $\Sigma(N)$  is the smallest  $\sigma$ -algebra on  $N$  to make measurable all mappings  $\phi \mapsto \phi(B)$ , as  $B$  runs through the bounded Borel sets of  $\mathcal{R}^n$ . Informally, a Point Process on  $\mathcal{R}^n$  can be thought of as a random pattern of points, scattered over  $\mathcal{R}^n$ .

Consider the measurable space  $(\mathcal{R}^n, \mathcal{B}(\mathcal{R}^n))$  and a measure  $\Lambda$  on  $\mathcal{B}(\mathcal{R}^n)$  such that for all bounded  $B \in \mathcal{B}(\mathcal{R}^n)$  the measure of  $B$ ,  $\Lambda(B)$ , is finite. The measure  $\Lambda$  is referred to as a **Radon measure** [15, p32]. If, in addition,  $\Lambda$  gives zero mass to each point (like the Lebesgue measure) then  $\Lambda$  is called a **Diffuse Radon measure** [15, p33].

**Definition 2** [15, p50] *A Poisson Point Process (PPP),  $\Phi$ , on  $\mathcal{R}^n$ , with diffuse Radon measure (or mean measure)  $\Lambda$  on  $\mathcal{B}(\mathcal{R}^n)$  is a point process which is completely specified by the following two properties*

(i) *Poisson distribution of point counts; the number of points in a bounded set  $B \in \mathcal{B}(\mathcal{R}^n)$  has a Poisson distribution with mean  $\Lambda(B)$*

$$P(\Phi(B) = m) = \frac{(\Lambda(B))^m e^{-\Lambda(B)}}{m!}, \quad m = 0, 1, 2, \dots$$

(ii) *Independent Scattering; the number of points in  $k$  disjoint Borel sets form  $k$  independent random variables.*

If the diffuse Radon measure,  $\Lambda$ , is absolutely continuous (admits a density) with respect to the Lebesgue measure, then it can be written as

$$\Lambda(B) = \int_B \lambda(z) dz, \quad \forall B \in \mathcal{B}(\mathcal{R}^n)$$

The density  $\lambda(z) \geq 0, \forall z \in \mathcal{R}^n$ , is called the *intensity* of the general PPP. Henceforth we make the assumption that  $\Lambda$  is absolutely continuous with respect to the Lebesgue measure (and, therefore, the intensity  $\lambda(z), z \in \mathcal{R}^n$ , exists) and use it freely throughout the rest of the paper. In case  $\lambda(z) = \lambda = \text{const}, \forall z \in \mathcal{R}^n$ , we have the special case of a *Stationary PPP (SPPP)*. In the following we set up our observation model.

## 2.1 Observation Model and Hypothesis Testing for PPP's

We observe a *single realization* of a PPP,  $\Phi$ , on  $\mathcal{R}^n$ . The actual observations are over  $B$ , a compact Borel subset of  $\mathcal{R}^n$ . Asymptotic results are understood to be valid as the process is observed over a sequence of compact Borel subsets of  $\mathcal{R}^n$  increasing to  $\mathcal{R}^n$ . The case of a single realization of a point process on a compact Borel subset of  $\mathcal{R}^n$  is of theoretical and practical interest, but it is notoriously difficult! [4, p97].

### Model

(a) Sample Space  $(\Omega, \Sigma(\Omega)) = (N, \Sigma(N))$ .

(b) We consider two candidate probability laws,  $P_{\Lambda_0}$  and  $P_{\Lambda_1}$ . Under  $P_{\Lambda_j}, j = 0, 1$  the PPP has diffuse Radon measure  $\Lambda_j$ .

(c) The data representing complete observation of  $\Phi$  over  $B$  are the  $\sigma$ -algebra

$$\Sigma(N_B) = \sigma(\phi(A), \forall A \in \mathcal{B}(\mathcal{R}^n) \cap B)$$

i.e. the smallest  $\sigma$ -algebra on  $N_B = \{\mathcal{N} \cap B, \mathcal{N} \in N\}$  to make measurable all mappings  $\phi \mapsto \phi(A)$ , as  $A$  runs through  $\mathcal{B}(\mathcal{R}^n) \cap B$ .

We assume that there is a single operative probability measure present. We wish to make a statistically “correct” decision upon which one of the two probability measures is actually operative, based on the observation model which we have just set forward. We cast this problem as a simple-vs-simple binary hypothesis test.

$$\begin{aligned} H_0 : & \quad \Lambda = \Lambda_0 \quad (\text{“null” hypothesis}) \\ \text{vs:} H_1 : & \quad \Lambda = \Lambda_1 \quad (\text{“alternative” hypothesis}) \end{aligned}$$

This problem has been studied extensively [4, p230]. Under our assumptions,  $\Lambda_0$  and  $\Lambda_1$  are both absolutely continuous with respect to the Lebesgue measure, that is

$$\Lambda_j(A) = \int_A \lambda_j(z) dz, \quad \forall A \in \mathcal{B}(\mathcal{R}^n), \quad j = 0, 1$$

Furthermore, assume that  $\Lambda_1$  is absolutely continuous with respect to  $\Lambda_0$  (henceforth denoted by  $\Lambda_1 \ll \Lambda_0$ ) on the  $\sigma$ -algebra  $\mathcal{B}(\mathcal{R}^n) \cap B$ , where  $B$  is the observation window. By the Radon-Nykodym theorem, this is equivalent to the following condition

$$\Lambda_0(A) = 0 \implies \Lambda_1(A) = 0, \quad \forall A \in \mathcal{B}(\mathcal{R}^n) \cap B$$

i.e.

$$\int_A \lambda_0(z) dz = 0 \implies \int_A \lambda_1(z) dz = 0, \quad \forall A \in \mathcal{B}(\mathcal{R}^n) \cap B$$

Then [4, p231],  $P_{\Lambda_1} \ll P_{\Lambda_0}$  on  $\Sigma(N_B)$ , and the log-likelihood ratio is given by

$$\log \frac{dP_{\Lambda_1}}{dP_{\Lambda_0}} = \int_B \left(1 - \frac{d\Lambda_1}{d\Lambda_0}\right) d\Lambda_0 + \int_B \log \frac{d\Lambda_1}{d\Lambda_0} d\Phi$$

i.e.

$$\log \frac{dP_{\Lambda_1}}{dP_{\Lambda_0}} = \int_B \left(1 - \frac{d\Lambda_1}{d\Lambda_0}\right) d\Lambda_0 + \sum_{z \in \Phi \cap B} \log \frac{d\Lambda_1}{d\Lambda_0}(z)$$

Under the absolute continuity assumptions

$$\log \frac{dP_{\Lambda_1}}{dP_{\Lambda_0}} = \int_B (\lambda_0(z) - \lambda_1(z)) dz + \sum_{z \in \Phi \cap B} \log \frac{\lambda_1(z)}{\lambda_0(z)}$$

Under our assumptions, all Bayesian tests can be shown to reduce to testing the log-likelihood ratio against a precomputable threshold, which depends on the particular choice of differential costs and prior probabilities, but *not* on the likelihood itself [11]. For example, assuming equal priors and equal differential costs the **Maximum Likelihood (ML)** test is given by

$$\text{decide } H_1 \text{ iff : } \int_B (\lambda_0(z) - \lambda_1(z)) dz + \sum_{z \in \Phi \cup B} \log \frac{\lambda_1(z)}{\lambda_0(z)} > 0 \quad (1)$$

There exists an interesting asymptotic result for our observation model.

**Theorem 1** [4, pp230-235] *Suppose that  $\Lambda_1 \ll \Lambda_0$  and  $\Lambda_0 \ll \Lambda_1$  on the entire space  $\mathcal{R}^n$  (In this case the two measures are called **equivalent**, and we denote this by  $\Lambda_1 \sim \Lambda_0$ ), and, furthermore,  $\Lambda_0(\mathcal{R}^n) = \Lambda_1(\mathcal{R}^n) = \infty$ . Then, on  $\Sigma(N)$ , either  $P_{\Lambda_0} \sim P_{\Lambda_1}$ , or  $P_{\Lambda_0} \perp P_{\Lambda_1}$  (in the later case, the two probability measures are called **singular***

$$\exists D \in \Sigma(N) \ni P_{\Lambda_0}(D) = P_{\Lambda_1}(D^c) = 0$$

*i.e. singular measures are concentrated on disjoint subsets) according to whether the integral*

$$\int_{\mathcal{R}^n} \left( 1 - \left( \frac{d\Lambda_1}{d\Lambda_0} \right)^{1/2} \right)^2 d\Lambda_0$$

*converges or diverges. In the former case, there exists a finite log-likelihood ratio*

$$\log \frac{dP_{\Lambda_1}}{dP_{\Lambda_0}} = \int_B \left( 1 - \frac{d\Lambda_1}{d\Lambda_0} \right) d\Lambda_0 + \int_B \log \frac{d\Lambda_1}{d\Lambda_0} d\Phi$$

*that can be used to perform Bayesian decision tests. In the later case, error-free discrimination is possible. For additional results refer to [4].*

### 3 Random Set preliminaries

In this section we formally develop the Boolean Random Set model directly from scratch. Let  $\mathcal{O}, \mathcal{F}, \mathcal{K}$  denote the space of open, closed and compact subsets of  $\mathcal{R}^n$ , for some  $n$ . Consider the following collections of sets in  $\mathcal{F}$

$$F_G = \{F \in \mathcal{F} \ni F \cap G \neq \emptyset\}, \quad G \in \mathcal{O}$$

$$F^K = \{F \in \mathcal{F} \ni F \cap G = \emptyset\}, \quad K \in \mathcal{K}$$

The former collection comprises of all sets in  $\mathcal{F}$  which “hit” an open set  $G$ , while the later collection comprises of all sets in  $\mathcal{F}$  which “miss” a compact set  $K$ . The collection of sets  $\{F_G, G \in \mathcal{O}\}$  and  $\{F^K, K \in \mathcal{K}\}$  generates a **Topology**,  $T(\mathcal{F})$ , on  $\mathcal{F}$ . This is known as the **hit-or-miss topology**, and it allows the study of convergence and continuity in  $\mathcal{F}$  [2]. By taking countable unions and intersections of the open sets of the topological space  $(\mathcal{F}, T(\mathcal{F}))$ , a  $\sigma$ -algebra,  $\Sigma(\mathcal{F})$ , is generated on  $\mathcal{F}$ .

**Definition 3** A **Random Set (RS)**,  $X$ , is a measurable mapping of a probability space  $(\Omega, \Sigma(\Omega), P)$  into the measurable space  $(\mathcal{F}, \Sigma(\mathcal{F}))$ .

**Definition 4** The **capacity functional**,  $T_X(K)$ , of a RS  $X$ , is defined by

$$T_X(K) = P_X(X \in F_K) = P_X(x \cap K \neq \emptyset), \quad K \in \mathcal{K}$$

The capacity functional  $T_X(K)$ , for all  $K \in \mathcal{K}$ , contains all the information about the RS  $X$ .

**Theorem 2** [2, Choquet] Given  $T_X(K)$ ,  $\forall K \in \mathcal{K}$ , there exists a unique probability measure,  $P_X$ , on  $\Sigma(\mathcal{F})$ , such that  $P_X(X \in F_K) = T_X(K)$ ,  $\forall K \in \mathcal{K}$

**Definition 5** Two compact sets  $K_1, K_2 \in \mathcal{K}$  are **separated** by a compact set  $K \in \mathcal{K}$  iff

$$L_{k_1, k_2} \cap K \neq \emptyset, \quad \forall (k_1, k_2) \ni k_1 \in K_1, k_2 \in K_2$$

Here,  $L_{k_1, k_2}$  denotes the straight line segment connecting the points  $k_1$  and  $k_2$ .

**Definition 6** A RS  $X$  is **convex** iff

$$P_X(X \in F_{K_1, K_2}^K) = 0$$

for all  $K_1, K_2, K \in \mathcal{K}$  such that  $K_1$  and  $K_2$  are separated by  $K$ . Here,  $F_{K_1, K_2}^K$  is the collection of all sets in  $\mathcal{F}$  which hit  $K_1$  and  $K_2$ , and miss  $K$ .

If the RS  $X$  is convex and  $W$  is a closed convex set then  $X \cap W$  is also a convex RS [2].



**Definition 7** The RS  $X$  defined by

$$X = \bigcup_{i=1,2,\dots} G_i \oplus \{p_i\}$$

where  $P = \{p_1, p_2, \dots\}$  is a Point Process and  $\{G_1, G_2, \dots\}$  is a set of non-empty, bounded RS's, is a **germ-grain RS**. The points  $p_i$ ,  $i = 1, 2, \dots$  are the **germs**, whereas the RS's  $G_i$ ,  $i = 1, 2, \dots$  are the **primary grains** of the germ-grain RS  $X$ .

**Definition 8** Let  $\Phi$  be a PPP with intensity  $\lambda(x) \geq 0$ ,  $\forall x \in \mathcal{R}^n$ , and let  $\{G_1, G_2, \dots\}$  be a set of non-empty, bounded, and convex i.i.d. RS's, each with capacity functional  $T_G(K)$ , which are statistically independent of  $\Phi$ . If

$$E \left[ \int_{\mathcal{R}^n} 1_{G_1 \oplus K}(x) \lambda(x) dx \right] < \infty, \forall K \in \mathcal{K}$$

then the resulting germ-grain RS  $X$  is called a **Boolean RS with convex primary grains**.

In the following, we further restrict the nature of the primary grains. We assume that the primary grain  $G$  is of *fixed shape* (e.g. disc, square, hexagon, octagon) and only its size varies. More formally, we assume that  $G = RH$ , where  $H$  is a non-empty, compact, convex, "simple", fixed shape of unit size that contains the origin ( $H$  is a special case of a *structuring element*, in the nomenclature of Mathematical Morphology). Here,  $R$  is a random variable that measures the size of the primary grain, i.e.

$$rH \triangleq \begin{cases} \{z \in \mathcal{R}^n \ni \frac{z}{r} \in H\}, & r > 0 \\ \{\bar{0}\}, & r = 0 \end{cases} \quad (2)$$

We assume that  $R$  is distributed according to the pdf  $f_R(r)$ , which is assumed to be a distribution of compact support over  $\mathcal{R}$ , whose support is bounded by  $(0, \bar{R})$ , i.e.  $\text{spr}t f_R \subseteq (0, \bar{R})$ . Under these conditions  $G$  is a non-empty, bounded, convex RS containing the origin, of the same shape as  $H$ , but of size  $R$  times the size of  $H$ .

A primary grain restricted as above is a **radial random set** [12]. The Boolean RS  $X$  produced by using these grains, strictly speaking, *is not* a radial random set (for example, it does not necessarily contain the origin). Nevertheless, with some abuse of terminology, we shall call the resulting Boolean RS a **Radial Boolean RS (RBRS)**, to emphasize the nature of the underlying grain process.

Even though it is not necessary, we shall restrict ourselves to **planar** RBRS's, i.e. RBRS's defined on  $\mathcal{R}^2$ . This choice is simply made for clarity of presentation and also because of the importance of this special case in many applications. The results easily extend to higher dimensional spaces. Such a planar RBRS is completely specified by the triple

$$(\lambda(x); x \in \mathcal{R}^2, H, f_R(r); r \in (0, \bar{R}))$$

where  $\lambda(x); x \in \mathcal{R}^2$  is the intensity of the germ PPP,  $H$  is the unit size structuring element, and  $f_R(r)$  is the density of the radius  $R$ . We will denote it by  $(\lambda, H, f_R)$ -RBRS. For the moment let us fix  $H$  and use the simplified notation  $(\lambda, f_R)$ -RBRS, where it is understood that  $H$  is given and  $\lambda$  is a function and not a constant. With all this background in place, we are now ready to proceed to the first major result.

**Theorem 3** *Assume we have a PPP,  $\Phi$ , on  $\mathcal{R}^3$ , with intensity  $\lambda(z)$ ,  $z \in \mathcal{R}^3$ ,  $z = (x, r)$ ,  $x \in \mathcal{R}^2$ ,  $r \in \mathcal{R}$ , satisfying the following conditions*

$$\lambda(x, r) = \lambda_s(x)f_R(r), \quad \forall x \in \mathcal{R}^2, \quad \forall r \in \mathcal{R}$$

with

$$\lambda_s(x) \geq 0, \quad \forall x \in \mathcal{R}^2$$

and

$$\int_W \lambda_s(x) dx < \infty, \quad \forall W \in \mathcal{B}(\mathcal{R}^2)$$

$$f_R(r) \begin{cases} \geq 0, & \forall r \in (0, \bar{R}) \\ = 0, & \text{elsewhere} \end{cases}$$

$$\int_0^{\bar{R}} f_R(r) dr = 1$$

(such an intensity will be called **space-size separable**) **Then  $\Phi$  induces a unique (for fixed  $H$ )  $(\lambda_s, f_R)$ -RBRS,  $X$ , on  $\mathcal{R}^2$ , via**

$$X = \bigcup_{(x,r) \in \Phi} rH \oplus \{x\} \tag{3}$$

**Conversely**, a  $(\lambda_s, f_R)$ -RBRS,  $X$ , on  $\mathcal{R}^2$ , induces a unique PPP,  $\Phi$ , on  $\mathcal{R}^3$ , with intensity measure

$$\lambda(x, r) = \lambda_s(x)f_R(r), \quad \forall x \in \mathcal{R}^2, \quad \forall r \in \mathcal{R}$$

via

$$\Phi = \bigcup_{rH \oplus \{x\} \in X} \{(x, r)\} \quad (4)$$

**Remark:** *There exists an apparent subtlety in this theorem. The reconstruction formula (4) is valid (i.e. the PPP  $\Phi$  is indeed unique) only if we know the exact locations of the germs and the corresponding radii of the primary grains (“marks”). Otherwise  $\Phi$  is not unique. In other words, here we treat  $X$  as a **marked point process**. We will discuss this point in detail later on.*

**Proof (of theorem 3)**

If, as remarked above, we treat  $X$  as a marked point process, the result is a well known one [14, p145]. For completeness purposes, we proceed here with a proof which is heavily based on the defining properties of a PPP and also those of a RBRS.

By definition, (see (3) above),  $X$  is a germ-grain RS, with corresponding germ point process  $\Psi$ , given by

$$\Psi = \bigcup_{(x,r) \in \Phi} \{x\}$$

First, we show that  $\Psi$  is a PPP on  $\mathcal{R}^2$  with intensity  $\lambda_s$ . Let  $B$  be any bounded Borel subset of  $\mathcal{R}^2$ . Let  $C_B$  denote the set  $\{(x, r) \ni x \in B, r \in \mathcal{R}\}$ . Clearly, for  $B_1, B_2$  bounded, Borel subsets of  $\mathcal{R}^2$  we have that

$$B_1 \cap B_2 = \emptyset \iff C_{B_1} \cap C_{B_2} = \emptyset \quad (5)$$

Let  $W$  be any Borel subset of  $\mathcal{R}^3$ . Let  $\Phi(W)$  denote the number of points of  $\Phi$  in  $W$ . Since  $\Phi$  is a PPP with intensity  $\lambda(x, r)$ ,  $\Phi(W)$  is a Poisson r.v. with mean  $\Lambda(W) = \int_W \lambda(x, r) d(x, r)$ . Clearly,  $C_B \in \mathcal{B}(\mathcal{R}^3)$ . Therefore,  $\Phi(C_B)$  is a Poisson r.v. with mean  $\Lambda(C_B) = \int_{C_B} \lambda(x, r) d(x, r)$ . Now, let  $\Psi(B)$  denote the number of points of  $\Psi$  in  $B$ .

**Claim:**

$$\Psi(B) \stackrel{M.S.}{=} \Phi(C_B) \quad (6)$$

**Proof (of claim)**

By the Radon-Nykodym theorem, the assumed absolute continuity of the diffuse Radon measure  $\Lambda$  with respect to the Lebesgue measure implies that  $\Lambda$  assigns zero mass to all events that have zero Lebesgue measure.

Therefore, given that  $(x^*, r^*) \in \Phi$  the event  $\Phi(\{(x, r) \ni x = x^*, r \in \mathcal{R}\}) > 1$  has zero measure. The validity of the representation as a limit in the mean square sense can be shown using a limiting procedure which parallels that used in [14, p144].

Thus, by (6),  $\Psi(B)$  is a Poisson r.v. with mean

$$\begin{aligned} \int_{C_B} \lambda(x, r) d(x, r) &= \int_{C_B} \lambda_s(x) f_R(r) d(x, r) \\ &= \int_B \lambda_s(x) dx \int_{\mathcal{R}} f_R(r) dr = \int_B \lambda_s(x) dx \end{aligned}$$

as expected (i.e. Poisson points with intensity  $\lambda_s$ ). Furthermore, the independence of  $\Psi(B_1)$  and  $\Psi(B_2)$ , for any two disjoint Borel  $B_1$  and  $B_2$ , is directly implied by the independence of  $\Phi(C_{B_1})$  and  $\Phi(C_{B_2})$ , which in turn is implied by the fact that  $\Phi$  is a PPP and (5) (independent scattering). Therefore, *the germ point process,  $\Psi$ , is a PPP on  $\mathcal{R}^2$ , with intensity  $\lambda_s$ .*

The fact that the radii of the primary grains are i.i.d. and independent of  $\Psi$  is a direct consequence of the fact that  $\Phi$  is a PPP, and  $\lambda$  is space-size separable

$$\lambda(x, r) = \lambda_s(x) f_R(r), \quad \forall x \in \mathcal{R}^2, \quad \forall r \in \mathcal{R}$$

The fact that they are non-empty, bounded and convex follows from the restrictions posed on  $f_R$  and  $H$ . For the same reasons,

$$E \left[ \int_{\mathcal{R}^2} 1_{G_1 \oplus K}(x) \lambda_s(x) dx \right] < \infty, \quad \forall K \in \mathcal{K}$$

Therefore,  $X$  is a  $(\lambda_s, f_R)$ -RBRS.

**Conversely:**

Assume  $X$  is a  $(\lambda_s, f_R)$ -RBRS on  $\mathcal{R}^2$ . By definition

$$\Phi = \bigcup_{rH \oplus \{x\} \in X} \{(x, r)\}$$

and, therefore,  $\Phi$  is a point process on  $\mathcal{R}^3$ . Let  $W$  be any bounded Borel subset of  $\mathcal{R}^3$ . Then

$$\Phi(W) = \left| W \cap \bigcup_{rH \oplus \{x\} \in X} \{(x, r)\} \right|$$

where  $|\cdot|$  denotes set cardinality. Then

$$\begin{aligned}\Phi(W) &= \left| \bigcup_{rH \oplus \{x\} \in X} \{(x, r)\} \cap W \right| = |\{(x, r) \in W \ni rH \oplus \{x\} \in X\}| \\ &= |\{x \in W_x \ni x \in \Psi \text{ and } (x, R(x)) \in W\}| \end{aligned}$$

where  $W_x$  is the projection of  $W$  onto the subspace spanned by  $x$ ,  $\Psi$  is the germ process of the RBRS  $X$  on  $\mathcal{R}^2$ , and the r.v.  $R(x)$  is the radius (size) corresponding to the site  $x \in \Psi$ . Clearly,  $\Phi(W)$  is a Poisson r.v. with parameter

$$\int_W \lambda_s(x) f_R(r) d(x, r)$$

In order to see this, we use the constructive definition for PPP's on bounded subsets, to set up the following, completely equivalent, experiment [14, p69]. First select a nonnegative integer  $N$ , according to the Poisson distribution, and with parameter  $\int_{W_x} \lambda_s(x) dx$ . Then, select  $N$  points, independently over  $W_x \times (0, \bar{R})$ , each with probability density function

$$\frac{\lambda_s(x) f_R(r)}{\int_{W_x} \lambda_s(x) dx}, \quad x \in W_x, \quad r \in (0, \bar{R})$$

Then [14, p69] the resulting point process is a PPP on  $W_x \times (0, \bar{R})$  with intensity  $\lambda_s(x) f_R(r)$ ,  $x \in W_x$ ,  $r \in (0, \bar{R})$ . Therefore,  $\Phi(W)$  is a Poisson r.v. with parameter

$$\int_W \lambda_s(x) f_R(r) d(x, r)$$

The independent scattering property of  $\Phi$  is implied by the properties of the underlying RBRS  $X$ . If  $W_1$  and  $W_2$  are two disjoint Borel subsets of  $\mathcal{R}^3$  such that  $W_{1_x} \cap W_{2_x} = \emptyset$  then the independence of the r.v.'s  $\Phi(W_1)$  and  $\Phi(W_2)$  is implied by the independence of  $\Psi(W_{1_x})$  and  $\Psi(W_{2_x})$  and the i.i.d. assumption on the corresponding radii. In case  $W_{1_x} \cap W_{2_x} \neq \emptyset$  then independence is guaranteed by the fact that  $\Psi(W_{1_x} \cap W_{2_x})$  is a Poisson r.v., and, given  $\Psi(W_{1_x} \cap W_{2_x}) = n$ , each site makes an independent decision in regard to its corresponding radius. To see this, set up the following, completely equivalent, experiment. Let  $N$  be a Poisson r.v. with parameter

$$\mu = \int_{W_{1_x} \cap W_{2_x}} \lambda_s(x) dx$$

Given  $N = n$ , select  $n$  real-valued, i.i.d. r.v.'s, each distributed in the interval  $(0, \bar{R})$  according to the pdf  $f_R$ . Then the resulting process is a PPP on  $(0, \bar{R})$ , with intensity  $\mu f_R$ , and, therefore, the independent scattering property holds for disjoint subsets.

### 3.1 Binary Hypothesis Testing for RBRS's

Consider the following simple-vs-simple binary hypothesis testing problem

$$\begin{aligned} H_0: & \quad Y \sim (\lambda_s^{(0)}, H, f_R^{(0)})\text{-RBRS} \quad (\text{“null” hypothesis}) \\ \text{vs:} H_1: & \quad Y \sim (\lambda_s^{(1)}, H, f_R^{(1)})\text{-RBRS} \quad (\text{“alternative” hypothesis}) \end{aligned}$$

Assume that we observe the true marked point process corresponding to  $X \triangleq Y \cap B$ , where  $B$  is bounded, Borel. According to our results, this is equivalent to observing the corresponding PPP on  $B \times (0, \bar{R})$ . Therefore, the log-likelihood ratio is given by

$$\begin{aligned} & \int_{B \times (0, \bar{R})} \left( \lambda_s^{(0)}(x) f_R^{(0)}(r) - \lambda_s^{(1)}(x) f_R^{(1)}(r) \right) d(x, r) + \sum_{rH \oplus \{x\} \in X} \log \frac{\lambda_s^{(1)}(x) f_R^{(1)}(r)}{\lambda_s^{(0)}(x) f_R^{(0)}(r)} \\ & = \int_B \left( \lambda_s^{(0)}(x) - \lambda_s^{(1)}(x) \right) dx \\ & + \sum_{r \ni \exists x \in B \ni rH \oplus \{x\} \in X} \sum_{x \in B \ni rH \oplus \{x\} \in X} \log \frac{\lambda_s^{(1)}(x) f_R^{(1)}(r)}{\lambda_s^{(0)}(x) f_R^{(0)}(r)} \end{aligned}$$

The first term is a constant that can be precomputed and does not affect the test statistic (only affects the choice of threshold). Let  $\{L_1, \dots, L_N\}$  denote the set of ordered lists of sites with corresponding radii  $\{r_1, \dots, r_N\}$ , where  $\bar{R} > r_1 > r_2 > \dots > r_N > 0$ . Let  $|L_i|$  denote the cardinality of  $L_i$ . Then the second term can be written as

$$\begin{aligned} & \sum_{r \in \{r_1, \dots, r_N\}} \sum_{x \in L_r} \log \frac{\lambda_s^{(1)}(x) f_R^{(1)}(r)}{\lambda_s^{(0)}(x) f_R^{(0)}(r)} \\ & = \sum_{r \in \{r_1, \dots, r_N\}} \left( \sum_{x \in L_r} \log \frac{\lambda_s^{(1)}(x)}{\lambda_s^{(0)}(x)} + \sum_{x \in L_r} \log \frac{f_R^{(1)}(r)}{f_R^{(0)}(r)} \right) \\ & = \sum_{r \in \{r_1, \dots, r_N\}} \left( \sum_{x \in L_r} \log \frac{\lambda_s^{(1)}(x)}{\lambda_s^{(0)}(x)} + |L_r| \log \frac{f_R^{(1)}(r)}{f_R^{(0)}(r)} \right) \\ & = \sum_{x \in \cup_{r \in \{r_1, \dots, r_N\}} L_r} \log \frac{\lambda_s^{(1)}(x)}{\lambda_s^{(0)}(x)} + \sum_{r \in \{r_1, \dots, r_N\}} |L_r| \log \frac{f_R^{(1)}(r)}{f_R^{(0)}(r)} \end{aligned}$$

Therefore, the information that is needed comprises of two parts: (1) the collection of sites, regardless of the corresponding radii, and, (2) the relative distribution (histogram data) for the measured radii, regardless of the places where they occur. The complete log-likelihood ratio is given by

$$\int_B \left( \lambda_s^{(0)}(x) - \lambda_s^{(1)}(x) \right) dx$$

$$+ \sum_{x \in \cup_{r \in \{r_1, \dots, r_N\}} L_r} \log \frac{\lambda_s^{(1)}(x)}{\lambda_s^{(0)}(x)} + \sum_{r \in \{r_1, \dots, r_N\}} |L_r| \log \frac{f_R^{(1)}(r)}{f_R^{(0)}(r)}$$

## 4 Morphological Skeleton and Shape Decomposition

In the previous section we have treated the RBRS  $X$  as a *marked point process*, i.e. we assumed that the observations consist of the visible germ locations and the corresponding radii. This assumption *is not* plausible, because in practice we are presented with a realization of the RBRS  $X$ , and the actual germ locations and corresponding radii are unknown. Except for a somewhat degenerate case, it is not possible to recover the germ locations and radii with certainty, and, therefore, we have to *estimate* these data from the available observations. This ambiguity is due to an inherent non-uniqueness in representing a set  $X$  as a union of “simple” subsets of the form

$$X = \bigcup_i L_i \oplus r_i H \quad (7)$$

even if  $X$  is a realization of a RBRS and the structuring element  $H$  is fixed. In fact, there exists a class of morphological transforms that allow for invertibility via dilation, as in (7). Reference [3] states a necessary and sufficient condition (for the discrete case) for invertibility via dilation for a fairly general class of morphological transforms. As if this was not enough, other approaches [10, 8] also allow for reconstruction via dilation. These different approaches represent the different ways in which we can perceive a realization of a RBRS with overlapping grains. The natural question, then, is which particular representation to pick (or, “believe”). From a statistical point of view, we should pick the “most probable” one. This is very difficult, both analytically, and computationally, because it implies the use of a spatially and/or radially adaptive representation scheme. For the sake of simplicity, a “nearly optimal” strategy needs to be pursued, but even this is not a trivial choice. For these reasons, we shall consider two possible representation schemes, the morphological skeleton transform, and morphological shape decomposition.

### 4.1 Basic morphological operators

Basic morphological operators are defined in terms of a “simple” set,  $H$ , which is bounded, and, for our purposes, convex. In morphological terms,  $H$  is called a *structuring element*, and, in our setup, it is exactly the unit size primary grain which is used to construct the RBRS.



**Definition 9** The **erosion**,  $X \ominus H^s$ , of a RS  $X$ , by a structuring element  $H$ , is defined as

$$X \ominus H^s = \bigcap_{h \in H} X_{-h} = \{z \in \mathcal{R}^n \ni H_z \subset X\}$$

**Definition 10** The **dilation**,  $X \oplus H^s$ , of a RS  $X$ , by a structuring element  $H$ , is defined as

$$X \oplus H^s = \bigcup_{h \in H} X_{-h} = \{z \in \mathcal{R}^n \ni H_z \cap X \neq \emptyset\}$$

**Definition 11** The **opening**,  $X_H$ , of a RS  $X$ , by a structuring element  $H$ , is defined as

$$X_H = (X \ominus H^s) \oplus H$$

**Definition 12** The **closing**,  $X^H$ , of a RS  $X$ , by a structuring element  $H$ , is defined as

$$X^H = (X \oplus H^s) \ominus H$$

If  $X$  is a RS then  $X \ominus H^s$ ,  $X \oplus H^s$ ,  $X_H$ ,  $X^H$  are RS's. Also, if  $X_1$  and  $X_2$  are RS's, then  $X_1 \cap X_2$ ,  $X_1 \cup X_2$  are RS's [2].

## 4.2 Morphological Skeleton Transform

The *Morphological Skeleton*,  $SK(X)$ , of a set  $X$ , with respect to a (fixed) structuring element,  $H$ , is given by

$$SK(X) = \bigcup_{r>0} S_r(X) = \bigcup_{r>0} [(X \ominus rH^s) - (X \ominus rH^s)_{drH}]$$

where  $H$  is an open, bounded, convex structuring element, and  $drH$  is the closure of a replica of  $H$  of infinitesimally small radius. Reconstruction from the skeleton subsets,  $\{S_r(X), r > 0\}$ , is possible, via

$$X = \bigcup_{r>0} [S_r(X) \oplus rH]$$

The morphological skeleton is the locus of the centers of the maximal inscribable replicas of the structuring element,  $H$ , inside the closed set  $X$ . A replica of the structuring element is maximal in  $X$ , if it is not properly contained in any other replica totally included in  $X$ .

There exist many excellent references for morphological skeletonization. However, at least within our scope, the most useful one is [3].

### 4.3 Morphological Shape Decomposition

The *Morphological Shape Decomposition* is another approach which has been proposed in [10]. It has subsequently been enhanced and used for pattern recognition purposes in [8]. A recursive formulation of the morphological shape decomposition algorithm is now presented.

$$X_i = \left( (X - \tilde{X}_{i-1}) \ominus r_i H^s \right) \oplus r_i H = (X - \tilde{X}_{i-1})_{r_i H}, \quad i = 1, 2, \dots$$

where

$$r_i = \sup \left\{ r \ni (X - \tilde{X}_{i-1}) \ominus r H^s \neq \emptyset \right\}$$

and the reconstruction is given by

$$\tilde{X}_i = \bigcup_{0 \leq j \leq i} X_j, \quad \tilde{X}_0 = \emptyset, \quad X = \bigcup_{j \geq 0} X_j$$

The  $X_i$ 's are the (clearly disjoint) subsets of the decomposition. Each  $X_i$  is the union of a finite number of primitives (translated and scaled replicas of the structuring element, or, primary grain), each of (fixed for each subset) size  $r_i$ . The first cluster of the decomposition,  $X_1$ , is the union of all maximal inscribable replicas of  $H$  in  $X$ .  $X_1$  is subsequently subtracted from  $X$ , and the procedure is executed recursively on the remaining set. Observe that, under our assumptions,  $r_1$  is *finite*. In fact,  $r_1 < \bar{R}$ .

Because of the special structure of the input RBRS  $X$ , each  $X_i$  can be written as

$$X_i = L_i \oplus r_i H$$

where  $L_i$  is a set of *isolated points* (i.e. the germ locations of those primary grains having size exactly  $r_i$ ). Each subset,  $X_i$ , is (morphologically) *open* set with respect to  $r_i H$  [7], i.e.  $(X_i)_{r_i H} = X_i$ . Therefore,  $L_i$  can be recovered from  $X_i$  via

$$L_i = X_i \ominus r_i H^s$$

Hence we have the following, completely equivalent, formulation of the morphological shape decomposition algorithm

$$L_i = \left( X - \bigcup_{0 \leq j \leq i-1} (L_j \oplus r_j H) \right) \ominus r_i H^s, \quad i = 1, 2, \dots$$

where

$$r_i = \sup \left\{ r \ni \left( X - \bigcup_{0 \leq j \leq i-1} (L_j \oplus r_j H) \right) \ominus r_i H^s \neq \emptyset \right\}$$

and

$$L_0 = \emptyset, \quad \tilde{X}_i = \bigcup_{0 \leq j \leq i} (L_j \oplus r_j H)$$

The subsets  $L_i$  are the “spine” subsets of the morphological shape decomposition.

#### 4.4 Comparative Discussion

The two representation schemes presented above map the input RBRS into a collection of ordered lists of sites,  $\{L_1, \dots, L_N\}$  (the skeleton subsets,  $S_r$ , or the morphological shape decomposition “spine” subsets  $L_i$ ), with corresponding radii  $\{r_1, \dots, r_N\}$ , where  $\bar{R} > r_1 > r_2 > \dots > r_N > 0$ . As we mentioned above, these two (generally distinct) mappings, correspond to two possible ways of viewing the input RBRS  $X$ . However, if the primary grains of the input RBRS  $X$  are not overlapping, then the two representations are identical and, in fact, map  $X$  to the true underlying marked point process. This is because the  $r^{\text{th}}$  skeleton subset of the union of non-overlapping sets is the union of the  $r^{\text{th}}$  skeleton subsets of the non-overlapping sets, and the same is true for morphological shape decomposition “spine” subsets (for the later, the union is taken over the spine subsets corresponding to the same (given) *radius* **not** the same *index*). Furthermore, the skeleton of  $rH \oplus \{x\}$  consists of the unique point  $x$  with corresponding radius  $r$ , and it is the same as its morphological shape decomposition. Therefore, the true log-likelihood ratio can be computed using either one of the two representations, and a truly Bayesian test can be formed. This assumption (i.e. that the primary grains are not overlapping) is sometimes plausible (e.g. when the intensity of the germ point process is low, and/or  $\bar{R}$  is small). It is common practice to make such an assumption when the RBRS  $X$  is used to model a degradation process [12]. If this is not the case, then a probabilistic treatment is appropriate. Since this later approach is more suited for the discrete case, we shall consider it in detail in the corresponding section.

## 5 The Discrete Case

Up to this point, our development has focused on the continuous case. In practical terms, though, we observe discretized data. The utility of the approach lies on its power to model the problem rigorously. First, the problems and the results are stated in terms of continuous domains that model

the applications. The accuracy of the implementation is then understood to depend on the sampling rate, with convergence as sampling rates increase. Nevertheless, the intrinsic continuity of the PPP model, plus an inherent incompatibility between the continuous and discrete cases in formalizing the notion of *size* of a structuring element, make it necessary to consider the discrete case in its own right.

The discrete “analog” of a planar PPP,  $\Psi$ , observed through a bounded Borel window,  $B$ , is defined in what follows.

**Definition 13** *Let  $B$  be a bounded subset of  $\mathcal{Z}^2$ . Let  $\Sigma(B)$  denote the power set (i.e. the set of all subsets) of  $B$ . A **Discrete Point Process (DPP)**,  $\Psi$ , on  $B$ , is a measurable mapping of a probability space  $(\Omega, \Sigma(\Omega), P)$  into the measurable space  $(B, \Sigma(B))$ .*

Informally,  $\Psi$  can be thought of as a random pattern of points scattered over  $B$ . The continuous case PPP can be derived from a generalization of the continuous case Binomial Point Process [15, pp36-38], by using a limiting argument. Alternatively, the continuous case PPP can also be derived from a generalization of the *Bernoulli Lattice Process (BLP)* [15, pp40-42], again by using a limiting argument. This, then, is the correct discrete-case analog of the PPP.

**Definition 14** *A **Generalized Bernoulli Lattice Process (GBLP)**,  $\Psi$ , on  $B$ , is a discrete point process on  $B$  which is constructively defined in the following manner. Each point  $x \in B$  is contained in  $\Psi$  with probability  $p\lambda_s(x)$ , independently of all others. Here,  $p \in (0, 1]$  and  $\lambda_s(x) \in [0, 1]$ ,  $\forall x \in B$ .*

Clearly, the GBLP enjoys the independent scattering property. In contrast with equation (2), in the discrete case the notion of size is formalized via the operation of set dilation

$$rH = \begin{cases} \{\bar{0}\} \oplus H \oplus H \oplus \cdots \oplus H, & (r \text{ dilations}), r = 1, 2, \dots \\ \{\bar{0}\} & , r = 0 \end{cases} \quad (8)$$

Observe that in the continuous case  $r \in \mathcal{R}_+$ ,  $r < \bar{R}$ , whereas in the discrete case  $r \in \mathcal{Z}_+$ ,  $r < \bar{R}$ . Further observe that in the discrete case (at least in principle) there is no distinction between random point processes and random sets; in fact discrete random sets *are* discrete random point processes, because the regularity conditions are automatically satisfied here (discrete random sets *are* locally finite and simple). Therefore, even though one could

mimic the continuous case construction of the RBRS, we will skip this part for compactness. It suffices to say that discrete RS's can be formally defined as discrete point processes. We have the following definition.

**Definition 15** Let  $\Psi$  be a GPLP on  $B$  with parameters  $(p, \lambda_s)$ . Let  $\{G_1, G_2, \dots\}$  be a set of nonempty, bounded and convex i.i.d. discrete RS's, on  $B' \subset B$ ,  $|B'| \ll |B|$ , each given by  $G_i = R_i H$ , where  $\{R_1, R_2, \dots\}$  form an i.i.d. sequence of  $\mathcal{Z}_+$ -valued r.v.'s which is independent of  $\Psi$ ,  $R_i < \bar{R}$ ,  $\forall i$ , and each  $R_i$  is distributed according to the pmf  $f_R(r)$ , which is compactly supported on  $\{0, 1, \dots, \bar{R} - 1\}$ . Define

$$X = \bigcup_{i=1,2,\dots} G_i \oplus \{y_i\}$$

where  $\Psi = \{y_1, y_2, \dots\}$ . Then  $X$  will be called a **Discrete Radial Boolean RS (DRBRS)**, with parameters  $(p, \lambda_s, H, f_R)$ , and will be denoted by  $(p, \lambda_s, H, f_R)$ -DRBRS.

**Remark:** Here we assume that  $\bar{R}$  and  $B'$  are sufficiently small relative to  $B$  such that all grains that are *visible* (i.e. intersect  $B$ ) are *entirely visible* (i.e. fall completely inside  $B$ ).

## 5.1 Binary Hypothesis Testing for the case of DRBRS's

The problem of hypothesis testing for DRBRS's is more directly amenable to analysis than the corresponding continuous case problem, mainly because it only involves pmf's instead of sample density functions. Therefore, standard tools are sufficient to characterize the solution. In particular, there is no need to map this problem to an equivalent binary hypothesis testing problem for GPLP's. We consider the simple vs simple hypothesis testing problem

$$\begin{aligned} H_0: X &\sim (p^{(0)}, \lambda_s^{(0)}, H, f_R^{(0)})\text{-DRBRS} \\ \text{vs } H_1: X &\sim (p^{(1)}, \lambda_s^{(1)}, H, f_R^{(1)})\text{-DRBRS} \end{aligned}$$

Here, again, we assume that we observe the corresponding marked point process, i.e. an ordered list of sites,  $\{L_1, \dots, L_N\}$  corresponding to radii  $\{r_1, \dots, r_N\}$  respectively, where  $\bar{R} > r_1 > r_2 > \dots > r_N \geq 0$ ,  $r_i \in \mathcal{Z}_+$ , where  $N \leq \bar{R}$ . Under hypothesis  $j$ ,  $j = 0, 1$ , look at the conditional probability  $Pr_j\{(L_1, r_1), \dots, (L_N, r_N)\}$ . The crucial observation is that the sets

$\{L_i\}_{i=1}^N$  are *disjoint*, because the underlying GBLP is a simple point process.

$$\begin{aligned}
Pr_j\{(L_1, r_1), \dots, (L_N, r_N)\} &= \\
&\prod_{x \in B \ni x \notin \cup_{i=1}^N L_i} (1 - p^{(j)} \lambda_s^{(j)}(x)) \prod_{i=1}^N \prod_{x \in L_i} p^{(j)} \lambda_s^{(j)}(x) f_R^{(j)}(r_i) \\
&= \prod_{x \in B \ni x \notin \cup_{i=1}^N L_i} (1 - p^{(j)} \lambda_s^{(j)}(x)) \prod_{i=1}^N \left\{ \left[ p^{(j)} f_R^{(j)}(r_i) \right]^{|L_i|} \prod_{x \in L_i} \lambda_s^{(j)}(x) \right\}
\end{aligned}$$

Therefore, the likelihood ratio

$$\frac{Pr_1\{(L_1, r_1), \dots, (L_N, r_N)\}}{Pr_0\{(L_1, r_1), \dots, (L_N, r_N)\}}$$

is equal to

$$\prod_{x \in B \ni x \notin \cup_{i=1}^N L_i} \frac{(1 - p^{(1)} \lambda_s^{(1)}(x))}{(1 - p^{(0)} \lambda_s^{(0)}(x))} \prod_{i=1}^N \left\{ \left[ \frac{p^{(1)} f_R^{(1)}(r_i)}{p^{(0)} f_R^{(0)}(r_i)} \right]^{|L_i|} \prod_{x \in L_i} \frac{\lambda_s^{(1)}(x)}{\lambda_s^{(0)}(x)} \right\}$$

The log-likelihood ratio

$$\log \frac{Pr_1\{(L_1, r_1), \dots, (L_N, r_N)\}}{Pr_0\{(L_1, r_1), \dots, (L_N, r_N)\}}$$

is given by

$$\begin{aligned}
&\sum_{x \in B \ni x \notin \cup_{i=1}^N L_i} \log \left( \frac{1 - p^{(1)} \lambda_s^{(1)}(x)}{1 - p^{(0)} \lambda_s^{(0)}(x)} \right) + \sum_{i=1}^N |L_i| \log \left( \frac{p^{(1)} f_R^{(1)}(r_i)}{p^{(0)} f_R^{(0)}(r_i)} \right) \\
&\quad + \sum_{i=1}^N \sum_{x \in L_i} \log \frac{\lambda_s^{(1)}(x)}{\lambda_s^{(0)}(x)}
\end{aligned}$$

## 5.2 Discrete Morphological Skeleton Transforms and Shape Decomposition

The definitions of the basic morphological operators for the discrete case are identical to the corresponding definitions for the continuous case, except for the obvious change of domain. Nevertheless, the transition from continuous skeletons to discrete skeletons is a troublesome one. Within our scope, the

difficulty arises from the fact that the discrete-case skeleton of a realization of a DRBRS does not comprise of isolated (in the 8-nearest neighbor sense) points, but it generally consists of very thin “stripes” whose width is smaller than the diameter of the unit size structuring element. In a sense, the width of these stripes corresponds to a resolution threshold (constraint) which is set by the size of the unit structuring element.

**The Discrete Morphological Skeleton** of a set  $X \in \mathcal{Z}^2$ , with respect to a (fixed) convex and bounded structuring element,  $H \in \mathcal{Z}^2$ , is given by

$$SK(X) = \bigcup_{n=0}^N S_n(X) = \bigcup_{n=0}^N [(X \ominus nH^s) - (X \ominus nH^s)_H]$$

where

$$N = \max \{n \ni X \ominus nH^s \neq \emptyset\}$$

Reconstruction from the skeleton subsets is possible via dilation

$$X = \bigcup_{n=0}^N [S_n(X) \oplus nH]$$

**The Reduced Discrete Morphological Skeleton** of a set  $X \in \mathcal{Z}^2$ , with respect to a (fixed) convex and bounded structuring element,  $H \in \mathcal{Z}^2$ , is given by

$$RSK(X) = \bigcup_{n=0}^N RS_n(X) = \bigcup_{n=0}^N [(X \ominus nH^s) - ((X \ominus nH^s)_H)^{nH}]$$

where

$$N = \max \{n \ni X \ominus nH^s \neq \emptyset\}$$

Reconstruction from the reduced skeleton subsets is possible via dilation

$$X = \bigcup_{n=0}^N [RS_n(X) \oplus nH]$$

**The Discrete Morphological Shape Decomposition** of a set  $X \in \mathcal{Z}^2$ , with respect to a (fixed) convex and bounded structuring element,  $H \in \mathcal{Z}^2$ , is given by

$$X_i = ((X - \tilde{X}_{i-1}) \ominus n_i H^s) \oplus n_i H = (X - \tilde{X}_{i-1})_{n_i H}, \quad i = 1, 2, \dots$$

where

$$n_i = \max \left\{ n \geq 0 \ni (X - \tilde{X}_{i-1}) \ominus nH^s \neq \emptyset \right\}$$

and the reconstruction formula

$$\tilde{X}_i = \bigcup_{j=0}^i X_j, \quad \tilde{X}_0 = \emptyset, \quad X = \bigcup_{j \geq 0} X_j$$

Again, each subset,  $X_i$ , is a (morphologically) *open* set with respect to  $n_i H$  [7], i.e.  $(X_i)_{n_i H} = X_i$ . Therefore,  $L_i$  can be recovered from  $X_i$  via

$$L_i = X_i \ominus n_i H^s$$

Hence we have the following, completely equivalent, formulation of the discrete morphological shape decomposition algorithm

$$L_i = \left( X - \bigcup_{0 \leq j \leq i-1} (L_j \oplus n_j H) \right) \ominus n_i H^s, \quad i = 1, 2, \dots$$

where

$$n_i = \max \left\{ n \geq 0 \ni \left( X - \bigcup_{0 \leq j \leq i-1} (L_j \oplus n_j H) \right) \ominus n_i H^s \neq \emptyset \right\}$$

and

$$L_0 = \emptyset, \quad \tilde{X}_i = \bigcup_{0 \leq j \leq i} (L_j \oplus n_j H)$$

The subsets  $L_i$  are the “spine” subsets of the discrete morphological shape decomposition.

Let  $Q$  denote the nine pixel *square* structuring element, depicted in figure (1). The following definition will be useful.

**Definition 16** [12] *Two sets  $X_1$  and  $X_2$  are said to be **disconnected sets** if*

$$X_1 \oplus Q \cap X_2 = \emptyset \iff X_1 \cap X_2 \oplus Q = \emptyset$$

This condition essentially guarantees that no two or more grains can compose to form a structuring element, or a connected (in the 8-nearest neighbor sense) component.

If the visible primary grains,  $\{G_i \oplus \{y_i\}\}$ , of the observed realization of the DRBRS  $X$  are pairwise disconnected, then there exists a *unique* representation i.e. all three representations above map  $X$  to the true marked



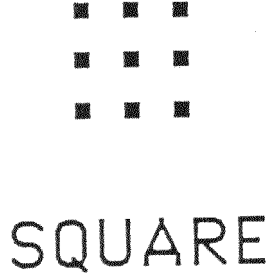


Figure 1: *Square* structuring element comprising of 9 pixels.

point process. A brief explanation is given below. The opening of the union of disconnected sets with respect to any given structuring element equals the union of the openings of the disconnected sets with respect to the given structuring element [12]. The  $n^{\text{th}}$  skeleton subset of the union of disconnected sets can be easily shown to be the union of the  $n^{\text{th}}$  skeleton subsets of the disconnected sets. The same is true for reduced skeleton subsets, or morphological shape decomposition “spine” subsets (for the later, the union is taken over the spine subsets corresponding to the same (given) *radius* **not** the same *index*). Finally, the skeleton of  $nH \oplus \{x\}$  consists of the unique point  $x$  with corresponding radius  $n$ , and it is the same as its reduced skeleton or its morphological shape decomposition.

### 5.3 Discussion

The three transformations presented above map the input DRBRS  $X$  into three (generally different) representations, except for the case when the grains of  $X$  are disconnected. In the later case, all three representations are identical, and map  $X$  to the true underlying marked point process. In all other cases, since the true marked point process is “hidden”, one would like to form an estimate of it, based on the observation  $X$ . The three transformations above correspond to three *nonlinear* estimates of the true marked point process, but, in fact, one can form many other estimates. Observe, though, that all three representations form *sufficient statistics*, in the sense that they all allow for complete reconstruction of the original observation (i.e. the realization of the input DRBRS  $X$ ).

In principle, one would like to base the test on the actual observation,

which is a realization of the DRBRS  $X$ . By the independence assumptions, this reduces to collectively testing for the connected components of  $X$ , i.e. the likelihood under a given hypothesis is equal to the product of the likelihoods for each connected component, under the same hypothesis. This approach is extremely difficult, both analytically, and computationally, because usually there exists a very large number of germ-grain configurations that can give rise to the same connected component. Furthermore, the connected components need to be found and specified in an efficient manner. Therefore, this approach is not practical.

The morphological shape decomposition provides an estimate which can be very unreasonable, especially in the case of two just slightly overlapping grains. In fact, the smaller one of the two grains will *never* be recovered, and instead it will be represented as a union of many more grains of smaller size. In effect, this introduces spurious germs (“responses”), and shifts the size distribution to the lower end of the pattern spectrum (the *pattern spectrum* is a transform that measures the *size distribution* of its input set. There exist many similarities between the pattern spectrum and frequency domain spectra. The lower end of the pattern spectrum corresponds to small sizes, while the upper end corresponds to bigger sizes. The pattern spectrum is related to the morphological skeleton transforms. Several important results appear in [3, 6]). Furthermore, the morphological shape decomposition is more computationally intensive than the skeleton transforms. For these reasons, this representation is not (generally speaking) a good choice.

The morphological skeleton gives a reasonable representation, and it avoids the problem discussed above. This is because maximal inscribable replicas of the structuring element are allowed to overlap with other maximal inscribable replicas, even though they can not be properly contained in other maximal inscribable replicas. A fast skeletonization algorithm has been proposed [5]. The morphological skeleton computes a “dense” estimate of the underlying germ point process, in the sense that it computes as many maximal grains as possible. The representation favors the upper end of the pattern spectrum (i.e. large grains), because it computes *maximal* grains.

Even though a statistical viewpoint has been adopted throughout this work, it is useful to pause here and elaborate on what we perceive as a “reasonable” representation. In gestalt psychology, the *law of simplicity* states that *every stimulus pattern is seen in such a way that the resulting structure is as simple as possible*. In an attempt to quantify this law, we model the realization of the input DRBRS  $X$  as a *minimal union of maximal inscribable replicas of the structuring element*, where *minimal* means

that the total number of replicas needed to represent  $X$  should be minimum. This implies the use of the so-called **Globally Minimal Skeleton Transform**. This transform does not possess several nice properties of the skeleton transform, and it is difficult to obtain. A compromise would be to use the reduced skeleton transform, which removes some, but not all, of the redundancy present in the skeleton transform. The reduced skeleton transform has another interesting feature: it is strongly related to the pattern spectrum [6, 3].

Assume that, under both hypotheses, the observed data are spatially sparse and the degree of overlap (or clustering) of the primary grains is low. This is plausible if the product  $p^{(j)}\lambda_s^{(j)}$ , for  $j = 0, 1$ , is uniformly small over the observation window  $B$ , and  $\bar{R}$  is relatively small. This assumption is justified if the DRBRS is used to model a degradation process. The stronger assumption that the primary grains are disconnected is frequently used in such cases [12]. Then a statistically good approach is to choose the representation that gives subsets of minimal cardinality. The reduced morphological skeleton subsets have the smallest cardinality among a fairly large class of morphological representations [3, 6]. Therefore, the reduced morphological skeleton is a statistically good choice, which also happens to be relatively consistent with the perceptual arguments of gestalt psychology.

Extensive simulations have shown that, on the average, the cardinality of the union of the morphological shape decomposition spine subsets of a given set is smaller than the cardinality of the union of the skeleton subsets of the same set [8]. This is because the former representation does not allow maximal inscribable replicas to overlap. Therefore, were it not for the fact that it sometimes produces unreasonable representations, the morphological shape decomposition would have been a good candidate too.

Some task-specific improvements can also be pursued. Depending on the particular combination of parameters that characterize a given problem, one can discard some reduced skeleton subsets, or weight the reduced skeleton subsets according to the assumed distribution for the radii. For example, if the radii are all fixed and equal to a constant, then one should only keep the reduced skeleton subset with radius equal to this constant. Another approach is to apply a *thinning operation* on the reduced skeleton subsets [13]. Finally, if  $\lambda_s^{(0)} = \lambda_s^{(1)}$  the only test statistic needed is the relative distribution (histogram data) of the radii, and therefore one need only compute the pattern spectrum of the DRBRS  $X$ .

In any case, these approximations result in a test which *is not* optimal.

If the degree of overlap (or clustering) of the primary grains is small, then using the reduced morphological skeleton will result in a test which is *almost optimal* (*almost Bayesian*). In case the primary grains are disconnected, the test is *truly optimal* (*truly Bayesian*).

Simulation results have demonstrated the power of this approach, even when the primary grains overlap a lot, and the morphological skeleton representation is used instead of the reduced morphological skeleton representation. The test is able to decide correctly for inputs that a human observer would probably find hard to distinguish. It seems that even when there exists a significant overlap of primary grains, the test is balanced by the influence of those grains that are entirely visible. The best case seems to be when  $f_R^{(0)} = f_R^{(1)}$ . In this case, the size distribution is of no concern, and the introduction of spurious germs actually *enhances* the distinguishing power of the test. The worst case is when  $\lambda_s^{(0)} = \lambda_s^{(1)}$ . In this case, the introduction of spurious germs can distort the size distribution significantly, and, therefore, the power of the test can be reduced. Extensive simulations are needed, in order to determine empirical rules for the power of the test, relative to the various representations and design parameters involved.

In the three figures that follow, (2), (3), (4), the 21 pixel discrete structuring element *octagon*, a realization of a DRBRS with parameters  $p = 0.001$ ,  $\lambda_s(i, j) = j/511$ ,  $H = \text{octagon}$ , and  $f_R = \text{uniform in } \{0, \dots, 9\}$ , and its morphological skeleton are shown. The resolution is  $512 \times 512$  pixels. As it can be seen, a significant amount of overlap is present. The hypothesis of  $\lambda_s(i, j) = j/511$ ,  $\forall i$  was tested vs. the hypothesis  $\lambda_s(i, j) = 0.5$ ,  $\forall i, j$ , where all the other parameters were assumed to be identical, as above. The morphological skeleton points were used as the estimate of the true underlying germ point process and the log-likelihood ratio was computed, based on these points. The result was then compared to 0. This procedure corresponds to a test which is “close” (in the sense outlined above) to the true Maximum Likelihood test. The test resulted in the correct decision, and did so with a significantly wide margin. The value of the log-likelihood ratio was of the order of  $10^6$ .

## 6 Conclusions and further research

We have employed several techniques from the areas of statistical inference for point processes, random set theory, and mathematical morphology, to come up with an almost Bayesian binary hypothesis testing procedure, for



Figure 2: *Octagon* structuring element comprising of 21 pixels

the case of Boolean random sets with radial convex primary grains. We have considered both the continuous and the discrete case, and we have derived explicit formulas for approximately Bayesian tests. In the continuous case, the tests are truly Bayesian if the primary grains are not overlapping, whereas, in the discrete case, the tests are truly Bayesian when the primary grains are disconnected. In case of limited overlap, the tests are quite robust, and very close to the optimal Bayesian tests.

The random set models that we have considered are “purely random” in the sense that no interaction is allowed between neighboring germs. Therefore, they are suitable for modeling “noise” processes. However, one usually wants to test the presence of a known signal (shape), which is hidden in such a noise process. One possible approach is to use some filtering scheme, such as those proposed in [12] to get rid of as much of the noise as possible, while preserving the essential characteristics of the hidden signal. A deterministic set-matching algorithm can be applied as the last stage of the analysis [8]. This approach is conceptually simple, but clearly non-optimal. A much more rigorous approach would be to look at the combined problem, and cast it within a statistical framework. Since this approach is difficult, a first step would be to incorporate statistical information in the matching process. A more general scheme would suggest the use of *feedback* from the matching process to the preprocessing filter. In this way, the filter parameters can be trimmed to improve the confidence margin of the matching process. Research is currently underway to pursue these ideas.



Figure 3: Realization of a  $(0.001, j/511, \text{octagon}, \text{uniform in } \{0, \dots, 9\})$ -DRBRS

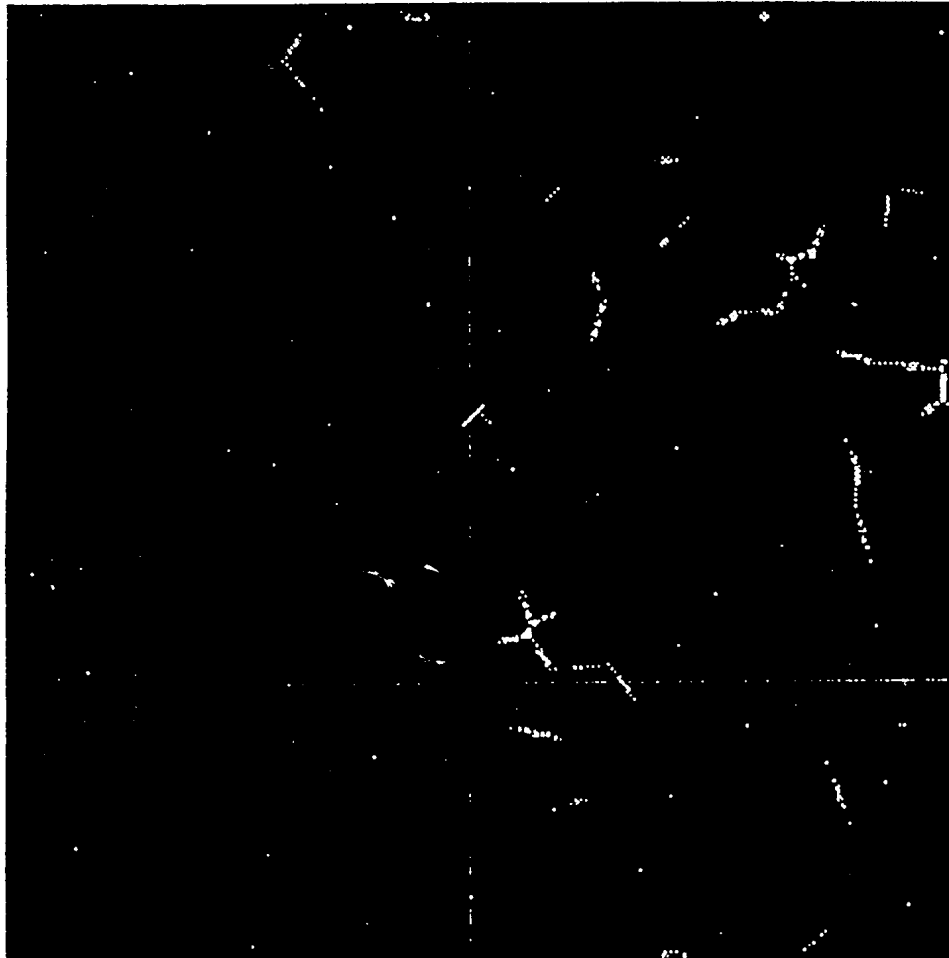


Figure 4: Morphological Skeleton of the DRBRS realization shown in the previous figure

## References

- [1] J. Serra Ed. *Image Analysis and Mathematical Morphology, vol. 2, Theoretical Advances*. Academic, San Diego, 1988.
- [2] J. Goutsias. Modeling random objects: An introduction to random set theory. *To appear in: Mathematical Morphology. Theory and Applications*, R. M. Haralick, Ed., Springer Verlag, 1991.
- [3] J. Goutsias and D. Schonfeld. Morphological representation of discrete and binary images. *To appear in: IEEE Trans. SP*, 1991.
- [4] A. F. Karr. *Point Processes and their Statistical Inference*. Marcel Dekker, New York and Basel, 1990.
- [5] P. Maragos. Optimal morphological approaches to image matching and object detection. In *Proc. of the IEEE second Int. Conf. on Computer Vision, Tampa, Florida*, pages 695–699, 1988.
- [6] P. Maragos. Pattern spectrum and multiscale shape representation. *IEEE Trans. Patt. Anal. Mach. Intell.*, 11(7):701–716, July 1989.
- [7] G. Matheron. *Random Sets and Integral Geometry*. Wiley, New York, 1975.
- [8] I. Pitas and N. D. Sidiropoulos. Pattern recognition of binary image objects using morphological shape decomposition. *To appear in: Computer Vision, Graphics, and Image Processing (Edited book)*, 1991.
- [9] I. Pitas and A. N. Venetsanopoulos. *Nonlinear Digital Filters: Principles and Applications*. Kluwer, Boston, 1990.
- [10] I. Pitas and A. N. Venetsanopoulos. Shape decomposition by mathematical morphology. *IEEE Trans. Patt. Anal. Mach. Intell.*, 12(1):38–46, January 1990.
- [11] H. V. Poor. *An introduction to Signal Detection and Estimation*. Springer-Verlag, New York, 1988.
- [12] D. Schonfeld and J. Goutsias. Optimal morphological pattern restoration from noisy binary images. *IEEE trans. Pattern Anal. Mach. Intell.*, 13(1):14–29, Jan. 1991.



- [13] J. Serra. *Image Analysis and Mathematical Morphology*. Academic, New York, 1982.
- [14] D. L. Snyder. *Random Point Processes*. Wiley, New York, 1975.
- [15] D. Stoyan, W.S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications*. Wiley, Berlin, 1987.