

# Expectations of Random Sets and Their Boundaries

## Using Oriented Distance Functions

Larissa I. Stanberry\* and Hanna K. Jankowski\*\*

\*School of Mathematics, University of Bristol.

\*\*Department of Mathematics and Statistics, York University.

March 10, 2009

### Abstract

Shape estimation and object reconstruction are common problems in image analysis. Mathematically, viewing objects in the image plane as random sets reduces the problem of shape estimation to inference about sets. Currently existing definitions of the expected set rely on different criteria to construct the expectation. This paper introduces new definitions of the expected set and the expected boundary, based on oriented distance functions. The proposed expectations have a number of attractive properties, including inclusion relations, convexity preservation and equivariance with respect to rigid motions. The paper introduces a special class of separable oriented distance functions for parametric sets and gives the definition and properties of separable random closed sets. Further, the definitions of the empirical mean set and the empirical mean boundary are proposed and empirical evidence of the consistency of the boundary estimator is presented. In addition, the paper gives loss functions for set inference in frequentist framework and shows how some of the existing expectations arise naturally as optimal estimators. The proposed definitions of the set and boundary expectations are illustrated on theoretical examples and real data.

## 1 Introduction

Boundary reconstruction and shape estimation are frequently encountered problems in image analysis. For example, it is often of interest to determine a characteristic shape of a cell, reconstruct tissue boundary in medical images, or estimate a probable area in the earthquake disaster zone. Mathematically, the objects of interest can be viewed as sets, whilst inherent stochasticity of the acquisition process turns them into random entities. The problems of boundary reconstruction and shape estimation thus reduce to inference about random sets.

Early results in the theory of random sets date back to Choquet (1953) with a thorough mathematical treatment appearing in Matheron (1975) and the most recent developments in the field presented in (Molchanov, 2005). Because the family of closed sets is nonlinear, there is no natural way to define the

expected set. Currently, there exist a number of definitions of the expectation, though neither can be spoken of as the best. As pointed out in Molchanov (2005), the definition of the expectation depends on set features that are important to emphasize.

Many of the existing definitions are not based on a random set directly, but use an embedding into a function space. For example, the Vorob’ev definition uses the representation of a set given by its characteristic function to construct an expectation that is optimal with respect to Lebesgue measure (Vorob’ev, 1984). The distance-average approach maps a set into the space of distance functions, so that the expected set is optimal among all of the level sets of the expected distance function with respect to a predefined metric (Baddeley and Molchanov, 1998).

Currently existing definitions of the expected set do not address the question of boundary estimation, which is often of primary interest in practice. In this paper, we introduce a new definition of the expected set and the expected boundary based on oriented distance functions (ODFs). As opposed to a conventional distance function, an ODF takes into account the set and its complement, providing a more informative representation of a set (Delfour and Zolésio, 2001). In comparison with currently existing definitions, the new expectations have a number of attractive theoretical properties. In particular, they satisfy inclusion relations, preserve convexity and remain equivariant with respect to rigid motions. We introduce a new class of ODFs for parametric sets and show the connection between the expectation of a parametric random closed set and the expected parameter. We also outline a general framework for set inference and show how different expectations arise as natural estimators for different loss functions.

## 2 Random Closed Sets

### 2.1 Notation

For set  $A \subset \mathbb{R}^d$ , denote by  $A^c, \partial A, \text{cl } A, \text{int } A, \lambda(A)$  its complement, boundary, closure, interior, and Lebesgue measure, respectively. We write  $x = (x_1, \dots, x_d)$  for point  $x \in \mathbb{R}^d$  and use  $|\cdot|$  for the standard Euclidean norm. Let  $\mathcal{F}$  (resp.  $\mathcal{K}$ ) be the family of closed (resp. compact) subsets of  $\mathbb{R}^d$  and let the tripple  $(\Omega, \mathcal{A}, \mathbf{P})$  denote the probability space.

### 2.2 Definitions and Examples

**Definition 2.1.** A random closed set is the mapping  $\mathbf{A} : \Omega \mapsto \mathcal{F}$  such that for every compact set  $K \in \mathcal{K}$ ,

$$\{\omega : \mathbf{A}(\omega) \cap K \neq \emptyset\} \in \mathcal{A}.$$

Here, we consider random closed sets taking values in  $\mathbb{R}^d$ . Throughout this paper, we write “r.c.s.” for a random closed set, though some texts prefer RACS. We write  $\mathbf{A}$  for an r.c.s.  $\mathbf{A}(\omega)$  and reserve capital roman letters for closed subsets of  $\mathbb{R}^d$ .

Random closed sets give rise to several random variables and random functions, including the characteristic function

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A, \end{cases} \quad (1)$$

the Lebesgue measure  $\lambda(A) = \int_{\mathbb{R}^d} \chi_A(x) dx$ , and the distance function from point  $x \in \mathbb{R}^d$  to  $A$ ,

$$d_A(x) = \begin{cases} \inf_{y \in A} \rho(x, y), & A \neq \emptyset, \\ +\infty, & A = \emptyset, \end{cases} \quad (2)$$

where  $\rho(x, y)$  is a metric on  $\mathbb{R}^d$ . Some examples of  $\rho$  include the  $L_2$  Euclidean metric, the  $L_\infty$  Chebyshev metric, and the  $L_1$  Manhattan metric. Throughout this paper, we use the Euclidean distance function with  $\rho(x, y) = |x - y|$  in (2).

### 2.3 Expectations of Random Closed Sets

Because the space  $\mathcal{F}$  of all closed subsets of  $\mathbb{R}^d$  is nonlinear, there is no natural way to define the expected set. In this section, we briefly review some of the existing definitions and refer for more details to (Aumann, 1965; Artstein and Vitale, 1975; Vorob'ev, 1984; Baddeley and Molchanov, 1998; Molchanov, 2005).

**Definition 2.2** (Selection expectation). A random element  $\xi$  is called the selection of  $\mathbf{A}$ , if it belongs to  $\mathbf{A}$  with probability one. A selection  $\xi$  is integrable, if  $E|\xi| < \infty$ . The selection expectation,  $E_s[\mathbf{A}]$ , is the closure of the set of the expectations of all integrable selections of  $\mathbf{A}$ ,

$$E_s[\mathbf{A}] = \text{cl} \{E[\xi] : \xi \text{ is a selection and } E|\xi| < \infty\}.$$

The selection expectation depends on the structure of the probability space (Molchanov, 2005, see Example 1.14, Section 2). In addition, if the probability space is nonatomic,  $E_s[\mathbf{A}]$  is necessarily convex, even for nonconvex deterministic sets.

In the context of image analysis, the Vorob'ev expectation is perhaps the most intuitive construction (Vorob'ev, 1984). Consider an r.c.s.  $\mathbf{A} \subset \mathcal{D}$ . For every point  $x \in \mathcal{D}$ , we assign a probability mass depending on whether or not  $x \in \mathbf{A}$ . For example, assigning one to all points in  $\mathbf{A}$  and zero to points in  $\mathbf{A}^c$  converts the observed set into a binary image. Averaging binary images over all realizations of  $\mathbf{A}$ , we obtain a gray-scale image, where the intensity at each pixel is the probability of the point being in  $\mathbf{A}$ . The gray-scale image is not a binary image, unless  $\mathbf{A}$  is deterministic. The Vorob'ev definition then provides a criterion to construct an expected set that is optimal with respect to Lebesgue measure.

More precisely, given the characteristic function  $\chi_{\mathbf{A}}$  of an r.c.s.  $\mathbf{A}$ , define the coverage function by

$$E[\chi_{\mathbf{A}}(x)] = \mathbf{P}(x \in \mathbf{A}). \quad (3)$$

The excursions sets of the coverage function are given by

$$\mathbf{A}_u = \{x \in \mathbb{R}^d : E[\chi_{\mathbf{A}}(x)] \geq u\}, \quad q \in [0, 1]. \quad (4)$$

**Definition 2.3** (Vorob'ev expectation). The expectation,  $E_v[\mathbf{A}]$ , of a random closed set  $\mathbf{A}$  is the excursion set  $\mathbf{A}_q$ , where  $q \in [0, 1]$  is such that

$$\lambda(\mathbf{A}_u) \leq E[\lambda(\mathbf{A})] \leq \lambda(\mathbf{A}_q) \quad \text{for all } u > q. \quad (5)$$

Note that the optimality criterion (5) ensures that the Vorob'ev expectation ignores sets of measure zero.

The distance-average expectation is based on a representation of a set given by some function  $f_A: \mathcal{D} \mapsto \mathbb{R}$  (Baddeley and Molchanov, 1998). We call  $f_A$  a representative function of the set  $A$ . Examples of  $f$  include the distance function, the ODF, the characteristic function and others. For an r.c.s.  $\mathbf{A}$  with representative function  $f_{\mathbf{A}}$ , let  $E[f_{\mathbf{A}}(x)]$  be the expected value of  $f_{\mathbf{A}}$  at  $x$ , assuming it exists.

**Definition 2.4** (Distance-average expectation). For a compact window  $\mathcal{W}$ , define a (pseudo-) metric  $\mathfrak{m}(\mathbf{A}, \mathbf{B}) = \mathfrak{m}_{\mathcal{W}}(f_{\mathbf{A}}(\cdot), f_{\mathbf{B}}(\cdot))$ . The distance-average expectation,  $E_{\text{DA}}[\mathbf{A}]$ , of an r.c.s.  $\mathbf{A}$  is the level set

$$\mathbf{A}_u = \{x \in \mathcal{W}: E[f_{\mathbf{A}}(x)] \leq u\}, \text{ where } u = \arg \inf_{s \in \mathbb{R}} \mathfrak{m}(E[f_{\mathbf{A}}], f_{\mathbf{A}_s}) \quad (6)$$

Examples of pseudometric  $\mathfrak{m}$  in (6) include  $L_q$  distances and their variates, Baddeley's  $\Delta^q$ -distance, or any custom definitions (Baddeley, 1992). Note that the pseudometric  $\mathfrak{m}$  is computed over the window  $\mathcal{W}$ , which can be the entire domain  $\mathcal{D}$  or its subset. For brevity, we omit the subscript  $\mathcal{W}$  whenever possible.

The distance-average expectation strongly depends on the choice of the representative function  $f$ , window  $\mathcal{W}$ , (pseudo-) metric  $\mathfrak{m}$  and parameters of  $\mathfrak{m}$  (Baddeley and Molchanov, 1998). The Vorob'ev expectation can be viewed as a special case of the distance-average definition with  $f_{\mathbf{A}}(x) = 1 - \chi_{\mathbf{A}}(x)$  and  $L^1$  distance  $\mathfrak{m}$  (Baddeley and Molchanov, 1998, see Example 5.14).

The linearization approach gives a unified view of the existing expectations (Molchanov, 2005). In particular, consider a mapping  $f: \mathcal{F} \mapsto \mathcal{Y}$ , where  $\mathcal{Y}$  is some Banach space. Let  $f_{\mathbf{A}}$  be the image of  $\mathbf{A}$  in  $\mathcal{Y}$  and assume that  $E[|f_{\mathbf{A}}(x)|] < \infty$ . If there exists a unique  $F \in \mathcal{F}$  such that  $f_F(x) = E[f_{\mathbf{A}}(x)]$ , then declare  $E[\mathbf{A}] = F$ . In general, the inverse image rarely exists, and we need to define the criterion so that the resulting expectation is optimal in some sense. This is done as follows. Let  $\mathfrak{d}$  be a pseudometric in  $\mathcal{Y}$ , then the expectation of an r.c.s.  $\mathbf{A}$  is given by

$$E[\mathbf{A}] = \arg \min_{F \in \mathcal{F}} \mathfrak{d}(E[\zeta(\mathbf{A})], \zeta(F)). \quad (7)$$

Minimizing over the entire family  $\mathcal{F}$  is often nontrivial. Therefore, we can consider optimizing  $\mathfrak{d}(\cdot, \cdot)$  over a subfamily  $\mathcal{H} \subset \mathcal{F}$  of candidate sets. Clearly, the resulting expectation depends on the choice of Banach space  $\mathcal{Y}$ , mapping  $f$ , metric  $\mathfrak{d}$ , subfamily  $\mathcal{H}$  and any implicit parameters that enter into the calculation.

For bounded closed convex sets, the selection expectation is an example of a linearization approach, where the embedding into a Banach space is realised by support functions of random sets on the unit sphere and the expected set is given by the support function of the selection expectation. Another example is the Vorob'ev definition, which is based on the mapping given by  $\chi_{\mathbf{A}}$  and gives the expected set  $E_{\text{V}}[\mathbf{A}]$  which is optimal with respect to Lebesgue measure; the family of candidate sets  $\mathcal{H}$  consists of excursion sets (4) of the coverage function. The distance-average expectation is based on the embedding into the space of representative function and the resulting average is optimal among the level sets of the expected representation with respect to metric  $\mathfrak{m}(\cdot, \cdot)$  (Baddeley and Molchanov, 1998; Molchanov, 2005).

## 2.4 Oriented Distance Functions

In comparison to a distance function (2), an ODF takes into account both the set and its compliment, thus reflecting the interior and exterior properties. In addition, the ODF provides a level-set description of the

boundary.

**Definition 2.5.** The ODF from point  $x$  to set  $A \subset \mathbb{R}^d$  with  $\partial A \neq \emptyset$  is defined by

$$b_A(x) = d_A(x) - d_{A^c}(x), \quad \text{for all } x \in \mathbb{R}^d. \quad (8)$$

Using basic properties of the distance function, the expression (8) can be written as

$$b_A(x) = \begin{cases} d_A(x) = d_{\partial A}(x), & x \in \text{int } A^c \\ 0, & x \in \partial A \\ -d_{A^c}(x) = -d_{\partial A}(x), & x \in \text{int } A. \end{cases} \quad (9)$$

Below we summarize some properties of the ODFs and refer for proofs and details to Delfour and Zolésio (2001).

**Proposition 2.1.** *Let  $A$  and  $B$  be some subsets of  $cl \mathcal{D}$  with  $\partial A, \partial B \neq \emptyset$ .*

1. *The ODF provides a level-set description of a set, i.e*

$$cl A = \{x : b_A(x) \leq 0\}, \quad \partial A = \{x : b_A(x) = 0\}.$$

2.  *$A \supset B$  implies  $b_A \leq b_B$ .*
3.  *$b_A \leq b_B$  in  $\mathcal{D}$  iff  $cl B \subset cl A$  and  $cl A^c \subset cl B^c$ .*
4.  *$b_A = b_B$  in  $\mathcal{D}$  iff  $cl A = cl B$  and  $\partial A = \partial B$ .*
5. *For a convex set  $A$  with  $\partial A \neq \emptyset$ ,  $b_A = b_{cl A}$  is a convex function in  $\mathbb{R}^d$ .*
6. *The function  $b_A(x)$  is uniformly Lipschitz in  $\mathbb{R}^d$  with constant one, i.e.*

$$|b_A(y) - b_A(x)| \leq |y - x| \quad \text{for any } x, y \in \mathbb{R}^d.$$

*In addition,  $b_A(x)$  is (Fréchet) differentiable with  $|\nabla b_A(x)| \leq 1$  a.e. in  $\mathbb{R}^d$ . When it exists, the gradient of the ODF coincides with the outward unit normal to the boundary.*

Note that the ODF identifies the set up to its closure and boundary. In other words, ODF describes equivalence classes of sets with identical closure and boundary.

**Example 2.1.** For a singleton  $A = \{\theta\}$ ,  $\theta \in \mathbb{R}^d$ , the ODF coincides with the distance function and  $b_A(x) = d_A(x) = |x - \theta|$  for all  $x \in \mathbb{R}^d$ . Note that  $b_A(x) > 0$  for all  $x \neq \theta$  and the isocontours of  $b_A(x)$  are spheres centered at  $\theta$ ; Figure 1, 1st plot.

**Example 2.2.** Consider a closed ball in  $\mathbb{R}^d$  centered at the origin with radius  $\theta$ , i.e.  $A = \{x \in \mathbb{R}^d : |x| \leq \theta\}$ . The distance function of  $A$  is  $d_A(x) = |x| - \theta$ , for  $x \in A$ , and zero, elsewhere. Hence, the ODF has a form  $b_A(x) = |x| - \theta$  for all  $x \in \mathbb{R}^d$ . The isocontours of  $b_A(x)$  are spheres centered at the origin; Figure 1, 2nd plot.

**Example 2.3.** The distance function of a half plane set  $A = \{x: x_1 \leq \theta_1\}$  is given by  $d_A(x) = x_1 - \theta_1$  for  $x \in A^c$ . The ODF has a form  $b_A = x_1 - \theta_1$  for all  $x \in \mathbb{R}^2$  with isocontours given by vertical lines parallel to the boundary; Figure 1, 3rd plot.

**Example 2.4.** The ODF of an upper-plane set  $A = \{x: x_2 \geq x_1 \tan \theta\}$  in  $\mathbb{R}^2$  is  $b_A(x) = |x| \sin(\theta - \omega)$ , where  $\omega = \arcsin(x_2|x|^{-1})$ ; Figure 1, 4th plot.

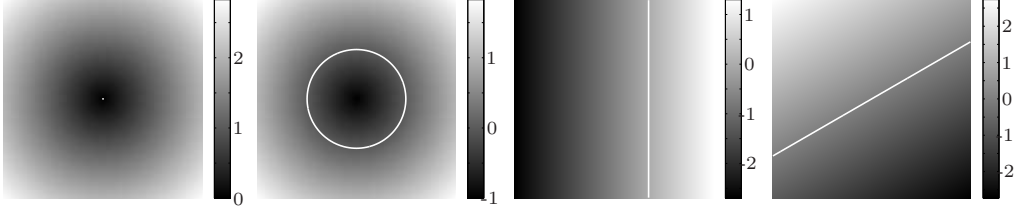


Figure 1: From left to right: The ODF of a singleton  $\{\theta\}$ , a unit disc, a half-plane and an upper half-plane in  $\mathbb{R}^2$ ; the zero-level isocontour (white) shows the boundary of the set.

### 3 Expectations of Random Sets via ODFs

#### 3.1 Definition and Properties

For an r.c.s.  $\mathbf{A}$  with  $\partial\mathbf{A} \neq \emptyset$ , the ODF  $b_{\mathbf{A}}(x)$  at point  $x \in \mathbb{R}^d$  is a random variable taking values in  $\mathbb{R}$ . Denote by  $E[b_{\mathbf{A}}(x)]$  the expected value of  $b_{\mathbf{A}}(x)$ , assuming it exists. We begin by proving some properties of the expected ODF.

**Proposition 3.1.** *Consider an r.c.s.  $\mathbf{A}$  such that  $\partial\mathbf{A} \neq \emptyset$  a.s. Then,*

1. *The expected ODF is uniformly Lipschitz with constant 1.*
2. *If  $E[|b_{\mathbf{A}}(x_0)|] < \infty$  for some  $x_0 \in \mathcal{D}$ , then  $E[|b_{\mathbf{A}}(x)|] < \infty$  for any  $x \in \mathcal{D}$ .*
3. *If  $b_{\mathbf{A}}(x)$  is differentiable at some  $x_0 \in \mathcal{D}$  a.s., then  $\nabla E[b_{\mathbf{A}}(x_0)]$  exists and  $|\nabla E[b_{\mathbf{A}}(x_0)]| \leq 1$ .*

*Proof.* (1) For any  $x, y \in \mathbb{R}^d$ ,  $|E[b_{\mathbf{A}}(y)] - E[b_{\mathbf{A}}(x)]| \leq E[|b_{\mathbf{A}}(y) - b_{\mathbf{A}}(x)|] \leq |y - x|$ .

(2) Using the triangle inequality, boundedness of  $\mathcal{D}$  and Lipschitz continuity of  $E[b_{\mathbf{A}}(x)]$ , we obtain  $E[|b_{\mathbf{A}}(x)|] \leq |x - x_0| + E[|b_{\mathbf{A}}(x_0)|] < \infty$ .

(3) Because  $\nabla b_{\mathbf{A}}(x_0)$  exists a.s. and  $|\nabla b_{\mathbf{A}}(x_0)| \leq 1$ , we can interchange the differentiation on the space  $\Omega$  and integration with respect to  $x$ , so the existence of  $\nabla E[b_{\mathbf{A}}(x_0)]$  follows. From  $\nabla b_{\mathbf{A}}(x_0) \leq 1$ , we have  $|\nabla E[b_{\mathbf{A}}(x_0)]| \leq E[|\nabla b_{\mathbf{A}}(x_0)|] \leq 1$ .  $\square$

**Definition 3.1** (ODF expectation). Let  $\mathbf{A}$  be a random closed set such that  $\partial\mathbf{A} \neq \emptyset$  a.s. and suppose that  $E[|b_{\mathbf{A}}(x_0)|] < \infty$  for some  $x_0 \in \mathcal{D}$ . Then the expectation of an r.c.s.  $\mathbf{A}$  is the zero-level set of the expected

ODF and expectation of the boundary of  $\mathbf{A}$  is the zero-level isocontour of the expected ODF, i.e.

$$E[\mathbf{A}] = \{x : E[b_{\mathbf{A}}(x)] \leq 0\}, \quad (10)$$

$$E[\partial\mathbf{A}] = \{x : E[b_{\mathbf{A}}(x)] = 0\}. \quad (11)$$

We refer to  $E[\mathbf{A}]$  as “the expected set” or “the ODF expectation” of  $\mathbf{A}$ . Note, that the expected set defined by (3.1) can be an empty set. Later we give examples of such random sets and discuss the appropriateness of the outcome.

**Theorem 3.1** (Inclusion properties). *Consider r.c.s.  $\mathbf{A}, \mathbf{B} \subset cl \mathcal{D}$  with  $\partial\mathbf{A}, \partial\mathbf{B} \neq \emptyset$  and assume that  $E[b_{\mathbf{A}}(x)]$  and  $E[b_{\mathbf{B}}(x)]$  exist for some  $x \in cl \mathcal{D}$ . That is, suppose  $E[\mathbf{A}]$  and  $E[\mathbf{B}]$  are well-defined. Then,*

1. *The expectation  $E[\mathbf{A}]$  is a closed set.*
2. *If  $\mathbf{A} = A$  a.s., for some deterministic  $A \in \mathcal{F}$ , then  $E[\mathbf{A}] = A$  and  $E[\partial\mathbf{A}] = \partial A$ .*
3. *Suppose that an r.c.s.  $\mathbf{A}$  satisfies  $\partial\mathbf{A} = B$  a.s. for some deterministic set  $B$ , then  $E[\partial\mathbf{A}] \supset B$ .*
4. *For an r.c.s.  $\mathbf{A}$ ,  $\partial E[\mathbf{A}] \subset E[\partial\mathbf{A}]$ .*
5. *If  $E[b_{\mathbf{A}}(x)] \geq E[b_{\mathbf{B}}(x)]$  a.s., then  $E[\mathbf{A}] \subset E[\mathbf{B}]$ .*
6. *If  $\mathbf{A} \subset \mathbf{B}$  a.s., then  $E[\mathbf{A}] \subset E[\mathbf{B}]$ .*

*Proof.* (1) The closedness of  $E[\mathbf{A}]$  follows from the continuity of  $E[b_{\mathbf{A}}]$ .

(2) is straightforward.

(3) It suffices to note that for all points  $x \in B$ ,  $b_{\mathbf{B}}(x) = b_{\mathbf{A}}(x) = 0$ , so that  $E[b_{\mathbf{A}}(x)] = 0$  and  $B \subset E[\partial\mathbf{A}]$ .

(4) For any  $x \in \partial E[\mathbf{A}]$ ,  $E[b_{\mathbf{A}}(x)] = 0$  and hence  $x \in E[\partial\mathbf{A}]$ . Example 3.2 shows that the reverse relation does not hold.

(5) For any  $x \in E[\mathbf{A}]$ ,  $0 \geq E[b_{\mathbf{A}}(x)] \geq E[b_{\mathbf{B}}(x)]$ . Thus  $x \in E[\mathbf{B}]$  and  $E[\mathbf{A}] \subset E[\mathbf{B}]$ .

(6) The proof follows immediately from the previous statement and Property 3 of the ODFs, which implies that for  $\mathbf{A} \subset \mathbf{B}$ ,  $b_{\mathbf{A}}(x) \geq b_{\mathbf{B}}(x)$  in  $\mathcal{D}$  and hence,  $E[b_{\mathbf{A}}(x)] \geq E[b_{\mathbf{B}}(x)]$  a.s.  $\square$

**Corollary 3.1.** *If an r.c.s.  $\mathbf{A} \supset U$  a.s. for some deterministic set  $U \subset cl \mathcal{D}$ , then  $E[\mathbf{A}] \supset cl U$ . Similarly, if an r.c.s.  $\mathbf{A} \subset W$  a.s. for some deterministic set  $W \subset cl \mathcal{D}$ , then  $E[\mathbf{A}] \subset cl W$ .*

*Proof.* Follows from statements (2) and (6) of Theorem 3.1.  $\square$

Although the inclusion properties of Theorem 3.1 and Corollary 3.1 seem natural, they do not necessarily hold for other definitions, including the selection expectation (Molchanov, 2005) and the distance-average definition. The latter does not satisfy a weaker version of inclusion in the sense that the resulting expectation is not necessarily contained in the convex hull of the union of all possible realizations of a random set (Baddeley and Molchanov, 1998).

**Example 3.1.** (random singleton) Let  $\xi$  be a random variable and consider an r.c.s.  $\mathbf{A} = \{\xi\}$ . Then, for any  $x \in \mathcal{D}$ ,  $E[b_{\mathbf{A}}(x)] = E[|x - \xi| > 0]$  and hence  $E[\mathbf{A}] = \emptyset$ .

The last result does not agree with suggestions in Molchanov (2005), where  $E[\{\xi\}] = \{E[\xi]\}$  is listed as one of the desirable properties of the expected set. Indeed, the selection expectation satisfies the suggested property and so does the distance-average definition, for certain choices of  $\mathbf{m}$ ,  $f$  and  $W$  (Baddeley and Molchanov, 1998, see Example 5.4). However, the ODF definition of the expected set was primarily motivated by problems in image analysis, where random singletons are often regarded as noise. Example 3.1 highlights the natural denoising property of the ODF expectation in a sense that any speckles, unless observed a.s., are not present in the average reconstruction. We argue that if random singletons are the sets of primary interests, they should be viewed as random vectors rather than random sets, in which case conventional definitions of the expectation and dispersion are applicable.

More generally, Example 3.1 can be extended to random sets with zero Lebesgue measure. In particular, for a r.c.s.  $\mathbf{A}$  with  $\text{int } \mathbf{A} = \emptyset$  and  $\lambda(\mathbf{A}) = 0$ , we have  $b_{\mathbf{A}}(x) > 0$  for all  $x \in \mathbf{A}^c$  and  $b_{\mathbf{A}}(x) = 0$  for all  $x \in \partial\mathbf{A}$ . Thus, unless  $x \in \partial\mathbf{A}$  almost surely,  $E[b_{\mathbf{A}}(x)] > 0$  and the expected set  $E[\mathbf{A}] = \{x: x \in \mathbf{A} \text{ almost surely}\}$ .

In the rest of this section, we discuss equivariance properties of the ODF expectation. To begin with, recall that the translation of a set  $A \subset \mathbb{R}^d$  by  $a \in \mathbb{R}^d$  is the set  $a + A = \{a + x: x \in A\}$ . Similarly, the homothety of  $A$  by a scalar  $\alpha$  is given by  $\alpha A = \{\alpha x: x \in A\}$ . The homothety is a dilation for  $\alpha \geq 1$ , a contraction for  $\alpha \in [0, 1)$ , and a reflection for  $\alpha = -1$ . For negative  $\alpha$ , the homothety is a dilation or a contraction of the reflection for  $|\alpha| < -1$  and  $|\alpha| \in (0, 1)$ , respectively. Rigid motion transformations  $g: \mathbb{R}^d \mapsto \mathbb{R}^d$  are given by  $g(x) = \Gamma x + a$ , where  $\Gamma \in O(n)$  is an orthogonal matrix and  $a \in \mathbb{R}^d$ . In  $\mathbb{R}^d$ , these transformations are isometries, i.e.  $|x - y| = |g(x) - g(y)|$ , and form the Euclidean group  $E(d)$  with respect to the composition.

**Theorem 3.2** (Equivariance properties). *The expected set and the expected boundary are equivariant with respect to:*

1. Homothety, i.e. for a fixed scalar  $\alpha \neq 0$ ,

$$E[\alpha\mathbf{A}] = \alpha E[\mathbf{A}] \quad \text{and} \quad E[\partial\alpha\mathbf{A}] = \alpha E[\partial\mathbf{A}]. \quad (12)$$

2. The group of rigid motions, i.e. for any  $g \in E(d)$ ,

$$E[g\mathbf{A}] = gE[\mathbf{A}] \quad \text{and} \quad E[g\partial\mathbf{A}] = gE[\partial\mathbf{A}]. \quad (13)$$

*Proof.* (1) We begin by establishing the relation between the ODFs of a set and its homothety. In particular,  $d_{\alpha A}(x) = \inf_{y \in \alpha A} |y - x| = \inf_{y \in A} |\alpha y - x| = |\alpha| d_A(x/\alpha)$ . Similarly,  $d_{(\alpha A)^c}(x) = |\alpha| d_{A^c}(x/\alpha)$ , so that  $b_{\alpha A}(x) = |\alpha| b_A(x/\alpha)$ .

For any  $y \in E[\alpha\mathbf{A}]$ ,  $E[b_{\alpha\mathbf{A}}(y)] \leq 0$ . Writing  $y = \alpha x$  for some  $x$ , we obtain  $0 \geq E[b_{\alpha\mathbf{A}}(y)] = E[b_{\alpha\mathbf{A}}(\alpha x)] = |\alpha| E[b_{\mathbf{A}}(x)]$ . Thus,  $E[b_{\mathbf{A}}(x)] \leq 0$  and  $y = \alpha x$  with  $x \in E[\mathbf{A}]$ . Hence,  $y \in \alpha E[\mathbf{A}]$  and  $E[\alpha\mathbf{A}] \subseteq \alpha E[\mathbf{A}]$ .

For any  $y \in \alpha E[\mathbf{A}]$ ,  $y = \alpha x$  for some  $x \in E[\mathbf{A}]$ . Then  $E[b_{\alpha\mathbf{A}}(y)] = E[|\alpha| b_{\mathbf{A}}(y/\alpha)] = |\alpha| E[b_{\mathbf{A}}(x)] \leq 0$ . Thus,  $y \in E[\alpha\mathbf{A}]$  and  $E[\alpha\mathbf{A}] \supseteq \alpha E[\mathbf{A}]$ .

The equivariance property for the boundary is proved similarly and we omit the details.

(2) A transformation  $g(x) = \Gamma x + a$ , for some orthogonal matrix  $\Gamma \in O(d)$  and  $a \in \mathbb{R}^d$ , is a bijection and  $g^{-1}(x) = \Gamma^{-1}(x - a) = \Gamma^T(x - a)$ . For sets  $g\mathbf{A} = \{g(x): x \in \mathbf{A}\}$  and  $(g\mathbf{A})^c = \{g(x): x \in \mathbf{A}^c\}$ , we



obtain  $d_{gA}(x) = \inf_{y \in gA} |y - x| = \inf_{y \in A} |\Gamma y + b - x| = \inf_{y \in A} |y - \Gamma^T(x - b)| = d_A(g^{-1}(x))$ . Similarly,  $d_{(gA)^c}(x) = d_{A^c}(g^{-1}(x))$  and  $b_{gA} = b_A(g^{-1}(x))$ .

Recall that  $gE[\mathbf{A}] = \{g(x) : x \in E[\mathbf{A}]\}$  and  $E[g\mathbf{A}] = \{x : E[b_{g\mathbf{A}}(x)] \leq 0\}$ . Hence, for any  $y \in gE[\mathbf{A}]$ , there exists  $x \in E[\mathbf{A}]$  such that  $y = g(x)$ . From here,  $E[b_{g\mathbf{A}}(y)] = E[b_{\mathbf{A}}(g^{-1}(\Gamma x + b))] = E[b_{\mathbf{A}}(x)] \leq 0$ . Hence,  $y \in E[g\mathbf{A}]$  and  $E[g\mathbf{A}] \supseteq gE[\mathbf{A}]$ .

On the other hand, any  $y \in E[g\mathbf{A}]$  can be written as  $y = \Gamma x + b$ , where  $x = g^{-1}(y)$ . It then follows that  $0 \geq E[b_{g\mathbf{A}}(y)] = E[b_{\mathbf{A}}(g^{-1}(g(x)))] = E[b_{\mathbf{A}}(x)]$ . Thus, any  $y \in E[g\mathbf{A}]$  has a form  $y = \Gamma x + b$ , where  $x \in E[\mathbf{A}]$ . Hence,  $E[g\mathbf{A}] \subseteq gE[\mathbf{A}]$  and  $E[g\mathbf{A}] = gE[\mathbf{A}]$ .  $\square$

**Corollary 3.2.** *Note that any translation in  $\mathbb{R}^d$  can be described as an isometry with an identity matrix  $\Gamma$ . Thus, it immediately follows that the expected set and the expected boundary are translation-equivariant, i.e. for some fixed  $a \in \mathbb{R}^d$ ,*

$$E[a + \mathbf{A}] = a + E[\mathbf{A}] \quad \text{and} \quad E[\partial(a + \mathbf{A})] = a + E[\partial\mathbf{A}]. \quad (14)$$

**Theorem 3.3** (Convexity preservation). *Suppose that an r.c.s.  $\mathbf{A}$  is convex a.s.. Then the expectation  $E[\mathbf{A}]$  of  $\mathbf{A}$  is also convex.*

*Proof.* The a.s. convexity of  $\mathbf{A}$  implies that  $\text{cl } \mathbf{A}$  is convex and  $b_{\mathbf{A}} = b_{\text{cl } \mathbf{A}}$  is a convex function a.s. (Delfour and Zolésio, 2001, Theorem 7.1). Hence, for all  $\alpha \in [0, 1]$ ,  $b_{\mathbf{A}}(\alpha x + (1-\alpha)y) \leq \alpha b_{\mathbf{A}}(x) + (1-\alpha)b_{\mathbf{A}}(y)$  a.s.. Taking expectations, we obtain that  $E[b_{\mathbf{A}}]$  is convex, and so is  $E[\mathbf{A}]$ .  $\square$

We refer to Theorems 3.2 and 3.3 as the shape-preservation properties of the ODF expectation. These qualities are particularly desirable in shape and image analysis. In contrast, given a nonatomic probability space, the selection expectation of an r.c.s. is convex and coincides with the expectation of its convex hull. Notably, the convexification of  $E_s[\mathbf{A}]$  holds even for nonconvex deterministic sets (Molchanov, 2005). The distance-average expectation preserves the convexity, but only if the set, its representative function and the window are convex (Baddeley and Molchanov, 1998).

**Example 3.2** (Set and its boundary). Consider an r.c.s.  $\mathbf{A}$  such that  $\mathbf{A} = \{0, 1\}$  or  $\mathbf{A} = [0, 1]$  in  $\mathbb{R}$  with probability  $p$  and  $1 - p$ , respectively. Note that the boundary  $\partial\mathbf{A} = \{0, 1\}$  a.s. in  $\mathbb{R}$ , so essentially, we are observing either the set or its boundary. The expected ODF is given by

$$E[b_{\mathbf{A}}(x)] = \begin{cases} -x, & x \leq 0 \\ (2p - 1)x, & 0 < x \leq 0.5 \\ (2p - 1)(1 - x), & 0.5 < x \leq 1 \\ x - 1, & x > 1. \end{cases}$$

For  $p = 0.5$ ,  $E[\mathbf{A}] = E[\partial\mathbf{A}] = [0, 1]$ . For  $p < 0.5$ ,  $E[\mathbf{A}] = [0, 1]$  and  $E[\partial\mathbf{A}] = \{0, 1\}$ . For  $p > 0.5$ ,  $E[\mathbf{A}] = E[\partial\mathbf{A}] = \{0, 1\}$ . Note that for  $p = 0.5$ ,  $\partial\mathbf{A} = \{0, 1\}$  a.s. However,  $E[\partial\mathbf{A}] \neq \{0, 1\}$ , but rather  $\partial\mathbf{A} \subset E[\partial\mathbf{A}]$ , as in Proposition 3.1(3).

For  $p = 0.5$ , the distance-average expectation is  $E_{\text{DA}}[\mathbf{A}] = [-1/12, 1/6] \cup [5/6, 13/12]$ , for  $\mathbf{m}$  uniform, and  $E_{\text{DA}}[\mathbf{A}] = \{0, 1\}$ , for  $\mathbf{m} = \Delta_{\mathcal{W}}^p$  with sufficiently large  $\mathcal{W}$  (Baddeley and Molchanov, 1998).

More generally, consider an r.c.s.  $\mathbf{A}$  such that, for some deterministic  $B \in \mathcal{F}$ ,  $\mathbf{A} = B$  with probability  $p$  and  $\mathbf{A} = \partial B$ , otherwise. From

$$b_{\partial B}(x) = \begin{cases} b_B(x), & x \in B^c, \\ 0, & x \in \partial B, \\ -b_B(x), & x \in \text{int } B, \end{cases}$$

it follows that

$$\mathbb{E}[b_{\mathbf{A}}(x)] = \begin{cases} b_B(x), & x \in B^c, \\ 0, & x \in \partial B, \\ (1 - 2p)b_B(x), & x \in \text{int } B. \end{cases}$$

Hence,

$$\begin{aligned} \text{for } p &= 0.5, & \mathbb{E}[\mathbf{A}] &= \mathbb{E}[\partial \mathbf{A}] = B, \\ \text{for } p &< 0.5, & \mathbb{E}[\mathbf{A}] &= B, \mathbb{E}[\partial \mathbf{A}] = \partial B, \\ \text{for } p &> 0.5, & \mathbb{E}[\mathbf{A}] &= \mathbb{E}[\partial \mathbf{A}] = \partial B. \end{aligned}$$

**Example 3.3** (Ball with random radius). Consider a random closed ball  $\mathbf{A}$  in  $\mathbb{R}^d$  with a fixed center  $x_0$  and a random radius  $\Theta$ . Recall that the ODF of  $\mathbf{A}$  is given by  $b_{\mathbf{A}}(x) = |x - x_0| - \Theta$  for all  $x \in \mathbb{R}^d$ . Hence,  $\mathbb{E}[b_{\mathbf{A}}(x; \Theta)] = |x - x_0| - \mathbb{E}[\Theta]$  and the expected set is a closed ball centered at  $x_0$  with radius  $\mathbb{E}[\Theta]$ . The expected boundary  $\mathbb{E}[\partial \mathbf{A}]$  is a sphere centered at  $x_0$  with radius  $\mathbb{E}[\Theta]$ .

**Example 3.4** (Flashing discs). Let  $\Theta$  be a Bernoulli random variable with parameter  $p$ . Consider an r.c.s.  $\mathbf{A} \subset \mathbb{R}^2$  given by a disc of radius  $r$  centered at the origin, with probability  $p$ , and centered at a point  $a \in \mathbb{R}^2$ , otherwise. The ODF of  $\mathbf{A}$  is given by  $b_{\mathbf{A}}(x) = |x|\mathbb{I}(\Theta = 1) + |x - a|\mathbb{I}(\Theta = 0) - r$ , where  $\mathbb{I}(\cdot)$  is an indicator function. Taking the expectation with respect to  $\Theta$ , we obtain  $\mathbb{E}[b_{\mathbf{A}}(x)] = p|x| + (1 - p)|x - a| - r$ . The isocontours of the expected ODF are Cartesian ovals with foci  $\{0, a\}$ .

Figure 2 shows the expected ODF with the expected boundary superimposed in white, for  $p = 0.8$  and three different values of  $a$ . The observed discs  $\mathbf{A}$  are shown in gray. When the realizations of  $\mathbf{A}$  do not intersect, the expected boundary is nearly a circle contained in the leftmost disc (Figure 2, left). As the distance between the centers decreases, the expected set expands. In general, as  $p \rightarrow 1$ , the expected ODF approaches that of a disc centered at the origin with radius  $r$ . Alternatively, as  $p \rightarrow 1/2$ , the isocontours approach an elliptical shape.

For  $p = 0.5$ , the  $c$ -level isocontour of the expected ODF is an ellipse with foci  $(0, 0)$  and  $a$  and semimajor axis  $r + c$ . From  $\mathbb{E}[b_{\mathbf{A}}(x)] \geq |a|/2 - r$ , we obtain that if  $|a| > 2r$ , then  $\mathbb{E}[\mathbf{A}] = \emptyset$ ; if  $|a| = 2r$ , then  $\mathbb{E}[\mathbf{A}] = \{a/2\}$ ; if  $|a| < 2r$ , then  $\mathbb{E}[\mathbf{A}]$  is the closure of an ellipse with foci  $\{(0, 0), a\}$  and a semimajor axis  $r$ . Figure 3 shows the expected boundaries (white) overlaid on the expected ODF, for discs with  $r = 1$  and  $p = 0.5$ .

In comparison, the selection expectation  $\mathbb{E}_s[\mathbf{A}]$  is always a disc centered at  $a/2$  with radius  $r$ , irrespective of  $p$ . The Vorob'ev expectation  $\mathbb{E}_v[\mathbf{A}]$  for  $p \neq 0.5$  is given by the more likely of the two discs. For  $p = 0.5$ ,  $\mathbb{E}_v[\mathbf{A}]$  is the union of the two discs, provided the cardinality of the intersection does not exceed one, otherwise  $\mathbb{E}_v[\mathbf{A}]$  is the intersection of the two discs.

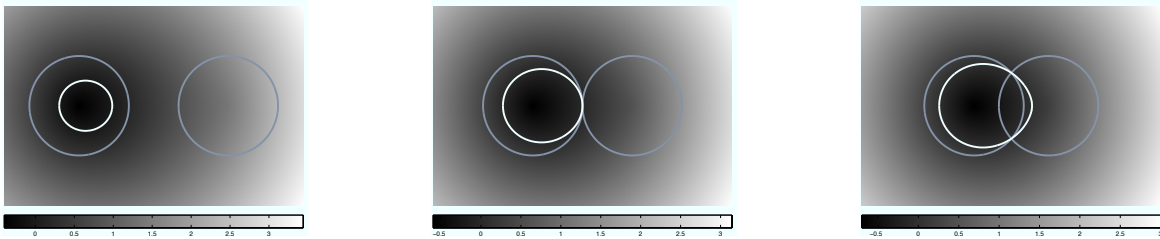


Figure 2: The expected boundaries (white) for  $\mathbf{A}$  with  $p = 0.8, r = 1$  and  $|a| = 3, 2, 1.5$  (from left to right), superimposed on the gray-scale image of the expected ODF. Realizations of  $\partial\mathbf{A}$  are shown in white.

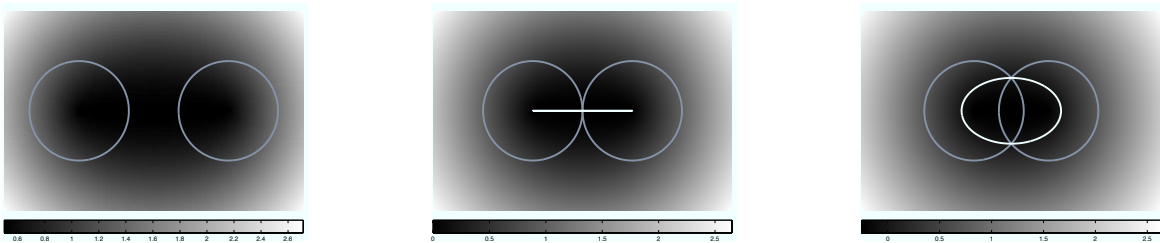


Figure 3: The expected boundaries (white) for random discs with  $p = 0.5, r = 1$  and  $|a| = 3, 2, 1.5$  (from left to right), superimposed on the gray-scale image of the expected ODF. Realized boundaries are shown in gray.

**Example 3.5** (Random half-plane). Let  $\Theta$  be a random variable and consider a random half-plane in  $\mathbb{R}^2$  given by  $\mathbf{A} = \{x: x_1 \leq \Theta\}$  with  $\partial\mathbf{A} = \{x: x_1 = \Theta\}$ . The ODF of  $\mathbf{A}$  is  $b_{\mathbf{A}}(x) = x_1 - \Theta$  and the expectation has a form  $E[b_{\mathbf{A}}(x; \Theta)] = x_1 - E[\Theta]$ . From here, the expected set is a half-plane intersecting the horizontal axis at point  $E[\Theta]$  and  $E[\partial\mathbf{A}] = \{x: x_1 = E[\Theta]\}$ . In this example,  $E_s[\mathbf{A}] = E_v[\mathbf{A}] = E[\mathbf{A}]$ .

**Example 3.6** (Random upper half-plane). Consider a random upper half-plane in  $\mathbb{R}^2$  given by  $\mathbf{A} = \{x: x_2 \geq x_1 \tan \Theta\}$  with  $\partial\mathbf{A} = \{x: x_2 = x_1 \tan \Theta\}$ . The ODF of  $\mathbf{A}$  is  $b_{\mathbf{A}}(x) = |x| \sin(\Theta - \omega)$ , where  $\omega = \arcsin(x_2|x|^{-1})$ . For  $\Theta \sim U[a, b]$  the expected ODF is  $E[b_{\mathbf{A}(\Theta)}(x)] = 2(b-a)^{-1} \sin((b-a)/2)|x| \sin((a+b)/2 - \omega)$ . From here, the expected set is a homothecy with coefficient  $\alpha = (b-a)^{-1}(2 \sin(b-a)/2)^{-1}$  of an upper half-plane with boundary angle  $(a+b)/2$ . Hence,  $E[\mathbf{A}]$  is an upper half-plane with boundary given by the line that passes through the origin and makes angle  $(a+b)/2$  with the positive horizontal axis. Again, we have  $E_s[\mathbf{A}] = E_v[\mathbf{A}] = E[\mathbf{A}]$ .

## 4 Separable Oriented Distance Functions

Consider a parametric closed set, whose geometry depends on the parameter  $\theta \in \mathbb{R}^p$ . We call  $\theta$  the generating parameter and write  $A(\theta)$  for a closed set generated by  $\theta$ . For example, a disc with a center  $\gamma$  and a radius  $\rho$  has generating parameters  $(\gamma, \rho)$ . We now extend the notion of parametric sets to random closed sets. Specifically, let  $\Theta$  be a  $p$ -dimensional ( $p \geq 1$ ) random variable defined on the probability space  $(\Omega, \mathcal{A}, \mathbf{P})$  with  $E|\Theta| < \infty$ . Consider a random closed set given by a mapping from  $\Omega$  to  $\mathcal{F}$ , so that the geometry of the set depends on the realized value of  $\Theta$ . Similarly, we call  $\Theta$  the generating parameter of a r.c.s.  $\mathbf{A}(\Theta)$ . For example, a disc with a random center  $C$  and a random radius  $R$  has generating parameters  $\Theta = (C, R)$ , whilst a disc with a fixed center and a random radius  $R$  has a generating parameter  $\Theta = R$ . The ODF  $b_{\mathbf{A}}(x)$  of  $\mathbf{A}(\Theta)$  is a function of random variable  $\Theta$  and so, is a random quantity. We again assume that  $b_{\mathbf{A}}(x)$  is integrable for some  $x_0 \in \mathcal{D}$  and hence is well-defined.

**Definition 4.1** (Separable ODF). The ODF of set  $A = A(\theta)$  with parameter  $\theta$  is separable, if it has a form  $b_A(x; \theta) = h^T(x)g(\theta)$ , where  $h(x) = (h_1(x), \dots, h_K(x))^T$  and  $g(\theta) = (g_1(\theta), \dots, g_K(\theta))^T$  for some functions  $h_i(x), g_i(\theta)$ ,  $i = 1, \dots, K$ .

**Definition 4.2** (Separable r.c.s). An r.c.s.  $\mathbf{A}$  is separable, if there exists a random variable  $\Theta$  such that  $b_{\mathbf{A}}(x; \Theta)$  is separable a.s.. That is  $b_{\mathbf{A}}(x; \Theta) = h^T(x)g(\Theta)$  a.s., with  $h(x) = (h_1(x), \dots, h_K(x))^T$  and  $g(\theta) = (g_1(\theta), \dots, g_K(\theta))^T$  for some functions  $h_i(x), g_i(\theta)$ ,  $i = 1, \dots, K$ .

The following theorem establishes the connection between the expectation of a set and its generating parameter.

**Theorem 4.1.** *Let  $\Theta$  be a generating parameter of a separable r.c.s.  $\mathbf{A}(\Theta)$  and assume that  $E[|g_i(\Theta)|] < \infty$  for every  $i = 1, \dots, K$ . Then,*

1. *If  $g(\theta)$  is convex in every argument, then  $E[\mathbf{A}(\Theta)] \subset A(E[\Theta])$ .*
2. *If  $g(\theta)$  is an affine function of  $\theta$  in every component  $g_i(\theta)$  a.s., then  $E[\mathbf{A}(\Theta)] = A(E[\Theta])$ .*

*Proof.* (1) It follows from Jensen's inequality that  $E[b_{\mathbf{A}}(x; \Theta)] = h^T(x) E[g(\Theta)] \geq h(x)^T g(E[\Theta])$ . The right-hand side of the equation is the ODF of a set with parameter  $E[\Theta]$ . For any  $x \in E[\mathbf{A}(\Theta)]$ , we have  $h^T(x)g(E[\Theta]) \leq E[b_{\mathbf{A}}(x; \Theta)] \leq 0$  and  $x \in \text{cl } A(E[\Theta]) = A(E[\Theta])$ . Thus,  $E[\mathbf{A}(\Theta)] \subset A(E[\Theta])$ .

(2) From Jensen's inequality,  $E[b_{\mathbf{A}}(x; \Theta)] = h^T(x)g(E[\Theta])$  iff the function  $g_i(\theta)$  is affine in  $\theta$  for every  $i = 1, \dots, K$ . The expression on the right-hand site is the ODF of a set with parameter  $E[\Theta]$ , hence the statement (2) of the theorem follows.

Note that although the linearity of  $g(\theta)$  is a necessary and sufficient condition for the form of  $E[b_{\mathbf{A}}(x; \Theta)]$ , it is only a sufficient condition for the separability of  $E[\mathbf{A}(\Theta)]$ . This is because the expected set depends on  $E[b_{\mathbf{A}}(x; \Theta)]$  up to a constant, so that for fixed  $c > 0$ ,  $E[b_{\mathbf{A}}(x; \Theta)]$  and  $cE[b_{\mathbf{A}}(x; \Theta)]$  define the same expected set.  $\square$

**Example 4.1.** In Example 3.3, the ODF of a closed ball with random radius  $b_{\mathbf{A}}(x) = |x| - \Theta$  is separable with  $h(x) = (|x|, -1)^T$  and  $g(\theta) = (1, \theta)^T$ , and the expected set is a closed ball with radius  $E[\Theta]$ .

In Example 3.4, the ODF of flashing discs is not separable, and the geometry of the expected set differs from that of the observed sets. However, note that both the random set and its expectation can be described as Cartesian ovals.

In Example 3.5, the ODF of a half-plane is separable with  $h(x) = (x_1, -1)^T$  and  $g(\theta) = (1, \theta)^T$ . Hence, the expected set is a half-plane with parameter  $E[\Theta]$ .

In Example 3.6, the ODF of an upper half-plane can be written as  $b_{\mathbf{A}}(x) = x_1 \sin \Theta - x_2 \cos \Theta$ , and so is separable with  $h(x) = (x_1, -x_2)^T$  and  $g(\theta) = (\sin \theta, \cos \theta)^T$ . The expected set is an upper half-plane whose boundary makes angle  $E[\Theta]$  with the positive horizontal axis. This example illustrates the second statement of Theorem 4.1, in that  $E[\mathbf{A}(\Theta)] = A(E\Theta)$ , even for nonlinear  $g$ .

**Remark 4.1.** Note that if  $g$  is invertible in every component, we can reparametrize the expected ODF as  $E[b_{\mathbf{A}}(x; \Theta)] = h^T(x)g(\eta)$ , where  $\eta_i = g_i^{-1}(E g_i(\Theta))$ ,  $i = 1, \dots, K$ . Hence, the expected set and the expected boundary have the same geometric form in terms of parameters  $\eta_i$ . The relation between  $\eta_i$  and the moments of  $\Theta$  is established on a case-by-case basis. In particular, for the ODFs of the form  $b_{\mathbf{A}}(x) = h(x) + g(\Theta)$  for some function  $h(x): \mathbb{R}^d \mapsto \mathbb{R}$  and an invertible function  $g(\theta): \mathbb{R}^p \mapsto \mathbb{R}$ , we obtain that  $E[b_{\mathbf{A}}(x; \Theta)] = h(x) + E[g(\Theta)] = h(x) + g(g^{-1}(E[g(\Theta)]))$  and  $E[\mathbf{A}(\Theta)] = A(g^{-1} E[g(\Theta)])$ .

**Remark 4.2.** In digital processing, data are stored as an array of pixels and so the ODF of an object in the image is discretised. This discretization implies that for realizations of an r.c.s.  $\mathbf{A}$  at times  $t = 1, 2, \dots$ , the ODFs of  $\mathbf{A}_t$  can be viewed as a random sample of arrays in the time domain. For a separable ODF, the space-time covariance matrix has a form

$$\text{cov}(b_{\mathbf{A}_t}(x), b_{\mathbf{A}_\tau}(y)) = h^T(x) \text{cov}(g(\Theta_t), g(\Theta_\tau)) h(y),$$

for lattice vertices  $x, y$  and time points  $t, \tau$ . This expression shows that the covariance of a separable ODF is a product of a deterministic spatial component and the covariance of the stochastic time component. In particular, for an ODF of the form  $b_{\mathbf{A}}(x) = h(x) + g(\Theta)$ , the space-time covariance is given by  $\text{cov}(b_{\mathbf{A}_t}(x), b_{\mathbf{A}_\tau}(y)) = \text{cov}(g(\Theta_t), g(\Theta_\tau))$ , and hence the dependence between the sets can be inferred from their generating parameters.

## 5 Further Developments and Applications

### 5.1 The Sample Mean Set

Let  $\mathbf{A}_1, \dots, \mathbf{A}_m$  be the observed realizations of an r.c.s.  $\mathbf{A}$ . For every  $i$ , we assume that  $\mathbf{A}_i \subset \text{cl } \mathcal{D}$  and  $\partial \mathbf{A}_i \neq \emptyset$  and denote by  $b_i(x)$  the ODF of  $\mathbf{A}_i$  at point  $x$ . Let  $\bar{b}_m(x) = m^{-1} \sum_{i=1}^m b_i(x)$  be the sample mean ODF at point  $x$ .

**Proposition 5.1.** *The sample mean ODF  $\bar{b}_m(x)$  is uniformly Lipschitz with constant one.*

*Proof.* The proof follows immediately from the Lipschitz continuity of  $b_i(x)$ .  $\square$

**Definition 5.1** (Sample mean set). The sample mean set  $\bar{\mathbf{A}}_m$  and the sample mean boundary  $\partial \bar{\mathbf{A}}_m$  are given by the zero-level set and by the zero-level isocontour of the empirical ODF, respectively, i.e.

$$\bar{\mathbf{A}}_m = \{x : \bar{b}_m(x) \leq 0\}, \quad \partial \bar{\mathbf{A}}_m = \{x : \bar{b}_m(x) = 0\}. \quad (15)$$

The sample mean set defined in (15) is a random closed set, since  $\{\bar{\mathbf{A}}_m \cap K \neq \emptyset\} = \{\inf_{x \in K} \bar{b}_m(x) \leq 0\}$  is measurable.

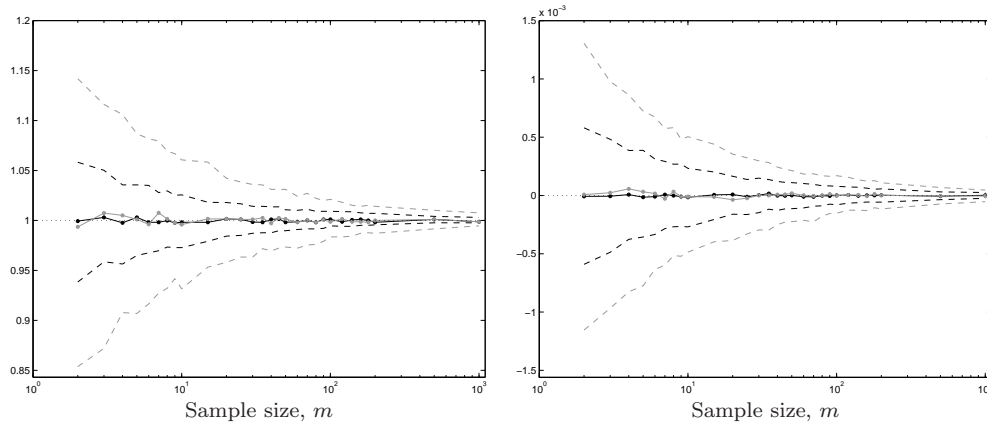


Figure 4: (Left) The ratio of the radii of  $\bar{\mathbf{A}}_m$  to that of the expected set  $E[\mathbf{A}]$  in Example 5.1 for  $\Theta \sim U[0.8, 1.2]$  (black) and  $\Theta \sim U[0.5, 1.5]$  (gray). (Right) The difference between the boundary angles of  $\bar{\mathbf{A}}_m$  and  $E[\mathbf{A}]$  in Example 5.2 with  $\Theta \sim U[\pi/8, 3\pi/8]$  (black) and  $\Theta \sim U[0, \pi/2]$  (gray). The median (solid) and the 25-th and 75-th percentile (dashed) values are based on 1000 simulations for each  $m$ .

**Example 5.1** (Disc with random radius). Recall that for a closed ball  $\mathbf{A}$  with a random radius  $\Theta$ ,  $E[\mathbf{A}(\Theta)] = A(E[\Theta])$  and  $E[\partial \mathbf{A}] = \partial A(E[\Theta])$ . For a sample of closed balls  $\mathbf{A}_1, \dots, \mathbf{A}_m$  with random radii  $\Theta_1, \dots, \Theta_m$ , the empirical ODF is  $\bar{b}_m(x) = |x| - \bar{\Theta}_m$  and hence  $\bar{\mathbf{A}}_m = A(\bar{\Theta}_m)$  and  $\partial \bar{\mathbf{A}}_m = \partial A(\bar{\Theta}_m)$ .

For an independent random sample of radii values  $\Theta_1, \dots, \Theta_m$ , we construct  $m$  random discs in  $\mathbb{R}^2$  centered at the origin. To compare the empirical boundary estimator with the theoretical average, we compute the ratio of the radii of  $E[\mathbf{A}]$  to  $A(\bar{\Theta}_m)$ ; see Figure 4, left. The radii values were sampled from the uniform distribution  $U[0.8, 1.2]$  (black) and  $U[0.5, 1.5]$  (gray). The median values of the ratios (solid curves)

and the 25-th and 75-th percentiles (dashed curves) are based on 1000 samples for each sample size  $m$ . The figure shows that the accuracy of the empirical estimate is higher for bigger sample sizes and for smaller variance values of the generating parameter.

**Example 5.2** (Random upper half-plane). Consider an r.c.s. in Example 3.6 given by  $\mathbf{A}(\Theta) = \{x \in \mathbb{R}^2: x_2 \geq x_1 \tan \Theta\}$ , where  $\Theta$  is the angle between the boundary of the plane and the positive horizontal axis. Recall that for  $\Theta \sim U[a, b]$ ,  $E[\mathbf{A}(\Theta)] = A(E[\Theta])$  and  $E[\partial \mathbf{A}] = \partial A(E[\Theta])$ .

For an independent random sample of the angle values  $\Theta_1, \dots, \Theta_m$ , we construct  $m$  upper half-planes. The empirical mean boundary is estimated by the zero-level isocontour of the empirical mean ODF. Figure 4 (right) shows the difference between the boundary angles for the empirical mean set and that of the expected set for  $\Theta \sim U[\pi/8, 3\pi/8]$  (black) and  $\Theta \sim U[0, \pi/2]$  (gray). The median values of the ratios (solid curves) and the 25-th and 75-th percentiles (dashed curves) are based on 1000 replicates for each value of  $m$ . As in the previous case, the boundary estimate is more accurate for bigger sample sizes and for smaller variance values of  $\Theta$ .

The empirical mean set and the empirical boundary provide a recipe for practical construction of the set and boundary estimators. Examples 5.1 and 5.2 suggest the consistency of the boundary estimator, based on the empirical ODF. Note that both the expected set  $E[\mathbf{A}]$  and the empirical mean set  $\bar{\mathbf{A}}_m$  are described as level sets of the continuous function  $b_{\mathbf{A}}$  and hence, under certain conditions, the consistency of  $\bar{\mathbf{A}}_m$  as an estimator of  $E[\mathbf{A}]$  can be inferred from (Molchanov, 1998, Theorem 2). The result of (Molchanov, 1998), however, does not apply to the boundary estimator  $\partial \bar{\mathbf{A}}_m$ . In our upcoming paper, we study the consistency of the boundary estimator in greater detail Jankowski and Stanberry (in preparation).

## 5.2 Loss Functions

Assume that the observed data  $\mathbf{A} \in \mathcal{F}$  follows a distribution  $F_{\mathfrak{A}}$  with mean parameter  $\mathfrak{A} = E[\mathbf{A}]$ ,  $\mathfrak{A} \in \mathcal{F}$ . We call  $\mathfrak{A}$  the parameter set. Here, data  $\mathbf{A}$  represent random closed sets with nonempty boundary. In practice, it is more common to acquire images rather than observe sets per se, so the observed  $\mathbf{A}$  can be seen as a result of the first-step analysis. The goal is to estimate the parameter set  $\mathfrak{A}$  given the observed data  $\mathbf{A}$ .

If set  $\mathfrak{A}$  is parametric with parameter  $\theta$ , we can construct an estimator of  $\mathfrak{A}$  using conventional loss functions for  $\theta$ . For example, if  $\mathbf{A}$  is a collection of deformed balls, the distribution  $F$  can be parametrized by the center and radius of a ball. Depending on the problem, we might be interested in testing the hypothesis about either the location or the size of the ball, or both. Then the optimal estimator  $\hat{\theta}$  is determined using standard loss functions, e.g. the indicator loss  $\ell(\theta, a) = \mathbb{I}(a = \theta)$ , the absolute error loss  $\ell(\theta, a) = |a - \theta|$  or the squared error loss  $\ell(\theta, a) = (a - \theta)^2$ . However, often, the sets of interest cannot be parametrized and at best can be described as “blobs”, in which case conventional loss functions are not applicable. Here, we relay loss functions for Bayesian inference about sets introduced in Stanberry and Besag (in preparation) in the frequentist framework.

To begin with, we consider a loss function based on representation of a set  $A$  given by its characteristic

function  $\chi_A(x) = 1_A(x)$ . The space of characteristic functions is complete with metric

$$d(A, B) = \|\chi_A - \chi_B\|_{L^q} = \left( \int_{\mathcal{D}} |\chi_A(x) - \chi_B(x)|^q \lambda(\mathrm{d}x) \right)^{1/q}. \quad (16)$$

Here and later, the integration domain  $\mathcal{D}$  is either the entire image or its subset. Let  $\widehat{\chi}$  be an estimator of  $\chi_A$ . Given the metric  $d$ , we define an  $L_1$  loss function as

$$\ell_1(\chi_{\mathfrak{A}}, \widehat{\chi}) = \|\chi_{\mathfrak{A}} - \widehat{\chi}\|_{L^1} = \int_{\mathcal{D}} |\chi_{\mathfrak{A}}(x) - \widehat{\chi}(x)| \lambda(\mathrm{d}x). \quad (17)$$

Because the estimator  $\widehat{\chi}$  is the characteristic function that minimizes the expected loss, the estimator  $\widehat{\mathfrak{A}}$  of  $\mathfrak{A}$  is immediately available. Strictly speaking,  $d(A, B)$  distinguishes between sets up to a set of Lebesgue measure zero, so we assume that the parameter set has no punctures and hence  $\widehat{\mathfrak{A}}$  is uniquely determined. Note that the equation (17) is the measure of the symmetric difference, since

$$\ell_1(\chi_{\mathfrak{A}}, \widehat{\chi}) = \lambda(\mathfrak{A} \Delta \widehat{\mathfrak{A}}) = \lambda((\mathfrak{A} \setminus \widehat{\mathfrak{A}}) \cup (\widehat{\mathfrak{A}} \setminus \mathfrak{A})) \quad (18)$$

and, hence,  $\ell_1(\chi_{\mathfrak{A}}, \widehat{\chi})$  is akin to the absolute error loss in conventional settings. The set that minimizes  $E[\lambda(\mathfrak{A} \Delta \widehat{\mathfrak{A}})]$  is the Vorob'ev median, i.e. is the median level set of the coverage function (3); (see Molchanov, 2005, p. 178). The Vorob'ev median is a set-analog of the ordinary median in the sense that, as the latter, the former minimizes the absolute error loss.

Moreover, the Vorob'ev median is optimal with respect to a much broader class of  $L_q$  loss functions,

$$\ell_q(\chi_{\mathfrak{A}}, \widehat{\chi}) = \|\chi_{\mathfrak{A}} - \widehat{\chi}\|_{L^q}, \quad 1 \leq q < \infty. \quad (19)$$

This is because  $\chi_{\mathfrak{A}}$  takes values in  $\{0, 1\}$  and the induced  $L_q$  topologies (19) are all equivalent for  $1 \leq q < \infty$ ; (see Delfour and Zolésio, 2001, Theorem 2.2).

Instead of optimizing the expected  $L_q$  loss over the set of characteristic functions, consider minimizing (19) over a larger set of functions with values in  $[0, 1]$ . For the squared  $L_2$  loss

$$\ell_2(\chi_{\mathfrak{A}}, \widehat{\chi}) = \|\chi_{\mathfrak{A}} - \widehat{\chi}\|_{L_2}^2, \quad (20)$$

the pointwise estimator  $\widehat{\chi}(x) \in [0, 1]$  is the coverage function (3). The estimator  $\widehat{\chi}(x)$  is a global minimum in  $L_2$  since  $E[\|\chi_{\mathfrak{A}} - E[\chi_{\mathfrak{A}}]\|_{L_2}^2] \leq E[\|\chi_{\mathfrak{A}} - \widehat{\chi}\|_{L_2}^2]$ . Note that  $\widehat{\chi}$  is essentially a gray-scale image with intensities reflecting the probability of a pixel being in  $\mathfrak{A}$ .

The coverage function, however, is not necessarily a characteristic function, unless  $\mathbf{A}$  is deterministic a.s.. Consider a set-valued estimator given by an excursion set of  $\widehat{\chi}(x)$  so that determining the optimal estimator  $\widehat{\mathfrak{A}}$  now reduces to choosing an appropriate threshold  $u$  for the excursion set (4). The Vorob'ev criterion (5) gives an estimator  $\widehat{\mathfrak{A}}$ , which is optimal with respect to Lebesgue measure in the sense that  $E[\lambda(\mathfrak{A} \Delta \widehat{\mathfrak{A}})] \leq E[\lambda(\mathfrak{A} \Delta \widehat{\mathfrak{A}})]$  for all measurable sets with measure  $E[\lambda(\mathbf{A})]$  (see Molchanov, 2005, Theorem 2.3, pg. 177). From (18) and the equivalence of the  $L_q$ -topologies on the space of characteristic functions, it follows that the Vorob'ev expectation minimizes the expected loss (19) for any  $q \in [1, \infty)$  over all of the sets with Lebesgue measure  $E[\lambda(\mathbf{A})]$ .



In general, given the representative functions  $f_A$  and  $f_B$  of sets  $A$  and  $B$ , the discrepancy between  $A$  and  $B$  can be taken in terms of some (pseudo-) metric

$$\mathfrak{m}(A, B) = \mathfrak{m}_{\mathcal{W}}(f_A, f_B), \quad (21)$$

defined on a compact window  $\mathcal{W} \subseteq \mathcal{D}$ . Further, we suppress the subscript  $\mathcal{W}$  whenever possible. Given the metric  $\mathfrak{m}(A, B) = \|f_A - f_B\|_{L_2}$ , the global estimator based on the squared  $L_2$  loss

$$\ell_2(f_{\mathfrak{A}}, \hat{f}) = \|f_{\mathfrak{A}} - \hat{f}\|_{L_2}^2 \quad (22)$$

is the average function  $E[f_{\mathfrak{A}}]$ . Hence, the distance-average expectation (6) is an optimal estimator of  $\mathfrak{A}$  with respect to (22), among all of the level sets of  $E[f_{\mathfrak{A}}]$ .

Note that although the distance-average expectation arises naturally for the  $L_2$  loss, it was originally defined for a generic (pseudo-) metric  $\mathfrak{m}(\cdot, \cdot)$  (Baddeley and Molchanov, 1998). Thus, in principle, we can construct an estimator as the level set of the expected representation for arbitrary choice of  $\mathfrak{m}$ . For that, we must specify the representative function  $f$ , select an appropriate (pseudo-)metric  $\mathfrak{m}$ , and choose the loss function  $\ell(f_{\mathfrak{A}}, \hat{f})$  based on  $\mathfrak{m}$ . The threshold  $u$  in (6) then determines the optimal estimator with respect to  $\ell(f_{\mathfrak{A}}, \hat{f})$  among all of the level sets of  $E[f_{\mathfrak{A}}(x)]$ .

We now consider a representation of a set given by its ODF. Recall that the space of ODFs of sets with non-empty boundary is complete with metric  $d(A, B) = \|b_B - b_A\|_{L_q}$ . For the loss function

$$\ell_2(b_{\mathfrak{A}}, \hat{b}) = \|b_{\mathfrak{A}} - \hat{b}\|_{L_2}^2 = \int_{\mathcal{D}} |b_{\mathfrak{A}}(x) - \hat{b}(x)|^2 \lambda(dx), \quad (23)$$

let  $\hat{b}$  be the estimator of  $b_{\mathfrak{A}}$  that minimizes the expectation of (23). Because the ODF gives a unique (up to the boundary) representation of a set, the estimator  $\hat{\mathfrak{A}}$  is given by the zero-level set of  $\hat{b}$ . However, optimizing  $E[\ell_2(b_{\mathfrak{A}}, \hat{b})]$  over the set of ODFs is nontrivial, so instead we consider unrestricted minimization over the set of functions defined on  $\mathcal{D}$ . In this case, the pointwise estimator is the expected ODF, which is also a global minimum. Because the estimator  $E[b_{\mathfrak{A}}(x)]$  is not necessarily an ODF itself, it does not uniquely determine the set. Hence, we take the set-valued estimator  $\hat{\mathfrak{A}}$  of  $\mathfrak{A}$  to be the zero-level set of  $E[b_{\mathfrak{A}}(x)]$ , so that  $\hat{\mathfrak{A}}$  is the expected set as defined in (3.1). The estimator of the boundary  $\partial\mathfrak{A}$  is given by the zero-level isocontour of  $E[b_{\mathfrak{A}}(x)]$ .

### 5.3 Image Averaging

In this section, we consider the example of image averaging originally discussed in Baddeley and Molchanov (1998). In image averaging, which relates to Bayesian image classification and reconstruction, the goal is to determine an average object or a typical shape from the collection of images of the same scene or objects of the same type.

To begin with, we quickly outline the image sampling procedure and refer for more details to Baddeley and Molchanov (1998). Figure 5 (left) shows the true binary image,  $I$ , of the scanned and thresholded newspaper fragment. After adding Gaussian noise, 15 independent realizations from the posterior were obtained using a Gibbs sampler with true noise parameters in the likelihood and an Ising prior. In the current notations,  $\mathcal{D} = \mathcal{W} = I$ ,

$\mathfrak{A}$  is the true newspaper text and  $\mathbf{A}_1, \dots, \mathbf{A}_{15}$  are the 15 independent reconstruction of  $\mathfrak{A}$ . The data for this example was downloaded from Baddeley (2009).



Figure 5: The true binary image of the newspaper fragment.

Figure 5 shows the ODF estimator  $\hat{\mathfrak{A}}_{ODF}$  (left) and the distance-average estimator  $\hat{\mathfrak{A}}_{DA}$  (right) with  $f_A(\cdot) = b_A(\cdot)$  and  $m(\cdot, \cdot) = L_2(\cdot, \cdot)$ ; see also Figure 6 in Baddeley and Molchanov (1998). The estimator  $\hat{\mathfrak{A}}_{ODF}$  appears less noisy and more accurate as compared to  $\hat{\mathfrak{A}}_{DA}$ , where the letters look overinflated.



Figure 6: The ODF estimator  $\hat{\mathfrak{A}}_{ODF}$  (left) and the distance-average estimator  $\hat{\mathfrak{A}}_{DA}$  (right).

To compare the results of the reconstruction, it is instructive to look at the residual images for  $\hat{\mathfrak{A}}_{ODF}$  and  $\hat{\mathfrak{A}}_{DA}$  in Figure 7, left and right, respectively. Note that black (resp. gray) pixels in the figure correspond to the set  $\hat{\mathfrak{A}} \setminus \mathfrak{A}$  (resp.  $\mathfrak{A} \setminus \hat{\mathfrak{A}}$ ). The residual image for  $\hat{\mathfrak{A}}_{DA}$  shows a clear spatial pattern, indicating that  $\hat{\mathfrak{A}}_{DA}$  tends to overestimate the set. In turn, the residuals of  $\hat{\mathfrak{A}}_{ODF}$  show little spatial clustering. In addition,  $\hat{\mathfrak{A}}_{ODF} \subset \hat{\mathfrak{A}}_{DA}$ , that is the distance-average estimator is inclusive of the ODF one, with discrepancies observed along the boundary and in the background.

To compare the quality of the estimators, it is common to use the fraction of misclassified pixels. Here, the error is 4.12% for  $\hat{\mathfrak{A}}_{ODF}$  and 10.68% for  $\hat{\mathfrak{A}}_{DA}$ . However, a small misclassification error does not necessarily imply similar looking images (Baddeley, 1992), therefore, we compare the two estimators by computing an

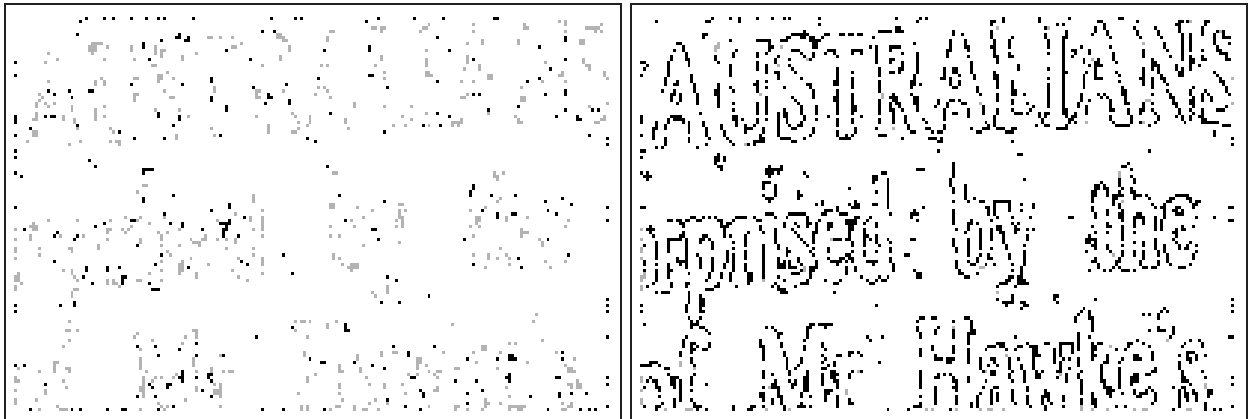


Figure 7: The residual images of the discrepancy between the estimated set and the truth for  $\widehat{\mathfrak{A}}_{ODF}$  (left) and  $\widehat{\mathfrak{A}}_{DA}$  (right). Black (resp. gray) pixels mark set  $\widehat{\mathfrak{A}} \setminus \mathfrak{A}$  (resp.  $\mathfrak{A} \setminus \widehat{\mathfrak{A}}$ ).

$L_2$  distance to the truth

$$d(\mathfrak{A}, \widehat{\mathfrak{A}}) = \left( \int_{\mathcal{D}} |b_{\mathfrak{A}}(x) - b_{\widehat{\mathfrak{A}}}(x)|^2 dx \right)^{1/2},$$

which is compatible with the distance-average expectation. Calculating the distances for the two estimators, we obtain  $d(\mathfrak{A}, \widehat{\mathfrak{A}}_{ODF}) = 1.78$  and  $d(\mathfrak{A}, \widehat{\mathfrak{A}}_{DA}) = 2.66$ , which shows that the ODF estimator outperforms the distance-average reconstruction.

Note that the reconstruction results using ODFs do not change if we reverse the color of pixels, as opposed to the distance-average estimator with  $f_A(\cdot) = d_A(\cdot)$ , where the swapping affects the quality of the reconstruction (see Baddeley and Molchanov, 1998, Figures 4 and 5).

## 5.4 Boundary Reconstruction in Noisy Images

In image analysis, it is often required to reconstruct the boundary of an object in a noisy image. Low level methods such as edge detection and intensity thresholding strongly depend on image characteristics and require a careful tuning of parameters to produce satisfactory results (Canny, 1986; Gonzalez and Woods, 2002). A number of methods for boundary reconstruction have been proposed in the literature. In active contour models, the boundary is represented by a parametric or a free-form curve and the reconstruction is driven by local image forces, with smoothness controlled via user-imposed regularity constraints (Blake and Isard, 1998; Brigger et al., 2000; El-Baz et al., 2005; Flickner et al., 1994; Kass et al., 1988; Menet et al., 1990; Rueckert and Burger, 1997; Tauber et al., 2004; Terzopoulos et al., 1987; Yuille et al., 1992). The use of local properties of the image implies strong dependence of the boundary estimate on initialization of the modeling curve, a contrast-to-noise ratio of the image and convexity of the object. Various extensions were proposed in the literature to improve the performance of active contours (Coughlan et al., 2000; El-Baz et al., 2005; Wang et al., 1996; Cohen, 1991; Xu and Prince, 1998).

Many Bayesian methods for boundary reconstruction use templates which represent local or global geometry of an object (Grenander, 1976; Tjelmeland and Besag, 1998; Amit et al., 1991; Grenander, 1996; Jain et al., 1996; Grenander and Keenan, 1993; Grenander, 1996; Jain et al., 1998; Chalmond, 2003; Geman et al.,

1990). The stochastic method in Qian et al. (1996) models the boundary by closed polygons. The method was then extended to allow for the dynamic estimation of the number of polygon vertices, thus balancing the complexity and the flexibility of the model (Pievatolo and Green, 1998). However, the reconstruction is given by a gray-scale image, where intensities reflect the posterior mean state of pixels, so that constructing a boundary estimator requires additional post-processing.

In Stanberry and Besag (in preparation), the authors proposed a new method to reconstruct a smooth connected boundary of an object in a noisy image using B-spline curves. The method is Bayesian and uses a Markov chain Monte Carlo algorithm to obtain curve samples from the posterior. Using the squared  $L_2$ -loss based on ODFs, the posterior estimator is given by the posterior expected boundary as defined in (10). The simulation study showed that the ODF estimator is more accurate as compared to the distance-average reconstruction and Vorob'ev sets.

## 5.5 Implementation

In addition to its appealing theoretical properties, the ODF expectation can be efficiently computed, using algorithms for the distance function (Breu et al., 1995; Freidman et al., 1977; Rosenfeld and Pfaltz, 1966). The ODF is computed by combining the distance function for the original image and that of the inverted image, i.e. the black and white colors are interchanged. The examples and results reported in this paper were coded in MATLAB (R2008A, The MathWorks), where the distance function to a set is computed using the `bwdist` command.

In comparison to the distance-average and Vorob'ev expectations, the ODF average is more efficient, because it requires no optimization since the threshold level is fixed at zero. This difference is particularly prominent in Bayesian reconstruction, where the estimator is based on thousands of samples from the posterior. In addition, the storage requirements for the ODF reconstruction in Bayesian framework are miniscule, as we only need to update the average after every sweep.

## 6 Discussion

In this paper, we present new definitions of the expected set and the expected boundary, using the ODF representation of a set. In conventional settings, where the parameter of interest is a scalar or a vector, it is straightforward to construct an estimator of the mean from the observed data. However, statistical inference about sets is nontrivial. When dealing with sets, one has to determine the features that are important to emphasize and develop inference methods based on them. These features may include the location, the size, or the orientation of a set. If a set has an analytic representation, the most natural approach to set inference is via inference about its parameters. This, however, is not applicable for sets with arbitrary geometry.

A number of existing definitions of the expected set are based on the linearization idea outlined in Molchanov (2005), where inference is made on the space of functions representing a set. The resulting estimator necessarily depends on the type of linearization and the choice of the representative function. In this paper, we use an approach similar in spirit to the linearisation idea and define the expectation based on the ODF representation of a set. However, we forgo the optimality criterion and simply choose zero as

a threshold. Whilst the choice might appear simplistic, it gives an expectation with attractive theoretical traits, including equivariance, convexity-preservation and inclusion properties.

The definition is particularly appealing for problems in image analysis. In particular, its denoising property implies that random specks will be averaged out unless observed with probability one. In comparison, the distance-average expectation is not empty by construction, and therefore identifies false positives in any noisy image. Furthermore, the distance-average reconstruction strongly depends on the choice of pseudo-metric, representative function, restriction window and any parameters of the above. In turn, the selection expectation depends on the structure of the probability space, whilst the Lebesgue measure criterion of the Vorob'ev expectation is not well suited for image analysis, although, the optimality of the Vorob'ev median is rather attractive.

In our upcoming paper, we study the asymptotic properties of the ODF estimators for the boundary and describe a method for constructing confidence sets (Jankowski and Stanberry, in preparation).

## References

- Y. Amit, U. Grenander, and M. Piccioni. Structural image restoration through deformable templates. *Journal of the American Statistical Association*, 86(414):376–387, 1991.
- Z. Artstein and R.A. Vitale. A strong law of large numbers for random compact sets. *Annals of Probability*, 3(5):879–882, 1975.
- R.J. Aumann. Integrals of set-valued functions. *Journal of Mathematical Analysis and Applications*, 12: 1–12, 1965.
- A. J. Baddeley. Errors in binary images and an  $L^p$  version of the Hausdorff metric. *Nieuw Archief voor Wiskunde*, 10:157–183, 1992. URL [citeseer.ist.psu.edu/baddeley92errors.html](http://citeseer.ist.psu.edu/baddeley92errors.html).
- A. J. Baddeley, 2009. URL <http://school.maths.uwa.edu.au/homepages/adrian/>.
- A. J. Baddeley and I. Molchanov. Averaging of random sets based on their distance functions. *Journal of Mathematical Imaging and Vision*, 8(1):79–92, 1998.
- A. Blake and M. Isard. *Active Contours: The Application of Techniques from Graphics, Vision, Control Theory and Statistics to Visual Tracking of Shapes in Motion*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1998.
- H. Breu, J. Gil, D. Kirkpatrick, and M. Werman. Linear time euclidean distance transform algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(5):529–533, May 1995. ISSN 0162-8828.
- P. Brigger, J. Hoeg, and M. Unser.  $B$ -spline snakes: a flexible tool for parametric contour detection. *IEEE Transactions on Image Processing*, 9(9):1484–1496, 2000.
- J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(6): 679–698, 1986. ISSN 0162-8828.

- B. Chalmond. *Modeling and Inverse Problems in Imaging Analysis*, volume 155 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2003.
- G. Choquet. Theory of capacities. *Ann. Inst. Fourier*, 5:131–295, 1953.
- L.D. Cohen. On active contour models and balloons. *Computer Vision, Graphics, and Image Processing. Image Understanding*, 53:211–218, 1991.
- J. Coughlan, A.L. Yuille, C. English, and D. Snow. Efficient deformable template detection and localization without user initialization. *Computer Vision and Image Understanding*, 78(3):303–319, 2000.
- M.C. Delfour and J.-P. Zolésio. *Shapes and Geometries*, volume 4 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001.
- A. El-Baz, S.E. Yuksel, H. Shi, A.A. Farag, M. Abou El-Ghar, T. Eldiasty, and M.A. Ghoneim. 2D and 3D shape based segmentation using deformable models. In *MICCAI (2)*, pages 821–829, 2005.
- M. Flickner, H. Sawhney, D. Pryor, and J. Lotspiech. Intelligent interactive image outlining using spline snakes. *Conference Record of the 28th Asilomar Conference on Signals, Systems and Computers.*, 1: 731–735 vol.1, 1994.
- J. H. Freidman, J. L. Bentley, and R. A. Finkel. An algorithm for finding best matches in logarithmic expected time. *ACM Trans. Math. Softw.*, 3(3):209–226, 1977. ISSN 0098-3500.
- D. Geman, S. Geman, C. Graffigne, and P. Dong. Boundary detection by constrained optimization. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(7):609–628, 1990. ISSN 0162-8828.
- R. C. Gonzalez and R. E. Woods. *Digital Image Processing (2nd Edition)*. Prentice Hall, January 2002. ISBN 0201180758.
- U. Grenander. *Pattern Synthesis*. Springer-Verlag, New York, 1976. Lectures in Pattern Theory, Vol. 1, Applied Mathematical Sciences, Vol. 18.
- U. Grenander. *Elements of Pattern Theory*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, 1996.
- U. Grenander and D.M. Keenan. On the shape of plane images. *SIAM Journal of Applied Mathematics*, 53(4):1072–1094, 1993.
- A.K. Jain, Y. Zhong, and S. Lakshmanan. Object matching using deformable templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(3):267–278, 1996.
- A.K. Jain, Y. Zhong, and M.-P. Dubuisson-Jolly. Deformable template models: a review. *Signal Processing*, 71(2):109–129, 1998.
- H.K. Jankowski and L.I. Stanberry. On boundary estimation. in preparation.

- M. Kass, A. P. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1988.
- G. Matheron. *Random Sets and Integral Geometry*. J. Wiley & Sons, New York-London-Sydney, 1975.
- S. Menet, P. Saint Marc, and G. Medioni. B-snakes: Implementation and application to stereo. In *Image Understanding Workshop*, volume 90, pages 720–726, 1990.
- I. Molchanov. A limit theorem for solutions of inequalities. *Scandinavian Journal of Statistics*, 25:235–242, 1998.
- I. Molchanov. *Theory of Random Sets*. Probability and its Applications (New York). Springer-Verlag London Ltd., London, 2005.
- A. Pievatolo and P.J. Green. Boundary detection through dynamic polygons. *Journal of the Royal Statistical Society, Series B, Statistical Methodology*, 60(3):609–626, 1998.
- W. Qian, D.M. Titterton, and J.N. Chapman. An image analysis problem in electron microscopy. *Journal of American Statistical Association*, 91(435):944–952, 1996.
- A. Rosenfeld and J.L. Pfaltz. Sequential operations in digital picture processing. *J. ACM*, 13(4):471–494, 1966. ISSN 0004-5411.
- D. Rueckert and P. Burger. Geometrically deformable templates for shape-based segmentation and tracking in cardiac MR images. In *EMMCVPR '97: Proceedings of the First International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 83–98, London, UK, 1997. Springer-Verlag.
- L.I. Stanberry and J. Besag. Boundary reconstruction in binary images using splines. in preparation.
- C. Tauber, H. Batatia, G. Morin, and A. Ayache. Robust B-spline snakes for ultrasound image segmentation. *Computers in Cardiology*, pages 325–328, 2004.
- D. Terzopoulos, A.P. Witkin, and M. Kass. Snakes: Active contour models. In *IEEE International Conference on Computer Vision*, pages 259–268, 1987.
- H. Tjelmeland and J. Besag. Markov random fields with higher-order interactions. *Scandinavian Journal of Statistics. Theory and Applications*, 25(3):415–433, 1998.
- O.Yu. Vorob'ev. *Srednemernoe Modelirovanie*. Nauka, Moscow, USSR, 1984.
- M. Wang, J. Evans, L. Hasebrook, and C. Knapp. A multistage, optimal active contour model. *IEEE Transactions on Image Processing*, 5(11):1586–1591, 1996.
- C. Xu and J.L. Prince. Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*, 7(3):359–369, 1998.
- A.L. Yuille, P.W. Hallinan, and D.S. Cohen. Feature extraction from faces using deformable templates. *Int. J. Comput. Vision*, 8(2):99–111, 1992.