Detection of an Abnormal Cluster in a Network

Arias-Castro, Ery *
University of California, San Diego, Mathematics
La Jolla, CA 92093, United States
E-mail: eariasca@math.ucsd.edu

Candès, Emmanuel J.
California Institute of Technology, Applied and Computational Mathematics
Pasadena, CA 91125, United States
E-mail: emmanuel@acm.caltech.edu

Durand, Arnaud
Université Paris-Sud 11, Mathématiques
Orsay, 91405, France
E-mail: arnaud.durand@math.u-psud.fr

Introduction

This paper discusses the model problem of detecting whether or not in a given network, there is a cluster of connected nodes which exhibit an "unusual behavior".

The model. For concreteness, we model the network as the d-dimensional lattice $\mathbb{Z}^d$. In the *spatial* setting, a random variable X_v is attached to each node $v \in \mathbb{Z}^d$. We observe $\{X_v : v \in \mathbb{Z}^d\}$ and want to decide between the following two hypotheses. For specificity again, we assume a normal location family: under the null, $X_v \sim^{\text{i.i.d.}} \mathcal{N}(0,1)$; under the alternative, there is a cluster (connected component) $K \subset \mathbb{Z}^d$ and $\mu > 0$ such that $X_v \sim^{\text{i.i.d.}} \mathcal{N}(\mu|K|^{-1/2}, 1)$ for $v \in K$, while $X_v \sim^{\text{i.i.d.}} \mathcal{N}(0,1)$ for $v \notin K$. ($|K|$ denotes the size of K.) The cluster K is unknown, but restricted to belong to a class of interest, denoted by $\mathcal{K}$. The calibration makes detecting clusters of different sizes of comparable difficulty. In the *spatio-temporal* setting, a time series $X_v(\cdot)$ is attached to each node $v \in \mathbb{Z}^d$. We observe $\{X_v(t) : v \in \mathbb{Z}^d, t \in \mathbb{T}\}$. Under the null, $X_v(t) \sim^{\text{i.i.d.}} \mathcal{N}(0,1)$ for all $v \in \mathbb{Z}^d$ and $t \in \mathbb{T}$; under the alternative, there is an emerging cluster (K_t) and $\mu > 0$ such that $X_v(t) \sim^{\text{i.i.d.}} \mathcal{N}(\mu|K_{t-t_0}|^{-1/2}, 1)$ for $v \in K_{t-t_0}$, while $X_v(t) \sim^{\text{i.i.d.}} \mathcal{N}(0,1)$ for $v \notin K_{t-t_0}$. The emerging cluster K_t is random and its distribution is known up to some parameters defining the growth model. The starting time t_0 is unknown.

Motivation. Such a model is relevant in a surprisingly wide array of applications, e.g. in surveillance based on sensor networks [11]. Specific examples include the detection of radioactive materials [8], as well as other types of hazardous substances, such as biological or chemical [10]; and target tracking [35]. As a digital camera may be seen as a sensor network, with CCD or CMOS pixel sensors, the setting also includes detection problems in images, for which the literature is quite extensive, spanning several decades, in particular in satellite imagery [16], computer vision [36] and medical imaging [28]. Disease outbreak detection is another area where the goal is to detect emerging epidemics based on a network of information incorporating data from hospital emergency visits, ambulance dispatch calls and pharmacy sales of over-the-counter drugs [25]. Diseases affect computers as well, in the form of viruses and worms spreading from host to host in a computer network [34]. The information network may also take the form of a field survey. For example in [29], water quality in a network of streams in Pennsylvania is being assessed by field biologists performing a variety of analyses at various locations along the streams.

*Corresponding author

Methods. By far, the most common test rejects for large values of the *scan statistic* [23]. The method is quite pervasive and appears under different names, such as *matched filters* or *deformable templates* in the engineering literature, and corresponds to the *generalized likelihood ratio test* under the normal model assumed here. The scan statistic was originally proposed for point processes [17]. In our particular model, the scan statistic takes the form of

$$\max_{K \in \mathcal{K}} |K|^{-1/2} \sum_{v \in K} X_v.$$

The scan statistic is often implicitly assumed to be near-optimal, which explains its popularity. Many variations have been proposed, motivated in part by the fact that it is challenging to compute, in particular when the class $\mathcal{K}$ is nonparametric. We mention the *level-sets* method proposed in [30], which extracts upper level sets of the node values and sums the node values over each connected component.

Literature. Although the detection problem formulated above seems of great practical relevance, the statistics literature is almost silent on the subject, with the notable exception of the sibling topics of change-point analysis [9] and sequential analysis [33]. Indeed, the former is a special instance of the spatial setting where the graph is the one-dimensional lattice; the latter is a special case of the spatio-temporal setting where the graph is reduced to a single node. The framework of multiple hypothesis testing [5, 15, 14] may also be seen as a special case of our spatial setting where the graph is the complete graph, essentially ignoring the spatial information.

The vast majority of publications addressing the task of detection in sensor networks are found the engineering literature. They tend to assume overly simplistic models where the alternative is simple; the papers instead focus on other aspects of sensor networks such as bandwith. In stark contrast, in all the applications described earlier, the set of alternatives is composite. On the other hand, most papers addressing composite alternatives assume a parametric model (e.g. discs) for clusters [23, 26, 19], for which theoretical results are available, especially in the case of images [3, 13, 21, 31, 7]. In particular, the scan statistic performs well in the sense that it is asymptotically minimax; this is shown in [3] in a slightly different context tailored to image processing applications. The scan statistic performs well in nonparametric settings as well; in [3], the scan statistic is indeed shown to be asymptotically minimax for the case of star-shaped clusters with 'smooth' boundaries. In [2], the focus is on clusters that are paths of a certain length, and again the scan statistic is asymptotically minimax when the graph is a complete, regular tree, and near-minimax for many other types of graphs, such as the d-dimensional lattice for $d \geq 3$.

Our contribution. In this paper, we further investigate theoretical detection thresholds and the performance of the scan statistic both in the spatial and spatio-temporal models. For the practitioner, the main implication of this work is to confirm that the scan statistic enjoys a good performance in a wide variety of settings. This is the short version of a longer paper to be published later on, where the reader will find the proofs of the results stated here, extensions to other types of graphs and other node distributions, numerical experiments exploring a variety of algorithms, some novel, and a larger bibliography.

The spatial setting

We implicitly assume that the graph $\mathbb{Z}^d$ is naturally embedded in $\mathbb{R}^d$. We let $B(x, r)$ denote the closed ball for the ℓ^1 (or any other) norm on $\mathbb{R}^d$ with center x and radius $r \geq 0$. (Note that the ℓ^1-norm coincides with the shortest path metric on $\mathbb{Z}^d$.) For a set of vertices $K \subset \mathbb{Z}^d$, define $X_K = \sum_{v \in K} X_v$. Let $|K|$ denote the size of K, i.e. the number of vertices belonging to K. We consider an asymptotic setting $(m \to \infty)$ relative to a sequence of problems defined by a sequence of clusters classes $(\mathcal{K}_m)$, and a sequence of positive numbers (μ_m); the resulting alternative hypothesis is denoted H_1^m. Often, we will focus on the subgraph $\mathbb{B}_m = \{-m, \ldots, m\}^d$. For two sequences of real numbers (a_m) and (b_m), $a_m \asymp b_m$ means that $a_m = O(b_m)$ and $b_m = O(a_m)$. For $a, b \in \mathbb{R}$, we use $a \vee b$ (resp. $a \wedge b$) to denote $\max(a, b)$ (resp. $\min(a, b)$).

A lower bound on the minimax detection rate

When the clusters in the class $\mathcal{K}_m$ are disjoint, namely $\mathcal{K}_m = \{K_1, \ldots, K_m\}$ with $K_i \cap K_j = \emptyset$ for $i \neq j$, the statistics $|K_1|^{-1/2} X_{K_1}, \ldots, |K_m|^{-1/2} X_{K_m}$, which are jointly sufficient, are i.i.d. standard normal under the null, while under an alternative one of them has mean μ_m. This is the classical 'Needle in a Haystack Problem' [20] and it is known that the two hypotheses are asymptotically inseparable (contiguous) if $\overline{\lim}_{m \to \infty} \mu_m (\log m)^{-1/2} < \sqrt{2}$. This is easily extended to the general case where the clusters may intersect. Let $\sqcup_m$ denote the maximum number of disjoint clusters in $\mathcal{K}_m$. Then, H_0 and H_1^m are asymptotically inseparable if $\overline{\lim}_{m \to \infty} \mu_m (\log \sqcup_m)^{-1/2} < \sqrt{2}$. When clusters intersect substantially, we employ the usual stratagem of defining a prior on the set of alternatives, here indexed by clusters in $\mathcal{K}_m$; this is for example done in [2].

Lemma 1 *Suppose there is a prior Ψ_m on $\mathcal{K}_m$ with positive constants k_m, a_m such that for all $K \in \mathrm{supp}(\Psi_m)$, $|K| \asymp k_m$ and, for $K, K' \sim \Psi_m$ independent,*

$$\mathbf{P}\left\{|K \cap K'| \geq \ell\right\} \leq B_m e^{-\ell/B_m}, \ \forall \ell \geq 0.$$

Then, H_0 and H_1^m are asymptotically inseparable if

$$\lim_{m \to \infty} \mu_m k_m^{-1/2} (B_m \log B_m)^{1/2} = 0.$$

An entropy bound for the scan statistic

Regarding the performance of the scan statistic, a simple application of Boole's inequality and standard bounds on the tail of the normal distribution shows that it asymptotically separates H_0 and H_1^m if $\underline{\lim}_{m \to \infty} \mu_m (\log \#\mathcal{K}_m)^{-1/2} > \sqrt{2}$, where $\#\mathcal{K}_m$ denotes the number of clusters in $\mathcal{K}_m$. A refinement is obtained in the usual way, by 'thinning out' the class of clusters using ε-nets; this is done in [4, 3] to provide an upper bound on the detection rate achieved by the scan statistic. We follow this line of thought here. Given a class of clusters $\mathcal{K}$ and $\varepsilon > 0$, we consider its ε-covering number with respect to the symmetric difference (we call ε-net an ε-covering of minimal size):

$$N_\varepsilon(\mathcal{K}) = \arg\min_n \left\{ K_1, \ldots, K_n \subset \mathbb{Z}^d : \quad \max_{K \in \mathcal{K}} \min_j \frac{|K \Delta K_j|}{|K| \wedge |K_j|} \leq \varepsilon \right\}.$$

The following result is similar to [3, Theorem 4.1].

Lemma 2 *The scan statistic asymptotically separates H_0 and H_1^m if there is a sequence (ε_m) such that*

$$\underline{\lim}_{m \to \infty} \mu_m (\log N_{\varepsilon_m}(\mathcal{K}_m) + \varepsilon_m \log \#\mathcal{K}_m)^{-1/2} > \sqrt{2}.$$

Moreover, if restricted to an ε_m-net, the scan statistic asymptotically separates H_0 and H_1^m if

$$\underline{\lim}_{m \to \infty} \mu_m (1 - \varepsilon_m)^{1/2} (\log N_{\varepsilon_m}(\mathcal{K}_m))^{-1/2} > \sqrt{2}.$$

The second part of Lemma 2 is particularly useful for large (nonparametric) classes of clusters, where scanning over all clusters in the class is often impractical.

Specific examples

Equipped with tools to obtain lower bounds on the minimax detection rate and upper bounds on the rate achieved by the scan statistic, we now consider specific classes of cluters. The variety of settings we examine is wide and large.

Classes of thick clusters. Consider the simple class $\mathcal{K}_m$ of d-dimensional hypercubes of side r_m that are within $\mathbb{B}_m$. Assuming that $r_m = o(m)$, there is a subset of disjoint clusters with cardinality $[m/r_m]^d$, so that we get the following lower bound on the minimax detection rate for $\mathcal{K}_m$:

$$\overline{\lim_{m\to\infty}}\ \mu_m(\log m)^{-1/2} < \sqrt{2d}.$$

In the context of pixel images, it is shown in [3] that the scan statistic achieves that rate for a wide range of (mostly parametric) clusters classes. The same holds in our setting, and we prove that for a larger class of 'thick' clusters. For $C \geq 1$, let $\mathcal{F}_{p,d}(C)$ be the subclass of bi-Lipschitz functions $f : [-1,1]^p \to [-1,1]^d$ satisfying

$$\sup_{x\neq y} \frac{\|f(x)-f(y)\|}{\|x-y\|} \leq C \inf_{x\neq y} \frac{\|f(x)-f(y)\|}{\|x-y\|}.$$

For $f : [-1,1]^p \to \mathbb{R}^d$, define

$$\mathrm{im}(f) = \{f(x) : x \in [-1,1]^p\}.$$

Proposition 1 *Let $\mathcal{K}_m$ be the class of clusters of the form $K_f = \mathrm{im}(mf) \cap \mathbb{Z}^d$, $f \in \mathcal{F}_{d,d}(C)$. Then, for $\varepsilon_m \to 0$ slowly enough, $\log N_{\varepsilon_m}(\mathcal{K}_m) \sim d \log m$; as a consequence, the scan statistic over an ε_m-net asymptotically separates H_0 and H_1^m if*

$$\underline{\lim}_{m\to\infty}\ \mu_m(\log m)^{-1/2} > \sqrt{2d}.$$

Classes of smooth, thin clusters. Here we consider thin clusters that are smooth. For $f : [-1,1]^p \to \mathbb{R}^d$ and $r > 0$, define

$$\mathrm{im}(f,r) = \{y \in \mathbb{R}^d : \min_{x\in[-1,1]^p} \|y - f(x)\| \leq r\}.$$

For a class of functions $\mathcal{F}$, let $M_\eta(\mathcal{F})$ be its η-covering number for the supnorm (we call η-net an η-covering of minimal size):

$$M_\eta(\mathcal{F}) = \arg\min_n \{\exists f_1,\ldots,f_n : \max_{f\in\mathcal{F}} \min_j \|f - f_j\|_\infty \leq \eta\}.$$

For most parametric classes, $\log M_\eta(\mathcal{F}) \sim q \log(1/\eta)$, where q is the number of parameters defining the class. For most nonparametric smoothness classes, such as Hölder, $\log M_\eta(\mathcal{F}) \asymp (1/\eta)^\gamma$ for some exponent $\gamma > 0$ [22].

Proposition 2 *Fix two sequences, $R_m = o(m)$ and $r_m \leq R_m$ with $r_m \to \infty$. Let $\mathcal{F}$ be a subclass of $\mathcal{F}_{p,d}(C)$, $p < d$, with each $f \in \mathcal{F}$ having Lipschitz constant bounded below by C^{-1}, and let $\mathcal{K}_m$ be the class of clusters K of the form $K_f = \mathrm{im}(mf, r) \cap \mathbb{Z}^d$, $f \in \mathcal{F}$, $r_m \leq r \leq R_m$. Let $f_1,\ldots,f_{M_m}$ be an η_m-net for $\mathcal{F}$, with $\eta_m = o(r_m/m)$. Then the scan statistic over $K_{f_1},\ldots,K_{f_{M_m}}$ asymptotically separates H_0 and H_1^m if*

$$\overline{\lim_{m\to\infty}}\ \mu_m(\log M_m)^{-1/2} > \sqrt{2}.$$

Classes of bands and paths. We continue with larger classes of thin clusters. A *band* of length m and width r is of the form $\bigcup_{j=0}^m B(v_j, r)$ where $(v_0,\ldots,v_m)$ forms a path in $\mathbb{Z}^d$ (of length m); we always assume that $r \leq m$. A band with zero width ($r = 0$) is an actual path. In the following, we exclude the case where $\mathbb{Z}^d$ is of dimension $d = 1$ as it is covered by Proposition 1.

Proposition 3 *Suppose $d \geq 2$, and let $\mathcal{K}_m$ be the class of bands of width at least r_m generated by paths in $\mathbb{Z}^d$ of length at most m of the form $(v_0, \ldots, v_m)$, starting at $v_0 = 0$. Then, H_0 and H_1^m are asymptotically inseparable if*

$$\lim_{m\to\infty} \mu_m m^{-1/2}(r_m \vee 1)\log(m)^{3/2}\log(r_m \vee \log m)^{1/2} = 0, \qquad \text{for } d = 2;$$
$$\lim_{m\to\infty} \mu_m m^{-1/2}(r_m \vee 1)\log(r_m \vee 1)^{(2d-1)/(2d-2)} = 0, \qquad \text{for } d \geq 3.$$

On the other hand, for any $d \geq 2$, $\log N_{1/2}(\mathcal{K}_m) \asymp m(r_m \vee 1)^{-1}$. As a consequence, the scan statistic over a corresponding ε_m-net asymptotically separates H_0 and H_1^m if

$$\varliminf_{m\to\infty} \mu_m m^{-1/2}(r_m \vee 1)^{1/2} \text{ is large enough.}$$

The lower bound on the minimax detection rate is actually valid for the subclass of bands of width exactly r_m and generated by nondecreasing paths in $(\mathbb{Z}^+)^d$, that is paths with transitions of the form $(n_1, \ldots, n_d) \to (n_1 + \varepsilon_1, \ldots, n_d + \varepsilon_d)$ with $\varepsilon_j \in \{0, 1\}$ and $\varepsilon_1 + \cdots + \varepsilon_d = 1$. Also, if the starting point is unknown, then the problem is of course at least as hard.

Classes of arbitrary clusters. By arbitrary cluster we simply mean a connected component. Arbitrary connected components in the square lattice are sometimes called *animals* or *polyominos*, which are well-studied objects in combinatorics. We mention in passing the results in [12] which provide a law of large numbers for the scan statistic under the null. Otherwise, such objects are fairly new to statistics. Detecting animals is of course harder than detecting paths, since paths are themselves animals; so in general the rate is nonparametric. However, using the simplest bounds on the detection rate and on the rate achieved by the scan statistic, we obtain a sharp, parametric rate for small, arbitrary clusters.

Proposition 4 *Let $\mathcal{K}_m$ be the class of animals of size $k_m = o(m)$ contained in $\mathbb{B}_m$. Then H_0 and H_1^m are asymptotically inseparable if*

$$\varlimsup_{m\to\infty} \mu_m (\log m)^{-1/2} < \sqrt{2d}.$$

On the other hand, if $\mathcal{K}_m$ is the class of animals of size not exceeding $k_m = o(\log m)$ contained in $\mathbb{B}_m$, the scan statistic asymptotically separates H_0 and H_1^m if

$$\varliminf_{m\to\infty} \mu_m (\log m)^{-1/2} > \sqrt{2d}.$$

Note that, in general, we can obtain a quick (naive) upper bound on the detection rate for large clusters by considering the simple test that rejects for large values of $\sum_{v\in\mathbb{B}_m} X_v$. This test asymptotically separates H_0 and H_1^m if $\varliminf_{m\to\infty} \mu_m k_m^{1/2} m^{-d/2} = \infty$, assuming the clusters in $\mathcal{K}_m$ have size bounded below by k_m.

The spatio-temporal setting

For concreteness, and because of its relevance to applications such as disease outbreak detection, we assume that the graph is the two dimensional lattice $\mathbb{Z}^2$. Though it is tempting to see this setting as a special case of the spatial setting in $\mathbb{Z}^3$, with time as the third dimension, we only consider space-time clusters that grow over time. Therefore, we only search through time starting at the most recent time and this only adds a negligible term in the detection rate. In the discussion that follows, we observe the process $\{X_v(t) : v \in \mathbb{Z}^2; t = 0, \ldots, m\}$.

Growth models

We start by describing the *Richardson's model* [32], which is perhaps the simplest. Using standard terminology, we say that a node is occupied if it belongs to the cluster. Given a parameter $p \in (0,1)$, the cluster grows as follows. At time 0, a few nodes are occupied. When a node is occupied, it remains so indefinitely. At time t, a node becomes occupied with probability p if one of its neighbors is occupied. All the decisions at different times and different locations are made independently of each other. Richardson's model is a special case of *threshold growth automata* [18,6], which have been used to model epidemics [1]. Here, a site is occupied with probability non-decreasing in the number of its neighbors already occupied.

Growth models with limiting shape

We say that a growth process K_t in $\mathbb{Z}^2$ has *limiting shape* $S \subset \mathbb{R}^2$ if there is $v_0 \in \mathbb{Z}^2$ such that

$$\lim_{t\to\infty} \mathrm{P}\left((1-\varepsilon)tS \cap \mathbb{Z}^d \ \subset K_t - v_0 \subset \ (1+\varepsilon)tS \cap \mathbb{Z}^d\right) = 1, \quad \forall \varepsilon > 0.$$

Threshold growth automata develop a polygonal limiting shape [18, 6]. Perhaps less relevant for modeling epidemics, *internal diffusion limited aggregation* is another growth model with a limiting shape [27]. The following proposition considers a setting similar to that of Proposition 1.

Proposition 5 *Suppose the growth process K_t has limiting shape $S = \mathrm{im}(f)$ for some $f \in \mathcal{F}_{2,2}(C)$, and that $K_t \subset \mathbb{B}_t$ for all times $t_0 \le t \le m$. Then, H_0 and H_1^m are asymptotically inseparable if*

$$\overline{\lim_{m\to\infty}}\ \mu_m (\log m)^{-1/2} < 2.$$

On the other hand, consider the scan statistic over space-time cones with base any spatial cluster in an ε_m-net implicitly defined in Proposition 1; it asymptotically separates H_0 and H_1^m if

$$\underline{\lim_{m\to\infty}}\ \mu_m (\log m)^{-1/2} > 2.$$

Near-isotropic growth models

The condition that the growth model has a limiting shape may be relaxed. For a constant $C > 1$, a growth model K_t is said to be *near-isotropic* if there is a node v_0 such that

$$\lim_{t\to\infty} \mathrm{P}\left(B(v_0, C^{-1}t) \subset K_t \subset B(v_0, Ct)\right) = 1.$$

For instance, growth models with asymptotic limiting shape corresponding to a bounded, open subset of $\mathbb{R}^2$ containing the origin, are near-isotropic.

Proposition 6 *Assume that the growth process K_t is near-isotropic, with $K_t \subset \mathbb{B}_t$ for all times $t_0 \le t \le m$. Then, H_0 and H_1^m are asymptotically inseparable if*

$$\overline{\lim_{m\to\infty}}\ \mu_m (\log m)^{-1/2} \textit{ is sufficiently small.}$$

On the other hand, consider the scan statistic over space-time cylinders with base $B(v,r) \subset \mathbb{B}_m$; it asymptotically separates H_0 and H_1^m if

$$\underline{\lim_{m\to\infty}}\ \mu_m (\log m)^{-1/2} \textit{ is sufficiently large.}$$

Of course, using cones instead of cylinders would provide even sharper (implicit) constants. Here we chose to use cylinders to draw a connection with the literature in disease outbreak detection [25, 24].

Acknowledgments

EAC was partially supported by a National Science Foundation grant (DMS-06-03890). EAC and EJC were partially supported by an Office of Naval Research grant (N00014-09-1-0258).

References

[1] S. Agur, O. Diekmann, H. Heesterbeek, J. Cushing, M. Gyllenberg, M. Kimmel, F. Milner, P. Jagers, and T. Kostova, editors. *Epidemiology, cellular automata, and evolution.* Elsevier Ltd, Oxford, 1999.

[2] E. Arias-Castro, E. J. Candès, H. Helgason, and O. Zeitouni. Searching for a trail of evidence in a maze. *Ann. Statist.*, 36(4):1726–1757, 2008.

[3] E. Arias-Castro, D. Donoho, and X. Huo. Near-optimal detection of geometric objects by fast multiscale methods. *IEEE Trans. Inform. Theory*, 51(7):2402–2425, 2005.

[4] E. Arias-Castro, D. L. Donoho, X. Huo, and C. A. Tovey. Connect the dots: how many random points can a regular curve pass through? *Adv. in Appl. Probab.*, 37(3):571–603, 2005.

[5] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B*, 57(1):289–300, 1995.

[6] T. Bohman and J. Gravner. Random threshold growth dynamics. *Random Structures Algorithms*, 15(1):93–111, 1999.

[7] M. V. Boutsikas and M. V. Koutras. On the asymptotic distribution of the discrete scan statistic. *J. Appl. Probab.*, 43(4):1137–1154, 2006.

[8] S. M. Brennan, A. M. Mielke, D. C. Torney, and A. B. Maccabe. Radiation detection with distributed sensor networks. *Computer*, 37(8):57–59, 2004.

[9] E. Carlstein, H.-G. Müller, and D. Siegmund, editors. *Change-point problems.* Institute of Mathematical Statistics Lecture Notes—Monograph Series, 23. Institute of Mathematical Statistics, Hayward, CA, 1994. Papers from the AMS-IMS-SIAM Summer Research Conference held at Mt. Holyoke College, South Hadley, MA, July 11–16, 1992.

[10] Y. Cui, Q. Wei, H. Park, and C. Lieber. Nanowire Nanosensors for Highly Sensitive and Selective Detection of Biological and Chemical Species. *Science*, 293(5533):1289–1292, 2001.

[11] D. Culler, D. Estrin, and M. Srivastava. Overview of sensor networks. *IEEE Computer*, 37(8):41–49, 2004.

[12] A. Dembo, A. Gandolfi, and H. Kesten. Greedy lattice animals: negative values and unconstrained maxima. *Ann. Probab.*, 29(1):205–241, 2001.

[13] A. Desolneux, L. Moisan, and J.-M. Morel. Maximal meaningful events and applications to image analysis. *Ann. Statist.*, 31(6):1822–1851, 2003.

[14] D. Donoho and J. Jin. Higher criticism for detecting sparse heterogeneous mixtures. *Ann. Statist.*, 32(3):962–994, 2004.

[15] B. Efron. Size, power and false discovery rates. *Ann. Statist.*, 35(4):1351–1377, 2007.

[16] D. Geman and B. Jedynak. An active testing model for tracking roads in satellite images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 18:1–14, 1996.

[17] J. Glaz, J. Naus, and S. Wallenstein. *Scan statistics.* Springer Series in Statistics. Springer-Verlag, New York, 2001.

[18] J. Gravner and D. Griffeath. Random growth models with polygonal shapes. *Ann. Probab.*, 34(1):181–218, 2006.

[19] A. Hobolth, J. Pedersen, and E. B. V. Jensen. A deformable template model, with special reference to elliptical templates. *J. Math. Imaging Vision*, 17(2):131–137, 2002. Special issue on statistics of shapes and textures.

[20] I. A. Ibragimov, A. S. Nemirovskiĭ, and R. Z. Khas'minskiĭ. Some problems of nonparametric estimation in Gaussian white noise. *Teor. Veroyatnost. i Primenen.*, 31(3):451–466, 1986.

[21] T. Jiang. Maxima of partial sums indexed by geometrical structures. *Ann. Probab.*, 30(4):1854–1892, 2002.

[22] A. N. Kolmogorov and V. M. Tihomirov. ε-entropy and ε-capacity of sets in functional space. *Amer. Math. Soc. Transl. (2)*, 17:277–364, 1961.

[23] M. Kulldorff. A spatial scan statistic. *Comm. Statist. Theory Methods*, 26(6):1481–1496, 1997.

[24] M. Kulldorff. Prospective time periodic geographical disease surveillance using a scan statistic. *J. Roy. Statist. Soc. Ser. A*, 164(1):61–72, 2001.

[25] M. Kulldorff, R. Heffernan, J. Hartman, R. Assuncao, and F. Mostashari. A Space-Time Permutation Scan Statistic for Disease Outbreak Detection. *PLOS MEDICINE*, 2(3):216, 2005.

[26] M. Kulldorff, L. Huang, L. Pickle, and L. Duczmal. An elliptic spatial scan statistic. *Stat Med*, 25(22):3929–43, 2006.

[27] G. Lawler, M. Bramson, and D. Griffeath. Internal diffusion limited aggregation. *Ann. Probab.*, 20(4):2117–2140, 1992.

[28] N. Moon, E. Bullitt, K. van Leemput, and G. Gerig. Automatic brain and tumor segmentation. In *MICCAI '02: Proceedings of the 5th International Conference on Medical Image Computing and Computer-Assisted Intervention-Part I*, pages 372–379, London, UK, 2002. Springer-Verlag.

[29] G. P. Patil, J. Balbus, G. Biging, J. Jaja, W. L. Myers, and C. Taillie. Multiscale advanced raster map analysis system: definition, design and development. *Environ. Ecol. Stat.*, 11(2):113–138, 2004.

[30] G. P. Patil and C. Taillie. Upper level set scan statistic for detecting arbitrarily shaped hotspots. *Environ. Ecol. Stat.*, 11(2):183–197, 2004.

[31] M. Perone Pacifico, C. Genovese, I. Verdinelli, and L. Wasserman. False discovery control for random fields. *J. Amer. Statist. Assoc.*, 99(468):1002–1014, 2004.

[32] D. Richardson. Random growth in a tessellation. *Proc. Cambridge Philos. Soc.*, 74:515–528, 1973.

[33] D. Siegmund. *Sequential analysis*. Springer Series in Statistics. Springer-Verlag, New York, 1985. Tests and confidence intervals.

[34] P. Szor. *The Art of Computer Virus Research and Defense*. Addison-Wesley Professional, 2005.

[35] Z. Zhang, X. Gao, J. Biswas, and J. K. Wu. Moving targets detection and localization in passive infrared sensor networks. *Information Fusion, 2007 10th International Conference on*, pages 1–6, July 2007.

[36] Y. Zhong, A. Jain, and M.-P. Dubuisson-Jolly. Object tracking using deformable templates. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2(5):544–549, 2000.