Robust nonparametric detection of objects in noisy images

Document status and date:
Published: 01/01/2010

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

**Robust nonparametric detection of
objects in noisy images.**

Mikhail Langovoy and Olaf Wittich

Robust nonparametric detection of objects in noisy images.

Mikhail Langovoy[*]

Mikhail Langovoy, Technische Universiteit Eindhoven,
EURANDOM, P.O. Box 513,
5600 MB, Eindhoven, The Netherlands
e-mail: `langovoy@eurandom.tue.nl`
Phone: (+31) (40) 247 - 8113
Fax: (+31) (40) 247 - 8190

and

Olaf Wittich

Olaf Wittich, Technische Universiteit Eindhoven and
EURANDOM, P.O. Box 513,
5600 MB, Eindhoven, The Netherlands
e-mail: `o.wittich@tue.nl`
Phone: (+31) (40) 247 - 2499

Abstract: We propose a novel statistical hypothesis testing method for detection of objects in noisy images. The method uses results from percolation theory and random graph theory. We present an algorithm that allows to detect objects of unknown shapes in the presence of nonparametric noise of unknown level and of unknown distribution. No boundary shape constraints are imposed on the object, only a weak bulk condition for the object's interior is required. The algorithm has linear complexity and exponential accuracy and is appropriate for real-time systems.

In this paper, we develop further the mathematical formalism of our method and explore important connections to the mathematical theory of percolation and statistical physics. We prove results on consistency and algorithmic complexity of our testing procedure. In addition, we address not only an asymptotic behavior of the method, but also a finite sample performance of our test.

Keywords and phrases: Image analysis, signal detection, image reconstruction, percolation, noisy image, nonparametric noise, robust testing, finite sample performance.

1. Introduction

Assume we observe a noisy digital image on a screen of $N \times N$ pixels. Object detection and image reconstruction for noisy images are two of the cornerstone problems in image analysis. In this paper, we propose a new efficient technique for quick detection of objects in noisy images. Our approach uses mathematical percolation theory.

Detection of objects in noisy images is the most basic problem of image analysis. Indeed, when one looks at a noisy image, the first question to ask is whether there is any object at all. This is also a primary question of interest in such

---
[*]Corresponding author.

imsart-generic ver. 2007/04/13 file: Detection_and_Percolation.tex date: November 23, 2010

diverse fields as, for example, cancer detection (Ricci-Vitiani et al. (2007)), automated urban analysis (Negri et al. (2006)), detection of cracks in buried pipes (Sinha and Fieguth (2006)), and other possible applications in astronomy, electron microscopy and neurology. Moreover, if there is just a random noise in the picture, it doesn't make sense to run computationally intensive procedures for image reconstruction for this particular picture. Surprisingly, the vast majority of image analysis methods, both in statistics and in engineering, skip this stage and start immediately with image reconstruction.

The crucial difference of our method is that we do not impose any shape or smoothness assumptions on the *boundary* of the object. This permits the detection of nonsmooth, irregular or disconnected objects in noisy images, under very mild assumptions on the object's interior. This is especially suitable, for example, if one has to detect a highly irregular non-convex object in a noisy image. This is usually the case, for example, in the aforementioned fields of automated urban analysis, cancer detection and detection of cracks in materials. Although our detection procedure works for regular images as well, it is precisely the class of irregular images with unknown shape where our method can be very advantageous.

Many modern methods of object detection, especially the ones that are used by practitioners in medical image analysis require to perform at least a preliminary reconstruction of the image in order for an object to be detected. This usually makes such methods difficult for a rigorous analysis of performance and for error control. Our approach is free from this drawback. Even though some papers work with a similar setup (see Arias-Castro et al. (2005)), both our approach and our results differ substantially from this and other studies of the subject. We also do not use any wavelet-based techniques in the present paper.

We view the object detection problem as a nonparametric hypothesis testing problem within the class of discrete statistical inverse problems. We assume that the noise density is completely unknown, and that it is not necessarily smooth or even continuous. It is even possible that the noise distribution doesn't have a density.

In this paper, we propose an algorithmic solution for this nonparametric hypothesis testing problem. We prove that our algorithm has linear complexity in terms of the number of pixels on the screen, and this procedure is not only asymptotically consistent, but on top of that has accuracy that grows exponentially with the "number of pixels" in the object of detection. The algorithm has a built-in data-driven stopping rule, so there is no need in human assistance to stop the algorithm at an appropriate step.

In this paper, we assume that the original image is black-and-white and that the noisy image is grayscale. While our focusing on grayscale images could have been a serious limitation in case of image reconstruction, it essentially does not affect the scope of applications in the case of object detection. Indeed, in the vast majority of problems, an object that has to be detected either has (on the picture under analysis) a color that differs from the background colours (for example, in roads detection), or has the same colour but of a very different intensity, or at least an object has a relatively thick boundary that differs in colour from the background. Moreover, in practical applications one often has some prior information about colours of both the object of interest and of the background. When this is the case, the method of the present paper is applicable after simple rescaling of colour values.

The paper is organized as follows. Section 2 describes in details our statistical model, and gives a necessary mathematical introduction into the percolation theory. In Section 3, the new statistical test is introduced and its consistency is proved. This section contains the main statistical result of this paper. The power of the test and the probability of false detection for the case of small images are studied in Section 4. In addition, in Subsections 4.1 and 4.5 we describe how to select a critical cluster size for any given finite screen size. Our approach uses the fast Newman - Ziff algorithm and is completely automatic. Our main algorithm for object detection is presented in Section 5. An example illustrating the connection between the asymptotic case and the small sample case is given in Section 6. Appendix is devoted to the proof of two auxiliary estimates.

2. Detection and percolation

2.1. Percolation theory

The key observation to understand our approach to signal detection is the following central result from percolation theory Kesten (1982):

Let $\mathcal{G}$ be an infinite graph consisting of *sites* $s \in \mathcal{G}$ and *bonds* between sites. The bonds determine the topology of the graph in the following sense: We say that two sites $s, s' \in \mathcal{G}$ are *neighbors* if there is a bond connecting them. We say that a subset $C \subset \mathcal{G}$ of sites is *connected* if for any two sites $s, s' \in C$ there are sites $s_1, ..., s_n$ such that s and s_1, s_n and s', and s_k and s_{k+1} are neighbors for all $k = 1, ..., n - 1$. Considering *site percolation* on the graph $\mathcal{G}$ means that we consider random configurations $\omega \in \{0, 1\}^{\mathcal{G}}$ where the probabilites are *Bernoulli*

$$P(\omega(s) = 1) = p, \quad P(\omega(s) = 0) = 1 - p$$

independently for each $s \in \mathcal{G}$ where $0 \leq p \leq 1$ is a fixed probability. If $\omega(s) = 1$, we say that the site s is *occupied*.

Then, under mild assumptions on the graph, there is a *phase transition* in the qualitative behaviour of cluster sizes. To be precise, there is a *critical percolation probability* p_c such that for $p < p_c$ there is no infinite connected cluster and for $p > p_c$ there is one.

This statement and the very definition p_c being the location of this phase transition are only valid for infinite graphs. We can not even speak of an infinite connected cluster for finite graphs. However, a qualitative difference of sizes of connected clusters of occupied sites can already be seen for finite graphs, say with $|\mathcal{G}| = N$ sites. In a sense that will be made precise below, the sizes of connected clusters are typically of order $\log N$ for small p and of order N for values of p close to one. This will yield a criterion to infer whether p is close to zero or close to one from observed site configurations. Intuitively, for large enough values of N the distinction between the two regimes is quite sharp and located very near to the critical percolation probability of an associated infinite lattice.

2.2. The general detection problem

Even though this paper deals mostly with a discussion of the problem for triangular lattices, we first want to sketch the problem in full generality, in order to emphasize that also other choices of the underlying lattice are possible in our processing of discretized pictures, and to make more transparent why we decided to work with six-neighborhoods in the present paper.

Let thus $\mathcal{G}$ denote a planar graph. We think of the sites $s \in \mathcal{G}$ as the pixels of a discretized image and of the graph topology as indicating neighboring pixels. We consider noisy signals of the form

$$Y(s) = \mathbf{1}_{\mathcal{G_0}}(\mathbf{s}) + \sigma\epsilon(\mathbf{s}) \tag{1}$$

where $\mathbf{1}_{\mathcal{G_0}}$ denotes the indicator function of a subset $\mathcal{G}_0 \subseteq \mathcal{G}$, the noise is given by independent, identically distributed random variables $\{\epsilon(s), s \in \mathcal{G}\}$ with $E\epsilon = 0$ and $V\epsilon = 1$, and $\sigma > 0$ is the *noise variance*. Thus, σ^{-1} is a measure for the *signal to noise ratio*.

Definition 1. *(The detection problem) For signals of the form (1), we consider the* detection problem *meaning that we construct a test for the following hypothesis and alternative:*

$\mathbf{H}_0$: $\mathcal{G}_0 = \emptyset$, *i.e. there is no signal.*
$\mathbf{H}_1$: $\mathcal{G}_0 \neq \emptyset$, *i.e. there is a signal.*

Remark. (i) We usually think of these signals as being *weak* in the sense that the signal-to-noise ratio is small, i.e. $\sigma^{-1} << 1$. (ii) Later in this paper we will specialize on the consideration of *symmetric* noise.

Our approach to the detection problem consists of translating the statements of hypothesis and alternative to statements from percolation theory. First of all, we choose a threshold $\tau \in \mathbb{R}$ and produce from Y a *thresholded signal*

$$Y_\tau(s) = \left\{ \begin{array}{ll} 1 & ,Y(s) > \tau \\ 0 & ,Y(s) \leq \tau \end{array} \right. . \tag{2}$$

To adapt the terminology to percolation theory, we call a site $s \in \mathcal{G}$ *occupied*, if $Y_\tau(s) = 1$. Under H_0, the probability that a site is occupied is given by

$$q = P(Y_\tau(s) = 1) = P\left(\epsilon(s) > \frac{\tau}{\sigma}\right)$$

independently for all sites $s \in \mathcal{G}$. That means, under H_0, the thresholded signal can be equivalently described by the occupied sites in a configuration of a *Bernoulli site percolation* on $\mathcal{G}$ with percolation probability q.

Under H_1, the occupation probabilities are different for the support $\mathcal{G}_0$ of the signal and its complement $\mathcal{G}_1 := \mathcal{G} - \mathcal{G}_0$. Namely,

$$P(Y_\tau(s) = 1) = \left\{ \begin{array}{ll} p_0 := P\left(\epsilon > \frac{\tau-1}{\sigma}\right) & ,s \in \mathcal{G}_0 \\ p_1 := P\left(\epsilon > \frac{\tau}{\sigma}\right) & ,s \in \mathcal{G}_1 \end{array} \right. . \tag{3}$$

The basic idea is now to choose a threshold τ such that

$$p_0 > p_c > p_1 \tag{4}$$

where p_c is a suitably chosen *critical probability* for the lattice $\mathcal{G}$. That means the occupation probability is supercritical on the support of the signal and subcritical outside.

Remark. Note that p_0 and p_1 in the present paper denote exactly the opposite to their meaning in Langovoy and Wittich (2009). We hope this causes no inconvenience for the reader, since this does not affect the main statements of this paper.

Remark. According to the intuitive picture described above, we believe that for *reasonably large* graphs, the critical probability p_c can be chosen close or equal to the critical percolation probability of a corresponding infinite graph. See for this also the example of the triangular lattice at the end of the subsequent section.

Under mild conditions on the noise distribution which will be specified below, this can always be arranged. This is the basic observation and the starting point for our approach to the problem stated above. The idea is now to make use of the fact that the global behavior of percolation clusters is qualitatively different depending on whether the percolation probabilities are sub- or supercritical meaning that there is a detectable difference in cluster formation of occupied sites according to whether $\mathcal{G}_0 = \emptyset$, or not.

2.3. The triangular lattice and the significance of $p_c = 1/2$

From the general description of the problem formulated, some immediate questions arise:

1. How do we have to choose the threshold ?
2. How does the test performance depend on the choice of the underlying lattice structure for the discretized picture, i.e. the choice of $\mathcal{G}$?
3. What can we say about the behavior of clusters of non-occupied sites in and outside of $\mathcal{G}_0$?

In particular, the first question is crucial. If we do not know how to properly choose the threshold to achieve super- and subcritical occupation probabilities in- and outside $\mathcal{G}_0$, we have to consider scales of thresholds to determine a proper one. This would at least slow down the detection. However, the second question also indicates that we have some freedom to choose a suitable structure for the underlying lattice. It will turn out that just by choosing $\mathcal{G}$ properly, i.e. as a *triangular lattice*, we obtain a universal answer to question (i) and a quite satisfying result concerning question (iii), too. But let's start from the beginning.

First of all, we will summarize the conditions on the noise distribution which will be valid throughout the entire remainder of the paper.

Noise Properties. For a given graph $\mathcal{G}$, the noise is given by random variables $\{\varepsilon(s) : s \in \mathcal{G}\}$ such that

1. the variables $\varepsilon(s)$ are independent, identically distributed with $E\varepsilon = 0$ and $V\varepsilon = 1$,

2. the noise distribution is *symmetric*,
3. the distribution of the noise is *non-degenerate* with respect to a critical probability p_c meaning that if F denotes the cumulative distribution function of the noise and we define

$$m_c^+ = \inf\{x \in \mathbb{R} : F(x) \geq 1 - p_c\}, \quad m_c^- = \sup\{x \in \mathbb{R} : F(x) \leq 1 - p_c\}$$

then we have $m_c^+ = m_c^-$ where we denote the common value by m, and *either*

$$F(m) > \lim_{h \to 0, h > 0} F(m - h), \tag{5}$$

or

$$F'(m) > 0. \tag{6}$$

The reason to assume symmetry will become clear below. The reason to assume non-degeneracy is the following simple observation.

Lemma 1. *Under the non-degeneracy condition, we can always find a threshold τ such that we have*

$$p_0 > p_c > p_1$$

for the probabilities p_0, p_1 defined in (3). This holds independently of the value $\sigma > 0$ of the signal to noise level.

Proof: By (3), we have $p_0 = 1 - F\left(\frac{\tau - 1}{\sigma}\right)$ and $p_1 = 1 - F\left(\frac{\tau}{\sigma}\right)$. Choosing now $\tau = \sigma m + 1/2$, we obtain

$$\frac{\tau - 1}{\sigma} < m < \frac{\tau}{\sigma}.$$

Under the conditions (5) and (6) above, that implies

$$F\left(\frac{\tau}{\sigma}\right) > 1 - p_c > F\left(\frac{\tau - 1}{\sigma}\right).$$

∎

Secondly, we change the terminology one last time. In the sequel, we call occupied sites *black* and non-occupied sites *white* for obvious reasons. The probability that a pixel (site) is white (not occupied) is given by

$$P(Y_\tau(s) = 0) = \begin{cases} 1 - p_0 = P\left(\epsilon \leq \frac{\tau - 1}{\sigma}\right) & , s \in \mathcal{G}_0 \\ 1 - p_1 = P\left(\epsilon \leq \frac{\tau}{\sigma}\right) & , s \in \mathcal{G}_1 \end{cases}.$$

To address question (iii) above, it would be favorable that the white pixels enjoy the same property as the black ones, just the other way round, namely that their probabilities are supercritical outside the support of the signal and subcritical inside. This will be the case if we can choose τ such that $p_0 > \max\{p_c, 1 - p_c\} \geq \min\{p_c, 1 - p_c\} > p_1$. A situation where this can be done and where, furthermore, we get a value for τ *for free* is provided by the following simple but crucial observation.

Proposition 1. *Let $\mathcal{G}$ be a lattice such that $p_c = 1/2$ and choose $\tau = 1/2$. If the distribution of the noise is* symmetric, *we have for the threshold signal Y_τ:*

1. *The probability that a given pixel s is black is subcritical for $s \in \mathcal{G}_1$ and supercritical for $s \in \mathcal{G}_0$.*
2. *The probability that a given pixel s is white is subcritical for $s \in \mathcal{G}_0$ and supercritical for $s \in \mathcal{G}_1$.*

Proof: Note first that due to the symmetry of the noise distribution, the non-degeneracy conditions (5) and (6) reduce to

$$F'(0) > 0, \quad \lim_{h \to 0, h > 0} F(-h) < F(0).$$

Thus, if $p_c = 1/2$ and the noise is symmetric we have

$$p_0 = P\left(\epsilon > \frac{\tau - 1}{\sigma}\right) = P(\epsilon > -1/2\sigma) = P(\epsilon < 1/2\sigma) > P(\epsilon \leq 0) \geq 1/2 = p_c.$$

Hence we also have $p_1 = P\left(\epsilon > \frac{\tau}{\sigma}\right) = P(\epsilon > 1/2\sigma) \leq 1 - P(\epsilon < 1/2\sigma) = 1 - p_0 < 1 - p_c = 1/2$ and thus

$$p_0 > \max\{p_c, 1 - p_c\} = p_c = \min\{p_c, 1 - p_c\} > p_1.$$

That proves the statement. ∎

Remark. Note that, so far, the considerations are completely non-parametric with respect to the noise. Apart from the condition for expectation and variance, we do not assume anything else about the cumulative distribution function F of the noise except the non-degeneracy condition and symmetry.

We round up the discussion in this section by the statement that there actually is an example for a lattice with $p_c = 1/2$.

Definition 2. *The* (infinitely elongated) planar triangular lattice *is the infinite graph $\mathcal{T} \subset \mathbb{C}$ in the complex plane with* edges (*sites*) *given by the elements of the additive subgroup $(S, +) \subset (\mathbb{C}, +)$ generated by the* sixth roots of unity

$$U := \{\rho \in \mathbb{C} : \rho^6 = 1\}.$$

Two sites $s, s' \in S$ are connected by a bond, *thus they are* neighbors, *if and only if their Euclidean distance in the plane is $d(s, s') = 1$.*

Proposition 2. *The critical percolation probability for the* planar triangular lattice *is given by $p_c = 1/2$.*

Proof: See Kesten (1982), p. 52 f. ∎

Remark. Please note that this theorem holds exclusively for the infinite triangular lattice but that we rely on the assumption that for finite triangular lattices of a *reasonable* size, the regime of logarithmic and linear cluster sizes will be separated quite clearly and the region of probabilities $0 < p < 1$ where the regimes change will be quite sharply located around the *critical percolation probability $p_c = 1/2$* of the infinite triangular lattice.

One particularly remarkable fact is that the basic receptor units (*ommatidia*) of an insect's eye are located at the sites of a triangular lattice. This may be caused by the requirement to have a densest possible packing of the units. However, once this discretization for visual perception is chosen, one may ask whether the analysis of the signals thus obtained use in some way or another the properties of the triangular lattice described above.

3. The Maximum Cluster Test and consistency

In this section, we will construct a test for the *detection problem* given in Definition 1 above and compute explicit upper bounds for the *type I and type II errors* under some mild condition on the shape of $\mathcal{G}_0$, called the *bulk condition*. We will not focus on completeness of the proofs for which we refer to Langovoy and Wittich (2009) but we attempt to make transparent the basic idea of the proof: We use known statements for the infinite triangular lattice $\mathcal{T}$ and transfer them to a finite lattice by arguments using a certain kind of monotonicity. The error bounds tend to zero as the lattice size aprroaches infinity. That actually provides us with a *consistency result*: If the image can be recorded with an unboundedly increasing resolution, the test will almost surely produce the right decision. The precise statement is given in Theorem 1 below which is the main result of this section.

Remark. Please note that consistency here only means that we asymptotically make the right decision. The support of the signal is not necessarily consistently reconstructed by the largest cluster. This is especially apparent when the object of interest consists of several connected components.

The setup is as follows: $\mathcal{T}^{(N)} \subset \mathcal{T}$ denotes the finite triangular lattice consisting of the N^2 sites $s \in \mathcal{T}$ and bonds which are contained in the subset

$$\{z \in \mathbb{C} \,:\, \Re(z) \le N + \frac{1}{2}, \Im(z) \le \frac{\sqrt{3}}{2}N\}.$$

By consistency we mean that the test will deliver the correct decision, if the signal can be detected with an arbitrarily high resolution. To be precise, we think of the signal as a subset $G_0 \subset [0,1]^2$ and write

$$G_0^{(N)} := \{(N+1/2)x + iN\sqrt{3}y/2 \,:\, (x,y) \in G_0\} \subset \mathbb{C}.$$

The model from equation (1) is now depending on N, and given by

$$Y^{(N)}(s) = \mathbf{1}_{\mathcal{G}_0^{(N)}}(s) + \sigma\,\varepsilon(s) \tag{7}$$

where the sites of the subgraph are given by $\mathcal{G}_0^{(N)} = \{s \in \mathcal{T} \,:\, s \in G_0^{(N)}\}$ and the bonds of the subgraph are all bonds in $\mathcal{T}$ that connect two points in $\mathcal{G}_0^{(N)}$.

We apply now the threshold as described above, i.e. we let $\tau = 1/2$ and

$$Y_\tau^{(N)}(s) = \left\{ \begin{array}{ll} 1 & , Y^{(N)}(s) > 1/2 \\ 0 & , Y^{(N)}(s) \le 1/2 \end{array} \right. .$$

We consider the following collection of black pixels

$$\hat{\mathcal{G}}_0^{(N)} := \{s \in \mathcal{T}^{(N)} \,:\, Y_\tau^{(N)}(s) = 1\}. \tag{8}$$

As a side remark, note that one can view $\hat{\mathcal{G}}_0^{(N)}$ as an (inconsistent) pre-estimator of $\mathcal{G}_0^{(N)}$. Now recall that we want to construct a test on the basis of this estimator for the hypotheses $\mathbf{H}_0^{(N)} : \mathcal{G}_0^{(N)} = \emptyset$ against the alternative $\mathbf{H}_1^{(N)} : \mathcal{G}_0^{(N)} \ne \emptyset$.

Definition 3. (The Maximum-Cluster Test) *Let $\phi(N)$ be a suitably chosen threshold depending on N. Let the test statistic T be the size of the largest connected black cluster $C \subset \hat{\mathcal{G}}_0^{(N)}$. We reject $\mathbf{H}_0^{(N)}$ if and only if $T \geq \phi(N)$.*

For this test, we have the following consistency result under the assumption that the support of the indicator function satisfies the following very weak type of a shape constraint.

Definition 4. (The Bulk Condition) *We say that the support $\mathcal{G}_0^{(N)}$ of the signal contains a square of side length $\rho(N) \leq N$ if there is a site $s \in \mathcal{G}_0^{(N)}$ such that $s + \mathcal{T}^{(\rho(N))} \subset \mathcal{G}_0^{(N)}$.*

Remark. If the subset $G_0 \subset [0,1]^2$ contains a square of side length $a > 0$, the respective support $\mathcal{G}_0^{(N)}$ contains squares of side lengths approximately $\rho(N) \approx aN$.

Theorem 1. *For the maximum cluster test, we have*

1. *There is some constant $K_0 > 0$ such that for $\phi(N) = K_0 \log N$, we have for the* type I error
$$\lim_{N \to \infty} \alpha(N) = 0.$$

2. *Let $\phi(N)$ be as above. Let the support $\mathcal{G}_0^{(N)}$ of the signal contain squares of side length $\rho(N)$. If $\rho(N) \geq K_0 \log N$, we have for the* type II error
$$\lim_{N \to \infty} \beta(N) = 0.$$

In particular, in the limit of arbitrary large precision of sampling, the test will always produce the right detection result.

Remark. Please note that the estimator $\hat{\mathcal{G}}_0$ itself is not a consistent estimator for the support of the signal. In the best case, it is an estimator for its largest component. But we will not pursue reconstruction issues at the moment.

By the previous result, we can detect a signal correctly for virtually every noise level if we can choose an arbitrary high resolution. To prove this result, we will have to collect some facts from percolation theory to make precise the intuitive idea explained in the preceding section: *cluster sizes are significantly larger in the supercritical regime.*

We begin with the classical Aizenman-Newman theorem and transfer it step by step to finite lattices.

Proposition 3. (Aizenman-Newman Theorem) *Consider percolation with subcritical probability $p < p_c = 1/2$ on the infinite triangular lattice $\mathcal{T}$. Then there is a constant $\lambda(p) > 0$ depending on p such that*

$$P(|C| \geq n) \leq e^{-n\,\lambda(p)} \tag{9}$$

for all $n \geq 1$ where C denotes the black cluster containing the origin.

Proof: See Kesten (1982). ∎

Note that this result holds for the infinite triangular lattice and that we have to investigate its consequences for the finite lattices $\mathcal{T}^{(N)}$. We do this by means of a classical monotonicity result known as the *FKG inequality*. To state this inequality, we first have to define what we mean by *increasing events*.

Definition 5. *We can introduce a partial ordering on the set* $\Omega = \{0,1\}^{\mathcal{T}}$ *of all* percolation configurations *by*

$$\omega_1 \preceq \omega_2 \quad :\Longleftrightarrow \quad \omega_1(s) \le \omega_2(s) \text{ for all } s \in \mathcal{T}.$$

Now we say that an event $A \subset \Omega$ *is* increasing *if we have for the corresponding indicator variable the inequality*

$$\mathbf{1}_A(\omega_1) \le \mathbf{1}_A(\omega_2)$$

whenever $\omega_1 \preceq \omega_2$. *The term* decreasing event *is defined analogously.*

Now, the FKG inequality reads as follows.

Proposition 4. (FKG inequality) *If A and B are both increasing (or both decreasing) events, then we have*

$$P(A \cap B) \ge P(A)\,P(B).$$

Proof: Fortuin et al. (1971) ∎

We will now apply this result to compute an estimate of the amount of wrong classifications of points caused by the noise for detection of $\mathcal{G}_0^{(N)} \subset \mathcal{T}^{(N)}$ in a finite lattice. To be precise, we consider the event $F^{(N)}(n)$ that among the sites of a configuration $\omega \in \Omega$ that are erroneously marked black on $\mathcal{T}^{(N)}$, i.e.

$$E^{(N)}(\omega) := \{s \in \mathcal{T}^{(N)} - \mathcal{G}_0^{(N)} \,:\, Y_\tau(s) = 1\}$$

and $C^{(N)}(\omega) \subset E^{(N)}(\omega)$ the largest connected cluster, then

$$F^{(N)}(n) := \{\omega \in \Omega \,:\, |C^{(N)}| \ge n\}. \tag{10}$$

Denote now by p_E the *error probability*

$$p_E := P(Y_\tau(s) = 1 \,|\, s \notin \mathcal{G}_0^{(N)}) \tag{11}$$

which only depends on the noise distribution and not on N or the particularly chosen site $s \notin \mathcal{G}_0^{(N)}$. The consequence of the Proposition above for the finite lattice reads now as follows

Proposition 5. *Suppose that $0 < p_E < 1/2$ is subcritical. Let $N \ge \phi(N)$ be a threshold value with*

$$\phi(N) = C \log N^2 \tag{12}$$

and $C\lambda(p_E) > 1$ where $\lambda(p_E) > 0$ is the value for the infinite lattice from (9). Then we have

$$P(F^{(N)}(\phi(N))) = N^{-2(C\lambda(p_E)-1)} + O(N^{-4(C\lambda(p_E)-1)})$$

as N tends to infinity.

Proof: Let $s \in \mathcal{T}^{(N)}$ and $C(s)$ the (possibly empty) largest connected cluster of occupied sites in $\mathcal{T}$ that contains s. The event $\{|C(s)| < n\}$ is obviously *decreasing* for all $s \in \mathcal{T}^{(N)}$. That implies by FKG inequality

$$
\begin{aligned}
P\left(\max_{s \in \mathcal{T}^{(N)}} |C(s)| < n\right) &= P\left(\cap_{s \in \mathcal{T}^{(N)}} \{|C(s)| < n\}\right) \\
&\geq \prod_{s \in \mathcal{T}^{(N)}} P\left(|C(s)| < n\right).
\end{aligned}
$$

By (9) and translation invariance on the infinite lattice, we thus obtain

$$
P\left(\max_{s \in \mathcal{T}^{(N)}} |C(s)| < n\right) \geq \left(1 - e^{-\lambda(p_E)n}\right)^{N^2}.
$$

Hence

$$
P(F^{(N)}(n)) \leq 1 - \left(1 - e^{-\lambda(p_E)n}\right)^{N^2}
$$

Let now $n = 2C \log N = \phi(N)$ with $C\lambda(p_E) > 1$. Then, by (14), we obtain

$$
P(F^{(N)}(\phi(N))) = N^{-2(C\lambda(p_E)-1)} + O(N^{-4(C\lambda(p_E)-1)}).
$$

∎

Definition 6. *Let $\mathcal{T}^{(N)} \subset \mathcal{T}$ be as above. A subset $\pi = \{s_1, ..., s_n\}$ of black sites $s_k \in \mathcal{T}^{(N)}$ with*

  1. s_k and s_{k+1} are neighboring sites for all $k = 1, ..., n-1$,
  2. $0 \leq \Re(s_1) \leq 1/2$,
  3. $N \leq \Re(s_n) \leq N + 1/2$,

is called a left-right crossing.

Proposition 6. *Consider site percolation on $\mathcal{T}^{(N)} \subset \mathcal{T}$ with supercritical percolation probability $p > 1/2$. Denote by A_N the event that there is some left-right crossing in $\mathcal{T}^{(N)}$. Then there is a constant $D(p) > 0$ such that*

$$
P(A_N) \geq 1 - Ne^{-D(p)N}.
$$

Proof: The proof is analogous to the one of the corresponding statement for bond percolation in Grimmett (1999). See Langovoy and Wittich (2009) for more information. ∎

Proof:(of Theorem 1) (i) By Proposition 1, the probability $p = p_E$ for a pixel to be black is subcritical under the null hypothesis. Let $\lambda(p)$ be the exponential factor for the infinite lattice in equation (9). Choosing now $K_0 = 2C$ with $C\lambda(p_E) > 1$ means that by Proposition 5 the type I error probability tends to zero as N tends to infinity.

(ii) Let the alternative be true and $\mathcal{G}_0^{(N)} \neq \emptyset$ contain a square of side length $\rho(N)$. Again by Proposition 1, the probability $p_B > 1/2$ that a site is (correctly) marked black is supercritical for sites inside the square. Therefore, the probability to find a left-right crossing π is given by Proposition 6 to be

$$
P(A_{\rho(N)}) \geq 1 - \rho(N)e^{-D(p_B)\rho(N)}.
$$

But every left-right crossing is a connected cluster of size at least $\rho(N)$. Hence

$$P(T \geq \rho(N)) \geq 1 - \rho(N)e^{-D(p_B)\rho(N)},$$

and together with $\rho(N) \geq K_0 \log N$ this implies that

$$\lim_{N \to \infty} P(T \geq K_0 \log N) \geq \lim_{N \to \infty} P(T \geq \rho(N)) = 1$$

and thus $\lim_{N \to \infty} \beta(N) = 0$. ∎

Finally, as an addition to Proposition 5, we prove that under the same conditions, a slightly changed arguments yields an exponential estimate for the tail probabilities of the cluster size distribution.

Proposition 7. *Suppose that $0 < p_E < 1/2$ is subcritical. Let $N \geq \phi(N)$ be a threshold value with*

$$\phi(N) = C \log N^2$$

and $C\lambda(p_E) > 1$ where $\lambda(p_E) > 0$ is the value for the infinite lattice from (9). Then there exists a constant $K^{(N)} > 0$ such that

$$P(F^{(N)}(n)) \leq e^{-K^{(N)}n}$$

for all $n \geq \phi(N)$.

Proof: Let $s \in \mathcal{T}^{(N)}$ and $C(s)$ the (possibly empty) largest connected cluster of occupied sites in $\mathcal{T}$ that contains s. The event $\{|C(s)| < n\}$ is obviously *decreasing* for all $s \in \mathcal{T}^{(N)}$. That implies by FKG inequality

$$
\begin{aligned}
P\left(\max_{s \in \mathcal{T}^{(N)}} |C(s)| < n\right) &= P\left(\cap_{s \in \mathcal{T}^{(N)}} \{|C(s)| < n\}\right) \\
&\geq \prod_{s \in \mathcal{T}^{(N)}} P\left(|C(s)| < n\right).
\end{aligned}
$$

By (9) and translation invariance on the infinite lattice, we thus obtain

$$P\left(\max_{s \in \mathcal{T}^{(N)}} |C(s)| < n\right) \geq \left(1 - e^{-\lambda(p_E)n}\right)^{N^2}.$$

Hence

$$P(F^{(N)}(n)) \leq 1 - \left(1 - e^{-\lambda(p_E)n}\right)^{N^2}$$

Let now $n \geq 2C \log N$ with $C\lambda(p_E) > 1$. Then, by essentially the same calculation as for (13), we obtain

$$
\begin{aligned}
P(F^{(N)}(n)) &\leq N^2 e^{-\lambda(p_E)n} \left(1 + e^{-\lambda(p_E)n}\right)^{N^2-1} \\
&\leq N^2 \left(1 + N^{-2C\lambda(p_E)}\right)^{N^2-1} e^{-\lambda(p_E)n}.
\end{aligned}
$$

Now, $n \geq C \log N^2$ and thus $\log N^2 \leq n/C$. That implies by $C\lambda > 1$ and

$$\lim_{N \to \infty} \left(1 + N^{-2C\lambda(p_E)}\right)^{N^2-1} = 1$$

that for N large enough, we have

$$\left(1 + N^{-2C\lambda(p_E)}\right)^{N^2-1} \le e^{\kappa n}$$

and therefore

$$\log N^2 + \kappa - \lambda(p_E)n \le \left(\frac{1}{C} - \lambda(p_E)\right)n = -K^{(N)}n$$

with $K^{(N)} := \lambda(p_E) - \kappa - C^{-1} > 0$. ∎

This proposition helps to strengthen Theorem 1 and to derive the actual rates of convergence for both types of testing errors. It is a remarkable fact that both types of errors in our method tend to zero exponentially fast in terms of the size of the object of interest.

Theorem 2. *Suppose assumptions of Theorem 1 are satisfied. Then there are constants $C_1 > 0$, $C_2 > 0$ such that*

1. *The type I error of the maximum cluster test does not exceed*

   $$\alpha(N) \le \exp(-C_2\phi(N))$$

   for all $N > \phi(N)$.

2. *The type II error of the maximum cluster test does not exceed*

   $$\beta(N) \le \exp(-C_1\rho(N))).$$

   for all $N > \rho(N)$.

Proof: Analogously to the proof of Theorem 1, only replacing Proposition 5 by Proposition 7 and using the inequality from Proposition 6 in a slightly modified fashion, just as was done in the proof of Theorem 1 in Langovoy and Wittich (2009). ∎

In other words, the maximum cluster test has power that goes to one exponentially, and the false detection rate that goes to zero exponentially, comparatively with the size of the object of interest.

4. Power, significance level and critical size of clusters

4.1. Type I error and the critical size of clusters

By the consistency result obtained above, we see that the maximum cluster test leads almost surely to the right test decision if we can sample the incoming signal with arbitrary precision. However, as frequently, the asymptotic result will not determine the exact threshold $\phi(N)$ for a given *finite* value of N when we assume some level of significance $\alpha > 0$. Since the significance level is based on the determination of the type one error, we can use the fact the under the null hypothesis $\mathbf{H}_0 : f = 0$, the black clusters on the graph are distributed according to a plain percolation with probability of a site to be black equal to p_E. In that

case, we can use the *Newman - Ziff algorithm* which is described in Wittich and Langovoy (2010b) to effectively simulate the distribution of the size of the largest cluster.

In the sequel, we present the results of these simulations for the corresponding critical regions of the maximum cluster size statistic T for a triangular lattice with $55 \times 55 = 3025$ sites. We consider *significance levels* of $\alpha = 0.05$ and $\alpha = 0.01$ and different possible values for p_E. We obtain the following tables for the lower bounds of the critical regions $\{T \geq c_0\}$ depending on values for p_E.

p_E	0.1	0.2	0.3	0.4	0.42	0.44	0.46	0.48	0.5
c_0	7	19	49	186	262	395	597	891	1184

Boundary of critical regions for $\alpha = 0.05$

p_E	0.1	0.2	0.3	0.4	0.42	0.44	0.46	0.48	0.5
c_0	9	23	62	247	352	524	765	1058	1302

Boundary of critical regions for $\alpha = 0.01$

These values are obtained from the quantiles of the simulated cluster size distributions according to the description given in Wittich and Langovoy (2010b). The site probabilities for which the simulations are run, are $p_E = 0.1, 0.2, 0.3,$ 0.4, 0.42, 0.44, 0.46, 0.48, 0.5, 0.52, 0.54, 0.56, 0.58, 0.6, 0.7, 0.8, and 0.9. They are displayed in the following graphics.
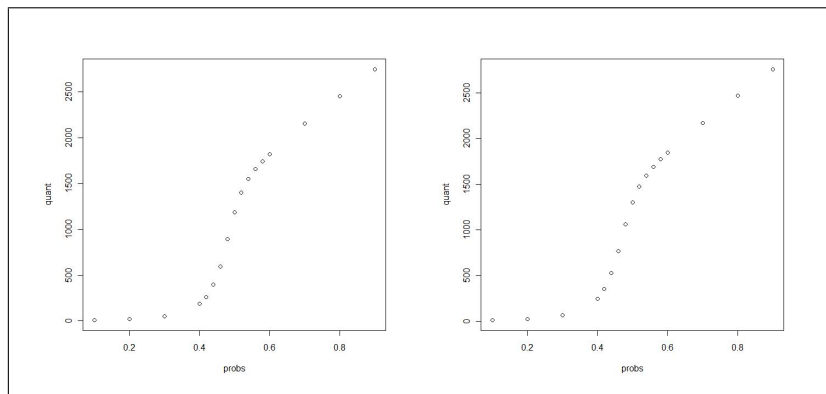

FIG 1. *0.95- and 0.99-quantiles of maximum cluster size as function of p_E*

4.2. Type II error – monotonicity

To compute upper and lower estimates of the type II probability, we will use the bulk condition together with the simple fact that the error probability of the test is monotonous in the support of the signal. To be precise, we have the following statement.

Lemma 2. *Let $\mathcal{G}_0 \subset \mathcal{G}'_0 \subset \mathcal{G}$ be two different supports for signals Y, Y' according to equation (1). Then, for any given level of significance, the type II error for signal Y is larger than for Y'.*

Proof: For every configuration $\omega \in \{0,1\}^{\mathcal{G}}$ after thresholding, we have using the probabilities $p_1 > p_0$ from (4) and the notation $\mathcal{G}_0, \mathcal{G}_1$ from(3)

$$
\begin{aligned}
P(\omega) &= p_0^{|\omega^{-1}(1)\cap\mathcal{G}_0|}(1-p_0)^{|\omega^{-1}(0)\cap\mathcal{G}_0|}p_1^{|\omega^{-1}(1)\cap\mathcal{G}_1|}(1-p_1)^{|\omega^{-1}(0)\cap\mathcal{G}_1|} \\
&= \frac{p_1^{|\omega^{-1}(1)\cap\mathcal{G}_1|-|\omega^{-1}(1)\cap\mathcal{G}'_1|}(1-p_0)^{|\omega^{-1}(0)\cap\mathcal{G}_0|-|\omega^{-1}(0)\cap\mathcal{G}'_0|}}{p_0^{|\omega^{-1}(1)\cap\mathcal{G}'_0|-|\omega^{-1}(1)\cap\mathcal{G}_0|}(1-p_1)^{|\omega^{-1}(0)\cap\mathcal{G}'_1|-|\omega^{-1}(0)\cap\mathcal{G}_0|}} \times \\
&\quad \times p_0^{|\omega^{-1}(1)\cap\mathcal{G}'_0|}(1-p_0)^{|\omega^{-1}(0)\cap\mathcal{G}'_0|}p_1^{|\omega^{-1}(1)\cap\mathcal{G}'_1|}(1-p_1)^{|\omega^{-1}(0)\cap\mathcal{G}'_1|} \\
&\leq p_0^{|\omega^{-1}(1)\cap\mathcal{G}'_0|}(1-p_0)^{|\omega^{-1}(0)\cap\mathcal{G}'_0|}p_1^{|\omega^{-1}(1)\cap\mathcal{G}'_1|}(1-p_1)^{|\omega^{-1}(0)\cap\mathcal{G}'_1|} \\
&=: P'(\omega)
\end{aligned}
$$

by $p_0 > p_1$ and

$$
|\omega^{-1}(1) \cap \mathcal{G}_1| - |\omega^{-1}(1) \cap \mathcal{G}'_1| = |\omega^{-1}(1) \cap \mathcal{G}'_0| - |\omega^{-1}(1) \cap \mathcal{G}_0|
$$

and where P denotes the the probability of the configuration associated to Y and P' the probability associated to Y'. That implies for the probability that the that the signal is not detected $P(T < n_0) \geq P'(T < n_0)$ where n_0 is a cluster size that is determined on the basis of the null hypothesis and the level of significance alone and does therefore not depend on the signal. That implies the statement. ∎

Thus, under the *bulk condition*, we obtain a conservative upper estimate for the type II error, if we simulate the type II error for the square contained in the support of the signal and a lower estimate if we compute it for a square containing the support. In the latter case, we will consider here the best case scenario of a signal that is supported by all of $\mathcal{G}$.

4.3. Type II error – lower bound for small images

In this paragraph we study numerical performance of our procedure in the case of small images. For illustrative purposes, we concentrate on the case $N = 55$, which is important for applications of our method in insect vision (see Wittich and Langovoy (2010a)). To estimate the type II error in this case, we compute the probability that a constant signal of the form $f(s) = a > 0$ (which is therefore positive on the whole graph) is not detected. So we actually consider the alternative $\mathbf{H}_1 : f = a > 0$. The probability that a given pixel is marked black is therefore supercritical given by $p_B > 1/2$. From the simulated distribution of the maximum cluster sizes, we thus obtain for the type II error the subsequent tables by just taking the value of the cumulative distribution function at the boundaries of the respective critical regions.

$p_E \rightarrow$ $p_B \downarrow$	0.1	0.2	0.3	0.4	0.42	0.44	0.46	0.48	0.5
0.52				10^{-6}	0.0005	0.01	0.09	0.31	0.66
0.54				$5\,10^{-9}$	10^{-5}	0.0007	0.01	0.08	0.23
0.56					$2\,10^{-9}$	$6\,10^{-7}$	0.0006	0.01	0.05
0.58							$7\,10^{-6}$	0.0009	0.006
0.6							$8\,10^{-9}$	$4\,10^{-5}$	0.0005
0.7									
0.8									
0.9									

Type II error for $\alpha = 0.05$, all omitted values are less than 10^{-9}

$p_E \rightarrow$ $p_B \downarrow$	0.1	0.2	0.3	0.4	0.42	0.44	0.46	0.48	0.5
0.52				0.0003	0.006	0.05	0.21	0.49	0.83
0.54				10^{-5}	0.0003	0.005	0.04	0.15	0.39
0.56					$2\,10^{-7}$	0.0002	0.004	0.03	0.09
0.58						$2\,10^{-6}$	0.0002	0.003	0.01
0.6							10^{-5}	0.0003	0.0009
0.7									
0.8									
0.9									

Type II error for $\alpha = 0.01$, all omitted values are less than 10^{-9}

Of course, we would ideally prefer to simulate the type II error for signals containing certain squares such as used for the consistency result. To be precise, we would like to consider the case where the real signal is of the form $a\,\mathbf{1}_Q$ where $Q \subset \mathcal{T}^{(N)}$ is a given square and $a > 0$. However, we can not use the Newman - Ziff algorithm directly for that since the site probabilities are inhomogeneous. To find an efficient simulation algorithm for the type II error in this case is therefore work in progress.

On the other hand, the above two tables show that our testing procedure can have very good power already for rather small images. Since it follows from Theorem 2 that both error probabilities tend to zero exponentially as the screen resolution increases, both the power and the level of the Maximum Cluster Test will improve rapidly even with a small growth of the image resolution.

4.4. Type II error – upper bound for small images

To simulate the type II error for a signal supported on a square sublattice, we use the modified version of the Newman - Ziff algorithm described in Wittich and Langovoy (2010b). The simulations are done for a 15×15 square lattice that is basically situated in the center of a 55×55 lattice. We expect that the simulated probabilities do depend on the location of the support but for us the simulated type II errors in this section only serve as a proof of principle. To further explore the type II error will take extensive simulations for different shapes and locations of the support, but we will not pursue that here.

4.5. Simulation of the maximum cluster size distribution

All the experiments described so far concerned the maximum cluster size distribution of a triangular lattice $\mathcal{T}(55)$ with $55 \times 55 = 3025$ sites for different site probabilities of $p_E = 0.1, 0.2, 0.3, 0.4, 0.42, 0.44, 0.46, 0.48, 0.5, 0.52, 0.54, 0.56, 0.58, 0.6, 0.7, 0.8,$ and 0.9. The results of these simulations are presented in Figures 2 and 3 below. We used the R-implementation `simulation` of the Newman - Ziff algorithm described in Wittich and Langovoy (2010b).

5. Detection algorithm

Once a threshold τ is fixed, we will base the test decision (i.e. *we detect a signal or not*) on the size of the maximal black cluster. It is therefore certainly an important point that there is a cluster search algorithm, the *Depth First Search* algorithm from Tarjan (1972) which is explained in details and implemented in R in Wittich and Langovoy (2010b), which is quite effective. That means, it is linear in the number of pixels. This is stated below. Please note that the R-implementation is rather slow compared to an implementation in C.

We now describe the detection algorithm and state the complexity result. The algorithm consists of the following steps:

1. Perform a τ-thresholding of the noisy picture $\hat{Y}$.
2. Run a *depth first search* algorithm on the graph of the thresholded signal until *either* a black cluster of size $|C| \geq c_0$ is found, *or* all black clusters are found.
3. If a black cluster of size $|C| \geq c_0$ was found, report that a signal was detected, otherwise do not reject $\mathbf{H}_0$.

The complexity result follows now from the linear complexity of Tarjan's algorithm (see Tarjan (1972)) and is given by the following statement.

Theorem 3. *The algorithm terminates in $O(N^2)$ steps, i.e. it is linear in the number of pixels.*

Proof: See Langovoy and Wittich (2009), Theorem 1. ∎

6. Application of asymptotic results to small images

Finally, we want to consider one (hopefully) instructive example. We assume that $\varepsilon \sim N(0,1)$ is standard normally distributed and that the true signal is given by the constant $a > 0$. We choose for the test the threshold $\tau > 0$ and obtain

$$p_E(\tau) = P(\varepsilon > \tau/\sigma) = 1 - \Phi(\tau/\sigma) < 1/2.$$

Given the true signal, the correct probability that a site is marked black is given by

$$p_B(\tau, a) = P(\varepsilon > (\tau - a)/\sigma) = \Phi((a - \tau)/\sigma).$$

The first observation is that p_B is supercritical, only if $\tau \in [0, a)$. This may help us to develop an algorithm to find good threshold values in the following way.
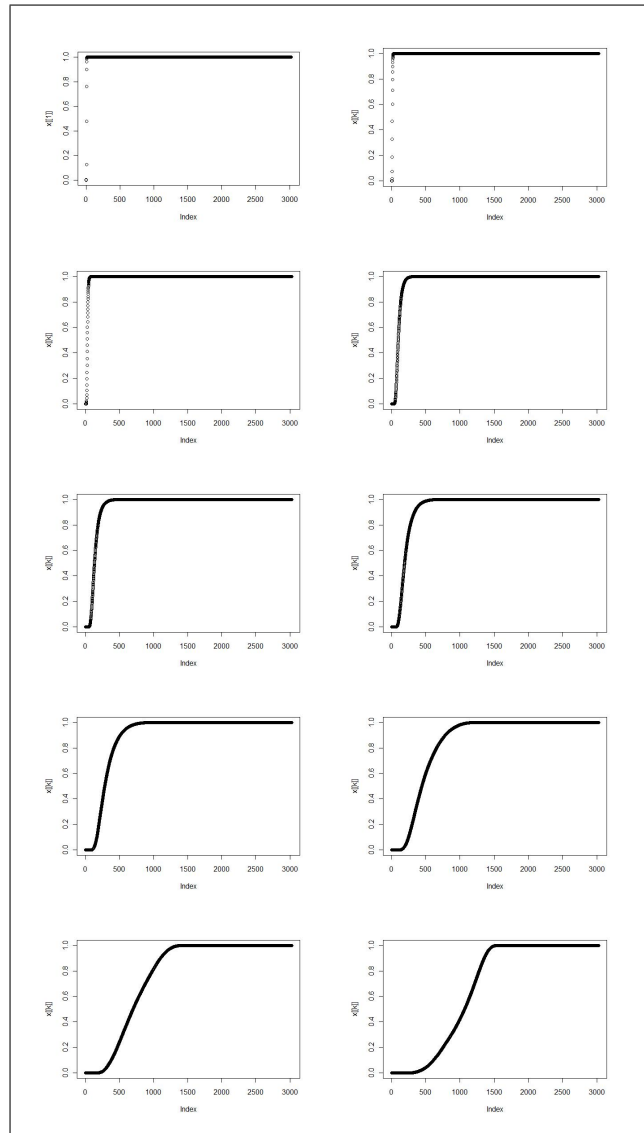
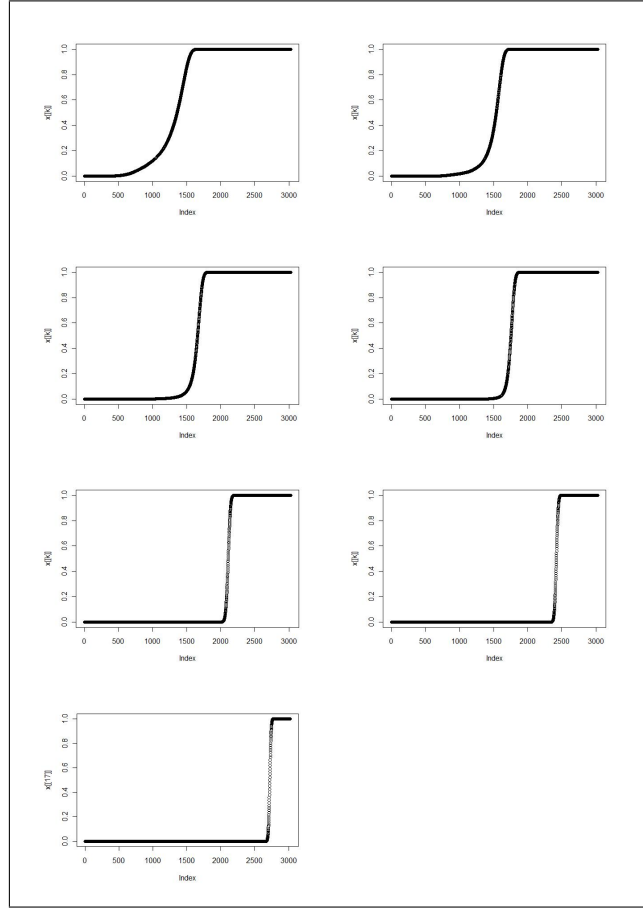FIG 2. *Maximum cluster size distribution for $p_E = 0.1, 0.2, 0.3, 0.4, 0.42, 0.44, 0.46, 0.48, 0.5, 0.52$*

FIG 3. *Maximum cluster size distribution for p_E = 0.54, 0.56, 0.58, 0.6, 0.7, 0.8, 0.9*

1. Start with some threshold value $\tau_1 > 0$ and perform the algorithm described above.
2. If $\mathbf{H}_0$ is not rejected, proceed with a smaller value τ_2.
3. Repeat step 2 until you reject $\mathbf{H}_0$ or $\tau_n < \tau_0$ where τ_0 is some minimal threshold given in advance.

To use such a minimal threshold makes sense due to the uncertainty relation, which in the context of a finite lattice with simulated maximum cluster size distribution corresponds to the fact that the type II error can get large. To be precise, for a given (small) $a > 0$ we have, because supercritical behavior can only be achieved for $\tau < a$, that

$$1/2 > p_E(\tau) > 1 - \Phi(a/\sigma), \ \ 1/2 < p_B(\tau) < \Phi(a/\sigma).$$

Remark. From the first sight, it seems obvious that given a signal strength $a > 0$, the threshold $\tau = a/2$ is optimal in separating the two probabilities p_E and p_B. But due to the asymmetry of the distribution functions around the critical probability $1/2$ visible in Figure 4.1, this may not be the case.

Moreover, this is only the case for certain symmetric lattices and symmetric noise distributions.

However, the inequalities above already yield that

$$1 - \Phi(a/\sigma) < p_E(\tau) < 1/2 < p_B(\tau) < \Phi(a/\sigma)$$

and that means for instance if $\alpha = 0.01$ and the signal to noise ratio is $\rho = a/\sigma = 0.05$, then $p_B \approx 0.52$, $p_E \approx 0.48$, we may infer from the second table about type II errors that $\beta \approx 0.49$. It may therefore not really make sense to consider threshold values below $\tau_0 = 0.05 \times \sigma$.

Acknowledgments. The authors would like to thank Laurie Davies and Remco van der Hofstad for helpful discussions.

References

Ery Arias-Castro, David L. Donoho, and Xiaoming Huo. Near-optimal detection of geometric objects by fast multiscale methods. *IEEE Trans. Inform. Theory*, 51(7):2402–2425, 2005. ISSN 0018-9448.

C. M. Fortuin, P. W. Kasteleyn, and J. Ginibre. Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.*, 22:89–103, 1971. ISSN 0010-3616.

Geoffrey Grimmett. *Percolation*, volume 321 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1999. ISBN 3-540-64902-6.

Harry Kesten. *Percolation theory for mathematicians*, volume 2 of *Progress in Probability and Statistics*. Birkhäuser Boston, Mass., 1982. ISBN 3-7643-3107-0.

M. Langovoy and O. Wittich. *Detection of objects in noisy images and site percolation on square lattices*. EURANDOM Report No. 2009-035. EURANDOM, Eindhoven, 2009.

M. Negri, P. Gamba, G. Lisini, and F. Tupin. Junction-aware extraction and regularization of urban road networks in high-resolution sar images. *Geoscience and Remote Sensing, IEEE Transactions on*, 44(10):2962–2971, Oct. 2006. ISSN 0196-2892. .

Lucia Ricci-Vitiani, Dario G. Lombardi, Emanuela Pilozzi, Mauro Biffoni, Matilde Todaro, Cesare Peschle, and Ruggero De Maria. Identification and expansion of human colon-cancer-initiating cells. *Nature*, 445(7123):111–115, Oct. 2007. ISSN 0028-0836.

Sunil K. Sinha and Paul W. Fieguth. Automated detection of cracks in buried concrete pipe images. *Automation in Construction*, 15(1):58 – 72, 2006. ISSN 0926-5805. .

Robert Tarjan. Depth-first search and linear graph algorithms. *SIAM J. Comput.*, 1(2):146–160, 1972. ISSN 0097-5397.

O. Wittich and M. A. Langovoy. Insect vision - a different image analysis paradigm. *In preparation*, 2010a.

O. Wittich and M. A. Langovoy. Computationally efficient algorithms for statistical image processing. implementation in r. *In preparation*, 2010b.

Appendix.

We will shortly explain how to obtain an asymptotic expression of the binomial sum in the proof of Proposition 5.

Lemma 3. *For $|x| < 1$, $1 \leq n \in \mathbb{N}$, we have*

$$|(1+x)^n - 1 - nx| \leq \frac{n(n-1)}{2}|x|^2(1+|x|)^{n-2}. \tag{13}$$

Proof: We have

$$
\begin{aligned}
|(1+x)^n - 1 - nx| &= \sum_{k=2}^{n} \binom{n}{k} |x|^k \\
&= \frac{n(n-1)}{2}|x|^2 \sum_{k=2}^{n} \frac{\binom{n}{k}}{\binom{n}{2}} |x|^{k-2} \\
&\leq \frac{n(n-1)}{2}|x|^2 \sum_{k=0}^{n-2} \binom{n-2}{k} |x|^k \\
&= \frac{n(n-1)}{2}|x|^2(1+|x|)^{n-2}.
\end{aligned}
$$

$\blacksquare$

Lemma 4. *Let $z_N = e^{-\lambda f(N)}$ and $f(N) = C\log(N^2) = 2C\log(N)$ where $\lambda > 0$ and $\lambda C > 1$. Then, as N tends to infinity, we have*

$$1 - (1-z_N)^{N^2} = N^{2(1-C\lambda)} + O\left(N^{4(1-C\lambda)}\right). \tag{14}$$

Proof: By inequality (13) above

$$|(1-z_N)^{N^2} - 1 + N^2 z_N| \leq \frac{N^2(N^2-1)}{2} z_N^2 (1+|z_N|)^{N^2-2}$$

By $C\lambda > 1$, we obtain

$$\lim_{N\to\infty} (1+|z_N|)^{N^2-2} = \lim_{N\to\infty} \left(1 + \frac{1}{N^{2C\lambda}}\right)^{N^2} = 1.$$

Thus, for $N > N_0$ large enough, there are constants $K^* > K > 1$ such that

$$|(1-z_N)^{N^2} - 1 + N^2 z_N| \leq K\frac{N^2(N^2-1)}{2} z_N^2 \leq K^* N^{4(1-C\lambda)}.$$

Together with

$$N^2 z_N = N^{2(1-C\lambda)},$$

that proves the assertion.

$\blacksquare$