INTENSITY ESTIMATION OF STATIONARY FIBRE PROCESSES FROM DIGITAL IMAGES WITH A LEARNED DETECTOR

PAOLA M.V. RANCOITA[1,2], ALESSANDRO GIUSTI[1] AND ALESSANDRA MICHELETTI[3]

[1]Istituto Dalle Molle di Studi sull'Intelligenza Artificiale, Galleria 2, 6928 Manno-Lugano, Switzerland;
[2]Institute of Oncology Research, via Vela 6, 6500 Bellinzona, Switzerland; [3]Dipartimento di Matematica, Università degli Studi di Milano, via Saldini 50, 20137 Milano, Italy
e-mail: paola@idsia.ch, alessandrog@idsia.ch, alessandra.micheletti@unimi.it

ABSTRACT

Stationary fibre processes are processes of curves in a higher dimensional space, whose distribution is translation invariant. In practical applications, they can be used to model several real objects, such as roots, vascular networks and fibres of materials. Often it is required to compare processes showing similar shape, thus a quantitative approach to describe their stochastic geometry is necessary. One of the basic geometric characteristics of these processes is the intensity (*i.e.*, mean total length per unit area or volume). Here, a general computational-statistical approach is proposed for the estimation of this quantity from digital images of the process, thus only planar fibre processes or projections of processes onto a plane are considered. Differently from approaches based on segmentation, it does not depend on the particular application. The statistical estimator of the intensity is proportional to the number of intersections between the process under study and an independent motion invariant test fibre process. The intersections are detected on the real digital image by a learned detector, easily trained by the user. Under rather mild regularity conditions on the fibre process under study, the method also allows to estimate approximate confidence intervals for the intensity, which is useful especially for comparison purposes.

Keywords: intensity estimator, intersection detector, machine learning, stationary fibre process.

INTRODUCTION

In many fields, such as biomedicine, material sciences, agronomy, remote sensing, structures of interest can be modeled as stationary fibre processes (curves/lines in a higher dimensional space). Here, we focus our attention only on planar fibre processes or on projections of 3D processes onto a plane. The characterization of such processes from digital images is a common task, which is frequently solved first by attempting an automated segmentation of the fibres (which is dependent on the particular application). If a trusted segmentation is available, then an appropriate fibre tracking algorithm can be applied and the process can be characterized by computing several geometric measures. For example, the intensity (*i.e.*, mean total length per unit area) can be estimated as the ratio between the total length of the fibres in the window of observation W (*i.e.*, the image) and the area of W. Unfortunately, in many practical scenarios the required segmentation can not be reliably obtained and often is not even a well-posed problem, for example in cases where fibres appear blurred or very thin. Other intensity estimators are usually based on Crofton-like formulas (see, *e.g.*, Ohser, 1981 and Stoyan *et al.*, 1995), but they have no asymptotic properties because they are obtained by means of the intersection with a finite deterministic test fibre system (such as a finite grid of circles or segments).

To overcome the request of a segmentation for the estimation of the intensity and to ensure good asymptotic properties of the estimation procedure, we propose a computational-statistical technique, which is based instead on a computationally simpler estimator, proportional to the number of intersections with an independent motion invariant (*i.e.*, invariant with respect to translations and rotations) test fibre process. The detector of the intersections is learned incrementally using a random forest classifier trained on few user inputs (*i.e.*, examples of intersections in the image), until the user is satisfied with the detector accuracy. Afterwards, the test fibre process is simulated several times on the image, obtaining an estimate of the intensity of the process under study and an approximation of the variance of the estimator. Thus, we can also provide approximate confidence intervals for the intensity of the process.

Due to its generality, the approach is successfully applied on images of very dissimilar structures, in a variety of application scenarios, avoiding the requirement of an ad-hoc segmentation method for the identification of the fibres.

In the following sections, we describe the

computational-statistical technique we propose for the estimation of the intensity of stationary fibre processes. The explanation of the procedure is divided into two sections: the former regarding the statistical estimator of the intensity (based on the intersection with an independent motion invariant test fibre process) and the estimation of its variance, the latter regarding the learned detector of the intersections from a digital image.

Finally, we present the results obtained on simulated and real data for the evaluation of the performance of the learned detector and the comparison between the methods for the variance estimation.

THE STATISTICAL ESTIMATOR OF THE INTENSITY AND THE ESTIMATION OF ITS VARIANCE

A planar fibre process Φ is a random variable taking values in the space of the systems of fibres in $\mathbb{R}^2$ (Stoyan *et al.*, 1995). A fibre can be defined as a curve of class $\mathscr{C}^1$ with finite length and a system of fibres as the union of at most countably many fibres that can have only the endpoints in common (see Stoyan *et al.* (1995) for a more formal definition and motivations for this choice). Thus, many processes represented by lines/curves in a two-dimensional space (such as, images of capillaries and roots) can be modeled as a planar fibre process. An example of fibre process is the Boolean fibre process,

$$\Phi = \bigcup_n \{X_n \oplus \Gamma_n\} \, , \qquad (1)$$

where $\oplus$ denotes the Minkowski sum (*i.e.*, $A \oplus B = \{a + b | \ a \in A, \ b \in B\}$), $\{X_n\}_n$ is a homogeneous Poisson point process with intensity λ (*i.e.*, mean number of points per unit area) and $\{\Gamma_n\}_n$ is a sequence of i.i.d. random fibres independent of $\{X_n\}_n$.

If the process is stationary (*i.e.*, its distribution is translation invariant), we can define the intensity L_A, as the constant such that $E[\mu_\Phi(B)] = L_A \nu_2(B)$, for all Borel sets B, where $\mu_\Phi(B)$ denotes the total length measure of the fibres of Φ in B (*i.e.*, $\mu_\Phi(B) := \nu_1(\Phi \cap B)$) and ν_d, $d = 1,2$, is the d-dimensional Lebesgue measure. The intensity L_A represents the mean total length of fibres per unit area; for example, in the case of the Boolean processes (Eq. 1) $L_A = \lambda m$, with $m = E[\nu_1(\Gamma_1)]$. As a consequence, the intensity is one of the basic characteristics of the stochastic geometry of a stationary fibre process.

As we mentioned in the Introduction, we can estimate unbiasedly the intensity with

$$\widehat{L}_A^{\text{measure}}(W) = \frac{\mu_\Phi(W)}{\nu_2(W)} \, ,$$

where W is a compact window of observation. If Φ is ergodic, then the estimator is strongly consistent, when enlarging the window of observation (from Corollary 10.2.V in Daley and Vere-Jones, 1998). Nevertheless, the computation of the estimator requires to measure the total length of the fibres in W. In practical applications, the sets under study are represented in a digital image, thus to calculate the total length of the fibres we need a reliable segmentation and an appropriate fibre tracking algorithm. To avoid these requirements, we can use another estimator which is unbiased, computationally simpler and has also good asymptotic properties,

$$\widehat{L}_{A,1}(W) = \frac{N_{\Phi_1 \cap \Phi_2}(W)}{\nu_2(W)} \frac{\pi}{2 L_{A,2}} \, , \qquad (2)$$

where Φ_1 is the fibre process under study, Φ_2 is a test independent motion invariant (*i.e.*, invariant with respect to translations and rotations) fibre process with intensity $L_{A,2}$, and $N_{\Phi_1 \cap \Phi_2}(.)$ is the counting measure associated to the the point process of the intersections $\Phi_1 \cap \Phi_2$, *i.e.*, $N_{\Phi_1 \cap \Phi_2}(W)$ is the number of the intersection points of Φ_1 and Φ_2 in W (Micheletti and Rancoita, 2009). Due to the stationarity and independence of Φ_1 and Φ_2, the point process $\Phi_1 \cap \Phi_2$ is a.s. stationary and, since Φ_2 is also isotropic, its intensity is $2L_{A,1}L_{A,2}/\pi$ (Lemma 3.2 in Mecke, 1981), which implies the unbiasedness of $\widehat{L}_{A,1}$.

$\widehat{L}_{A,1}$ is a generalization of the estimator defined by Ohser (1981), based on the intersection with a finite deterministic test fibre system. If the point process of the intersections $\Phi_1 \cap \Phi_2$ is ergodic, $\widehat{L}_{A,1}$ is strongly consistent when enlarging the window of observation (it follows from Corollary 10.2.V in Daley and Vere-Jones, 1998). If $\Phi_1 \cap \Phi_2$ has suitable mixing properties (such as the independence of the process in disjoint sets, having a distance greater than $l < \infty$), the estimator is also asymptotically normal (Jolivet, 1991; Rancoita, 2010). For example, if the point process of the centers of the fibres is independent in sets with distance greater than l and the fibres have maximum length smaller than l, with $l < \infty$, (like for Boolean fibre processes where the length of the fibres is smaller than $l < \infty$), then the fibre process is independent at distance l. It is easy to verify that if both Φ_1 and Φ_2 have this property, then it holds also for $\Phi_1 \cap \Phi_2$ and all the described asymptotic properties of the estimator hold. In Rancoita (2010) and Rancoita and Micheletti (2011), the authors showed the behavior of the estimator

on both simulated "continuous" and digital images (where the fibres are approximated with pixels), using different types of (even not isotropic) Boolean fibre processes as Φ_1, different isotropic Boolean fibre process as Φ_2 and different sizes of W. On both "continuous" and digital images, the approximation of the distribution of the estimator with the normal distribution was already good in square window of observation with side equal to 200–300 units or pixels. In case of digital images, the authors studied also the issue of over/underestimation of the number of intersections due to the pixel approximation: the eventual over/underestimation depended mainly on the shape of the fibres of the process Φ_1, thus the estimator can always be used for the comparison of the intensity of fibre processes with fibres showing the same shape (since the bias of the estimation is the same).

Under the assumptions needed for the asymptotic normality of $\widehat{L}_{A,1}$, if we can estimate the variance of $\widehat{L}_{A,1}$, we can provide an approximate confidence interval for the intensity. In practical applications, only a few or even one single image of the process under study is usually provided, thus it is not possible to estimate $\text{Var}(\widehat{L}_{A,1})$ with the sample variance of the estimator computed on the set of images. In this situation, we need an accurate way to approximate the variance of the point process of the intersections on a single image.

Heinrich and Prokešová (2010) proved that if $\mathscr{P}$ is a stationary point process, with milder mixing conditions than the ones required for the asymptotic normality of the corresponding counting measure, and if $\{W_n\}_n$ is a sequence of enlarging bounded convex windows of observation, then

$$\lim_{n \to \infty} \frac{\text{Var}(N_{\mathscr{P}}(W_n))}{v_2(W_n)} =: \sigma_{\mathscr{P}}^2 < \infty . \quad (3)$$

It follows that, if W is sufficiently large and $\mathscr{P} = \Phi_1 \cap \Phi_2$,

$$\text{Var}\left(\frac{N_{\mathscr{P}}(W)}{v_2(W)}\right) \approx \frac{\sigma_{\mathscr{P}}^2}{v_2(W)}$$

$$\implies \text{Var}(\widehat{L}_{A,1}(W)) \approx \left(\frac{\pi}{2L_{A,2}}\right)^2 \frac{\sigma_{\Phi_1 \cap \Phi_2}^2}{v_2(W)} . \quad (4)$$

Heinrich and Prokešová also defined a class of kernel estimators for $\sigma_{\mathscr{P}}^2$ and showed their good theoretical asymptotic properties. Choosing a cylinder kernel, the estimator becomes,

$$\hat{\sigma}_{\mathscr{P}}^2(W) = \frac{N_{\mathscr{P}}(W)}{v_2(W)} +$$

$$+ \sum_{x,y \in \mathscr{P} \cap W}^{\neq} \frac{1_{B(0,1)}\left(\frac{y-x}{b_w \sqrt{v_2(W)}}\right)}{v_2(T_{-x}(W) \cap T_{-y}(W))} -$$

$$- \pi(b_w \sqrt{v_2(W)})^2 \frac{N_{\mathscr{P}}(W)(N_{\mathscr{P}}(W) - 1)}{(v_2(W))^2} , \quad (5)$$

where b_W is a bandwidth (whose choice depends on the size of W, for convergence properties) and T is the translation operator (i.e., $T_x(B) = \{x + b \mid b \in B\}$). On simulations with small windows of observation, and different types of point process models, kernels and bandwidths, Heinrich and Prokešová (2010) obtained that the estimator, which achieved the lowest relative mean square error ($E[(\sigma_{\mathscr{P}}^2 - \hat{\sigma}_{\mathscr{P}}^2)^2]/\sigma_{\mathscr{P}}^4$), was the one with the cylinder kernel with a value of $b_w \sqrt{v_2(W)}$ in the interval [1,3]. Therefore, we can derive an estimator for $\text{Var}(\widehat{L}_{A,1})$ from Eqs. 4 and 5,

$$\widehat{\text{Var}}(\widehat{L}_{A,1}(W)) = \left(\frac{\pi}{2L_{A,2}}\right)^2 \frac{\hat{\sigma}_{\Phi_1 \cap \Phi_2}^2(W)}{v_2(W)} . \quad (6)$$

We will refer to this estimator as the Method 1 for the estimation of the variance.

Rancoita and Micheletti (2011) proposed another way for the approximation of the variance. Let us generate n i.i.d. test processes Φ_2 (which are also independent of Φ_1) and call $\{\widehat{L}_i\}_{i=1}^n$ the corresponding estimators $\widehat{L}_{A,1}(W)$, obtained on the same realization of Φ_1. If we assume that, $\forall i \neq j$, $\text{Cov}(\widehat{L}_i, \widehat{L}_j) \leq c_1 \text{Var}(\widehat{L}_i)$ with $c_1 < 1$, then an upper bound for the variance is given by,

$$\text{Var}(\widehat{L}_i) \leq \frac{1}{1-c_1} \frac{1}{n-1} \sum_{j=1}^n E\left[\left(\widehat{L}_j - \overline{\widehat{L}}\right)^2\right] ,$$

with $\overline{\widehat{L}} = \sum_{i=1}^n \widehat{L}_i / n$, and we can approximate the variance with the estimate of its upper bound,

$$\widehat{\text{Var}}(\widehat{L}_i) = \frac{1}{1-c_1} \frac{1}{n-1} \sum_{j=1}^n \left(\widehat{L}_j - \overline{\widehat{L}}\right)^2 . \quad (7)$$

We will call this method for the approximation of the variance Method 2. By using Lemma 2.3 in Mecke (1981), it is easy to prove that,

$$\text{Cov}(\widehat{L}_i, \widehat{L}_j) = \frac{\text{Var}(\mu_{\Phi_1}(W))}{(v_2(W))^2} = \text{Var}(\widehat{L}_A^{\text{measure}}(W)) ,$$

for any $i \neq j$, i.e., the covariance between the estimators computed on the same image depends only on the variance of process Φ_1 (in fact the two processes Φ_2, by which $\widehat{L}_i$ and $\widehat{L}_j$ are computed, are generated independently). Since the covariance among the estimators is equal to the variance of $\widehat{L}_A^{\text{measure}}(W)$, the following proposition and corollary

(see the Appendix for their proofs) allow us to show that the constant $c_1 < 1$ exists at least when the test fibre process Φ_2 is an isotropic Poisson segment process (*i.e.*, a Boolean fibre process with fibres that are segments uniformly oriented). In the applications, we will use this type of test fibre process.

Proposition 1 *Let Φ_1 be a stationary planar fibre process with intensity $L_{A,1}$ and Φ_2 be a test isotropic Poisson segment process with intensity $L_{A,2}$, then*

$$\mathrm{Var}(\widehat{L}_{A,1}(W)) \geq \mathrm{Var}(\widehat{L}_A^{\mathrm{measure}}(W)) + \frac{L_{A,1}\pi}{2L_{A,2}\nu_2(W)} \,,$$

for any compact window of observation W in $\mathbb{R}^2$.

Corollary 1 *Let Φ_1 be a stationary planar fibre process with intensity $L_{A,1}$ and Φ_2 be a test isotropic Poisson segment process with intensity $L_{A,2}$, then there exists at least one constant $c_1 < 1$ such that*

$$\mathrm{Var}(\widehat{L}_A^{\mathrm{measure}}(W)) \leq c_1 \mathrm{Var}(\widehat{L}_{A,1}(W)) \,,$$

for any compact window of observation W in $\mathbb{R}^2$.

In Rancoita (2010) and Rancoita and Micheletti (2011), the constant c_1 was found always strictly lower than one ($c_1 \in [0, 0.6]$), on both simulated "continuous" and digital images, by using different types of (even not isotropic) Boolean fibre processes as Φ_1, different isotropic Boolean fibre process as Φ_2 (including the isotropic Poisson segment process) and different sizes of W. Among the types of processes Φ_1, there were the anisotropic Boolean fibre processes whose fibres were either horizontal segments or arcs of parabola (with horizontal orientation) and thus resembled the real angiogenic processes that we will study in the applications.

In general, if we assume that a condition similar to Eq. 3 holds for Φ_1, *i.e.*,

$$\lim_{n \to \infty} \frac{\mathrm{Var}(\mu_{\Phi_1}(W_n))}{\nu_2(W_n)} =: \sigma_{\Phi_1}^2 < \infty \,, \qquad (8)$$

or even the following milder condition holds,

$$\limsup_{n \to \infty} \frac{\mathrm{Var}(\mu_{\Phi_1}(W_n))}{\nu_2(W_n)} \leq \sigma_{\Phi_1}^2 < \infty \,, \qquad (9)$$

then $\lim_{n \to \infty} \mathrm{Var}(\mu_{\Phi_1}(W_n))/(\nu_2(W_n))^2 = 0$, that is, for large windows, the covariance (and $\mathrm{Var}(\widehat{L}_A^{\mathrm{measure}}(W))$) is close to zero and we can approximate the constant c_1 with zero. This approximation is reasonable, since $\widehat{L}_A^{\mathrm{measure}}$ is already strongly consistent by assuming

only the ergodicity of Φ_1 (which is a milder condition than the ones required for the asymptotic normality of $\widehat{L}_{A,1}$). The hypothesis in Eq. 8 is verified, for example, for Boolean fibre processes whose fibres are boundaries of compact sets with a finite maximum diameter (Theorem 6.1 in Molchanov and Stoyan, 1994), and the hypothesis in Eq. 9 holds for general fibre processes independent at distance l ($l < \infty$) thanks to the following proposition and corollary (see the Appendix for their proofs).

Proposition 2 *Let Φ be a stationary planar fibre process independent at distance $l < \infty$. Then, for any bounded convex set B in $\mathbb{R}^2$,*

$$\frac{\mathrm{Var}(\mu_\Phi(B))}{\nu_2(B)}$$

$$\leq \left(3\frac{\sqrt{\mathrm{Var}(\mu_\Phi(Q_l))}}{l} + 6\sqrt{\mathrm{E}[\mu_\Phi^2(Q_l)]\left(\frac{1}{r(B)l} + \frac{1}{\nu_2(B)}\right)} \right)^2 ,$$

where $r(B)$ denotes the inradius of B and Q_l is any square of side l.

Corollary 2 *Let Φ be a stationary planar fibre process independent at distance $l < \infty$. Then, for any sequence of enlarging bounded convex windows $\{W_n\}_n$ such that $\lim_{n \to \infty} W_n = \mathbb{R}^2$,*

$$\limsup_{n \to \infty} \frac{\mathrm{Var}(\mu_\Phi(W_n))}{\nu_2(W_n)} \leq 9\frac{\mathrm{Var}(\mu_\Phi(Q_l))}{l^2} < \infty \,,$$

where Q_l is any square of side l.

Therefore, as general guideline, if we suppose that the hypotheses for the asymptotic normality of $\widehat{L}_{A,1}$ hold and the window of observation is sufficiently large, then we can reasonably approximate the constant c_1 with zero. If the window is small or if we want to be more conservative, we can estimate the value of c_1 either by simulating processes with characteristics similar to the process under study, or, in case Φ_1 is independent at distance $l < \infty$ and Φ_2 is an isotropic Poisson segment process, by using Proposition 2 together with the technique used to prove Corollary 1. In the applications, we will use $c_1 = 0$, since the side of the squared window will be greater than 100 units/pixels, which can be considered a big size, according to the simulations performed in Rancoita (2010) and Rancoita and Micheletti (2011), where the variance was estimated via replicates of Φ_1, and $c_1 = 0.6$, which was the maximum value obtained in the same simulations. Actually the optimal value for c_1 could be internal to the interval $[0, 0.6]$, thus more values in this interval should be tested. Anyway, both because the aim of this paper is to give general guidelines on the estimation procedure and because of space reason, we will restrict our analysis to the

170

extreme of the interval. We leave to subsequent papers the problem of identifying an optimal value for c_1, depending on the types of processes Φ_1 and Φ_2 and the size of W.

Note that the types of process Φ_1 used in the simulations in Rancoita (2010) and Rancoita and Micheletti (2011) comprise: the same processes that will be used in the simulations performed in this paper, and two non-isotropic processes with geometric characteristics similar to the angiogenic process that we will consider here as real application. Therefore, it is reasonable to assume that c_1 will belong to the interval $[0, 0.6]$ in both our simulated and real data.

THE LEARNED DETECTOR OF THE INTERSECTIONS

In order to apply the approach described above on real images, a system is needed to automatically detect intersections between fibres in Φ_1 (which are visible on the image) and fibres in Φ_2 (named *segments* in the following, since in the applications we will use an isotropic Poisson segment process).

In very simple scenarios, intersections may be detected easily, *e.g.*, as local maxima of image intensity along the segment. However, the majority of real applications requires more sophisticated image analysis approaches, and different types of images can have very different appearances of both the fibres and the background.

We propose a general approach for intersection detection based on a classifier learned from training data. In the current implementation we use a random forest classifier (Breiman, 2001). The classifier is applied for each pixel of the segment and uses the data (called also features) obtained from a neighborhood of the pixel. Namely, we look at a rectangular window with side lengths of 17 and 5 pixels, which is centered at the pixel under consideration and oriented such that the long sides are parallel to the segment (see Fig. 1). Pixel values from the original image are resampled at the real-valued locations, defined by a grid inside the window, by using bilinear interpolation. For dimensionality reduction, the features considered are the first 15 principal components of the resampled raw pixel values in the neighborhood. For each pixel of the segment, the classifier returns the probability to be an intersection.
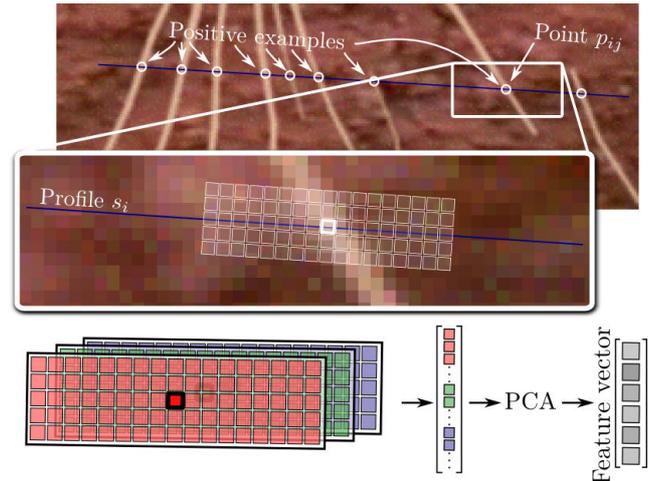

Fig. 1. *Overview of the intersection detector.*

The usage of raw pixel data rather than higher-level features for the classification allows us to remain completely scenario-neutral, since we do not need any assumption about the fibre appearance. The values of the side lengths of the rectangular window were set, in a very conservative way, based on the applications explained in the subsequent sections, but they are not critical parameters (*i.e.*, the use of other reasonable values would not significantly affect the performance of the detector). As general guidelines, the length of the long sides should be greater than the maximum width of the fibres we expect in the image (thus, the rectangular window contains also the boundaries of the fibre around the intersection point), while the value of the length of the short sides is even less crucial. A length of 1 pixel is sufficient in most cases and slightly larger values may improve the performance of the detector in datasets of noisy images.

Training the classifier. The classifier is trained by means of user-labeled segments obtained from a set of training images. Each segment is randomly generated (on a training image) and is shown to the user, who is asked to identify the intersections by clicking on them. All the clicked intersections are considered positive training examples for the classifier; negative examples are generated from the remaining pixels of the segment. The classifier is retrained as new user-labeled segments are available. After each retraining, the classifier performance on segments not belonging to the training set is visualized to the user, who may stop the training phase when accuracy is deemed sufficient. Once trained, the classifier can be applied to detect intersections in new images with a similar appearance to the ones used for training. As any supervised method, this approach works as long as the training set is representative of the testing set; appearance variations are correctly handled, assuming

that some examples for different appearances are available in training data.

Applying the classifier to a test segment. The number of intersections with a test segment is computed by applying the classifier to each pixel of the segment. The classifier returns, for each pixel, the probability of being an intersection and thus, on the whole, a profile of intersection probabilities along the segment: peaks of such profile are considered as candidate intersections with the corresponding probability (see Fig. 2). The candidate intersection points of a given segment are assumed to be conditionally independent Bernoulli random variables. The total number of intersections (with the segment) is estimated as the expected value of the sum of the candidate intersections.
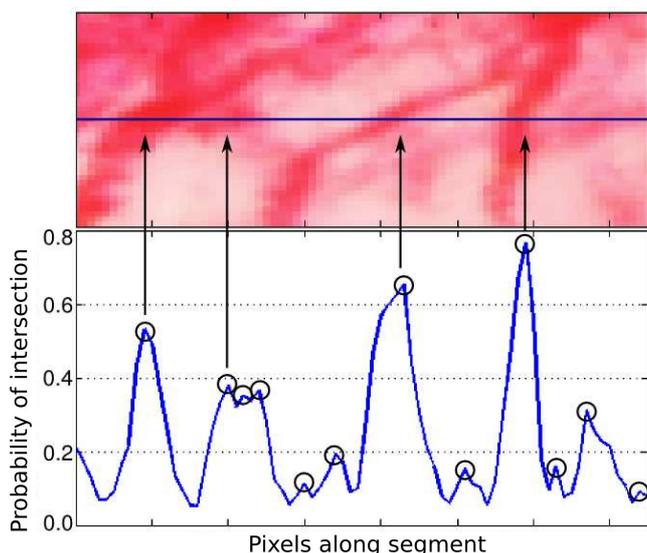

Fig. 2. *Example of profile of probability of being an intersection for all pixels in a segment.*

RESULTS ON SIMULATED AND REAL DATA

This section has three aims: (1) to compare the two methods for the estimation of the variance in order to select one of them, (2) to evaluate the performance of the learned detector, (3) to show a practical application of the proposed approach for intensity estimation. Some artificial data are used to address the first issue, while three real datasets are used for the second one. The practical application consists in a quantitative comparison between the intensities of real angiogenic processes, produced under the effect of different antiangiogenic treatments.

DATA DESCRIPTION

For the comparison of the two previously described methods for the variance estimation in situations similar to real applications, we use artificial datasets of digital images, generated with straightforward computer graphics techniques: namely, the fibres are rasterized as curves with a 1-pixel-wide stroke without antialiasing. In the simulations, both Φ_1 and Φ_2 are Boolean fibre processes (see Eq. 1), with intensity of the Poisson point process λ_1 and λ_2, respectively. The window of observation W is squared and its side is called dimension (or dim). The possible values of the parameters used in the simulations are:

- $\lambda_1 = 0.004$ and $\lambda_2 \in \{0.002, 0.004, 0.008\}$;

- the fibres of Φ_1 have been generated as: either horizontal segments with length $l_1 = 20$ (Poisson horizontal segment process), or segments with uniform orientation and length $l_1 = 20$ (isotropic Poisson segment process) or circles of radius $r_1 = 10$ (Poisson circle process);

- the fibres of Φ_2 are segments with uniform orientation and length $l_2 \in \{20, 60\}$;

- the dimension of W is 500.

For each type of process Φ_1, we generated 100 datasets (*i.e.*, images) and, to compute the variance with Method 2, for each image of Φ_1 and type of process Φ_2, we simulated 100 times the test process. To see the improvement of the estimations when enlarging the window of observation, for each image, we computed $\widehat{L}_{A,1}$ and its variance (with the two methods) also in subwindows of dimension 100, 200, 300 and 400.

For the evaluation of the performance of the learned detector of the intersections, we consider three datasets:

- *Simulated roots*: 49 simulated realistic root images (Dowdy *et al.*, 1998);

- *DRIVE*: 20 training and 20 testing images of retina blood vessels (Staal *et al.*, 2004);

- *Angiogenesis*: 16 images of angiogenic processes of mouse cornea (Corada *et al.*, 2002).

All datasets provide a ground truth (*i.e.*, a reference segmentation of the image made by an expert) for all images and, for the 20 testing images of the dataset *DRIVE*, a second ground truth is also available. The dataset *Angiogenesis* will be also used to show a real application of the intensity estimation. Examples of images of these datasets are given in Fig. 3.
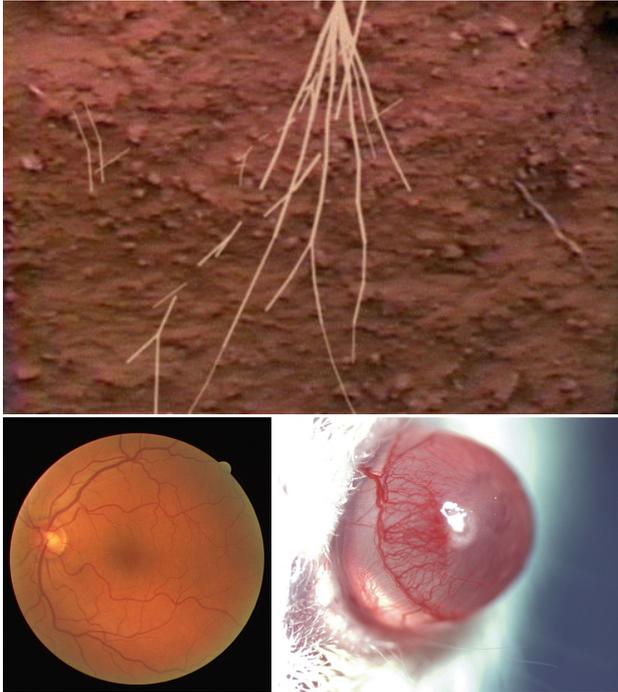
Fig. 3. *Examples of images from the three datasets Simulated roots, DRIVE and Angiogenesis, respectively.*

COMPARISON BETWEEN THE METHODS FOR VARIANCE ESTIMATION

We compared the two methods for the estimation of $\mathrm{Var}(\widehat{L}_{A,1}(W))$ on datasets of simulated digital images (both the methods and the data were described previously). Regarding the estimator described in Eq. 6, since in the experiments in Heinrich and Prokešová (2010) the optimal bandwidth b_W was found such that $d := b_w\sqrt{v_2(W)} \in [1,3]$, we chose b_W such that either $d=2$ or $d=3$. In fact, due to the pixel approximation, we can choose only integer values for d (in Eq. 5 we need to count the pairs of intersections (x,y) such that $|x-y| <= d$) and $d=1$ provides too narrow neighborhood to detect pairs of intersections. We computed the mean square error (MSE) and the relative mean square error (relMSE), as in Heinrich and Prokešová (2010), to compare the quality of the estimation. For the relMSE, we used as reference (true) variance of the estimator the sample variance of the estimator calculated over the 100 independent images of the specific pair (Φ_1, Φ_2). In Fig. 4, we show, as an example, the results obtained when Φ_1 is the Poisson horizontal segment process and Φ_2 has the lowest and highest intensity (among those used in the simulations).
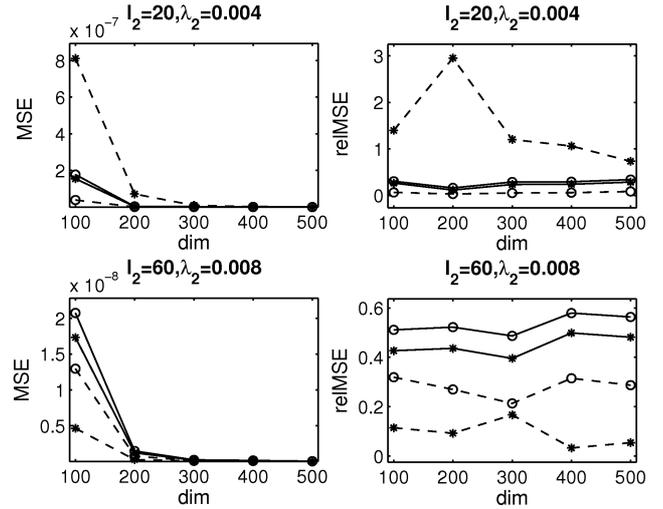

Fig. 4. *Mean square error and relative mean square error of the estimated* $\mathrm{Var}(\widehat{L}_{A,1}(W))$*, where* Φ_1 *is a Poisson horizontal segment process (*$\lambda_1 = 0.004$ *and* $l_1 = 20$*) and* Φ_2 *is a isotropic Poisson segment process. For the estimation we used: Method 1 with* $d = 2$ *(solid-circle line) or* $d = 3$ *(solid-star line) and Method 2 with* $c_1 = 0$ *(dash-circle line) or* $c_1 = 0.6$ *(dash-star line).*

For all methods, fixed a type of Φ_1 and Φ_2, the MSE was decreasing as the dimension increased, since also the variance itself decreases with the dimension for the convergence of the intensity estimator. Moreover, the MSE of the four methods are usually similar, except for the smallest considered side of the window ($dim = 100$). In that case, usually Method 2 with $c_1 = 0$ achieved the lowest error and with $c_1 = 0.6$ the highest. Regarding the relMSE, the error was almost constant as the dimension of W increased, since it is less sensitive to changes in the value of the variance. Furthermore, Method 1 with $d = 3$ achieved only a slightly smaller error than with $d = 2$ and, with both values of d, the error slightly increased with $L_{A,2}$. This effect might be due to the pixel approximation of the image, so that we find a smaller number of pairs of intersection points (x,y) such that $|x-y| <= d$, than what expected. In fact, we obtained almost always an underestimation of the variance and thus the error represents a quantification of this underestimation. Instead, Method 2 with $c_1 = 0.6$ almost always overestimated it, due to its formulation (Eq. 7). Moreover, the error decreased as $L_{A,2}$ increased, because the value of $c_1 = 0.6$ is an approximation obtained heuristically for high values of the intensity of Φ_2. Using $c_1 = 0$, the properties of the estimator of the variance are similar to the ones of Method 1, but with a lower relMSE and a smaller underestimation. In conclusion, when Eq. 8 or Eq. 9 hold, it is better to estimate $\mathrm{Var}(\widehat{L}_{A,1}(W))$

with Method 2 using $c_1 = 0$ (when the window of observation is sufficiently large), otherwise it is better to use a test process with high intensity $L_{A,2}$ and (over)estimate the variance with Method 2, using $c_1 = 0.6$ (for processes with characteristics similar to ours) or trying to estimate it, since in this case Method 1 leads to an underestimation of the variance, and, for comparison purposes, it is preferable to be more conservative and obtain larger approximated confidence intervals.

VALIDATION OF THE LEARNED DETECTOR

We evaluated the accuracy of the learned detector by verifying that the number of intersections computed for each segment estimates well the reference number of intersections, which is computed from the ground truth segmentation. For the validation, we used the images of the three real datasets described previously. For each dataset, we generated 1000 random segments and, for each segment, we computed the number of intersections from the ground truth and, using the detector, from the corresponding real image. We observed that the p-value of the test on the Pearson's correlation coefficient was always lower than 10^{-15} and the slope in the linear regression model was always positive and lower than one. Thus, even in challenging datasets, the detector found a number of intersections well-correlated to the one in the ground truth, with a slight expected underestimation, as tangent intersections might be ignored by the detector but counted in the ground truth. In Fig. 5, we can also observe that the relative absolute error is similar in all datasets, with a median close to 0.2 even in the more difficult case. Instead, for example, we would have obtained a greater performance variability among the datasets, if simply we had considered as intersections all peaks with probability greater than 0.5. In the DRIVE dataset, which is regarded as a standard benchmark for vessel segmentation, the detector performance was comparable to state of the art segmentation methods (designed specifically for such dataset), in the task of finding the number of fibre-segment intersections.

Beyond the good estimation of the number of intersections, we also qualitatively verified that the detected intersections were at the expected positions. Fig. 6 demonstrates this by showing that the spatial distribution of the intersections detected in many segments (white-background image labeled *Det. Inters.*) matches the actual position of the capillaries.
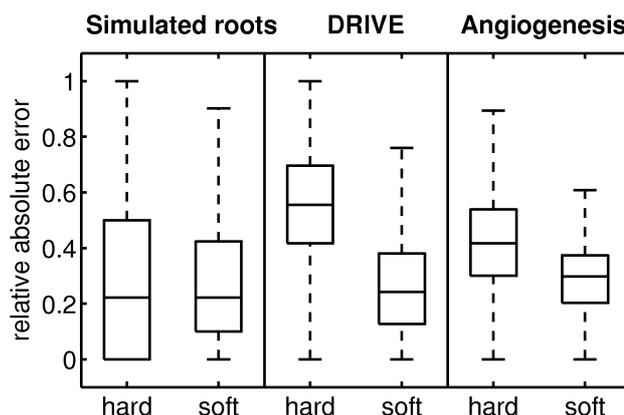

Fig. 5. *Boxplot of the relative absolute error between the number of intersections in the ground truth and the corresponding computed either with our method (here called soft) or by identifying the intersections as peaks with probability greater than 0.5 (hard).*
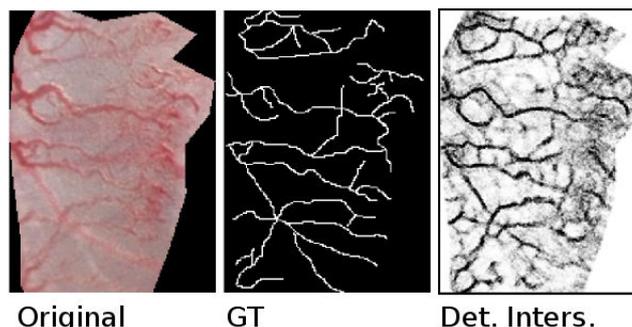

Fig. 6. *Original image (Original), ground truth segmentation (GT) and spatial distribution of detected intersections (Det. Inters.) for an image in the dataset Angiogenesis.*

REAL APPLICATION OF THE INTENSITY ESTIMATION

We show an application of the intensity estimation using the real dataset *Angiogenesis*. The images were produced at IFOM (FIRC Institute of Molecular Oncology Foundation, Milan), by the research group of Prof. Dejana, during a research project regarding the inhibition of angiogenic processes (*i.e.*, formation of new capillary blood vessels) for cancer therapy. They considered several antibodies against the protein VE-Cadherin and studied their ability of inhibiting the angiogenesis on mouse corneas. In each experiment, first they implanted a pellet containing hrFGF-2

(which induces the formation of capillaries and thus simulates a tumor) in the cornea of a mouse and then they treated the mouse with an antibody either putting it in the pellet together with the angiogenic factor (non-systemic treatment) or injecting it intraperitoneally starting from the day after the pellet implantation (systemic treatment). The four mice used as control were treated (two of them non-systemically and the others systemically) with the antibody nonimmune rat IgG (Rat-IgG), which has no antiangiogenic effect. For each type of treatment, two experiments were executed (*i.e.*, they used the treatment on two mice) and thus two eye images were available. The evaluation of the "performance" of the antibody was done by the biologists, by comparing the images of the specific antibody with the control ones of Rat-IgG.

In Fig. 7 we show a qualitative comparison between the estimation of the intensity and of the variance (with Method 2 and both $c_1 = 0$ and $c_1 = 0.6$), using the ground truth and the real image. All estimates are nicely correlated, with a slight underestimation in the real image (due to the difficulties encountered by the detector). Moreover, we can see that differences between the estimates of the variance are lower by using $c_1 = 0$ than $c_1 = 0.6$.
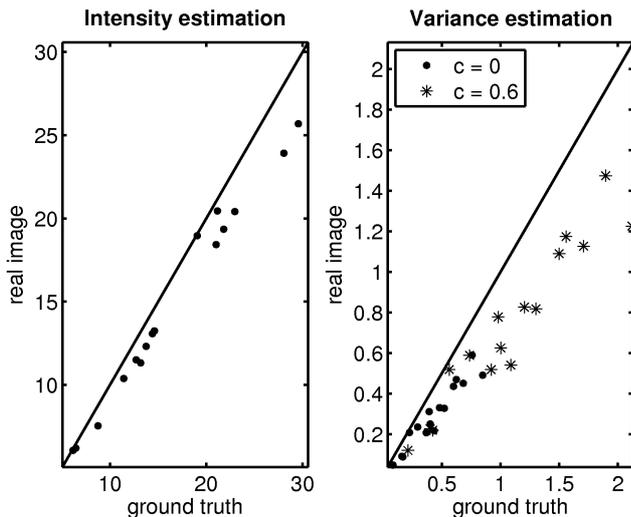

Fig. 7. *Comparison of the intensity and variance estimation by using the ground truth and the real image in the dataset Angiogenesis.*
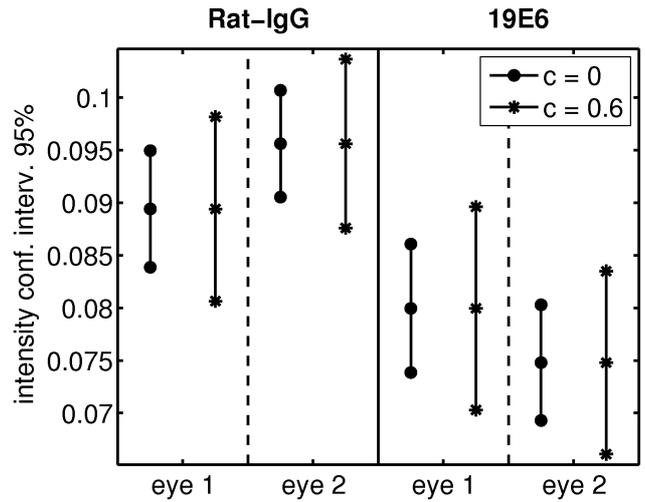

Fig. 8. *Confidence interval of the intensities of the mice non-systemically treated with antibodies Rat-IgG and 19E6.*
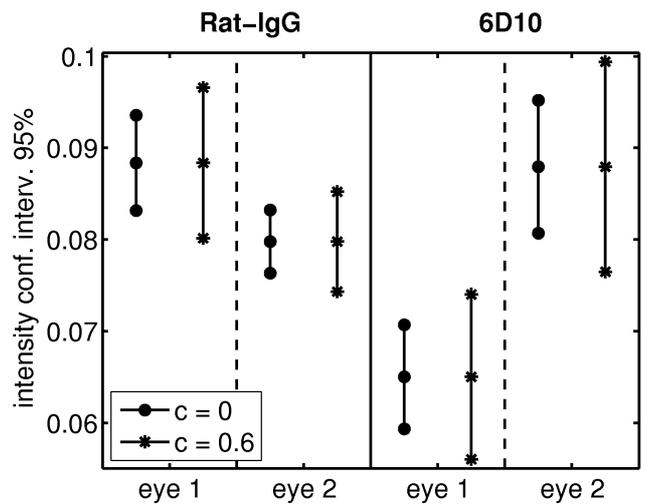

Fig. 9. *Confidence interval of the intensities of the mice systemically treated with antibodies Rat-IgG and 6D10.*

To evaluate the ability of the antibodies in inhibiting the angiogenic process, we computed the approximate confidence interval of the intensity for each image of eye, separately, and compared them with the corresponding ones obtained on the control eyes. The variance was estimated with both $c_1 = 0$ and $c_1 = 0.6$. Figs. 8 and 9 show two examples of comparison. In the former, we can observe that in the eyes of mice treated with antibody 19E6 the intensity looks significantly lower than in the control

eyes. In the latter, we cannot say that the intensity corresponding to the antibody 6D10 is significantly different from the one in the control eyes. In fact, in this case the treatment was systemic, that is the antibody was given the day after the implantation of the pellet, when the network of vessel was already partially formed, and the major differences are in the area occupied by the vessels and in their width, more than in the intensity. Moreover in Fig. 9 we can also note the high biological variability of the two replicates corresponding to antibody 6D10, which may suggest that systemic treatments give less reliable effects with respect to the non-systemic ones.

DISCUSSION

We propose a computational-statistical technique for the estimation of the intensity of stationary planar fibre processes from their digital images. The method is based on a statistical estimator ($\widehat{L}_{A,1}$) proportional to the number of intersections between the process under study Φ_1 and an independent motion invariant test fibre process (simulated on the real digital image of Φ_1). The points of intersection in the image are identified by a learned detector (based on a classifier), which is trained on few user inputs (examples of intersection and non-intersection points). The main advantage of such approach is that it can work on any suitable dataset after a simple and quick training by the user. Moreover, the resulting detector works well also in hard datasets, such as *DRIVE* and *Angiogenesis*, where a clear segmentation is not easy to determine even visually. Uncertainty and ambiguity in training data is well-handled by exploiting the probabilistic information returned by the detector, rather than forcing binary outputs. Both the statistical and the computational components of the method have good characteristics so that it performs well even on real challenging data.

The intensity estimator $\widehat{L}_{A,1}$ is asymptotically normal, under suitable mixing conditions (*e.g.*, the independence of the point process of the intersections in set separated by a distance greater than l, $l < \infty$) and, in the applications, these assumptions are often satisfied. Due to the asymptotic normality, the method can also provide an approximate confidence interval for the intensity, once computed the variance of the estimator. In case only one image of the process is available, we considered two possible ways for the approximation of the variance and, based on the results with simulated data, we identified which of them is more suitable for applications with digital images (where the curves are approximated with pixels). Thus, the approximate confidence interval can be calculated even in case of one single image of the process.

In real applications, also other geometric characteristics can be of interest, such as the distribution of the orientation of the fibres and their width (by considering the objects under study as two-dimensional random sets, instead of random fibres, which are one-dimensional). We leave to subsequent papers the study of estimators of these quantities and the extension of the detector to provide also the information required for their computation. Furthermore, the performance of the detector can be improved by implementing an active learning technique for its training. In this case, the segments shown to the user are chosen automatically in order to provide the largest amount of useful information to the classifier. Finally, we intend to release soon an open source code of our method in Matlab.

APPENDIX

In this section we provide the proofs of all propositions and corollaries stated in the article.

Proof of Proposition 1. As in Weiss and Nagel (1994), let us assume that the segments of Φ_2 are numbered and, for any $x \in \Phi_1 \cap \Phi_2$, $n(x)$ denotes the (a.s. unique) number of the segment of Φ_2 to which x belongs. Thus, we can rewrite the second moment of $N_{\Phi_1 \cap \Phi_2}(.)$ as,

$$
\begin{aligned}
E[N_{\Phi_1 \cap \Phi_2}^2(W)] = &\, E\left[\sum_{x \in \Phi_1 \cap \Phi_2} \mathbb{I}_W(x)\right] \\
&+ E\left[\sum_{\substack{x,y \in \Phi_1 \cap \Phi_2: \\ n(x) \neq n(y)}} \mathbb{I}_W(x)\mathbb{I}_W(y)\right] \\
&+ E\left[\sum_{\substack{x,y \in \Phi_1 \cap \Phi_2: \\ n(x) = n(y)}} \mathbb{I}_W(x)\mathbb{I}_W(y)\right],
\end{aligned}
$$

where $\mathbb{I}_W(.)$ is the indicator function of the set W. The first term is the first moment of $N_{\Phi_1 \cap \Phi_2}(.)$ and thus it is equal to $E[N_{\Phi_1 \cap \Phi_2}(W)] = 2L_{A,1}L_{A,2}\nu_2(W)/\pi = 2L_{A,2}E[\mu_{\Phi_1}(W)]/\pi$ (see Lemma 3.2 in Mecke, 1981). The second term can be explicitly computed by using Theorem 4.1 in Weiss and Nagel (1994) and it is equal to $E[\mu_{\Phi_1}^2(W)](2L_{A,2}/\pi)^2$. Finally, since the third term is greater or equal than zero, we obtain that

$$
E[N_{\Phi_1 \cap \Phi_2}^2(W)] \geq E[\mu_{\Phi_1}^2(W)]\left(\frac{2L_{A,2}}{\pi}\right)^2 + \frac{2L_{A,1}L_{A,2}\nu_2(W)}{\pi},
$$

that is,

$$\mathrm{Var}(N_{\Phi_1 \cap \Phi_2}(W)) \geq \mathrm{Var}(\mu_{\Phi_1}(W))\left(\frac{2L_{A,2}}{\pi}\right)^2$$
$$+ \frac{2L_{A,1}L_{A,2}\nu_2(W)}{\pi},$$

because $\mathrm{E}[N_{\Phi_1 \cap \Phi_2}(W)] = 2L_{A,2}\mathrm{E}[\mu_{\Phi_1}(W)]/\pi$. The thesis can now be derived, by using the definition of the estimators $\widehat{L}_{A,1}$ and $\widehat{L}_A^{\mathrm{measure}}$.

Proof of Corollary 1. From Proposition 1,

$$\mathrm{Var}(\widehat{L}_{A,1}(W)) \geq \mathrm{Var}(\widehat{L}_A^{\mathrm{measure}}(W)) + \frac{L_{A,1}\pi}{2L_{A,2}\nu_2(W)},$$

for any compact window of observation W in $\mathbb{R}^2$. Thus, for any constant $c > 0$ such that

$$c \leq \frac{L_{A,1}\pi}{2L_{A,2}\nu_2(W)\mathrm{Var}(\widehat{L}_A^{\mathrm{measure}}(W))},$$

we have that

$$\mathrm{Var}(\widehat{L}_A^{\mathrm{measure}}(W)) \leq \frac{1}{1+c}\mathrm{Var}(\widehat{L}_{A,1}(W)).$$

By setting $c_1 = 1/(1+c)$, we obtain the thesis and $c_1 < 1$ since $c > 0$.

Proof of Proposition 2. Let B be a bounded convex set in $\mathbb{R}^2$, then we can define a finite grid of squares of side l that covers B and can have only sides or vertices in common, or, more in general, intersections with null ν_2-measure: $\mathcal{Q} = \{Q_l^i : Q_l^i \cap B \neq \emptyset, Q_l^i \cap Q_l^j = A \text{ with } \nu_2(A) = 0, \forall i \neq j\}$, where Q_l^i denotes a square of side l. We can divide the collection of squares into two groups whether they lie completely inside B or not. Let us define the corresponding sets of indices as $I_1 = \{i : Q_l^i \in \mathcal{Q}, Q_l^i \subseteq B\}$ (with cardinality $|I_1| = N_1$) and $I_2 = \{i : Q_l^i \in \mathcal{Q}, Q_l^i \nsubseteq B\}$ (with $|I_2| = N_2$). Moreover, for each index $i \in I_k$ ($k = 1, 2$), let us define the set of indices in I_k corresponding to squares adjacent to Q_l^i, i.e., $H_i^k = \{j \in I_k : j \neq i, d(Q_l^i, Q_l^j) = 0\}$ (where $d(.,.)$ is the Euclidean distance between sets). Obviously, $|H_i^k| \leq 8$, for all i, k.

Now, we can decompose $\mu_\Phi(B)$ as $\sum_{i \in I_1} \mu_\Phi(Q_l^i) + \sum_{i \in I_2} \mu_\Phi(Q_l^i \cap B) =: X + Y$ and it is easy to verify that

$$\mathrm{Var}(\mu_\Phi(B)) \leq \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}, \tag{10}$$

by a suitable majorization of $\mathrm{Cov}(X, Y)$. In order to obtain an upper bound of the right-hand side of Eq. 10, firstly let us derive an upper bound for $\mathrm{E}[X^2]$,

$$\mathrm{E}[X^2] = \sum_{i \in I_1} \mathrm{E}[\mu_\Phi^2(Q_l^i)] + \sum_{i \in I_1} \sum_{j \in H_i^1} \mathrm{E}[\mu_\Phi(Q_l^i)\mu_\Phi(Q_l^j)]$$

$$+ \sum_{i \in I_1} \sum_{j \in I_1 \setminus \{H_i^1 \cup i\}} \mathrm{E}[\mu_\Phi(Q_l^i)]\mathrm{E}[\mu_\Phi(Q_l^j)]$$
$$\leq N_1 \mathrm{E}[\mu_\Phi^2(Q_l^1)] + \sum_{i \in I_1} |H_i^1| \mathrm{E}[\mu_\Phi^2(Q_l^1)]$$
$$+ \sum_{i \in I_1} (N_1 - 1 - |H_i^1|)\mathrm{E}[\mu_\Phi(Q_l^1)]^2$$
$$\leq 9N_1 \mathrm{Var}(\mu_\Phi(Q_l^1)) + (\mathrm{E}[X])^2,$$

by using the stationarity of the fibre process, the independence of the process at distance l, the Cauchy-Schwarz inequality and the inequality $|H_i^k| \leq 8$. Therefore, $\mathrm{Var}(X) \leq 9N_1 \mathrm{Var}(\mu_\Phi(Q_l^1))$. Regarding $\mathrm{E}[Y^2]$, by using a similar procedure we obtain,

$$\mathrm{E}[Y^2] = \sum_{i \in I_2} \mathrm{E}[\mu_\Phi^2(Q_l^i \cap B)]$$

$$+ \sum_{i \in I_2} \sum_{j \in H_i^2} \mathrm{E}[\mu_\Phi(Q_l^i \cap B)\mu_\Phi(Q_l^j \cap B)]$$
$$+ \sum_{i \in I_2} \sum_{j \in I_2 \setminus \{H_i^2 \cup i\}} \mathrm{E}[\mu_\Phi(Q_l^i \cap B)]\mathrm{E}[\mu_\Phi(Q_l^j \cap B)]$$
$$\leq 9N_2 \mathrm{E}[\mu_\Phi^2(Q_l^1)] + (\mathrm{E}[Y])^2,$$

since $\mu_\Phi^2(Q_l^i \cap B) \leq \mu_\Phi^2(Q_l^i)$ a.s., for any i, and $\mu_\Phi(.) \geq 0$ a.s.. Thus, $\mathrm{Var}(Y) \leq 9N_2 \mathrm{E}[\mu_\Phi^2(Q_l^1)]$. By applying all these results to Eq. 10,

$$\frac{\mathrm{Var}(\mu_\Phi(B))}{\nu_2(B)} \leq 9\left(\sqrt{\frac{N_1 \mathrm{Var}(\mu_\Phi(Q_l^1))}{\nu_2(B)}} + \sqrt{\frac{N_2 \mathrm{E}[\mu_\Phi^2(Q_l^1)]}{\nu_2(B)}}\right)^2. \tag{11}$$

In order to properly majorize the ratios $N_1/\nu_2(B)$ and $N_2/\nu_2(B)$, let us consider some properties of convex sets. Let us define: ∂B as the boundary of B, $p(B)$ as its perimeter, $d(B)$ as its diameter, $r(B)$ as its inradius and $Q_s(x)$ as a square of side s centered at x. Because of the definition of the sets of indices I_1, we have that $\bigcup_{i \in I_1} Q_l^i \subseteq B$, i.e., $N_1 l^2 \leq \nu_2(B)$. Moreover, since B can be always inscribed in a square of side $N_B l := (\lceil d(B)/l \rceil + 1)l$ (where $\lceil x \rceil = \min\{k \in \mathbb{Z} : x \leq k\}$, $x \in \mathbb{R}$), our grid of squares cannot be greater than a squared grid consisting of $N_B \times N_B$ squares of side l. Therefore, the number of squares that cover ∂B cannot exceed $4(N_B - 1)$ (see Corollary 2 in Lassak , 1988), and thus

$$N_2 \leq 4(N_B - 1) \qquad \leq 4\left(\frac{d(B)}{l} + 1\right)$$
$$< 4\left(\frac{p(B)}{2l} + 1\right) \leq 4\left(\frac{\nu_2(B)}{r(B)l} + 1\right),$$

(since $2d(B) < p(B)$ and $p(B) \leq 2\nu_2(B)/r(B)$ from Bonnesen and Fenchel, 1934). These facts together with Eq. 11 prove the thesis.

Proof of Corollary 2. From Proposition 2, for any bounded convex set W_n of the sequence $\{W_n\}_n$,

$$\frac{\text{Var}(\mu_\Phi(W_n))}{\nu_2(W_n)}$$

$$\leq \left(3 \frac{\sqrt{\text{Var}(\mu_\Phi(Q_l))}}{l} + 6 \sqrt{\text{E}[\mu_\Phi^2(Q_l)] \left(\frac{1}{r(W_n)l} + \frac{1}{\nu_2(W_n)} \right)} \right)^2,$$

Since $\lim_{n\to\infty} W_n = \mathbb{R}^2$ for hypothesis, then $\lim_{n\to\infty} r(W_n) = +\infty$ and $\lim_{n\to\infty} \nu_2(W_n) = +\infty$, and we obtain the thesis.

REFERENCES

Bonnesen T, Fenchel W (1934). Theorie der konvexen Körper. Berlin: Springer-Verlag.

Breiman L (2001). Random forests. Mach Learn 45:5–32.

Corada M, Zanetta L, Orsenigo F, Breviario F, Lampugnani MG, Bernasconi S, Liao F, Hicklin DJ, Bohlen P, Dejana E (2002). A monoclonal antibody to vascular endothelial-cadherin inhibits tumor angiogenesis without side effects on endothelial permeability. Blood 100:905–11.

Daley DJ, Vere-Jones D (1998). An introduction to the theory of point processes. New York: Springer-Verlag.

Dowdy RH, Smucker AJM, Dolan1 MS, Ferguson JC (1998). Automated image analyses for separating plant roots from soil debris elutrated from soil cores. Plant Soil 200:91–4. http://rootimage.msu.edu/.

Heinrich L, Prokešová M (2010). On estimating the asymptotic variance of stationary point processes. Methodol Comput Appl Probab 12:451–71.

Jolivet E (1991). Moment estimation for stationary point processes in $\mathbf{R}^d$. In: Spatial Statistics and Imaging, Lecture Notes–Monograph Series, Vol. 20. Hayward, Calif.: Institute of Mathematical Statistics, 138–49.

Lassak M (1988). Covering the boundary of a convex set by tiles. Proc Am Math Soc 104:269–72.

Mecke J (1981). Formulas for stationary planar fibre processes III - Intersections with fibre systems. Math Operationsforsch Stat Ser Stat 12:201–10.

Micheletti A, Rancoita PMV (2009). Estimators of the intensity of stationary planar fibre processes. In: Capasso V, Micheletti A, Aletti G, eds. Stereology and Image Analysis. ECS10 – Proc 10th Eur Congr ISS. The MIRIAM Project Series. Bologna: Esculapio. 131–6.

Molchanov I, Stoyan D (1994). Asymptotic properties of estimators for parameters of the Boolean model. Adv Appl Probab 26:301–23.

Ohser J (1981). A remark on the estimation of the rose directions of fibre processes. Math Operationsforsch Stat Ser Stat 12:581–5.

Rancoita PMV (2010). Stochastic methods in cancer research. Applications to genomics and angiogenesis. PhD Thesis, Università degli Studi di Milano.

Rancoita PMV, Micheletti A (2011). Estimators of the intensity of fibre processes and applications. In: Progress in Industrial Mathematics at ECMI 2010. Heidelberg: Springer, in press.

Staal J, Abràmoff MD, Niemeijer M, Viergever MA, van Ginneken B (2004). Ridge based vessel segmentation in color images of the retina. IEEE Trans Med Imaging 23:501–9.

Stoyan D, Kendall WS, Mecke J (1995). Stochastic geometry and its applications. Second Edition. Chichester: John Wiley & Sons.

Weiss V, Nagel W (1994). Second-order stereology for planar fibre processes. Adv Appl Probab 26:906–18.