Statistical modelling of the geometry of prostatic capillaries on the basis of stationary Strauss hard-core processes

Torsten Mattfeldt[1], **Stefanie Eckel**[2], **Frank Fleischer**[3], **Volker Schmidt**[2]

[1]Department of Pathology, University of Ulm
[2]Institute of Stochastics, University of Ulm
[3]Institute of Applied Information Processing & Department of Stochastics, University of Ulm

Keywords. Gibbs process, Markov chain Monte Carlo methods, Metropolis-Hastings algorithm, pair correlation function, point process, simulation, spatial statistics, stereology, Strauss hard-core process.

Running title: Statistical modelling of capillary geometry

Address for correspondence: Prof. Dr. T. Mattfeldt, Department of Pathology, Oberer Eselsberg M23, D-89081 Ulm, Germany.
E-mail: torsten.mattfeldt@medizin.uni-ulm.de. Fax: ++49 731 5002 3884.

Summary

In a recent study, the capillarization of normal prostatic tissue and prostatic carcinoma tissue was characterized by means of explorative methods of spatial statistics. In the present paper, an attempt was made to go beyond the explorative approach and to characterize the point process of the capillary profiles on sections by means of a parametric model. For this purpose, the flexible class of Gibbs processes was considered. Specifically, stationary Strauss hard-core processes were fitted to the observed point process data. The goodness of fit achieved by the model was checked by simulations with the Markov chain Monte Carlo method using the Metropolis-Hastings algorithm. Model fitting and simulations were performed with the help of the *spatstat* package under R. The observed point patterns were in some cases compatible with realizations of stationary Strauss hard-core processes for all ranges of spatial interaction. However, deviations from the model were found for one or more domains of ranges in other cases. In the tumour tissue, a highly significant decrease of the interaction parameter of the Strauss hard-core process could be found as compared to the normal prostatic tissue. This finding is discussed in terms of a loss of the normal lobular architecture of the glands in the tumour tissue.

1. Introduction

Capillaries are the smallest blood vessels whose wall consists only of endothelial cells and a basement membrane. Geometrically, capillaries may be considered as fibre processes, i.e. random processes of thin tubular, thread-like structures in space. In microscopical applications, planar point processes emerge from sections with fibres. Planar point process statistics, such as the intensity and various explorative summary statistics, are informative for the purposes of quantitation and statistical comparison in 2D (Mattfeldt, 2005). In a recent paper, it could be shown that an additional stereological interpretation of second-order statistics of the observed planar point process of the profiles is possible (Krasnoperov & Stoyan, 2004). For this purpose, the ordinary planar pair correlation function of the observable point process may be used as an estimator of the reduced pair correlation function of the fibre process in 3D space (Krasnoperov & Stoyan, 2004). The observable point pattern of the capillary profiles allows an inference of the true geometry of the fibre process, which lives in 3D space, under mild and realistic model assumptions. This is not automatically true for all observable microscopical point processes, e.g. it does not hold for general particles such as cells or cell nuclei of arbitrary shape (Mattfeldt, 2005).

In a subsequent paper, it was explored how this approach may be extended to explorative statistical analysis from multiple specimens of different groups and with replicated observations (Mattfeldt et al., 2006). Specifically, the capillarization of 12 specimens of normal prostatic tissue was compared with the capillarization of 12 specimens of prostatic carcinomas. Significant differences between the mean pair correlation functions of the capillary profile point processes of the normal and tumorous cases could be demonstrated for certain domains of interaction distances. In the present paper, we attempted to go beyond this explorative approach and performed a parametric modelling of the capillary point process data of the aforementioned paper. This approach was chosen from two reasons: (i) It was intended to obtain a better insight into the nature of the geometrical changes that occur in the capillaries of the tumour tissue as compared to the normal tissue. The estimated model parameters have intuitive meanings, which make an understanding of the findings easier. (ii) An explorative second-order statistics, such as the pair correlation function, reduces the complexity of the original sets of point patterns, but still remains rather high dimensional (Mattfeldt et al., 2006). Parametric modelling leads to a much stronger condensation of the informa-

tion to a few model parameters. In the case of a stationary Strauss hard-core process, every image may be characterized by only four quantities: the intensity and three model parameters. Thus, the second purpose of modelling was a drastic dimensional reduction of the information, while still preserving as much as possible of the information given in the data.

2. Materials and Methods

2.1. Brief review of some basic concepts of planar point processes

When dealing with planar point processes, it is important in the first instance to obtain clarity on two fundamental properties of the process, namely, whether they are isotropic and stationary (Stoyan & Stoyan, 1994; Stoyan et al., 1995; Diggle, 2003). Isotropy means that the distribution of the point process is invariant under rotations. If this is not the case, the point process is called anisotropic. Stationarity means that the distribution of the point process is invariant under translations. Otherwise the point process must be considered as nonstationary. Stationarity implies that the intensity, i.e. the expected number of points per unit area of the point process is constant for an arbitrarily chosen sampling window. Nonstationarity allows for deterministic local variation of the intensity within the (x, y)-plane. For example, gradients along the x- and y-axis may occur (Hahn et al., 1999; Fleischer et al., 2006).

With some experience, a first qualified guess whether a point process is isotropic and stationary is possible on the basis of a visual inspection of the patterns. Anisotropy may be inferred when the points are arranged preferentially along certain directions, whereas other directions are systematically underrepresented. Nonstationarity is obvious in drastical cases, if most or all points are concentrated in certain regions of the window, whereas other regions are only sparsely occupied (Hahn et al., 1999). For our data sets, visual inspection suggested isotropy and stationarity, as preferential directions and gradients were not obvious in the point patterns (see Mattfeldt et al., 2006, Fig. 1–3). Nevertheless, it may be difficult to distinguish visually between nonstationarity in the strict sense and random clustering, which can occur in stationary Gibbs processes and may look very similar (see e.g. Diggle, 2003).

As a null model for isotropic and stationary point processes, the model of a stationary Poisson point process is usually considered. For our data it was obvious that the observed point processes were not compatible with a planar Poisson point process (Mattfeldt et al., 2006). A hard-core property was evident, emerging from the fact that the capillaries have a positive diameter.

Moreover, a repulsive pattern was found for some short distances, whereas clustering of the points was observed at longer distances. These observations suggested to use the flexible class of *Gibbs processes (Markov point processes)* for the modelling (Stoyan & Stoyan, 1994; Stoyan et al., 1995; van Lieshout, 2002; Diggle, 2003; Møller & Waagepetersen, 2004; Baddeley & Turner, 2000). Gibbs processes are models for point processes with interaction. In the present context, we decided to use the *stationary Strauss hard-core process* as a candidate model, because it takes into account a hard-core property and allows repulsion as well as clustering of the points at different domains of r-values, depending on the model parameters (see Takacs & Fiksel, 1986; Goulard et al., 1996; Baddeley & Turner, 2000 and references therein).

The classical *Strauss model* and the *Strauss hard-core model* are both examples of Gibbs processes. These processes may be defined in terms of their pair potential or of their probability density (Takacs & Fiksel, 1986; Goulard et al., 1996; Baddeley & Turner, 2000). The probability density of the stationary Strauss hard-core process in a sampling window W (e.g. of unit area) is given by

$$f(\mathbf{x}|(r_0, R, \beta, \gamma)) = \alpha\beta^{n(\mathbf{x})}\gamma^{s_R(\mathbf{x})}\mathbf{1}_{\{\forall\{\mathbf{x_i},\mathbf{x_j}\}\subseteq\mathbf{x},\, \mathbf{x_i}\neq\mathbf{x_j}:|\mathbf{x_i}-\mathbf{x_j}|>\mathbf{r_0}\}}, \qquad (1)$$

where $n(x)$ is the number of points in the pattern, $s_R(x)$ is the number of distinct unordered pairs of points that have a distance to each other that is less or equal to R. Note that α is a (usually unknown) normalizing constant and that the Strauss hard-core process can be completely defined by its four model paramenters r_0, R, β, and γ, describing the hard-core distance, the interaction range, the intensity, and the strength of interaction, respectively. The meaning of the indicator function $\mathbf{1}$ in eq. (1) is that it becomes zero if the point configuration $\mathbf{x}$ contains at least one point pair $\{x_i, x_j\} \subseteq \mathbf{x}$ with a distance less or equal to r_0, otherwise the indicator function equals one. For point pairs of a distance between r_0 and R we can observe that, if $\gamma > 1$ we have an attraction effect, whereas for $\gamma < 1$ there is a repulsion tendency. If $\gamma = 1$ a classical hard-core process is obtained. For point pair distances larger than R, there is no more pairwise interaction. Note that, contrary to the case of a classical Strauss process, for Strauss hard-core processes the interaction parameter γ can assume any nonnegative value, in particular a value larger than 1.

One of the most popular functions of explorative spatial point pattern analysis is Ripley's K-*function* $K(r)$ (reduced second moment function) (Ripley, 1988; Stoyan et al., 1995). Intuitively, $K(r)$ is the mean ($= E =$ expected) number

of other points of the process lying within a circle of radius r, centred at a typical point (x, y) of the process, divided by the intensity λ of the process:

$$K(r) = \frac{E(\text{number of other points with distance} \leq r \ |\text{point at } (x, y))}{\lambda}, \quad (2)$$

where the symbol '|' denotes 'conditional to'. In analogy to a probability density function, which is the derivative of a cumulative distribution function, there is a counterpart to the K-function, namely *the pair correlation function* $g(r)$, which may be obtained after differentiation of $K(r)$:

$$g(r) = \frac{1}{2\pi r} \frac{dK(r)}{dr}. \quad (3)$$

It may also be defined as the product density of the point process, divided by the square of the intensity for the purpose of normalization (Stoyan & Stoyan, 1994; Stoyan et al., 1995; Mattfeldt et al., 2006). In the case of a planar Poisson point process, we have

$$g(r) = 1 \quad (4)$$

for all r. Values below 1 indicate repulsion, values above 1 indicate clustering effects for point pairs of such a distance. A hard-core-effect leads to an initial segment with zero values of $g(r)$.

2.2. Parametric modelling of empirical data sets

A vast amount of literature on methods for the statistics of planar point process data is available, including methods for explorative statistical analysis and for parametric fitting and simulation (Geyer & Møller, 1994; Stoyan & Stoyan, 1994; Stoyan et al., 1995; Møller, 1999; van Lieshout, 2002; Diggle, 2003; Møller & Waagepetersen, 2003, 2004; Baddeley & Turner, 2000). For practical modelling, the software package *spatstat* (Baddeley & Turner, 2005, 2006) was used with R 2.2.0 under Linux. The capillary profile midpoint coordinates of all 48 images (see below) were read into computer, and the fitting and simulations mentioned below were performed for each individual image. The *spatstat* package allows the modelling of a given planar point process in the presence of trend (nonstationarity) as well as modelling in the stationary case.

2.2.1. Fitting of the interaction component

Having decided to use the stationary Strauss hard-core process as a candidate model, it follows from eq. (1) that, given a fixed number of points, three model parameters have to be estimated: the hard-core distance r_0, the interaction radius R and the interaction parameter γ (see Takacs & Fiksel, 1986; Goulard et al., 1996; Baddeley & Turner, 2000). The parameters r_0 and R are denoted as irregular, i.e. they must be estimated in separate worksteps, whereas the regular parameter γ is found automatically during the fitting procedure itself (Baddeley & Turner, 2005, 2006). The hard-core distance r_0 was estimated for each visual field as the minimum value of the interpoint distances (Mattfeldt et al., 2006). The interaction radius R was estimated according to the profile pseudolikelihood method (Baddeley & Turner, 2005, 2006). This procedure was used to find the value of R between 20 and 100 pixels in steps of 1 pixel with the maximum pseudolikelihood for a given image. For the quadrature scheme, a 30 × 30 point grid plus 4 corner points was applied, and edge correction was performed by translation. The estimated value of R was then used together with that of r_0 for the subsequent model fitting, which ultimately yielded an estimate of γ (Baddeley & Turner, 2005, 2006).

2.2.2. Trend analysis

A trend analysis was performed by fitting harmonic polynomials of degree 2 to each image. This part of the analysis yielded 4 trend coefficients a_1–a_4 for the harmonic polynomials and an intercept value a_0. Essentially, the trend function models gradients of the intensity within the (x, y)-plane. The trend function can be visualized as a function $z = f(x, y)$, where hills and valleys correspond to regions with higher and lower intensities:

$$\lambda(x, y) = \exp\,(a_0 + a_1 x + a_2 y + a_3 xy + a_4(x^2 - y^2)). \qquad (5)$$

The coefficients were fitted to the data using *spatstat* according to the principle of least squares. An example is shown as perspective plot in Fig. 1.

2.3. Simulation studies

2.3.1. Model validation

The aim of this part of the study was a validation of our model of a stationary Strauss hard-core process on the basis of simulations. The null hypothesis was that the original image was compatible with a realization of a stationary Strauss hard-core model. After fitting of the model parameters as described in section 2.2.1, 999 simulations were performed per image, as described in

section 2.3.2. Each simulation yielded an image with a realization of a point process. From each simulated point process, the g-function was computed using *spatstat,* in which eq. (3) is used as estimator; edge correction was performed by translation, and an Epanechnikov kernel with a bandwidth of $0.1/\sqrt{\lambda}$ was used for smoothing (Krasnoperov & Stoyan, 2004; Mattfeldt et al., 2006). For each r-value, this yielded altogether 1000 g-values, including the g-value from the true sample. The data analysis was then performed in a global and in a local approach (see Schladitz et al, 2003; Mattfeldt et al., 2006). For the global approach, the 999 areas between the simulated g-functions and the mean g-function of the simulations, and the single area between the g-function estimated from the true image and the expected g-function of the simulations, were computed. Then these 1000 areas per image were sorted by size, and it was looked which rank was occupied by the area corresponding to the true image. If the rank of this area was ≥ 951, it was considered as significantly different from the expectation under the model at a level of 5%. (Note that in this case, the very small ranks are not significant, but indicate very little deviations from the model. Hence it would be wrong in this context to consider both the ranks 1–25 and 976–1000 as significant). In order to find out in more detail where the deviations occurred in terms of the interaction distances, the 1000 g-values were sorted by size for each r-value (i.e. locally), and it was determined which rank was occupied by the g-value of the true sample. A significant deviation was noted for those values of r where the g-value of the true sample occupied a rank between 1–25 or 976–1000.

2.3.2. Details of the algorithm used for the simulations

In this section some basic principles of the algorithm used for our simulations are outlined briefly and informally. In general, only very simple point processes can be simulated by classical Monte Carlo methods, where the simulated point process is obtained 'at once' after drawing a set of independent random numbers; this holds e.g. for simulation of the stationary Poisson point process. To simulate more complex point processes, in particular of Gibbs-type, it is usually necessary to use Markov chain Monte Carlo (MCMC) methods. Here the simulation is performed in a large number of dependent iterations, where it is hoped that the distribution of the simulated point processes converges stepwise to the distribution specified in the model. This holds for all kinds of Gibbs processes, whose simulation thus always implies the application of computer-intensive methods (Mattfeldt & Fleischer, 2006).

For details on MCMC methods for the simulation of planar point processes, the reader is referred to the special literature (Geyer & Møller, 1994; Møller & Waagepetersen, 2003, 2004; Baddeley & Turner, 2005, 2006).

In the present study, the *Metropolis-Hastings algorithm* as implemented within *spatstat* was used for the simulations (Chib & Greenberg, 1995; Baddeley & Turner, 2005, 2006). Essentially, simulations of Gibbs processes can be performed in two different manners: either the number of points is kept fixed, or it is a random variable (Geyer & Møller, 1994; Møller, 1999; Møller & Waagepetersen, 2003, 2004; Baddeley & Turner, 2005, 2006). In this study, conditional simulation was chosen, i.e. the number of points was kept fixed and identical to the number of points in the original image. The starting pattern consisted of a stationary Poisson process with the same number of points as the original image in a rectangular window with the same edges. Each iteration of the algorithm begins with a proposal, which principally means that a new point may be added, a point may be deleted, or a point of the pattern may be shifted at random. Due to the fact that the total number of points was chosen to be constant, shifts always were selected with probability 1, hence births or deaths of points did not occur. After the proposal has been made, an update of the point process follows. If the proposal is accepted, a new point pattern (where a single point has been moved) results; otherwise, the *status quo* persists. This is decided according to a Bernoulli experiment with a certain success probability, which is freshly performed at each iteration. Note that this method leads always to simulated point processes, in which the desired intensity and the desired hard-core distance are guaranteed.

The final aim of this MCMC method is convergence to a point process with a target density which is the same as the density of the required model, here: a stationary Strauss hard-core model with the specified parameters. When using MCMC methods, it is important to check whether the simulations converge fast enough to the target density (Geyer & Møller, 1994; Gelman, 1996; Møller, 1999; Møller & Waagepetersen, 2003, 2004; Baddeley & Turner, 2005, 2006). For a similar but different algorithm, Ripley suggested to use ≈ 100 $\times$ the number of points as the number of iterations n_{rep}, which would mean only about ≈ 3500 iterations for our data (Ripley, 1977). On the other hand, $n_{\text{rep}} = 10000$ is the suggested standard value in the *spatstat* handbook (Baddeley & Turner, 2005, 2006). After some pilot experiments with 5000–1000000 iterations, it was judged that 100000 iterations (steps) of the algorithm should lead to sufficient stability. Hence this number of iterations was used in all the

subsequent simulations. The increase in computation time needed for simulating processes of some hundred observed points with 100000 iterations instead of 10000 was not dramatic on a dual processor Linux workstation. The convergence of the algorithm was checked graphically, i.e. by visual comparison of simulated point processes with the original pattern, and by visual comparison of g-functions of the simulations with the g-function of the original pattern (see Fig. 2).

2.3.3. The stationarity assumption

In the aforementioned fitting procedure of a Strauss hard-core model, we used the tacit assumption of stationarity. As mentioned in section 2.1, this assumption seemed plausible for the capillary data, but ought to be checked. For this purpose the following idea was put into practice. In principle, Strauss hard-core processes are not restricted to the assumption of stationarity. For example, the *spatstat* package allows to fit models of stationary and non-stationary Strauss hard-core processes to empirical planar point processes (Baddeley & Turner, 2005, 2006). If the point process were not stationary, the goodness of fit obtained by the nonstationary variant of the point process should be distinctly better, than the fit obtained by the stationary variant. In order to check this, both model variants were fitted to all images, and for both variants the appropriate parameters were estimated. These were the intensity plus three parameters in the case of the stationary variant, as shown in detail above, and the intensity plus 8 parameters for the nonstationary variant, i.e. the 3 parameters r_0, R and γ as for the stationary variant, plus the 4 trend coefficients a_1–a_4 and the intercept value a_0 described in section 2.2.2. On the basis of these sets of parameters, 999 simulations per model were performed for each image per variant. For each simulated image the g-function was computed using *spatstat.* Thus, each series of simulations led to 999 g-values per interaction distance r, which were locally averaged per r-value. The resulting series of locally averaged values of $g(r)$ was considered as expectations under the model. They had to be computed by simulation, as analytical expressions are not known for $g(r)$ in the case of our model. Finally, squared residuals were computed r-wise between these mean g-values and the g-values from the real images. The mean of the squared residuals was computed for all r-values from 1–200. After averaging over the two images, this yielded two mean square residuals per case, corresponding to the two models with stationarity and nonstationarity. As there were 2 groups, this setting allowed to perform 2 tests on paired comparisons, each with 12 pairs

of values. For this purpose a paired t-test was used. Furthermore, the root of the mean square residual per case was computed. It can be seen roughly as an indicator of the mean absolute deviation of the g-values of the true sample from the mean values of the 999 simulations.

2.4. Practical methods

The practical methods with respect to case selection, preparation, staining, microscopy and image analysis have been presented in detail in (Mattfeldt et al., 2006). Briefly, twelve routine cases of prostatic cancer were compared to a control group of 12 cases with normal prostatic tissue. Paraffin sections were stained using a routine immunhistochemical marker for endothelial cells (CD34). The technically best two images of these series were selected. This approach provided two rectangular visual fields per case with 1240 $\times$ 1000 pixels, in which 61–341 capillary profiles could be found per field. The centres of the capillary profiles were detected interactively using light microscopy. The coordinates of these points were stored. The resulting data set, which was investigated in an explorative approach in (Mattfeldt et al., 2006), was now reanalyzed with the purpose of parametric modelling.

3. Results

3.1. Model parameters of the groups

The mean values of the model parameters of the two case groups are shown in Table I. The mean intensity of the capillary profile process, Q_A, was increased in the prostatic cancer group by 47% as compared to tumour-free tissue (probability value $p < 0.01$) as reported in (Mattfeldt et al., 2006). The mean interaction parameter γ showed a highly significant difference between the normal and the carcinoma group. In the normal group it amounted to 1.792 ($SD = 0.982$), whereas in the carcinoma group it was 0.968 ($SD = 0.288$) ($p < 0.025$). There were no significant differences w.r.t. the hard-core distance r_0 and the interaction radius R.

3.2. Model validation by simulation

Let us consider now the results of the simulations of the stationary Strauss hard-core process. Using the global approach as described in section 2.3.1., it was found that 34/48 images showed no significant deviation from an appropriately parametrized Strauss-hard core model in terms of the g-function. However, in 5/24 images from tumour-free tissue and in 9/24 images from tumour tissue, the area between the observed g-function and the expected

g-function exceeded 95% of the corresponding areas of the model simulations, hence these 14 images were considered as significantly different from the model. Three types of outcomes could be discerned for the 48 images in the local approach: (i) all values of $g(r)$ for the original image fell within the 95% central region of the simulations; (ii) one contiguous series of values of $g(r)$ fell outside of it, or (iii) multiple contiguous series $g(r)$ fell outside of it. Results are shown for selected cases in Fig. 3a–d. It never happened that only isolated points of $g(r)$ lay outside of it. The roots of the mean squared residuals of the g-values amounted to 0.148 in the tumour group, and to 0.185 in the normal group.

3.3. Check for stationarity

The mean values of the squared residuals between the g-values of the empirical patterns and the mean g-values of the simulations are given for the stationary and the nonstationary model for the groups of the normal cases and the tumour cases (Table II). The mean squared residuals were nearly the same when the nonstationary or the stationary model was used for the fitting to the point patterns of the normal cases. For this group, a paired t-test between the two mean residuals per case was not significant. In the tumour group, the fitting with the nonstationary model led to a slight reduction of the mean squared residual per case in the order of magnitude of 18%. This improvement was significant in a paired t-test ($p < 0.02$). Nevertheless, this improvement is only marginal, when one considers that the nonstationary model requires the fitting of 5 additional parameters as compared to the stationary model. Hence, we concluded that the consideration of nonstationarity leads to an excessive increase of dimensionality without a relevant gain in the goodness of fit, and it was decided to stick to the simpler stationary model.

4. Discussion

4.1. General aspects of modelling by Strauss hard-core processes

In the present study, point processes of profiles of blood capillaries of the prostate gland were parametrically modelled on the basis of stationary Strauss hard-core processes. The goodness of fit was tested by simulations. These showed that an important second order statistics of the images, the values of $g(r)$, either fell into the central region occupied by 95% of the simulations, or deviated from it for one or multiple short contiguous series of values (Fig. 3a–d). This finding held for the normal specimens as well as for the tumour tissue. It can be assumed that in the cases with contiguous deviations, the

true g-function of the process is systematically different from the stationary Strauss hard-core model in these regions. It must be admitted that the series of the points lying outside the central region of the simulations, as well as the well discernible oscillating behaviour of some g-functions, is not sufficiently explained by the model. The intensity, the hard-core distance and the general course of the g-function with flattening to $g \approx 1$ at longer distances was, however, generally well reflected. From our point of view the Strauss hard-core process yields not an absolutely perfect fit, but may be considered as a realistic approximation for practical purposes, which seems often to be superimposed by unknown other processes. This conclusion is supported by the results of the global approach to the analysis of the simulation data, which showed that $34/48 \approx 71\%$ images showed no significant deviation from an appropriately parametrized Strauss-hard core model in terms of the g-function.

The stationary Strauss hard-core process model, which is specified by the intensity and three additional parameters, is more flexible and realistic than the standard model of a stationary Poisson point process, which is completely specified by its intensity and allows no hard-core property. In histological applications the latter is however needed. The reason is, that in practically all applications in the biomedical field, one is not confronted with mathematically ideal points, but rather with dots of a positive size, which are regarded as points only for the sake of simplicity. These dots are sections of tubular or particulate structures, which cannot overlap from biological reasons. For example, separate nuclei, cells, nerves and blood vessels do not coalesce in reality, because of their bodily nature. Thus, a hard-core property was found for the capillaries in both groups of specimens in the present study, and also in the study on thyroid capillaries (Krasnoperov & Stoyan, 2004).

On the other hand, the data of this study show also that it is not enough to account only for the hard-core effect without allowing for further interaction. This follows from the data of the group of the normal specimens. Knowing the mean value of γ and its standard deviation for this group as given in Table I, it is possible to give parametric bounds of a 95% confidence interval of γ for this group on the basis of the t-distribution. With a standard error of $0.982/\sqrt{12} = 0.283$ and 11 degrees of freedom, these bounds amount to [1.17, 2.41]. Hence the lower bound of γ lies distinctly above 1 for this group. However, in the case of a classical hard-core process without further interaction terms, one would expect $\gamma = 1$ (see section 2.1, eq. 1). Thus, the data speak against a classical hard-core process for the capillary profile process in the normal group. In

agreement with the explorative data, they favour the existence of an ordering structure in addition to the hard-core property. Presumably a combination of a hard-core phenomenon with ordering at longer distances is not untypical for biological materials. For example, it was found by explorative methods also in a previous study on point processes of intramembranous particles on freeze fracture replicas (Schladitz et al., 2003; Mattfeldt, 2005). Besides, this statement holds only for the normal case group and not for the carcinoma group. Here a parametric confidence interval of γ would include the value $\gamma = 1$, compatible with a classical hard-core point process.

4.2. Findings in tumour tissue

In carcinomatous tissue, the intensity of the point process of the capillary profiles increased significantly as compared to the normal tissue. This finding was already reported in the previous explorative study and is not very surprising. It is known from many studies on the vascularization of tumour tissues and may be ascribed to an increased angiogenesis, i.e. formation of new vessels in the tumour, because the more rapidly proliferating tumour needs more blood supply and produces diverse angiogenic substances. Interestingly, two of the three model parameters, namely the hard-core distance r_0 and the interaction radius R, remained essentially unchanged in the tumour specimens. The first of these findings has a simple explanation: the hard-core distance is directly related to the physical size of the capillaries. These are in turn generally adapted to the size of red blood cells, which do not vary between normal and malignant tissue. Hence one would not expect a difference here. The interaction radius R indicates the distance until which a point exerts an influence on other points. Our finding may indicate that this parameter is truly unchanged in the tumour group, but one must also take methodological aspects into consideration. The quantity R belongs to the irregular parameters of the Strauss hard-core process. It is found by a complex procedure, the profile pseudolikelihood method, whereas the estimation of r_0 as minimum value of the interpoint distance is rather straightforward. Hence the estimate of R may show more fluctuation from statistical reasons than r_0. This view is supported by the high standard deviation of R in relation to its mean value, as compared to r_0 (see Table I).

The most striking biomedical finding was a drastic decrease of the mean value of the interaction parameter γ from > 1 in normal tissue to < 1 in the carcinoma group. The interaction parameter γ is related to 'attraction'. If $\gamma < 1$, the model describes an 'inhibitive' pattern. If $\gamma > 1$, the model

shows an 'attraction' between points lying at distances between the hard-core distance r_0 and R (see Takacs & Fiksel, 1986; Goulard et al., 1996; Baddeley & Turner, 2000). For biological structures, it may be better to say that certain distances are favoured by the process by the presence of an ordering principle, that makes such distances more probable. These favoured inter-point distances are distinctly diminished for distances between r_0 and R in the tumour tissue. To find an explanation, a biomedical consideration may help. The capillaries are often located near the surface of the glands, where the stroma normally cannot enter. This means a certain preponderance for distances between capillary profiles related to the diameter of the profiles of the normal glands. This normal glandular structure is largely lost in the tumour tissue. In well-differentiated (tubular) adenocarcinomas, the glandular profiles are still convex but much smaller than those in the tumours, which is a well-known diagnostic criterium in histopathological practice. Moreover, in the poorly differentiated 'cribriform' (sieve-like) variant of prostatic carcinoma, which was also found in our case series, irregular stroma bridges may radiate into the glands. This may lead to the occurrence of capillary profiles inside glandular complexes, where they do not appear in the normal state. Hence the normal grouping along intermediate distances is lost in both well and poorly differentiated tumour types, which explains that the interaction parameter γ decreases.

On the whole, the changes found by the explorative approach in the g-function are concordant with the changes of the parameters of the Strauss hard-core process. According to the explorative approach, we found significant attenuations of the g-function in the tumour tissue in two domains of interaction: in the ranges $r = 9$–40 pixels and $r = 60$–77 pixels (Mattfeldt et al., 2006). The decrease of γ predicts a decrease of $g(r)$ in the range between r_0 and R, i.e. in the range between 15–17 and 51–62 pixels (Table I). This corresponds roughly to the declines of the g-function between 9 and 77 pixels. One must here take into consideration that the values of $g(r)$ in the explorative study were obtained by using a kernel for smoothing, while kernel smoothing is not performed when the model parameters of the Strauss hard-core model were fitted. This may explain some discrepancies. For example, the lower bound of $r = 9$ pixels of the first domain of significant differences, is due to kernel smoothing, as it lies below r_0 (see Mattfeldt et al., 2006).

4.3. Stereological aspects

The present study consisted in the parametric modelling of point processes of

intersected fibres found on histological sections. In the case of stationary and isotropic fibre processes, some properties of the planar images are directly inherited from or closely linked to properties of the fibre process, from which they originate (Krasnoperov & Stoyan, 2004; Mattfeldt, 2005; Mattfeldt et al., 2006). The expectation of λ is $(1/2)L_V$, i.e. half of the intensity of the spatial fibre process. The hard-core distance of the point process of the fibre profiles on sections may be interpreted in 3D as the minimal distance between the longitudinal axes of the capillaries (Krasnoperov & Stoyan, 2004). The observable pair correlation function $g(r)$ itself is a direct estimator of the reduced pair correlation function $g_3(r)$ of the spatial fibre process. Intuitively, this suggests that the other parameters of the Strauss model should also have stereological interpretations, at least in a *qualitative* manner and for the purpose of comparisons.

The quantity R is defined as a radius of interaction in the sectional plane, i.e. the radius of a circle within which a point exerts an influence on other points. Here R is 'a radius beyond which influence is inconceivable' (Stoyan e al., 1995, p. 171). When extrapolating this concept to the 3D fibre process, this leads conceptionally to a sphere of interaction, within which a fibre still exerts an influence on other fibres. Let us denote the radius of this hypothetical sphere of interaction as R'. The planar interaction parameter γ indicates the strength of interaction at distances above the hard-core distance and below the radius of interaction between points in the plane. The spatial analogon would be a hypothetical interaction parameter γ', which indicates the strength of interaction in the sphere of interaction with radius R'. Clearly there is currently no way to estimate R' from R and γ' from γ quantitatively in the same way as L_V can be unbiasedly estimated from Q_A. Nevertheless, we think that changes of R' and γ' should lead *qualitatively* to changes of R and γ into the same direction. If the radius of interaction R' between the spatial fibres is drastically increased, this should lead also to an increase of the radius of interaction R between the points, as these are the traces of the same fibres. From the same reason, it seems plausible that an increase in the strength of interaction of the fibres in 3D leads qualitatively also to an increased strength of interaction in the point pattern of the profiles. Hence we infer qualitatively from the data that the radius of interaction between the capillaries in 3D has not changed in the tumour tissue, whereas the strength of interaction between the capillaries is reduced at some formerly favoured distances. In other words, the normal preference of capillaries around glandular lobules in 3D space is

reduced in the tumour tissue. This stereological diction is slightly stronger than the conclusion of section 4.2, where we said merely that the preference of capillary profiles around profiles of glandular lobules is reduced.

Likewise, stationarity and nonstationarity of the observed point process of the fibre profiles will somehow reflect stationarity and nonstationarity of the underlying fibre process. If the fibres show gradients of intensity in space, for example, this will also lead to gradients of the resulting point process on the sections. Hence the stationarity of the profile point process indicates stationarity of an isotropic fibre process, too. Physiologically, stationarity is also plausible: living tissue needs blood everywhere, hence larger regions cannot survive without capillaries. Necrotic regions, however, where not found in our specimens. Evidently the finding of stationarity is also very desirable from the stereological viewpoint, because stationarity is one of the critical assumptions when extrapolating the pair correlation function from 2D to 3D. Remember that the estimator of Krasnoperov & Stoyan (2004) is only valid for isotropic fibre processes that are also stationary. Hence, the check on planar stationarity delineated in sections 2.2.2 and 2.3.3 may also be considered as a test on a basic stereological model assumption.

Acknowledgment

Stefanie Eckel is supported by a grant of the graduate college 1100 at the University of Ulm.

References

Baddeley, A. & Turner, R. (2000) Practical maximum pseudolikelihood for spatial point patterns (with discussion). Australian and New Zealand Journal of Statistics 42(3), 283-322.

Baddeley, A. & Turner, R. (2005) spatstat: An R package for analyzing spatial point patterns. J. Statist. Software 12(6), 1–42.

Baddeley, A. & Turner, R. (2006) Modelling spatial point patterns in R. In: Baddeley, A., Gregori, P., Mateu, J., Stoica, R. & Stoyan, D. (editors): Case Studies in Spatial Point Pattern Modelling. Lecture Notes in Statistics 185, 23–74. New York, Springer.

Baddeley, A., Turner, R., Møller, J. & Hazelton, M. (2005) Residual analysis for spatial point processes (with discussion). J. Royal Statist. Soc. Series B 67, 617-666.

Chib, S. & Greenberg, E. (1995) Understanding the Metropolis-Hastings algorithm. Am. Statistician 49, 327–335.

Diggle, P.J. (2003) Statistical Analysis of Spatial Point Patterns. Second edition. London, Arnold.

Diggle, P.J., Fiksel, T., Grabarnik, P., Ogata, Y., Stoyan, D. & Tanemura, M. (1994) On parameter estimation for pairwise interaction point processes. Int. Statist. Rev. 62, 99–117.

Fleischer, F., Beil, M., Kazda, M. & Schmidt, V. (2006) Analysis of spatial point patterns in microscopic and macroscopic biological image data. In: Baddeley, A., Gregori, P., Mateu, J., Stoica, R. & Stoyan, D. (editors): Case Studies in Spatial Point Pattern Modelling. Lecture Notes in Statistics 185, pp. 23–74. New York, Springer.

Gelman, A. (1996) Inference and monitoring convergence. In: Gilks, W.R., Richardson, W.R., Spiegelhalter, D.J. (Editors): Markov Chain Monte Carlo in Practice. Boca Raton: Chapman Hall, 131–143.

Geyer, C.J. & Møller, J. (1994) Simulation procedures and likelihood inference for spatial point processes. Scand. J. Statist. 21, 359–373.

Goulard, M., Särkkä, A. & Grabarnik, P. (1996) Parameter estimation for marked Gibbs point processes through the maximum pseudo-likelihood method. Scand. J. Stat. 23, 365–379.

Hahn, U., Micheletti, A., Pohlink, R., Stoyan, D. & Wendrock, H. (1999) Stereological analysis and modelling of gradient structures. J. Microsc. 195, 113–124.

Krasnoperov, R.A. & Stoyan, D. (2004) Second-order stereology of spatial fibre systems. J. Microsc. 216, 156–164.

Mattfeldt, T. (2005) Explorative statistical analysis of planar point processes in microscopy. J. Microsc. 220, 131–139.

Mattfeldt, T. & Fleischer, F. (2005) Bootstrap methods for statistical inference from stereological estimates of volume fraction. J. Microsc. 218, 160–170.

Mattfeldt, T. & Fleischer, F. (2006) Computer-intensive methods for statistical inference from stereological data. Proc. Int. Conf. Spat. Statist. Stoch. Geom. (S4G), Prague. Ed. R. Lechnerova, I. Saxl & V. Beneš, 349–360.

Mattfeldt, T. Eckel, S., Fleischer, F. & Schmidt, V. (2006) Statistical analysis of reduced pair correlation functions of capillaries in the prostate gland. J. Microsc. 223, 107–119.

Møller, J. (1999) Markov chain Monte Carlo and spatial point processes. In: Barndorff-Nielsen, O., Kendall, W. & Lieshout, M.N.M.: Stochastic Geometry. Likelihood and Computation. Boca Raton: Chapman Hall,

141–172.

Møller, J. & Waagepetersen, R.P. (2003) An introduction to simulation-based inference for spatial point processes. In: Møller, J. (Editor): Spatial Statistics and Computational Methods, Lecture Notes in Statistics 173, 143–198. New York, Springer.

Møller, J. & Waagepetersen, R.P. (2004) Statistical Inference and Simulation for Spatial Point Processes. Boca Raton, Chapman and Hall.

Ripley, B.D. (1977) Modelling spatial patterns (with discussion). J. Royal Statist. Soc. B 39, 172–212.

Ripley, B.D. (1988) Statistical Inference for Spatial Processes. Cambridge: Cambridge University Press.

Schladitz, K., Särkkä, A., Pavenstädt, I., Haferkamp, O. & Mattfeldt, T. (2003) Statistical analysis of intramembranous particles using freeze fracture specimens. J. Microsc. 211, 137–153.

Stoyan, D. & Stoyan, H. (1994) Fractals, Random Shapes and Point Fields. Methods of Geometrical Statistics. Chichester: Wiley.

Stoyan, D., Kendall, W.S. & Mecke, J. (1995) Stochastic Geometry and its Applications, 2nd ed. Wiley, Chichester.

Takacs, R. & Fiksel, T. (1986) Interaction pair potentials for a system of ants' nests. Biometr. J. 28, 1007–1013.

Van Lieshout, M.N.M. (2002) Markov Point Processes and Their Applications. Singapore, World Scientific.