Modelling the Bivariate Spatial Distribution of Amacrine Cells

Peter J Diggle[1], Stephen J. Eglen[2], and John B. Troy[3]

[1] Lancaster University and Johns Hopkins University School of Public Health
   `p.diggle@lancaster.ac.uk`
[2] University of Cambridge `S.J.Eglen@damtp.cam.ac.uk`
[3] Northwestern University, USA `j-troy@northwestern.edu`

1 Introduction

1.1 Biological background

Humans and many vertebrates have a very highly specialised visual system that allows us to perceive the world. Our capacity to see begins at the back of the eye, where a neural structure called the retina converts light into electrical activity. The retina is a three-dimensional structure, composed of several types of cell (Figure 1). The light is first converted into neural activity by the photoreceptors, which then pass their signals through several types of interneuron. Eventually the activity reaches the retinal ganglion cells, which then send the signals to the brain.

There are many different types of neuron in the retina; with a few exceptions, each type of neuron is arranged in a regular fashion so that the visual world is systematically sampled, without leaving any 'holes' in visual space. In this chapter, we will focus on the spatial positioning of two types of retinal neuron, known as the cholinergic amacrine cells (Famiglietti, 1983; Tauchi and Masland, 1984). These interneurons modulate the pattern of visual information as it passes through the retina, and are thought to play an important role in the detection of motion in particular directions (Euler, Detwiler and Denk, 2002). There are two types of cholinergic amacrine cell, depending on the depth within the retina at which the cell body is found. Cells found within the inner nuclear layer are termed "off" cells here, whilst cells found in the ganglion cell layer are termed "on" cells.

We are interested in studying the spatial dependency between the positions of on and off cells because we hope this will tell us something about how the two cell types emerge during development: do the two cell types emerge from a single undifferentiated population, or do they develop independently of each other? Also, in more general terms, this question that we ask here about the cholinergic amacrine cells could be asked of other cell types. In the special

case when the two types of retinal neuron are in different layers, existing approaches (Diggle, 1986) may be suitable to test for independence. However, these techniques are *a priori* invalid when both cell types occur in the same layer, because in these circumstances the physical space required by each cell formally precludes statistical independence of the two component arrays.
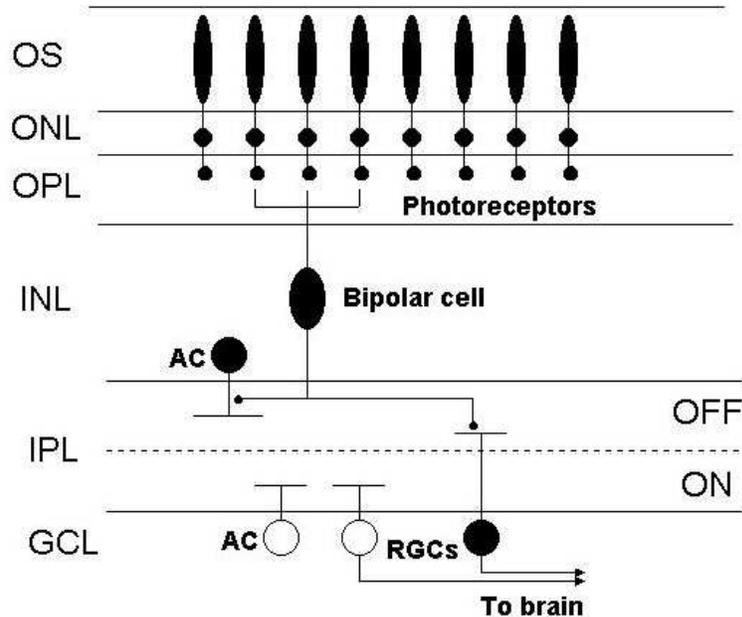

Fig. 1. Cross-section through the different layers of the retina. Layers are named to left, for reference. (OS: outer segments; ONL: outer nuclear layer; OPL: outer plexiform layer; INL: inner nuclear layer; IPL: inner plexiform layer; GCL: ganglion cell layer). Light enters the eye through the front (at bottom) and travels through the retina where it is converted to electrical activity by the photoreceptors. Two main cell types can be classified into "on" (open circles) or "off" (filled circles) depending on where their dendritic processes terminate within the IPL. Cholinergic amacrine cells (AC) are found at two different layers, whereas retinal ganglion cells (RGCs) are normally all located within the GCL. RGCs are the only cells that send their information along the optic nerve to the brain. Many cell types have been omitted from this diagram for simplicity.

The data that we shall analyse are shown in Figure 2. This shows a single, bivariate spatial point pattern taken from the retina of a rabbit, in which the two types of point correspond to the positions of the centres of 152 "on" and 142 "off" amacrine cells; these data are from Wieniawa-Narkiewicz (1983)

and were kindly made available to us by Prof. Abbie Hughes. For a general discussion of the biological background to these data, see Hughes (1985).

The pattern has been recorded within a rectangular section of the retina, of dimension 1060 by 662 μm. Visually, both on and off cells exhibit patterns which are more regular than would be the case for completely random patterns, i.e. realisations of homogeneous Poisson point processes. In particular, there is a pronounced inhibitory effect, meaning that no two on cells, and no two off cells, can be located arbitrarily close together. The inhibitory effect is much less pronounced between cells of opposite type. For example, the minimum observed distance between any two on cells is 21.4μm, between any two off cells is 15.8μm, and between any pair of on and off cells is 5μm. We shall use 1μm as the unit of distance throughout.
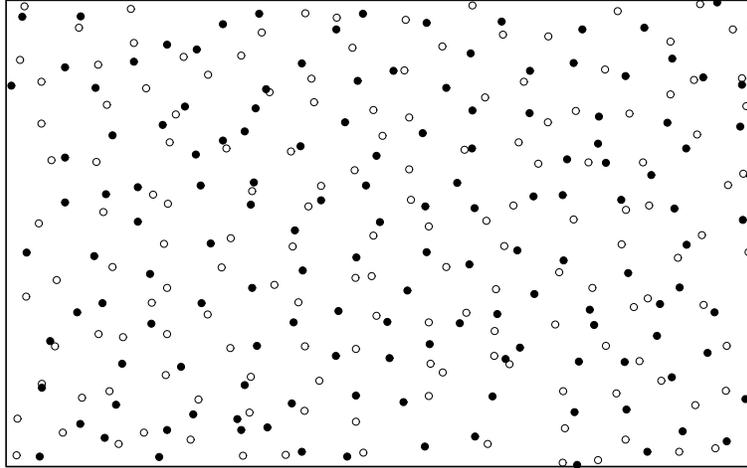
Fig. 2. The cholinergic amacrine data. On and off cells are shown as open and closed circles, respectively. The rectangular region on which the cells are observed has dimension 1060 by 662 μm. Cell bodies are drawn to scale (10 μm diameter). Cells of opposite polarity (on vs off) can partially overlap, since they are located in different layers, but cells of like polarity never overlap.

Previous analyses of the data have been reported by Diggle (1986), where non-parametric methods led to the conclusion that the two component patterns were approximately independent, and by Diggle and Gratton (1984) and Diggle (2003) who used the data to illustrate the fitting of univariate models by *ad hoc* and likelihood-based methods, respectively. Our goal in the current chapter is to demonstrate how recently developed Monte Carlo methods for conducting likelihood-based analysis of realistic point process models can lead to sharper inferences about the bivariate structure of the data. In par-

ticular, we will formulate and fit a bivariate pairwise interaction model for the amacrines data, and will argue that likelihood-based inference within this model is both statistically more efficient and scientifically more relevant than *ad hoc* testing of benchmark hypotheses such as independence.

One major limitation of the analysis reported here is that the data are unreplicated, i.e. they consist of a single point pattern. The literature on the statistical analysis of replicated spatial point pattern data is surprisingly sparse. Diggle, Lang and Benes (1991) and Baddeley, Moyeed, Howard and Boyde (1993) consider methods based on pooled estimates of non-parametric functional summary statistics such as the K-function (Ripley, 1976, 1977). Diggle, Mateu and Clough (2000) compare parametric and non-parametric approaches to testing for differences between replicated patterns in two or more experimental groups. We are assembling a collection of replicated patterns of retinal cells and intend to analyse these using parametric, likelihood-based methods of the kind described in the current chapter. We will report separately on the analyses of these data in due course.

2 Pairwise interaction point processes

2.1 Univariate pairwise interaction point processes

Markov point processes were introduced by Ripley and Kelly (1977). Van Lieshout (2000) discusses their construction, properties and uses as statistical models for spatial point patterns.

Pairwise interaction point processes are perhaps the most widely used sub-class of Markov point processes, In particular, they provide a flexible, parsimonious class of models for point patterns which display varying degrees of spatial regularity, as exhibited by our data.

Let $X = \{x_i : i = 1, ..., n\}$ be an observed spatial point pattern on a planar region A, hence each $x_i \in A$ and all points in A are observed. We call the points of the process *events* to distinguish them from arbitrary points $x \in A$. In a pairwise interaction point process, the likelihood ratio for X with respect to a homogeneous Poisson process of unit intensity takes the form

$$c\beta^n \prod_{i=2}^{n} \prod_{j=1}^{i-1} h(||x_i - x_j||). \tag{1}$$

In (1), $|| \cdot ||$ denotes Euclidean distance, $h(u) : u \geq 0$ is the pairwise interaction function, β reflects the intensity of the process and c is a normalising constant whose analytic form is typically intractable.

The essence of the model (1) lies in the interaction function, $h(\cdot)$. When $h(u) = 1$ for all u, the process is a homogeneous Poisson process of intensity β. When $h(u) = 0$ for $0 \leq u \leq \delta$, no two events can occur less than a distance δ apart and the process is said to display *strict inhibition*. Values

of $h(u)$ intermediate between zero and one correspond to non-strict forms of inhibition in which close pairs of events are relatively unlikely, but not ruled out completely. The smallest distance ρ such that $h(u) = 1$ for all $u > \rho$ is called the *range* of the process. Models with $h(u) > 1$ for certain ranges of u are potentially invalid because the likelihood ratio (1) may not be integrable over A; an early example is the Strauss (1975) model for clustering, subsequently shown by Kelly and Ripley (1976) to be invalid. In theory, models with $h(u) = 0$ for $u \leq \delta$ and $h(u) > 1$ for $\delta < u < \rho$ could be used to model aggregated spatial patterns, but in practice such models are not very useful because they correspond to very extreme forms of spatial aggregation, in a sense made precise by Gates and Westcott (1986).

In (1), conditioning on the observed number of events in A leads to a joint probability density function for X, proportional to

$$f(X) = \prod_{i=2}^{n} \prod_{j=1}^{i-1} h(||x_i - x_j||). \qquad (2)$$

In the general inhibitory case, i.e. when $h(u) \leq 1$ for all u, and when n is large, the distinction between processes with a fixed or random number of events in A is relatively unimportant. In what follows, we shall consider only the case of fixed n. Hence, we do not attempt to make inferences about the intensity of the process, but only about the form of the interaction function $h(u) = h(u; \theta)$. The log-likelihood for θ is then given by

$$\log L(\theta) = \log f(X; \theta) + \log c(\theta) \qquad (3)$$

where $c(\theta)$ is the normalising constant for (2). Figure 3 shows a realisation of a process with interaction function $h(u) = 1 - \exp(-u/\phi)$ for each of $\phi = 0.01, 0.05, 0.10, 0.15$ and, in each case, $n = 100$ events on the unit square. The progressive development of spatial regularity as the value of ϕ increases is clear. The simulations were generated on a toroidal region which was then unwrapped to form the unit square A; this counteracts a tendency for events to be artificially concentrated near the edge of A when the model is strongly inhibitory, i.e. in the present context, when ϕ is large.

2.2 Bivariate pairwise interaction point processes

A bivariate spatial pattern consists of two sets of locations corresponding to two distinguishable types of event, which in our application are the on and off cells.

Let $X_1 = \{x_{1i} : i = 1, ..., n_1\}$ and $X_2 = \{x_{2i} : i = 1, ..., n_2\}$ represent a bivariate spatial point pattern of events in a region A. A bivariate pairwise interaction model is specified by three interaction functions, $h_{11}(\cdot)$, $h_{22}(\cdot)$ and $h_{12}(\cdot)$, which operate between pairs of events of type 1, pairs of events of type 2, and pairs of events of opposite type, respectively. Then, if we condition on
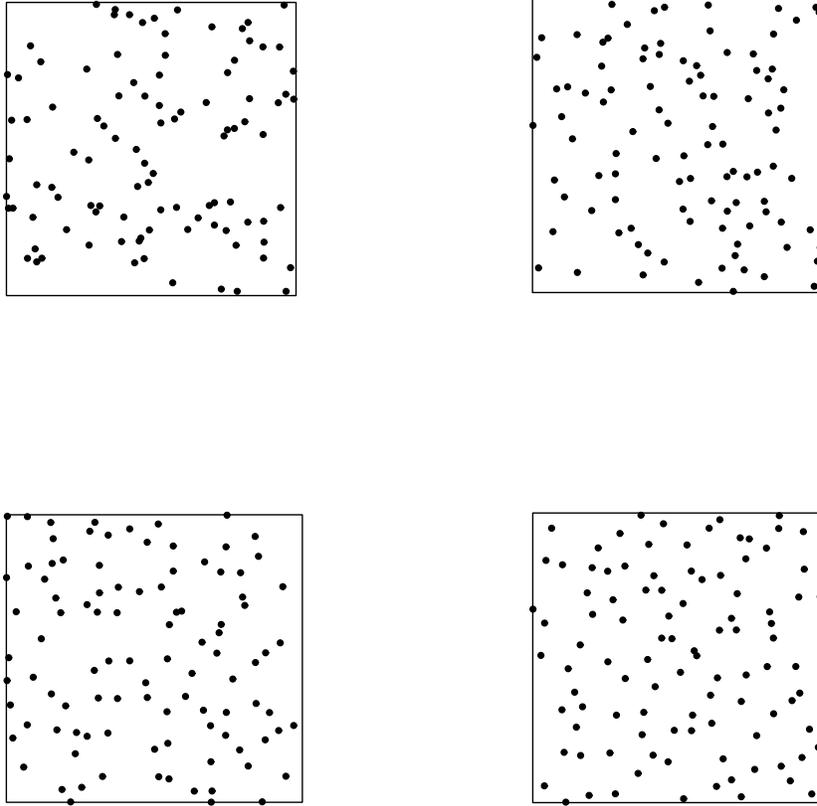
Fig. 3. Simulated realisations of pairwise interaction point processes each with 100 events on the unit square and interaction function $h(u) = 1 - \exp(-u/\phi)$. The values of ϕ are 0.01 (top-left), 0.05 (top-right), 0.1 (bottom-left) and 0.15 (bottom-right)

the observed numbers of events, n_1 and n_2, the probability density of (X_1, X_2) is proportional to

$$f(X_1, X_2) = P_{11}P_{22}P_{12}, \tag{4}$$

where

$$P_{11} = \prod_{i=2}^{n_1} \prod_{j=1}^{i-1} h_{11}(||x_{1i} - x_{1j}||), \tag{5}$$

$$P_{22} = \prod_{i=2}^{n_2} \prod_{j=1}^{i-1} h_{22}(||x_{2i} - x_{2j}||), \tag{6}$$

and

$$P_{12} = \prod_{i=1}^{n_1} \prod_{j=1}^{n_2} h_{12}(||x_{1i} - x_{2j}||). \tag{7}$$

Equation (4) is a natural bivariate counterpart of (2). An important feature of the bivariate model is that its marginal properties depend on all three interaction functions. To illustrate this, we use the family of simple inhibitory interaction functions,

$$h_{ij}(u) = \begin{cases} 0 : u < \delta_{ij} \\ 1 : u \geq \delta_{ij} \end{cases} \tag{8}$$

and specify $\delta_{11} = \delta_{22} = 0.025$. If we also specify $\delta_{12} = 0$, then the two component processes are independent copies of a univariate simple inhibition process. The left-hand panel of Figure 4 shows a realisation of this bivariate process. The two univariate components each display spatial regularity because of the inhibition effect but, because the two components are independent, arbitrarily close pairs of opposite type can and do occur. If we now introduce a strongly inhibitory interaction between events of opposite type by specifying $\delta_{12} = 0.1$, the effect is very different, as shown in Figure 4. The cross-inhibitory effect between events of opposite type leads to component patterns which are marginally spatially aggregated, albeit with a clearly discernible local inhibitory effect, and jointly spatially segregated.

3 Monte Carlo likelihood inference

The generally agreed "gold standard" for statistical estimation and hypothesis testing is to use likelihood-based methods; specifically, within a classical inferential framework, estimates should be maximum likelihood estimates and tests should be likelihood ratio tests.

The difficulty with applying this gold standard to our model is that the normalising constant for the joint probability density of (X_1, X_2), and hence the likelihood function for θ, is intractable. Geyer and Thompson (1992) provided an ingenious solution to this problem, which allows us to use simulations of the process at any fixed value θ_0 to compute an approximation to the likelihood ratio with respect to θ_0 for any value of θ. In the present context of pairwise interaction point processes, the argument runs as follows – we describe only the univariate case explicitly, but the extension to bivariate processes is obvious.

Let $c(\theta)$ be the normalising constant associated with the model (2), hence

$$c(\theta)^{-1} = \int f(X; \theta)dX$$
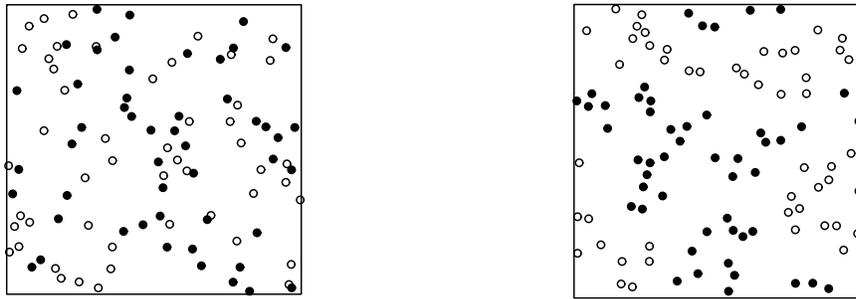
Fig. 4. Simulated realisations of bivariate pairwise interaction point processes each with 50 events of either type on the unit square and simple inhibitory interaction functions. In both panels, the minimum permissible distance between any two events of the same type is 0.025. In the left-hand panel, the two component patterns are independent. In the right-hand panel, the minimum permissible distance between any two events of opposite types is 0.1.

Now, for any fixed θ_0, write

$$c(\theta)^{-1} = \int f(X; \theta) \times \frac{c(\theta_0)}{c(\theta_0)} \times \frac{f(X; \theta_0)}{f(X; \theta_0)} dX, \qquad (9)$$

define $r(X; \theta, \theta_0) = f(X; \theta)/f(X; \theta_0)$ and re-arrange the right-hand-side of (9) to give

$$c(\theta)^{-1} = c(\theta_0)^{-1} \mathrm{E}_{\theta_0}[r(X; \theta, \theta_0)].$$

Hence, the normalised joint density for X can be expressed as

$$g(X; \theta) = c(\theta_0) f(X; \theta)/\mathrm{E}_{\theta_0}[r(X; \theta, \theta_0)].$$

Since θ_0 is a constant, it follows that the maximum likelihood estimator $\hat{\theta}$ maximises

$$L_{\theta_0}(\theta) = \log f(X; \theta) - \log \mathrm{E}_{\theta_0}[r(X; \theta, \theta_0)]. \qquad (10)$$

The Monte Carlo method replaces the expectation on the right-hand-side of (10) by a Monte Carlo estimate, computed from s replicate simulations. Hence the Monte Carlo maximum likelihood estimate maximises

$$L_{\theta_0, s}(\theta) = \log f(X; \theta) - \log s^{-1} \sum_{j=1}^{s} [r(X_j; \theta, \theta_0)], \qquad (11)$$

where the $X_j : j = 1, ..., s$ are simulated realisations with $\theta = \theta_0$.

Whilst the computations needed to secure a sufficiently accurate approximation can be time-consuming, the implication of Geyer and Thompson's

work is that in principle there is no obstacle to using likelihood-based inference rather than the more *ad hoc* methods which are traditionally used to analyse spatial point pattern data. For a more detailed account, see Geyer (1999).

Note that (11) defines a whole family of estimation criteria according to the choices made for θ_0 and s, and that for given s the extent of the stochastic variation introduced by the Monte Carlo simulation depends crucially on the choice of θ_0. In practice, the method works best when θ_0 is close to $\hat{\theta}$. Our approach has been to conduct a sequence of numerical optimisations of (11), updating θ_0 to the current maximising value after each stage until no further material change occurs, and increasing s until the Monte Carlo component of variance is negligible compared with the inherent uncertainty in $\hat{\theta}$ as measured by the Hessian matrix.

Whilst we favour Monte Carlo likelihood-based methods for formal parametric inference, in our opinion more *ad hoc* methods still have a useful role to play in the overall analysis. We use them to provide good initial values of θ_0 for the Monte Carlo likelihood calculations, and as checks on the goodness-of-fit of the final models produced by the likelihood-based analysis.

4 Analysis of the amacrines data

4.1 Exploratory analysis

A standard tool for exploratory analysis of spatial point pattern data is the K-function, introduced by Ripley (1976, 1977) and, in the bivariate case, by Lotwick and Silverman (1982). In its basic form, the K-function describes the second-order properties of a *stationary* spatial point process. Baddeley, Moller and Waagepetersen (2000) extend its definition to include processes with spatially varying intensities. For the current application, we shall assume stationarity.

In the stationary bivariate case, if λ_j denotes the intensity, or expected number per unit area, of type j events, then $\lambda_j K_{ij}(u)$ represents the expected number of additional type j events within distance u of an arbitrary type i event. If the component pattern of type j events is a homogeneous Poisson process, then $K_{jj}(u) = \pi u^2$. Spatially aggregated and spatially regular processes typically have $K_{jj}(u) > \pi u^2$ and $K_{jj}(u) < \pi u^2$, respectively. If type 1 and 2 events are generated by independent processes, then $K_{12}(u) = \pi u^2$. If type 1 and 2 events are generated by randomly labelling the events of a univariate process, i.e. types are assigned by the results of a sequence of independent Bernoulli trials, then $K_{11}(u) = K_{22}(u) = K_{12}(u)$. In particular, independence and random labelling are equivalent if and only if the component processes are homogeneous Poisson processes. These results explain why we favour plotting estimates $\hat{K}_{ij}(u) - \pi u^2$, rather than the $\hat{K}_{ij}(u)$ themselves. Several different estimators have been proposed, whose principal differences

concern their method of dealing with edge-effects; see, for example, Ohser and Stoyan (1981), Stein (1991) or Baddeley (1999). We use the original estimators proposed by Ripley (1976, 1977) and, in the bivariate case, by Lotwick and Silverman (1982).

Figure 5 shows the estimates $\hat{K}_{ij}(u) - \pi u^2$ for the amacrines data. Note firstly that $\hat{K}_{11}(u)$ and $\hat{K}_{22}(u)$ are close together, suggesting that they may be generated by the same underlying process. Also, both estimates follow the parabola $-\pi u^2$ at small distances, i.e. $\hat{K}_{11}(u) = \hat{K}_{22}(u) = 0$, confirming the visual impression of a strict inhibitory effect within each of the component patterns. In contrast, $\hat{K}_{12}(u) - \pi u^2$ fluctuates around zero at small u. This behaviour, coupled with the fact that the sampling variance of $\hat{K}_{12}(u)$ increases with u, is consistent with the component processes being approximately independent. Note also that the magnitude of the difference between $\hat{K}_{11}(u)$ and $\hat{K}_{22}(u)$ derives from the combination of sampling variation in the estimates and the difference, if any, between the two underlying theoretical functions; it therefore provides an informal upper bound for the sampling variation, and on this basis we can conclude that the much larger difference between the $\hat{K}_{jj}(u)$ and $\hat{K}_{12}(u)$ is incompatible with random labelling.
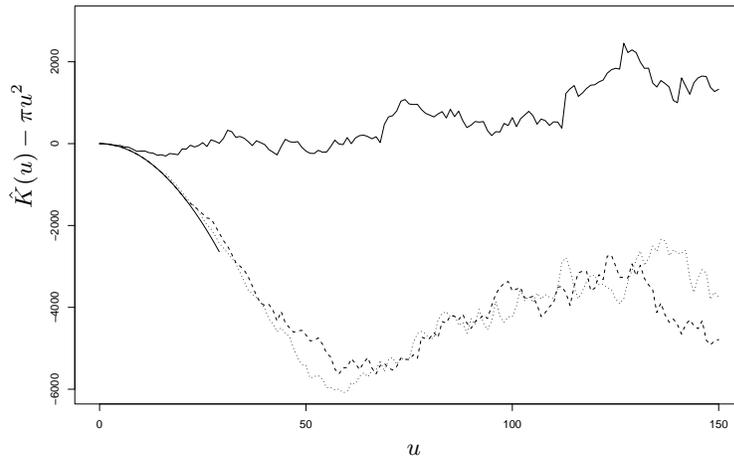
Fig. 5. Estimates of the K-functions for the on and off cells. Each plotted function is $\hat{K}(u) - \pi u^2$. The dashed line corresponds to $\hat{K}_{11}(u)$ (on cells), the dotted line to $\hat{K}_{22}(u)$ (off cells) and the solid line to $\hat{K}_{12}(u)$. The parabola $-\pi u^2$ is also shown as a solid line.

4.2 Structural hypotheses for the amacrines data

The exploratory analysis suggests that, purely from a statistical perspective, an inhibitory, bivariate pairwise interaction process with independent components and a common underlying model for the two components may provide a reasonable fit to the data. For many retinal cells, the hypothesis of *statistical independence* is strictly implausible because their cell bodies lie in the same cellular layer and two cells cannot occupy the same space. A more appropriate benchmark hypothesis, which we shall call *functional independence* is that the only form of interaction between type 1 and type 2 events is a simple inhibitory effect due to the physical size of the cells, i.e. an interaction function $h_{12}(u)$ of the form given by (8), with the value of δ_{12} no greater than the typical size of an individual cell.

A second hypothesis which is of some biological interest is *common components*, by which we mean that the data are generated by a bivariate model with $h_{11}(u) = h_{22}(u)$. Our analysis will therefore include formal tests of statistical independence, structural independence and common components.

4.3 Non-parametric estimation

We use the method of maximum pseudo-likelihood (Besag, Milne and Zachary, 1982; Baddeley and Turner, 2000) to obtain non-parametric estimates of the interaction functions $h_{11}(u)$ and $h_{22}(u)$. Formally, this is achieved by fitting a deliberately over-parameterised model in which the interaction function is assumed to be piecewise constant, with the heights of the pieces as its parameters.

Figure 6 shows the results. The estimates of the two interaction functions are quite similar, adding weight to the evidence for a common components model.

Figure 6 also suggests what approximate shape a more parsimonious parametric model for the interaction functions would need to accommodate. We shall use functions h_{ij} within the parametric family $h(u, \theta)$ where $\theta = (\delta, \phi, \alpha)$ and

$$h(u; \theta) = \begin{cases} 0 : u \leq \delta \\ 1 - \exp[-\{(u - \delta)/\phi\}^{\alpha}] : u > \delta \end{cases} \tag{12}$$

This allows a wide range of inhibitory interactions within and between types by varying the corresponding parameter vectors θ_{11}, θ_{22} and θ_{12} so as to define the corresponding interaction functions $h_{ij}(u) = h(u; \theta_{ij})$.

Because a large value for the parameter α allows $h(u)$ to take values close to zero even for relatively large values of u, we might expect the parameters of (12) to be poorly identified. Our response to this, following the discussion in Section 4.2, is to treat the values of δ_{11} and δ_{22} as fixed constants with a common value 10, corresponding to the approximate physical size of the cells (Famiglietti, 1985; Brandon, 1987). Of course, the model is at best an approximation to nature, and we should not over-interpret this precise value;
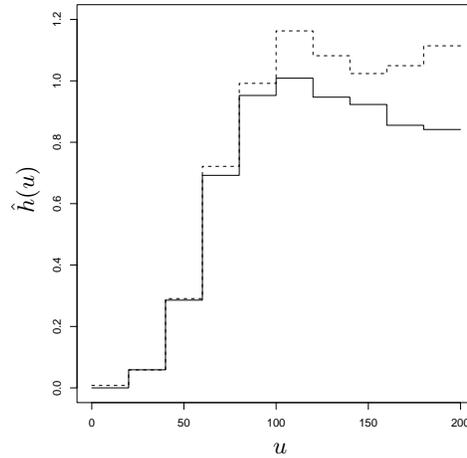
Fig. 6. Non-parametric maximum pseudo-likelihood estimates of the pairwise interaction functions for on cells (solid line) and for off cells (dashed line).

rather, it represents a plausible lower limit on the physical size of the cells. As we shall see, the model can still capture an effective inhibition distance between cells which is substantially greater than 10.

It is harder to argue for an *a priori* fixed value of δ_{12} because of the vertical displacement between the mature on and off cells. The on cells lie somewhat deeper than the off cells and a pair of cells of opposite type could in principle be almost co-located in the planar projection of the data. We shall therefore treat δ_{12} as a parameter to be estimated; as discussed earlier, inference concerning δ_{12} is of some biological interest in its own right.

4.4 Univariate parametric analysis

Under the working assumption of statistical independence, we can analyse the two patterns separately and investigate whether a common set of parameters provides a good fit to both. This analysis is also useful as a prelude to a bivariate analysis, whether or not the independence hypothesis is sustainable.

To obtain initial values for numerical optimisation of the Monte Carlo log-likelihood, we fitted the parametric form of $h(u; \theta)$ to each non-parametric estimate of $h(u)$ shown in Figure 6 by ordinary least squares. We then obtained Monte Carlo maximum likelihood estimates of ϕ and α separately for each of the two patterns, progressively increasing the number of Monte Carlo samples from 10 to 1000, until the estimates stabilised.

To test whether a common set of parameters fitted both patterns, we repeated the optimisation process, but now maximising a pooled Monte Carlo

log-likelihood with common parameter values for the two component patterns. The resulting log-likelihood ratio test statistic was $D = 1.36$ on two degrees of freedom, corresponding to $p = 0.244$. We therefore accepted the common components hypothesis, which gave us the parameter estimates shown in Table 1. Approximate standard errors, and the correlation between $\hat{\phi}$ and $\hat{\alpha}$, were derived from the estimated Hessian matrix of the pooled Monte Carlo log-likelihood at its maximum. All optimisations used the built-in `optim()` function within R; for details, see `http://www.r-project.org`. Figure 7 compares the fitted, common parametric form of $h(u)$ with the two non-parametric estimates. The fit appears to be satisfactory, but we postpone a formal goodness-of-fit assessment until we have fitted a bivariate model.

Table 1. Monte Carlo maximum likelihood estimates, standard errors and correlation, assuming independence between on and off amacrine cell patterns and common parameter values.

Parameter	Estimate	Std Error	Correlation
ϕ	49.08	2.51	
α	2.92	0.25	-0.06

PSfrag replacements
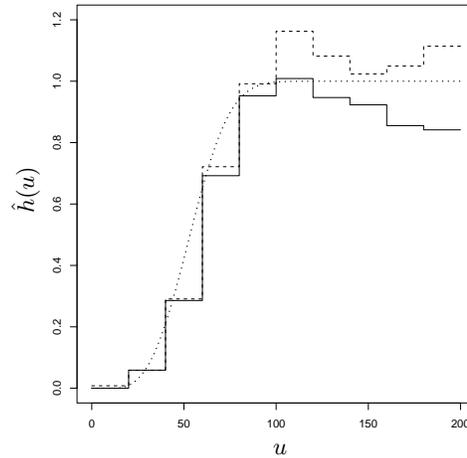
Fig. 7. Non-parametric maximum pseudo-likelihood estimates of the pairwise interaction functions for on cells (solid line) and for off cells (dashed line), together with parametric fit assuming common parameter values for both types of cell (dotted line).

4.5 Bivariate analysis

The first stage in the bivariate analysis is a simple likelihood ratio rest of *statistical independence* against *functional independence*. To do this, we first begin by estimating δ_{12}, obtaining the maximum likelihood estimate $\hat{\delta}_{12} = 4.9$. We then construct a likelihood ratio test of any fixed value of δ_{12} against $\delta_{12} = 4.9$. The set of values not rejected at the 5% level defines a Monte Carlo 95% confidence interval for δ_{12}. Note that all values of δ_{12} greater than 5, the smallest observed distance between a pair of cells of opposite type, are automatically excluded according to the likelihood criterion, because all such values are incompatible with the data. The resulting 95% confidence interval is $2.3 \leq \delta_{12} < 5.0$. In particular, this interval excludes zero, implying that statistical independence is rejected at the conventional 5% level; more precisely, the attained significance level of the likelihood ratio test of statistical independence against functional independence is $p = 0.021$ (test statistic $D = 5.30$, $P(\chi_1^2 > 5.30) = 0.021$).

We next investigate whether there is any further degree of dependence between the on and off cells by introducing additional parameters ϕ_{12} and α_{12}, holding the remaining parameters fixed at $\phi_{11} = \phi_{22} = 49.08$, $\alpha_{11} = \alpha_{22} = 2.92$, $\delta_{11} = \delta_{22} = 10$ and $\delta_{12} = 4.9$. The likelihood ratio test statistic to compare functional independence against the general bivariate model is $D = 0.30$ on 2 degrees of freedom, corresponding to $p = 0.861$. Hence, functional independence is not rejected.

To assess the goodness-of-fit to the bivariate, functional independence model we again use the K-function (Ripley, 1976, 1977). We define three test statistics

$$T_{ij} = \sum_{u=1}^{150} [\{\hat{K}_{ij}(u) - \bar{K}_{ij}(u)\}/u]^2 \tag{13}$$

where $\hat{K}_{ij}(u)$ is the estimate of $K_{ij}(u)$ calculated from the data and $\bar{K}_{ij}(u)$ the corresponding mean of estimates from 99 simulations of the fitted model. The three statistics of interest are T_{11} (on cells), T_{22} (off cells) and T_{12} (dependence between on and off cells). The attained significance levels of the three Monte Carlo tests were 0.11, 0.05 and 0.25 respectively, indicating a reasonable overall fit; an admittedly conservative bound for the combined significance level is $0.05 \times 3 = 0.15$. Figure 8 shows the three estimated K-functions together with the pointwise envelopes from 99 simulations of the fitted model. Although the estimated functions drift briefly outside the simulation envelopes at large values of u, the estimates themselves are imprecise at large values of u, as indicated by the widths of the simulation envelopes. This also explains why we have chosen to discount progressively the influence of estimates $\hat{K}_{ij}(u)$ at large values of u in our construction of the test statistics (13).

PSfrag replacements
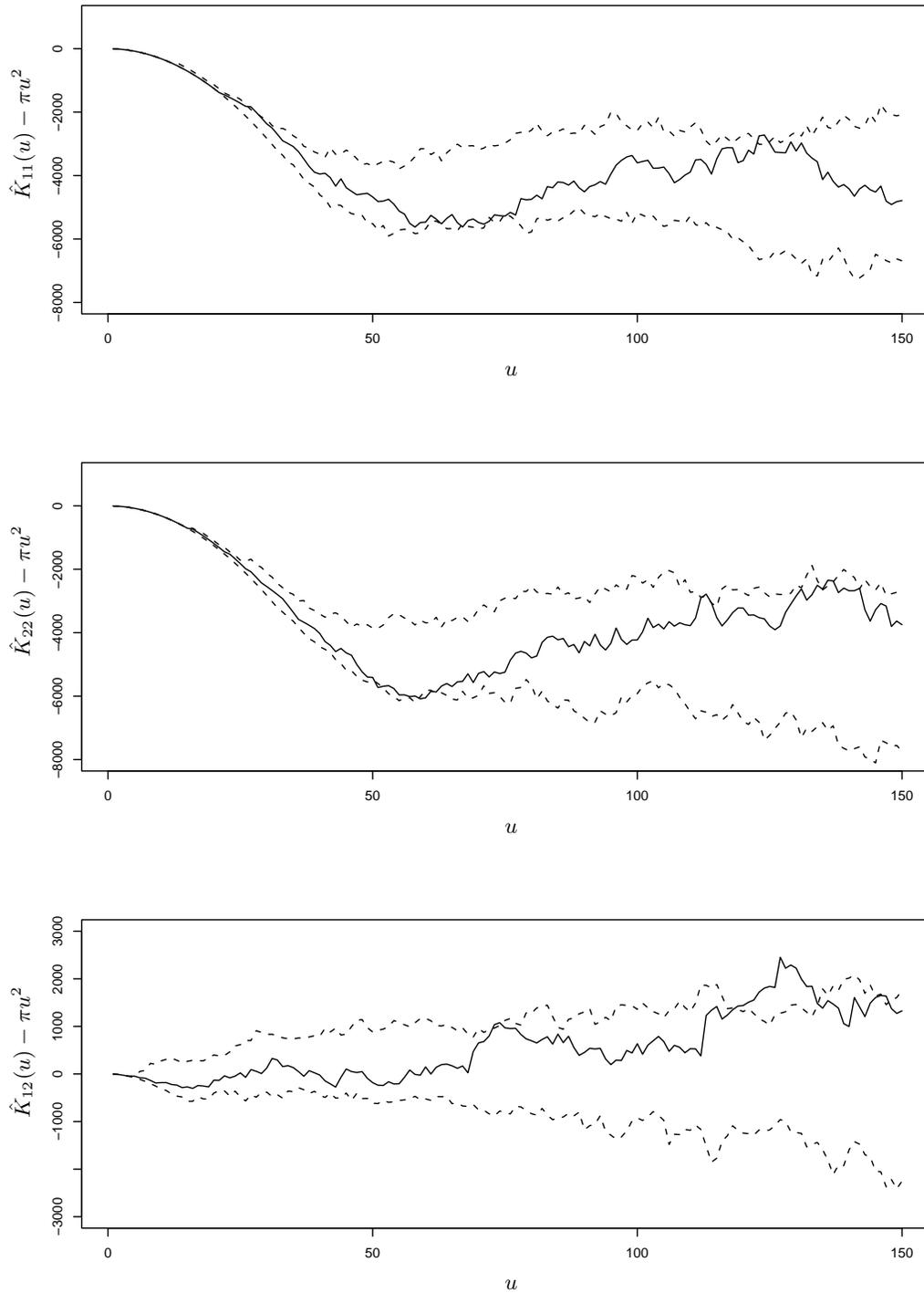
Fig. 8. Goodness-of-fit assessment for the fitted pairwise interaction process with common parameter value for on and off cells. The solid lines are the estimates $\hat{K}_{ij}(u) - \pi u^2$ for the data, the dashed lines are the pointwise envelopes from 99 simulations of the fitted model.

5 Conclusions

5.1 Statistical summary

The bivariate pattern of the displaced amacrine cells is well described by a pairwise interaction point process with functional independence between the two component processes, i.e. the interaction between the on and off cells has a simple inhibitory form, $h(u) = 0$ for $u < \delta_{12}$, $h(u) = 1$ for $u \geq \delta_{12}$, with estimated value $\hat{\delta}_{12} = 4.9\mu$m.

The two component patterns can be fitted with a common interaction function of the form (12) with estimated parameter values $\hat{\delta}_{jj} = 10.0\mu$m, $\hat{\phi} = 49.1\mu$m, $\hat{\alpha} = 2.92$. The resulting fitted interaction function represents a strongly inhibitory interaction within each component pattern, with an effective inhibition distance of about 20μm and an effective range of about 90μm.

5.2 Biological implications

The results from the bivariate analysis indicate that there is a small spatial dependency between the positioning of the on and off cells, since one of our conclusions is that δ_{12} is non-zero. This may appear to conflict with earlier assumptions of independence between the two types (Diggle, 1986). However, one advantage of the likelihood-based analysis over the earlier approach is that we can create 95% confidence intervals (here $2.3 \leq \delta_{12} < 5.0$). Hence, the interaction distance between the on and off cells is around 5 μm at most, which is smaller than the typical cell diameter ($\sim 10\mu$m). Therefore, any dependency in the positioning of the two types is quite weak. However, this dependency might reflect some early positioning constraints between the two cell types, as would occur if, for example, immature cells were positioned in the same layer before migrating to separate layers at a later developmental stage. By repeating the analysis on many data sets of cholinergic amacrine cells, we aim to determine how consistent is the evidence for this weak dependence between patterns formed by the two types of cell.

References

Baddeley, A.J. (1999). Spatial sampling and censoring. In *Stochastic Geometry: Likelihood and Computation*, eds. O.E. Barndorff-Nielsen, W.S. Kendall and M.N.M. Van Lieshout, 37–78. London : Chapman and Hall.

Baddeley, A.J., Moller, J. and Waagepetersen, R. (2000). Non and semi parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, **54**, 329–50.

Baddeley, A.J., Moyeed, R.A., Howard, C.V. and Boyde, A. (1993). Analysis of a three-dimensional point pattern with replication. *Applied Statistics*, **42**, 641–68.

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

Besag, J., Milne, R. and Zachary, S. (1982). Point process limits of lattice processes. *Journal of Applied Probability*, **19**, 210–6.

Brandon, C. (1987). Cholinergic neurons in the rabbit retina: dendritic branching and ultrastructural connectivity. *Brain Research*, **426**, 119-30.

Diggle, P.J. (1986). Displaced amacrine cells in the retina of a rabbit : analysis of a bivariate spatial point pattern. *Journal of Neuroscience Methods*, **18**, 115–25.

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

Diggle, P.J. and Gratton, R.J. (1984). Monte Carlo methods of inference for implicit statistical models (with discussion). *Journal of the Royal Statistical Society*, B **46**, 193-227.

Diggle, P.J., Lange, N. and Benes, F.M. (1991). Analysis of variance for replicated spatial point patterns in clinical neuroanatomy. *Journal of the American Statistical Association*, **86**, 618–25.

Diggle, P.J., Mateu, J. and Clough, H.E. (2000). A comparison between parametric and non-parametric approaches to the analysis of replicated spatial point patterns. *Advances in Applied Probability*, **32**, 331–43.

Euler, T., Detwiler, P.B. and Denk, W. (2002). Directionally selective calcium signals in dendrites of starburst amacrine cells. *Nature*, **418**, 845–52.

Famiglietti, E.V. (1983). "Starburst" amacrine cells and cholinergic neurons: mirror-symmetric on and off amacrine cells of rabbit retina. *Brain Research*, **261**, 138–44.

Famiglietti, E.V. (1985). Starburst amacrine cells: morphological constancy and systematic variation in the anisotropic field of rabbit retinal neurons. *Journal of Neuroscience*, **5**, 562-77.

Gates, D.J. and Westcott, M. (1986). Clustering estimates in spatial point distributions with unstable potentials. *Annals of the Institute of Statistical Mathematics*, **38 A**, 55-67.

Geyer, C. (1999). Likelihood inference for spatial point processes. In *Stochastic Geometry: likelihood and computation*, ed O.E. Barndorff-Nielsen, W.S. Kendall and M.N.M. van Lieshout, 79–140.

Geyer, C.J. and Thompson, E.A. (1992). Constrained Monte Carlo maximum likelihood for dependent data (with Discussion). *Journal of the Royal Statistical Society*, B **54**, 657–99.

Hughes, A. (1985) New perspectives in retinal organisation. *Progress in Retinal Research* **4** 243–314.

Kelly, F. P. and Ripley, B. D. (1976). A note on Strauss' model for clustering. *Biometrika*, **63**, 357–60.

Lotwick, H.W. and Silverman, B.W. (1982). Methods for analysing spatial processes of several types of points. *Journal of the Royal Statistical Society*, B **44**, 406–13.

Moller, J. and Waagepetersen, R.P. (2004). Statistical Inference and Simulation for Spatial Point Processes. London : Chapman and Hall.

Ohser, J. and Stoyan, D. (1981). On the second-order and orientation analysis of planar stationary point processes. *Biometrical Journal*, **23**, 523–33.

Ripley, B.D. (1976). The second-order analysis of stationary point processes. *Journal of Applied Probability*, **13**, 255–66.

Ripley, B.D. (1977). Modelling spatial patterns (with discussion). *Journal of the Royal Statistical Society*, B **39**, 172–212.

Ripley, B.D. and Kelly, F.P. (1977). Markov point processes. *Journal of the London Mathematical Society*, **15**, 188–92.

Stein, M.L. (1991). A new class of estimators for the reduced second moment measure of point processes. *Biometrika*, **78**, 281–6.

Strauss, D. J. (1975). A model for clustering. *Biometrika*, **62**, 467–75.

Tauchi, M. and Masland, R.H. (1984). The shape and arrangement of the cholinergic neurons in the rabbit retina. *Proceedings of the Royal Society, London*, B **23**, 101–19.

Van Lieshout, M.N.M. (2000). *Markov Point Processes and their Applications*. London : Imperial College Press.

Wienawa-Narkiewicz, E. (1983). *Light and Electron Microscopic Studies of Retinal Organisation*. Doctoral Thesis, Australian National University, Canberra.