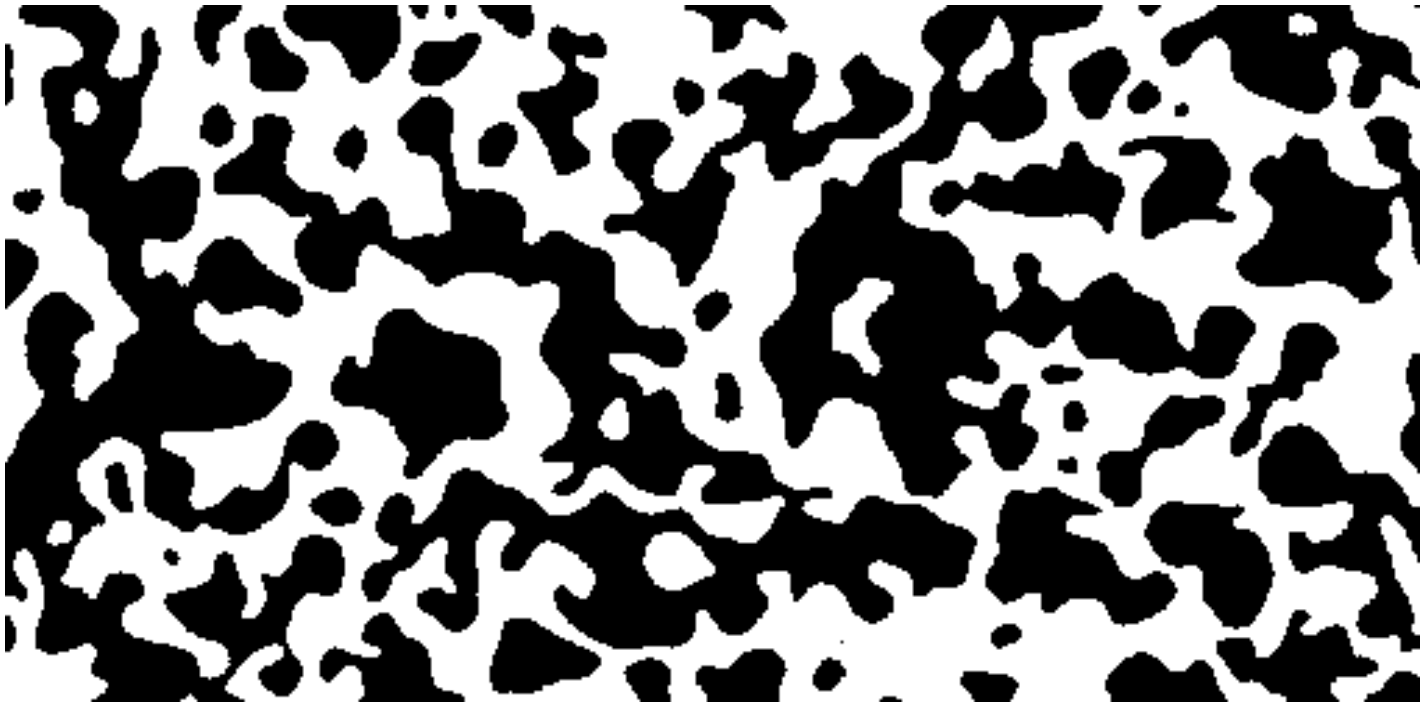
Likelihood inference and computational problems for a random set model

Jesper Møller, Aalborg University, Denmark

March 20, 2009

Heather dataset (Diggle, 1981)



Heather plants in a 10×20m window W (Jädraås, Sweden).

Grow from seedlings into (roughly) hemispherical bushes $\Rightarrow$ $\cup$ heather plants "seen from above" $=$ union of discs, $\mathcal{U}_X$, where X is a disc process.

Observe only $\mathcal{U}_X \cap W$.

$\mathcal{U}_X$ = germ-grain model:

General germ-grain model:

$$\mathcal{U}_X = \bigcup_i (u_i + K_i)$$

where

- the *germs* $\{u_i\} \subset \mathbb{R}^d$ form a (locally finite) point process,
- the *primary grains* $K_i \subset \mathbb{R}^d$ are random compact sets.

Heather dataset: $K_i = b(0, r_i)$ and

$$\mathcal{U}_X = \bigcup_i b(u_i, r_i).$$

Remarks

- *Theorem*: Any random closed set

  (i.e., a locally finite union of compact convex random sets)

  $\sim$ germ-grain model with convex and compact grains.

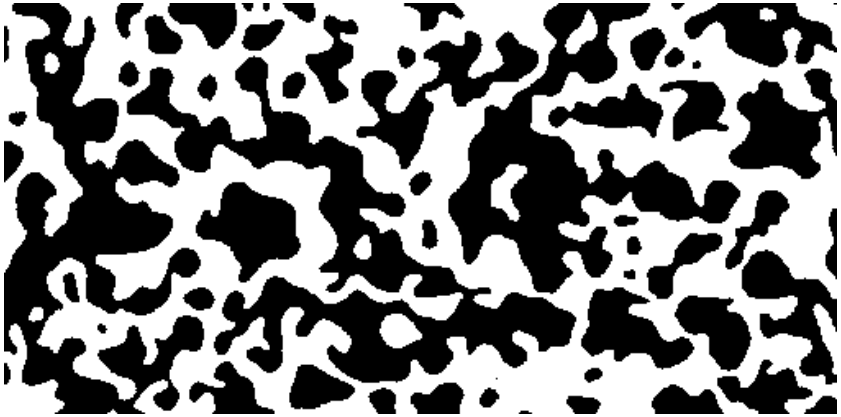
- However,

  - in practice, need a much smaller class of models;
  - often a Poisson disc process (Boolean model).

Poisson disc process

- $\{u_i\} \sim$ Poisson point process
- $\{r_i\}$ IID and independent of $\{u_i\}$
- Advantages: well-studied (moment results).
- Disadvantages/complications:
  - lack of interaction;
  - grains unobservable $\Rightarrow$ density/likelihood for $\mathcal{U}_X$???
  - MCMC missing data approach, simulating $X|\mathcal{U}_X \cap W$:
    not an option: grains only $\approx$ discs & digital image
    $\Rightarrow$ difficult to indentify circular structures.

Fitting a Poisson disc model (Diggle, 1981)



- Stationary; mark distribution = truncated Weibull.
- Fitted by a minimum contrast method *.
- Simulations: visual impression not good.
- "A model incorporating interaction may be appropriate".

* Further (non-likelihood-based) approaches:
Dupač (1980), Serra (1980), Hall (1985, 1988), Ripley (1988), Cressie (1993), Stoyan *et al.* (1995), Molchanov (1997), ...

Our (non-Poisson) model (Møller & Helisova, 2008a,b)

- Pragmatic approach:

a) Specify and fit a Poisson disc process.

b) Extend it to a certain interacting disc process model...

- ... which makes it possible to handle

(1) edge effects;

(2) individual grains are unobservable;

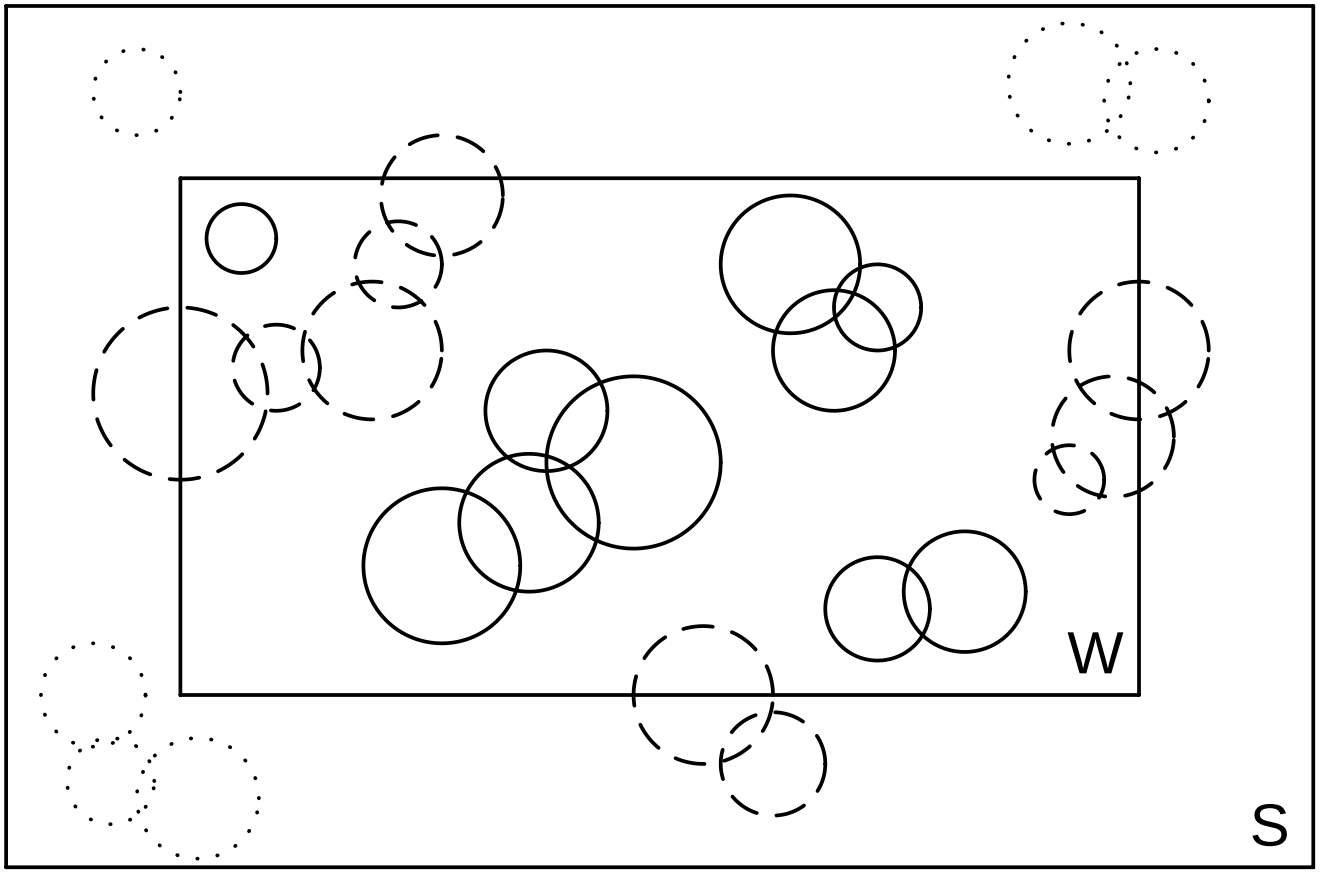
- ... and which provides (the first work on) simulation-based likelihood inference for a germ-grain model.

Some notation

Identify a finite collection of discs with a marked point pattern

$$x = \{(u_1, r_1), \ldots, (u_n, r_n)\} \subset S \times [0, \infty)$$

where $S \supset W$ is the (unknown!) bounded region for heather plant centers.



The density

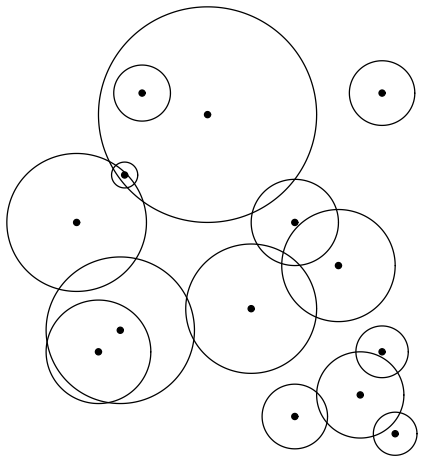
with respect to reference/fitted Poisson disc process,

$$f_\theta(x) = \frac{1}{c_\theta} \exp\left(\theta_1 A(\mathcal{U}_x) + \theta_2 L(\mathcal{U}_x) + \theta_3 N_{\mathrm{cc}}(\mathcal{U}_x) + \theta_4 N_{\mathrm{h}}(\mathcal{U}_x) + \ldots\right)$$

where

$$A = \text{area}, \quad L = \text{perimeter},$$

$$N_{\mathrm{cc}} = \#\ \text{connected components}, \quad N_{\mathrm{h}} = \#\ \text{holes}.$$



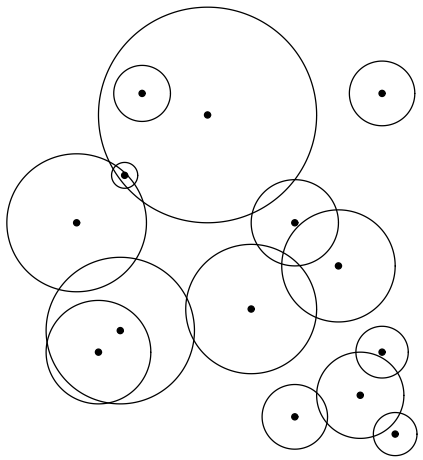
A connected component Markov process

(Baddeley & Møller, 1989)

$$f_\theta(x) = \frac{1}{c_\theta} \prod_{y \subseteq x \text{ con. comp.}} \phi_\theta(y)$$

where

$$\phi_\theta(y) = \exp\left(\theta_1 A(\mathcal{U}_y) + \theta_2 L(\mathcal{U}_y) + \theta_3 + \theta_4 N_{\mathrm{h}}(\mathcal{U}_y)\right)$$

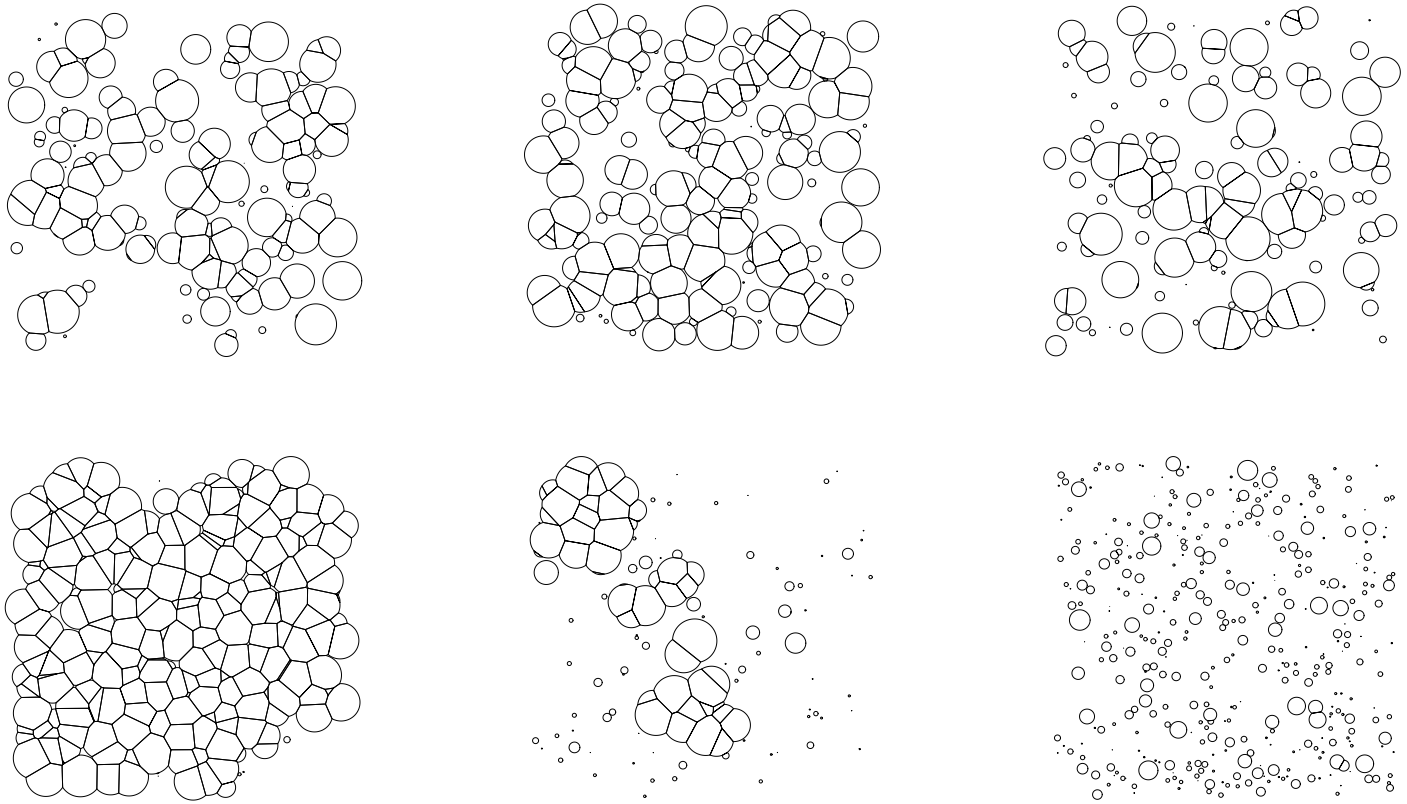


Special cases

- *Area-interaction process* (with varying radii) if $\theta_2 = \theta_3 = \theta_4 = 0$ and $\theta = \theta_1$:
$$f_\theta(x) = \frac{1}{c_\theta} \exp\left(\theta A(\mathcal{U}_x)\right)$$
(Widom & Rowlinson, 1970; Baddeley & Van Lieshout, 1995).

- *Quermass-interaction process* if $\theta_3 + \theta_4 = 0$ and $\theta = (\theta_1, \theta_2, \theta_3)$:
$$f_\theta(x) = \frac{1}{c_\theta} \exp\left(\theta_1 A(\mathcal{U}_x) + \theta_2 L(\mathcal{U}_x) + \theta_3 \chi(\mathcal{U}_x)\right), \quad \chi = N_{\mathrm{cc}} - N_{\mathrm{h}}$$
(Kendall, Van Lieshout & Baddeley, 1999).

- *Continuum random-cluster model* (with varying radii) if $\theta_1 = \theta_2 = \theta_4 = 0$ and $\theta = \theta_3$:
$$f_\theta(x) = \frac{1}{c_\theta} \exp\left(\theta N_{\mathrm{cc}}(\mathcal{U}_x)\right)$$
(Klein, 1982).



1) Reference Poisson disc process: $S = [0, 30] \times [0, 30]$, intensity $= 0.2$, IID radii $\sim$ Uniform$[0, 2]$.
2) $(\theta_1, \theta_2, \theta_3, \theta_4) = (0.1, 0, 0, 0)$.
3) $(\theta_1, \theta_2, \theta_3, \theta_4) = (-0.1, 0, 0, 0)$.
4) $(\theta_1, \theta_2, \theta_3, \theta_4) = (0.6, -1, 1, 0)$.
5) $(\theta_1, \theta_2, \theta_3, \theta_4) = (0.6, -1, 2, 0)$.
6) $(\theta_1, \theta_2, \theta_3, \theta_4) = (0.6, -1, 5, 0)$.

Parameter space

Under mild conditions

$$f_\theta(x) = \frac{1}{c_\theta} \exp\left(\theta_1 A(\mathcal{U}_x) + \theta_2 L(\mathcal{U}_x) + \theta_3 N_{\mathrm{cc}}(\mathcal{U}_x) + \theta_4 N_{\mathrm{h}}(\mathcal{U}_x)\right)$$

is a regular exponential family model with parameter space

$$\Theta = \mathbb{R}^4 \quad \text{if the radii are bounded,}$$

or in general

$$\Theta = \{(\theta_1, \theta_2, \theta_3, \theta_4) \in \mathbb{R}^4 : \int \exp\left(\pi\theta_1 r^2 + 2\pi\theta_2 r\right) Q(\mathrm{d}r) < \infty)\}$$

where Q is the mark distribution under the reference Poisson disc process.

A fundamental characteristic

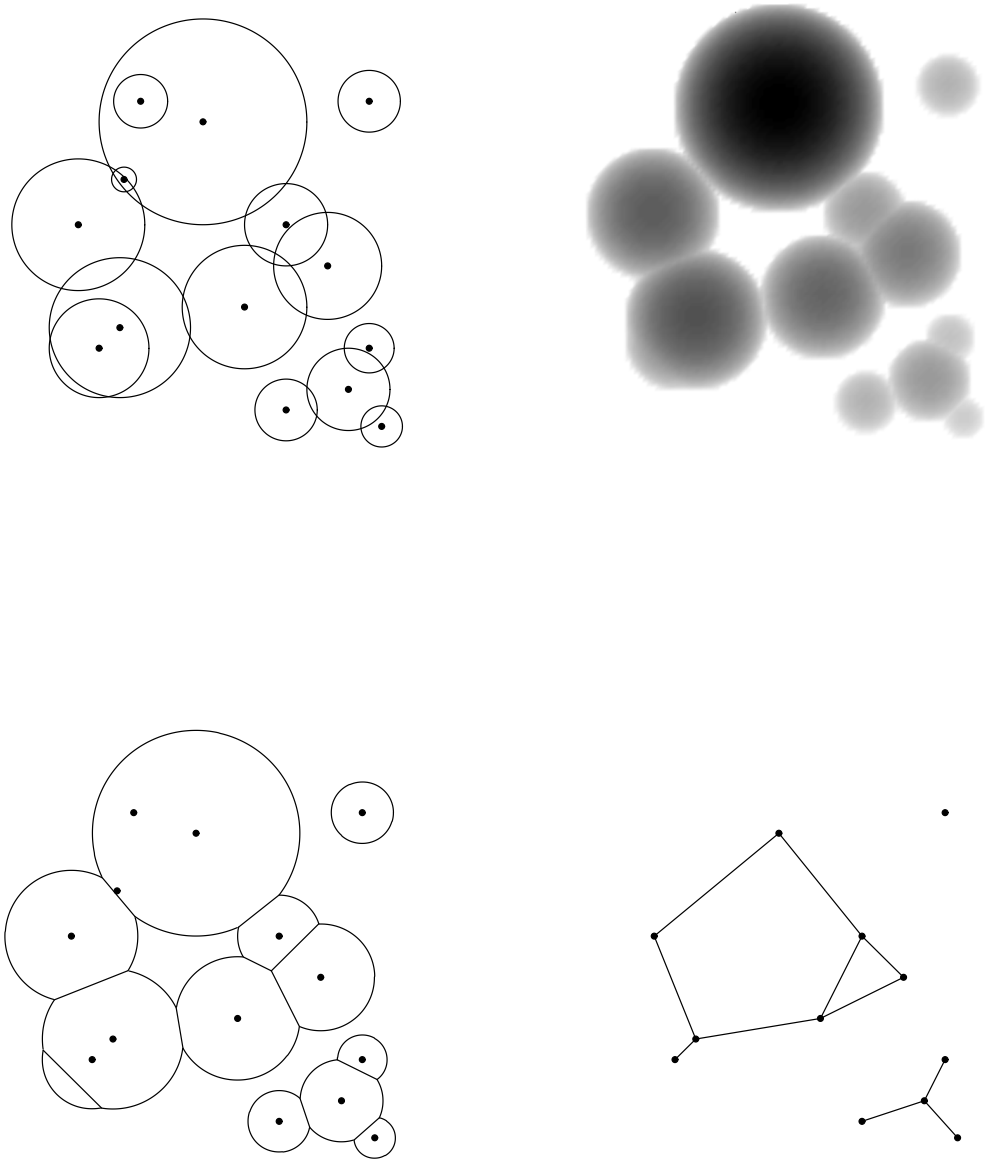
(Papangelou) conditional intensity:

$$\lambda_\theta(x,(u,r)) = f_\theta(x \cup \{(u,r)\})/f_\theta(x)$$

for $x = \{(u_1,r_1),\ldots,(u_n,r_n)\} \subset S \times [0,\infty)$, $(u,r) \in S \times [0,\infty)$.

- $\lambda_\theta(x,(u,r)) =$ "local characteristic";
  $-\log \lambda_\theta(x,(u,r)) =$ "local energy".
- Important because
  - $\lambda_\theta \leftrightarrow f_\theta$ but λ_θ does not involve c_θ;
  - specifies local Markov properties;
  - appears in the Hastings ratio of Metropolis-Hastings birth-death algorithm (Geyer & Møller, 1994; Green, 1995).
- In many cases, local stability and hence geometric ergodicity.

Key tool (for many things incl. λ_θ): Power tessellation



Book keeping in Metropolis-Hastings birth-death algorithm

(www.math.aau.dk/∼jm/Codes.union.of.discs)

Keep track on $(A, L, N_{\mathrm{cc}}, N_{\mathrm{h}})$ by also keeping track on

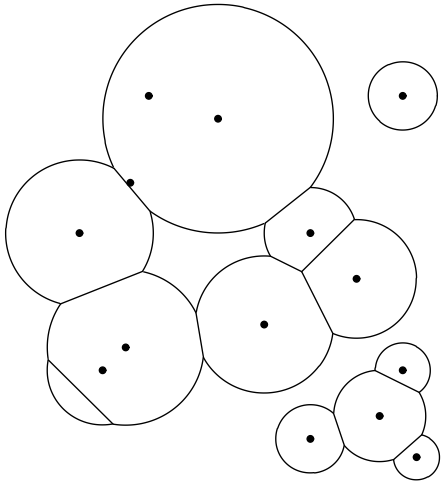
$N_{\mathrm{c}} = \#$ cells, $\quad N_{\mathrm{ie}} = \#$ interior edges, $\quad N_{\mathrm{iv}} = \#$ interior vertices,

and using

$$A = \sum \text{areas of cells}$$

$$L = \sum \text{lengths of arcs}$$

$$N_{\mathrm{h}} = N_{\mathrm{cc}} - N_{\mathrm{c}} + N_{\mathrm{ie}} - N_{\mathrm{iv}}$$

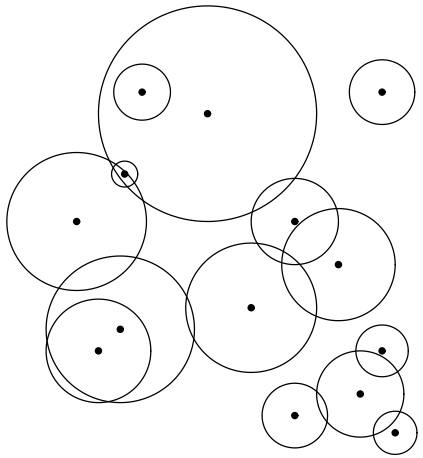


Local Markov property

$$\lambda_\theta(x,(u,r)) =$$

$$\exp\Big[\theta_1\left(A(\mathcal{U}_{x\cup\{(u,r)\}}) - A(\mathcal{U}_x)\right) + \theta_2\left(L(\mathcal{U}_{x\cup\{(u,r)\}}) - L(\mathcal{U}_x)\right) + \theta_3\left(N_{\mathrm{cc}}(\mathcal{U}_{x\cup\{(u,r)\}}) - N_{\mathrm{cc}}(\mathcal{U}_x)\right) + \theta_4\left(N_{\mathrm{h}}(\mathcal{U}_{x\cup\{(u,r)\}}) - N_{\mathrm{h}}(\mathcal{U}_x)\right)\Big]$$

depends on x only through those $\mathcal{U}_x$-cc which intersect $b(u,r)$.

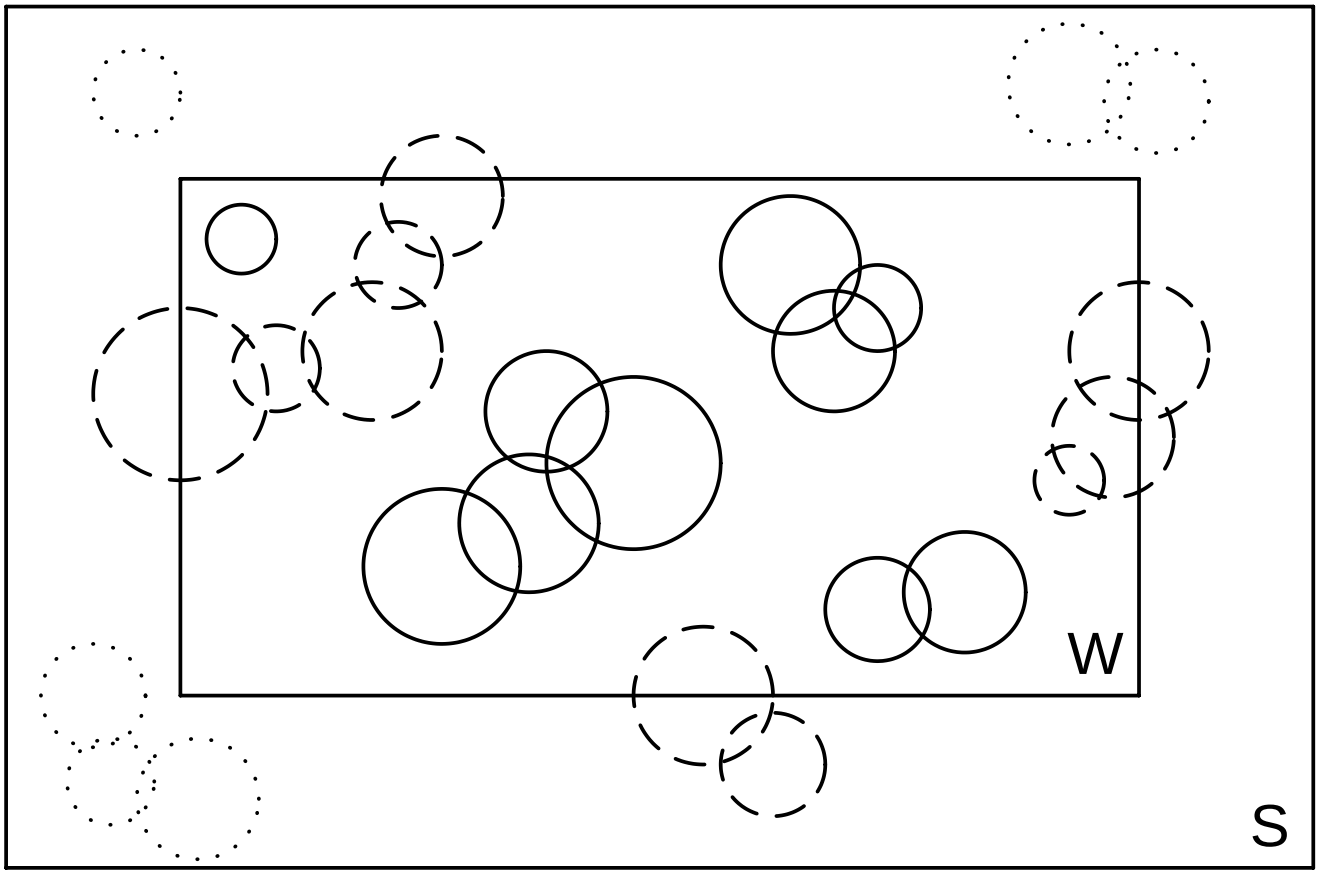


If $\theta_3 = \theta_4$: depends only on those $b(u_i, r_i)$ which intersect $b(u,r)$.

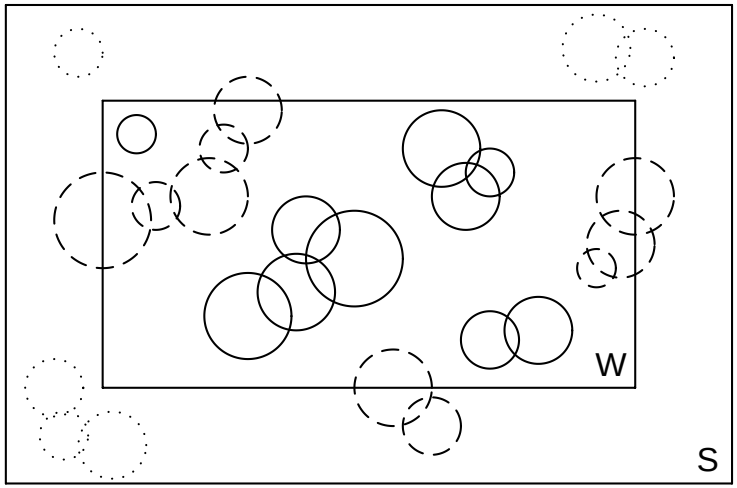
Handle edge effects by exploting a spatial Markov property

Split X into

$X^{(a)}$ (full circles), $X^{(b)}$ (dashed circles), $X^{(c)}$ (dotted circles)



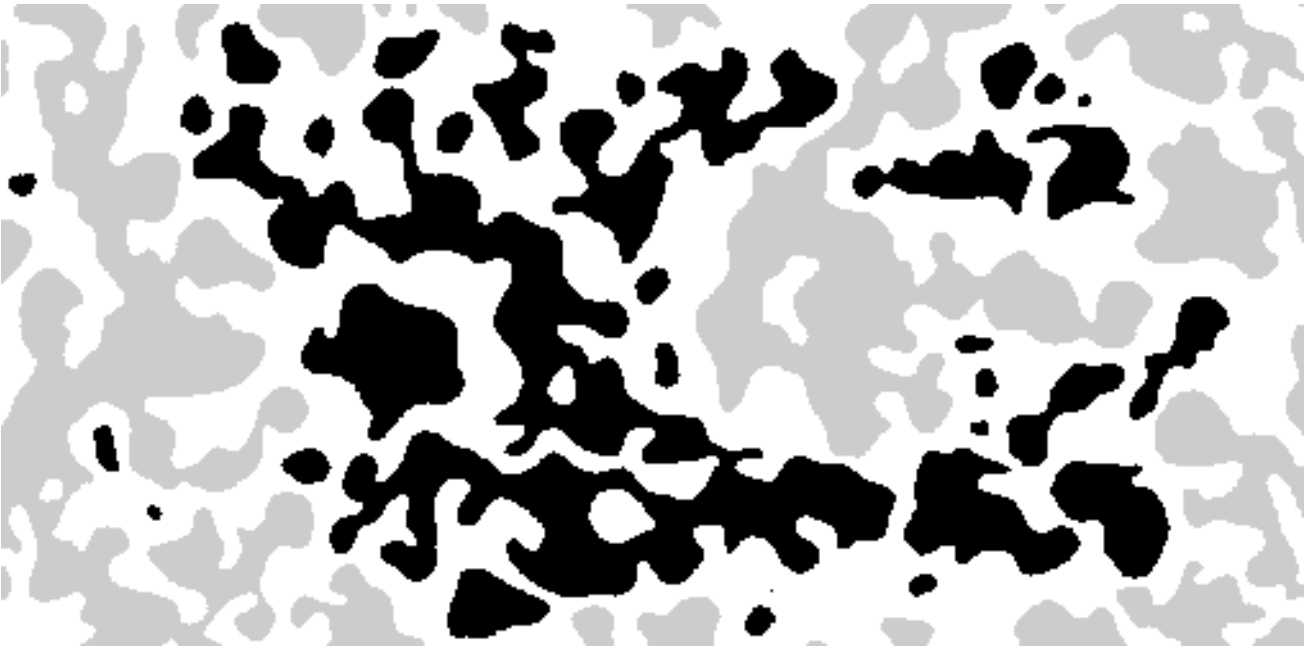
Spatial Markov property



- $X^{(a)}$ and $X^{(c)}$ are conditionally independent given $X^{(b)} = x^{(b)}$;
- $X^{(a)}|x^{(b)} \sim X^{(a)}|V$, where $V := W \cap \mathcal{U}_{x^{(b)}}$;
- for feasible $x^{(a)}$ (i.e. $\mathcal{U}_{x^{(a)}} \subseteq W$ and $\mathcal{U}_{x^{(a)}} \cap V = \emptyset$),

$$f_\theta(x^{(a)}|V) = \frac{1}{c_\theta(V)} \exp\left(\theta_1 A(\mathcal{U}_{x^{(a)}}) + \theta_2 L(\mathcal{U}_{x^{(a)}}) + \theta_3 N_{\mathrm{cc}}(\mathcal{U}_{x^{(a)}}) + \theta_4 N_{\mathrm{h}}(\mathcal{U}_{x^{(a)}})\right).$$

Heather data: conditional log likelihood

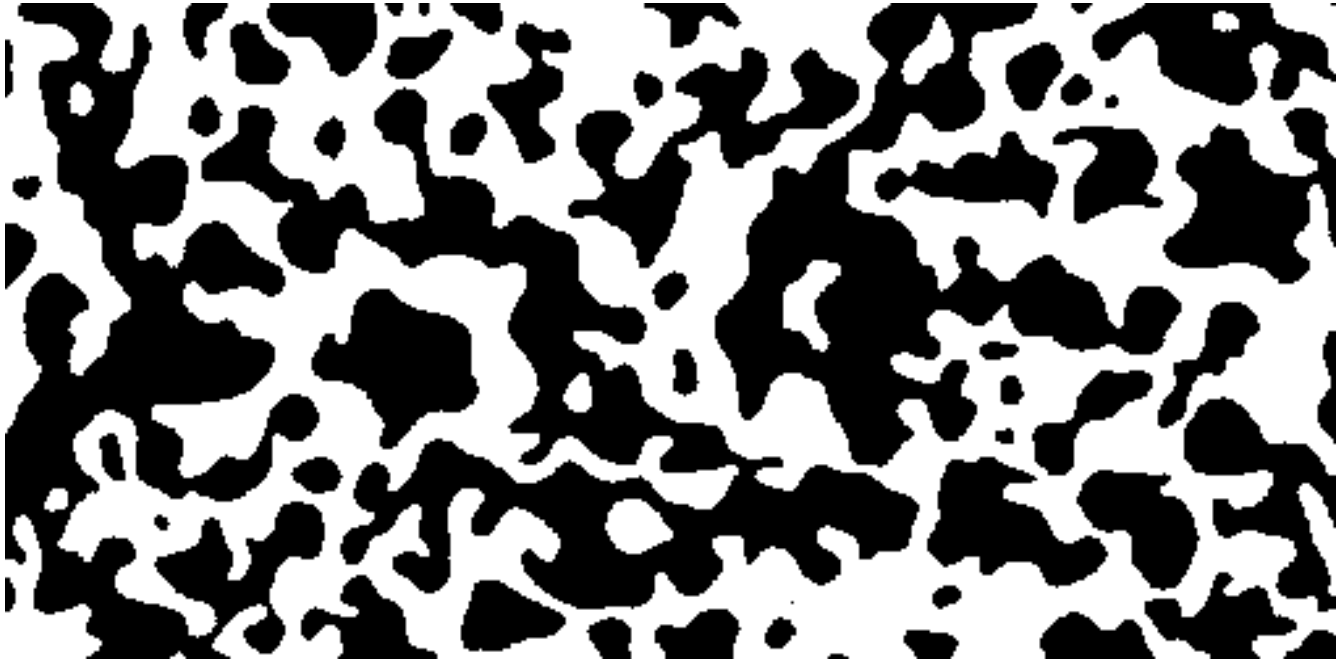


$$
\begin{aligned}
L_c(\theta) &= \log f_\theta(x^{(a)}|V) \\
&= \theta_1 A(\mathcal{U}_{x^{(a)}}) + \theta_2 L(\mathcal{U}_{x^{(a)}}) + \theta_3 N_{\mathrm{cc}}(\mathcal{U}_{x^{(a)}}) + \theta_4 N_{\mathrm{h}}(\mathcal{U}_{x^{(a)}}) - \log c_\theta(V)
\end{aligned}
$$

with $c_\theta(V)$ approximated by MCMC methods.

Unconditional log likelihood

Pretend $S = W$. Let $Y = \cup_i b(u_i, r_i) \cap W$.



Approximate the log (unconditional) likelihood by *ignoring edge effects*,

$$L_u(\theta) = \theta_1 A(Y) + \theta_2 L(Y) + \theta_3 N_{\mathrm{cc}}(Y) + \theta_4 N_{\mathrm{h}}(Y) - \log c_\theta$$

with c_θ approximated by MCMC methods.

Some remarks

- Rather different "conditional" and "unconditional" MLE's.
- Preferable(?) to use $L_c(\theta)$ (accounts for edge effects); the effect of using $L_u(\theta)$ is less well-understood.
- Loss of information in using $L_c(\theta)$:

  $$A(Y) = 100.28, \quad L(Y) = 382.82, \quad N_{\mathrm{cc}}(Y) = 56, \quad N_{\mathrm{h}}(Y) = 6,$$

  are about twice as large as

  $$A(\mathcal{U}_{x^{(a)}}) = 45.6, \quad L(\mathcal{U}_{x^{(a)}}) = 190, \quad N_{\mathrm{cc}}(\mathcal{U}_{x^{(a)}}) = 32, \quad N_{\mathrm{h}}(\mathcal{U}_{x^{(a)}}) = 2.$$

- Specification of reference Poisson disc process is important. (We compared results using 3 different reference processes.)

- But the same overall conclusion (Wald tests; summaries):

$$\theta_1 < 0, \quad \theta_2 > 0, \quad \theta_3 < 0, \quad \theta_4 = 0,$$

so

$$f_\theta(x) = \frac{1}{c_\theta} \exp\left(\theta_1 A(\mathcal{U}_x) + \theta_2 L(\mathcal{U}_x) + \theta_3 N_{\mathrm{cc}}(\mathcal{U}_x)\right).$$

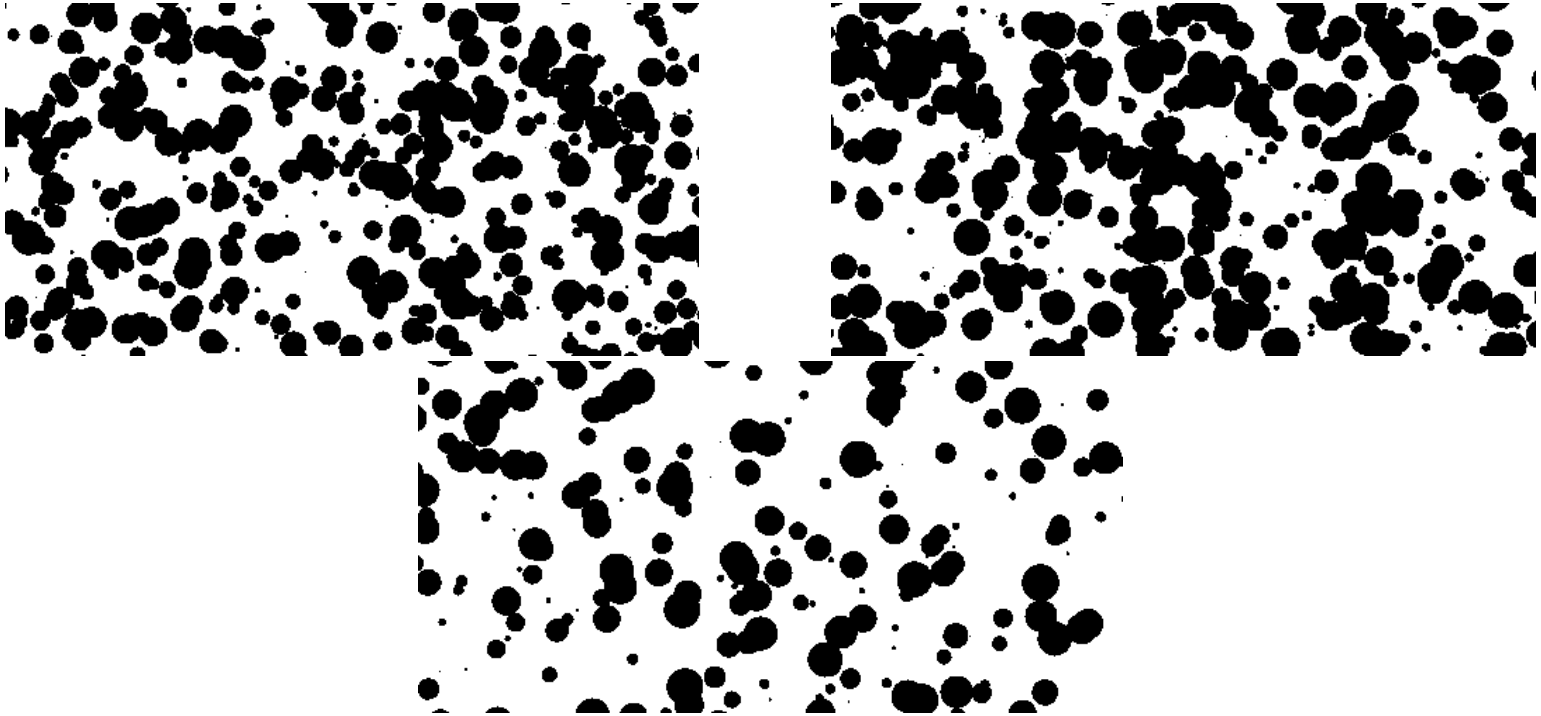
- The quermass-interaction process, i.e. $\theta_3 + \theta_4 = 0$ and

$$f_\theta(x) = \frac{1}{c_\theta} \exp\left(\theta_1 A(\mathcal{U}_x) + \theta_2 L(\mathcal{U}_x) + \theta_3 \chi(\mathcal{U}_x)\right), \quad \chi = N_{\mathrm{cc}} - N_{\mathrm{h}}$$

was rejected.

Some simulations of 3 reference Poisson disc processes

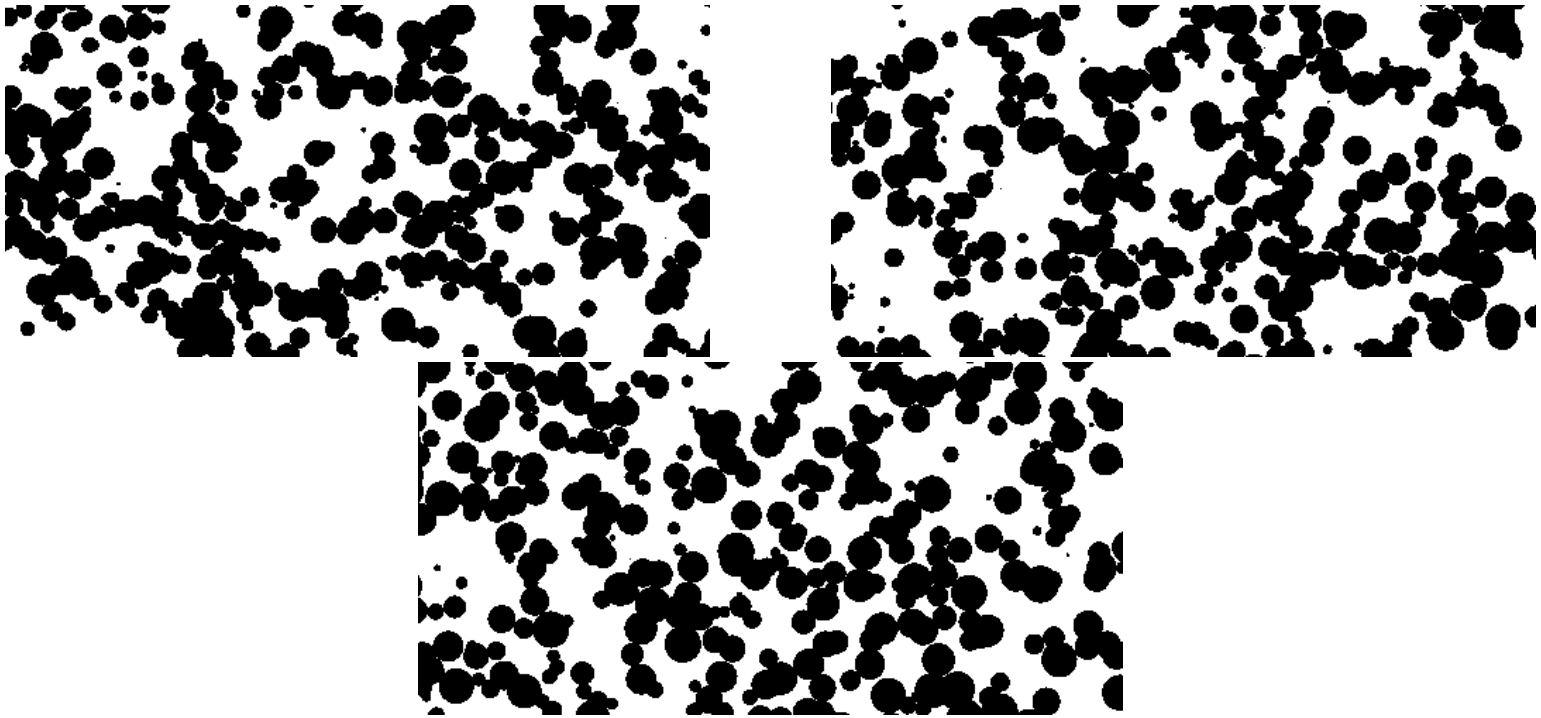
First two: fitted by a method from Hall (1985)
(using a truncated normal and uniform mark dist.).
Third: (partly) taken from Laslett *et al.* (1985)



Misfit in all 3 cases: too many small components.

Third case: too low area fraction.

Some simulations of 3 fitted (A, L, N_{cc})-interaction models



Model control based on various contact distribution functions and covariance function: no misfit for the two first models.

Shape characteristics: some misfit, possibly since the heather data are rather smooth while a disc process is naturally more rugged.

References

Møller, J. & Helisova, K. (2008). Power diagrams and interaction processes for unions of discs. *Advances in Applied Probability*, **40**, 321-347.

Møller, J. & Helisova, K. (2008). Likelihood inference for unions of interacting discs. Research Report R-2008-18, Department of Mathematical Sciences, Aalborg University. *Submitted*.

Software: www.math.aau.dk/∼jm/Codes.union.of.discs