Submitted to the Annals of Applied Statistics
2017 Vol. , No. , 1–8
arXiv: arXiv:0000.0000

SUPPLEMENT B: THE MCMC SAMPLER

By Jan-Otto Hooghoudt and Rasmus Waagepetersen

In this supplementary material to "Towards Bayesian Inference of the Spatial Distribution of Proteins from Three-Cube FRET Data" we give a detailed description of the Markov chain Monte Carlo sampler used to draw samples from the posterior distribution (equation (3.1) in the main article) in the first section. In the second section we briefly discuss the likelihood expressions we apply in the case the three-cube FRET channel intensities can not consist of negative values (i.e. empirical data).

1. Steps in the MCMC sampler. The steps involved in the sampler are illustrated in Figure 1 and are described in detail below.

Step 1: Generate initial configuration for the posterior point pattern. An initial point pattern $\mathbf{X}_0 = (\mathbf{X}_D^0, \mathbf{X}_A^0)$ from which the chain starts is generated from the prior distribution by using the R-software (R Core Team, 2014) and the package *spatstat* (Baddeley and Turner, 2005). The initial point pattern is simulated on a square W corresponding to the area covered by the three-cube FRET channel data. The generated point pattern is stored by writing the coordinates x, y of the points as well as the marks of each point—1 for a donor, 2 for an acceptor—to disk.

Step 2: Generate channel data from the initial posterior point pattern. Channel data conditional on the point pattern $\mathbf{X}_0$ is computed based on equations (2.1)–(2.3) in the main article. To employ these equations the square area W is defined as a union of square pixels C_i, i.e. $W = \cup_{i\in\mathcal{G}} C_i$, with the C_i equal to the pixels in the three-cube FRET channel dataset indexed by $\mathcal{G}$.

To compute the μ_{DD}^i and μ_{DA}^i (defined in Section 2.2. of the main article) the key component is to specify P_{da} for donor and acceptor pairs in $\mathbf{X}$. To keep the computation of transfer probabilities feasible, only those acceptors that reside within $4R_0$ of a donor are taking into account as a possible path for energy transfer for the donor. This important simplification will not lead to any significant difference in posterior results. By replacing P_{dA} in the equation for μ_{DD}^i by its definition $P_{dA} = \sum_{a\in\mathbf{X}_A} P_{da}$, and taking into

account the cut-off radius of $4R_0$, the equations for μ^i_{DD} and μ^i_{DA} become

$$\mu^i_{DD} = M_D \sum_{d \in \mathbf{X}_D \cap C_i} (1 - \sum_{\substack{a \in \mathbf{X}_A \\ r_{da} < 4R_0}} P_{da}), \tag{1}$$

$$\mu^i_{DA} = M_D G \sum_{a \in \mathbf{X}_A \cap C_i} \sum_{\substack{d \in \mathbf{X}_D \\ r_{da} < 4R_0}} P_{da}, \tag{2}$$

with r_{da} the distance between a donor d and an acceptor a and (as previously defined in Section 2.1 of the main article)

$$P_{da} = \frac{(R_0/r_{da})^6}{1 + S_d} \quad \text{with} \quad S_d = \sum_{\substack{\widetilde{a} \in \mathbf{X}_A \\ r_{d\widetilde{a}} < 4R_0}} (R_0/r_{d\widetilde{a}})^6.$$

For clarity and later use we further restate here the equation for μ^i_{AA} as previously defined in Section 2.2 of the main article

$$\mu^i_{AA} = M_D/K \sum_{a \in \mathbf{X}_A \cap C_i} 1. \tag{3}$$

To benefit from the approach of excluding transfer probabilities P_{da} for donor and acceptors pairs which are further than $4R_0$ from each other, it is necessary to store donors and acceptors pixel wise. We have used so called *linked lists* to implement this. A linked list can be viewed as a list containing boxes and each box stores the values of some variables as well as a *pointer* to the next box. In our program for each pixel there are two linked-lists available, one storing the information concerning the donor points within the pixel and the other storing the information concerning the acceptor points within the pixel. To each point (donor or acceptor) corresponds precisely one box in the corresponding linked list. For an acceptor point a the box contains the x, y coordinates of a. For a donor point d also the value of S_d is stored. Storing of S_d gives the possibility to compute the channel data very efficiently when a proposal update for the acceptor point pattern is made. This is further discussed under step 6.

In our program the sums on the right of the factors M_D, $M_D G$ and M_D/K in, respectively, equations (1), (2) and (3) are available and stored at every

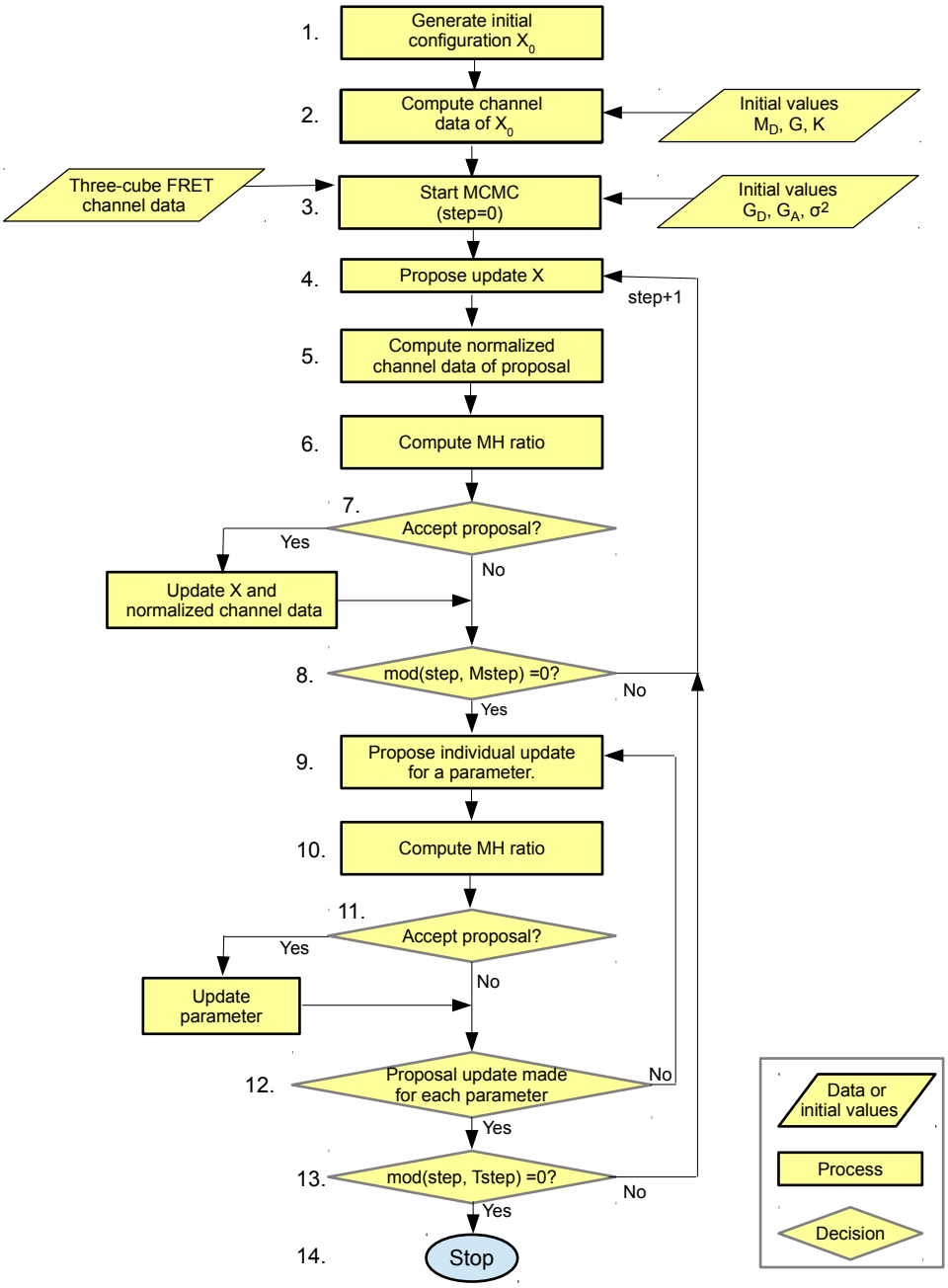

FIG 1. *Flow chart of the steps involved in the MCMC sampler.*

step in the three matrices XNDD, XNDA and XNA containing the elements

$$\begin{aligned}
\mathrm{XNDD}^i &= \sum_{d\in\mathbf{X}_D\cap C_i}(1-\sum_{\substack{a\in\mathbf{X}_A\\ r_{da}<4R_0}}P_{da})\\
&= \sum_{d\in\mathbf{X}_D\cap C_i}(1+S_d)^{-1},\\
\mathrm{XNDA}^i &= \sum_{a\in\mathbf{X}_A\cap C_i}\sum_{\substack{d\in\mathbf{X}_D\\ r_{da}<4R_0}}\frac{(R_0/r_{da})^6}{1+S_d},\\
\mathrm{XNA}^i &= \sum_{a\in\mathbf{X}_A\cap C_i}1,
\end{aligned}$$

for $i = 1,\ldots,n$ and n the total number of pixels. The X in these names refers to the posterior pattern $\mathbf{X}$ and the N in XNDD and XNDA stands for *normalized* as the values of XNDD^i and XNDA^i correspond to the situation where every donor within pixel i is excited exactly one time. For XNA it is appropriate to think of the N in its name to refer to *number*—and we have used only a single A at the end of the name— as XNA is the matrix that stores the number of acceptors within each pixel of the pattern $\mathbf{X}_A$. Equivalent to the latter matrix, also a matrix XND is available within the program which stores the number of donors within each pixel in the pattern $\mathbf{X}_D$. The values of μ^i_{DD}, μ^i_{DA} and μ^i_{AA} are now available at each step by multiplication of XNDD^i, XNDA^i and XNA^i by, respectively, the factors: M_D, M_DG and M_D/K, as specified in (1), (2) and (3).

For the initial posterior point pattern $\mathbf{X}_0$ the value of each XNDD^i, XNDA^i and XNA^i is now computed as follows:

1. The initial point pattern is read from disk and depending on the mark (1 or 2) and the coordinates x, y a point is added to the corresponding donor or acceptor linked lists. For each donor point the value of S_d is set to zero. Within the process the total number of acceptors as well as donors that reside within a pixel i is counted and these values are stored, respectively, in the matrix elements XNA^i and XND^i.
2. An element XNDD^i (initially set to zero) is computed by looping over all donors d within pixel i. For each donors d the distance r_{da} to acceptors a which reside in the same pixel or directly neighboring pixels is calculated.[1] If the distance r_{da} is within $4R_0$, then the corresponding

[1]For square pixels the approach of including only points residing within nearest neighbor pixels in the computation is valid for pixels with a side length $\geq 4R_0$. A condition always satisfied for empirical FRET datasets.

value of $(R_0/r_{da})^6$ is added to S_d (which initially is set to zero). After looping over all acceptors residing in the same or directly neighboring pixels the value S_d is stored and the value of $(1+S_d)^{-1}$ is added to XNDD^i.

3. An element XNDA^i (initially set to zero) is computed by looping over all acceptors a within pixel i and for each acceptor a the distance r_{da} to donors d which reside in the same pixel or directly neighboring pixels is calculated. If the distance r_{da} is within $4R_0$, then the corresponding value of $P_{da} = (R_0/r_{da})^6/(1+S_d)$ is added to XNDA^i.

Step 3: Start MCMC. By the initialization procedure, steps 1-2, and specifying initial values for the microscope parameters the following information is available within the program:

1. the point pattern $\mathbf{X}_0$ is stored as two sets of linked lists containing the coordinates of the donor and acceptor points.
2. the mean channel intensities $\mu^i_{DD}, \mu^i_{DA}, \mu^i_{AA}$ related to the initial point pattern $\mathbf{X}_0$ are available in the form of the three matrices XNDD, XNDA and XNA and the initial values of M_D, G and K.

Further, by specifying the prior distributions of the microscope and point process parameters and making a three-cube FRET data set available in the form of three matrices YDD, YDA and YAA containing the intensity values of respectively the DD-, DA- and AA-channel, the sampling procedure can start.

Step 4: Propose update pattern. A proposal is made to update the point pattern. With probability 1/2 the donor pattern is updated, otherwise the acceptor pattern is updated. Then it is proposed to add or remove a point, each with probability 1/2.

- In the case a point is added, random coordinates for x and y are generated from the uniform distribution on W.
- In the case a donor (acceptor) point is removed from the pattern, a random integer is drawn between 1 and the total number of donor (acceptor) points in the current pattern. It is then proposed to remove the donor (acceptor) point that is labeled by this number. The latter method is implemented by implicitly labeling each donor and acceptor point. As an example, consider for instance a donor point which is stored in linked list number 4 box 8. This point is implicitely labelled by the integer value resulting from adding 8 (of box 8) to the total number of donors in linked lists 1 to 3.

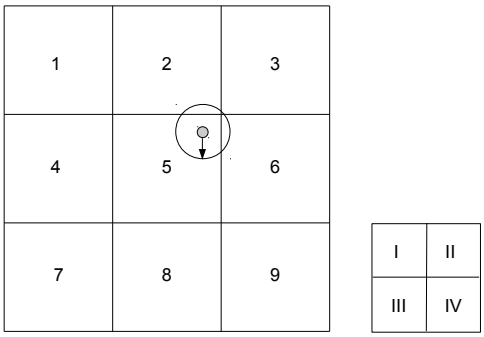

FIG 2. *Adding a donor d_0 (small gray circle) to the point pattern. The larger circle represents the radius of $4R_0$ around the donor. The figure is drawn on a scale with $R_0 = 6$ and the side length of the square pixels equal to 100. The small figure on the right shows how the quadrants within pixel 5 are labeled as used in the text.*

Step 5: Calculate channel data for a proposal update of the point pattern. In order to gain computational speed, we have implemented a rather sophisticated procedure to compute the channel data for a proposal update of the point pattern. The implemented procedure distinguishes between computation of the channel data when an update for the donor pattern $\mathbf{X}_D$ or the acceptor pattern $\mathbf{X}_A$ is proposed. First, we discuss the algorithm for the donor case in terms of proposing to add a donor (proposing to remove a donor is similar). Secondly, we discuss the algorithm for the acceptor case in terms of proposing to add an acceptor (proposing to remove an acceptor is similar).

Compute proposal channel data: Adding a donor. In Figure 2 a schematic representation of adding a donor d_0 (small gray circle) to the current pattern is shown. From the figure it is clear that if the new donor is placed in the second quadrant of a pixel only acceptors that reside within this same pixel, or in the pixels $2, 3$ and 6 can be affected by the newly placed donor.[2] Thereby, the only values of XNDA^i that need to be recomputed are for $i \in \{2, 3, 5, 6\}$. The new value XNDA_p^i for each of the possibly affected XNDA^i is derived by computing the change $\Delta\mathrm{XNDA}^i$ that arises due to adding d_0, and adding it to the current value XNDA_c^i, i.e.

$$\mathrm{XNDA}_p^i = \mathrm{XNDA}_c^i + \Delta\mathrm{XNDA}^i.$$

[2]For square pixels the approach of only including acceptors within pixels which are nearest neighbors of the quadrant where the donor is added is valid for pixels with a side length $\geq 8R_0$. A condition always satisfied for empirical FRET datasets.

The value of each ΔXNDA^i for $i \in \{2,3,5,6\}$ is defined by

$$\Delta\text{XNDA}^i = \sum_{\substack{a \in \mathbf{X}_A \cap C_i \\ r_{d_0 a} < 4R_0}} P_{d_0 a},$$

and easily obtained by: i) computing S_{d_0} for donor d_0 by looping over the acceptors in the pixels $i \in \{2,3,5,6\}$, while storing $r_{d_0 a}$ for those a's for which $r_{d_0 a} < 4R_0$ and ii) adding $P_{d_0 a}$ for each stored a to ΔXNDA^i (which initially was set to zero) when a is located in pixel i.

The sole element of XNDD that is affected by the new donor is the element corresponding to the pixel where the donor is added, that is $i = 5$. The new proposed value for XNDD^5 is available as

$$\text{XNDD}_p^5 = \text{XNDD}^5 + \Delta\text{XNDD}^5,$$

with

$$\Delta\text{XNDD}^5 = 1 - \sum_{i \in \{2,3,5,6\}} \Delta\text{XNDA}^i.$$

Compute proposal channel data: Adding an acceptor. Computing the channel data when an acceptor is added to the current pattern is rather more involved than for the donor case. We explain the procedure by example and in direct relation to the situation depicted in Figure 3 in which an acceptor a_0 is added to a current pattern. The new acceptor point is placed in the second quadrant of pixel 5 and only donors that reside within the pixels $2, 3, 5$ and 6 can be affected by the new acceptor. Donors that reside within the solid circle are affected by the presence of the acceptor as they get an extra path for de-excitation by energy transfer to the newly placed acceptor. Now assume that the donor d_1 (small solid circle in pixel 3) which resides within the radius of $4R_0$ to a_0, has currently the possibility of energy transfer to a number of acceptors a ($r_{da} < 4R_0$) and denote one of these acceptors by a_1. Then in the current situation (a_0 not added) the probability of energy transfer from donor d_1 to acceptor a_1 is defined by

$$P_{d_1 a_1}^{\text{current}} = \frac{(R_0/r_{d_1 a_1})^6}{1 + S_{d_1}},$$

while in the proposed situation (acceptor a_0 added) this becomes

$$P_{d_1 a_1}^{\text{proposed}} = \frac{(R_0/r_{d_1 a_1})^6}{1 + S_{d_1} + (R_0/r_{d_1 a_0})^6}.$$

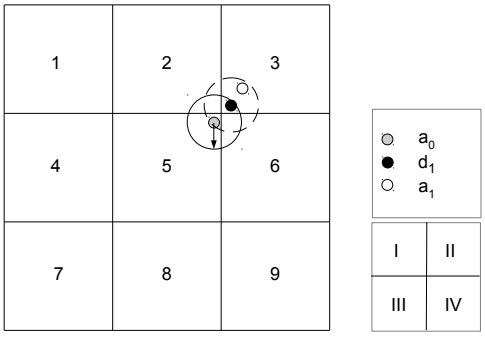

FIG 3. *Adding an acceptor a_0 (small gray circle) to the point pattern. The larger circles around a_0 and d_1 (small solid circle) have a radius of $4R_0$. The figure is drawn on a scale with $R_0 = 6$ and the side length of the square pixels equal to 100. The small figure on the right shows how the quadrants within pixel 5 are labeled as used in the text*

So, by placing a_0 within $4R_0$ of d_1, the probability $P_{d_1a_1}$ for energy transfer from donor d_1 to acceptor a_1 decreases. Clearly, also for other acceptors residing within $4R_0$ of d_1 (other than a_0 and a_1), their respective probabilities to receive energy transfer from donor d_1 will decrease when acceptor a_0 is added. The algorithm for updating the possibly affected elements of XNDA is now as follows:

1. Compute for donors residing within possibly affected pixels, i.e. $d \in \mathbf{X}_D \cap C_i, i \in \{2,3,5,6\}$, the distance r_{da_0} to a_0. If $r_{da_0} < 4R_0$, then store for each of these donors its coordinates as well as the value of $(R_0/r_{da_0})^6$ in a list.
2. Loop pixel wise over the acceptors residing in the possibly affected pixels, i.e. $a \in \mathbf{X}_A \cap C_i, i \in \{2,3,5,6\}$.[3] Determine for each acceptor a the distance r_{da} to any of the donors in the list made in 1. If the distance $r_{da} < 4R_0$, then the value of XNDA_c^i will decrease due to a and d by

$$\frac{(R_0/r_{da})^6}{1 + S_d + (R_0/r_{da_0})^6} - \frac{(R_0/r_{da})^6}{1 + S_d}.$$

   Taking into account all the donors and acceptors of which the transfer probabilities change due to adding a_0, the algorithm computes

[3]For square pixels the approach of only including acceptors within pixels which are nearest neighbors of the quadrant where the proposed acceptor is placed, is valid for pixels with a side length $\geq 16R_0$. A condition in general satisfied for empirical three-cube FRET datasets.

ΔXNDA^i as

$$\Delta\text{XNDA}^i = \sum_{a\in\mathbf{X}_A\cap C_i} \sum_{\substack{j=\\2,3,5,6}} \sum_{\substack{d\in\mathbf{X}_D\cap C_j\\ r_{da_0}<4R_0\\ r_{da}<4R_0}} \Big(\frac{(R_0/r_{da})^6}{1+S_d+(R_0/r_{da_0})^6} - \frac{(R_0/r_{da})^6}{1+S_d}\Big),$$

for $i \in \{2,3,5,6\}$. The values for ΔXNDA^i obtained by this equation are negative or zero as they compute the total decrease in the probability of energy transfer from donors d to acceptors $a \in \mathbf{X}_A \cap C_i$, due to adding a_0 to the current pattern. However, in the pixel where a_0 is added (pixel 5), there will be a possible increase in ΔXNDA, as a_0 can receive transfers from donors d possibly residing within $4R_0$ of it, which has to be added to ΔXNDA^5, i.e.

$$\Delta\text{XNDA}^5 = \Delta\text{XNDA}^5(\text{previous equation}) + \sum_{\substack{j=\\2,3,5,6}} \sum_{\substack{d\in\mathbf{X}_D\cap C_j\\ r_{da_0}<4R_0}} (\frac{R_0}{r_{da_0}})^6.$$

Computation of the difference ΔXNDD^i between the current values XNDD^i and proposed values XNDD^i_p for the possibly affected elements $i \in \{2,3,5,6\}$ is much simpler than for ΔXNDA^i. Because the current value of S_d is stored for each donor d, no looping over the acceptors within the current pattern has to be carried out and

$$\Delta\text{XNDD}^i = \sum_{\substack{d\in\mathbf{X}_D\cap C_i\\ r_{da_0}<R_0}} \frac{1}{1+S_d+(r_{da_0}/R_0^6)} - \frac{1}{1+S_d};\ i \in \{2,3,5,6\}.$$

We note that when an acceptor is added (removed) to (from) the current pattern this leads to to a redistribution of normalized pixel intensities between the DD- and DA-channel. Therefore the following equality holds

$$\sum_{i\in\{2,3,5,6\}} \Delta\text{XNDD}^i + \sum_{i\in\{2,3,5,6\}} \Delta\text{XNDA}^i = 0. \tag{4}$$

This observation provides a convenient way to check the proper implementation of the algorithms used to compute ΔXNDA^i and ΔXNDD^i. Computing both sums on the left-hand side of (4) and adding them should give the value of zero within numerical precision.

Step 6: Calculate Metropolis-Hastings ratio. We present here the steps involved to calculate (the logarithm of) the MH-ratio (Section 3.2 of

the main article) in terms of removing an accepter. The computation when adding an acceptor or adding/removing a donor is similar.

In case it is proposed to remove a point $u \in \mathbf{X}_A$, the Metropolis-Hastings ratio becomes

$$\frac{p(y|\mathbf{x}_D, \mathbf{x}_A \backslash \{u\}, \psi)}{p(y|\mathbf{x}_D, \mathbf{x}_A, \psi)} \frac{p(\mathbf{x}_A \backslash \{u\}|\theta_A)}{p(\mathbf{x}_A|\theta_A)} \frac{n(\mathbf{x}_A)}{\theta_A|W|},$$

with the logarithm of first two ratios specified below:

1. the likelihood ratio term

$$\begin{aligned} &\ln\left[\frac{p(y|\mathbf{x}_D, \mathbf{x}_A \backslash \{u\}, \psi)}{p(y|\mathbf{x}_D, \mathbf{x}_A, \psi)}\right] = \\ &\frac{1}{2}\sum_{i \in N_c}\Big(\ln[\frac{G_D\mu_{DD}^{i\mathbf{c}} + \sigma^2}{G_D\mu_{DD}^{i\mathbf{p}} + \sigma^2}] + \frac{(y_{DD}^i - \mu_{DD}^{i\mathbf{c}})^2}{G_D\mu_{DD}^{i\mathbf{c}} + \sigma^2} - \frac{(y_{DD}^i - \mu_{DD}^{i\mathbf{p}})^2}{G_D\mu_{DD}^{i\mathbf{p}} + \sigma^2} \\ &\quad + \ln[\frac{G_A\mu_{DA}^{i\mathbf{c}} + \sigma^2}{G_A\mu_{DA}^{i\mathbf{p}} + \sigma^2}] + \frac{(y_{DA}^i - \mu_{DA}^{i\mathbf{c}})^2}{G_A\mu_{DA}^{i\mathbf{c}} + \sigma^2} - \frac{(y_{DA}^i - \mu_{DA}^{i\mathbf{p}})^2}{G_A\mu_{DA}^{i\mathbf{p}} + \sigma^2} \\ &\quad + \ln[\frac{G_A\mu_{AA}^{i\mathbf{c}} + \sigma^2}{G_A\mu_{AA}^{i\mathbf{p}} + \sigma^2}] + \frac{(y_{AA}^i - \mu_{AA}^{i\mathbf{c}})^2}{G_A\mu_{AA}^{i\mathbf{c}} + \sigma^2} - \frac{(y_{AA}^i - \mu_{AA}^{i\mathbf{p}})^2}{G_A\mu_{AA}^{i\mathbf{p}} + \sigma^2}\Big). \end{aligned} \tag{5}$$

   Here the summation is over the pixels that are direct neighbors of the pixel c where a point was added of removed, and defined as the neighborhood N_c.[4]
2. The Poisson process ratio term

$$\ln\left[\frac{p(\mathbf{x}_A \setminus \{u\}|\theta_A)}{p(\mathbf{x}_A|\theta_A)}\right] = -\ln[\theta_A].$$

Step 7: Accept proposal? With the logarithm of the MH-ratio available as MHR, we draw a uniform random number u between 0 and 1. If $u < \exp(\text{MHR})$ then the proposal is accepted and otherwise it is declined. If the proposal is accepted:

1. the current point pattern $\mathbf{X}$ is updated such as proposed, i.e. adding or removing a donor or acceptor point to or from the appropriate linked list.
2. XNA^i or XND^i are updated according to the accepted proposal.

[4] In the case that the channel data set can not consist of negative values (i.e. empirical data), and $y_k^i = 0$ is observed, then the three terms in (5) related to channel k should be replaced by a so called truncated likelihood expression as discussed in Section 2.

3. The channel data matrices are updated by adding ΔXNDD^i and ΔXNDA^i to the values currently stored in XNDD^i and XNDA^i.

Step 8: mod(step,Mstep=0?) In order to draw approximately independent realizations for the point process and microscope parameters, a large number Mstep of sequential updates of the point pattern are made between every parameter proposal update step.

Step 9: Propose update microscope parameter(s). We discuss this step assumming that all six microscope parameters and both the Poisson point process parameters are included in the Bayesian inference. At every Mstep'th step a proposal update is made for each of the parameters in a random order. As the support of the microscope parameters is on $\mathbb{R}^+$ we use as proposal distribution for a microscope parameter ω_j the lognormal distribution, effectively ensuring that a proposal value $\omega_j^{\mathbf{p}}$ is always strictly positive. Within the MCMC procedure a proposal $\omega_j^{\mathbf{p}}$ is now generated by drawing a random normal number $\varepsilon \sim N(0, \tau_j^2)$ and setting

$$\omega_j^{\mathbf{p}} = \omega_j^{\mathbf{c}} \exp(\varepsilon),$$

with τ_j^2 a tuning parameter, tuned such that for each of the microscope parameters the acceptance probability is around 30%.

For the Poisson point process parameters θ_D and θ_A independent updates are generated by a Gibbs step. If we wish to draw samples for θ_D conditional on the current state we need to specify $p(\theta_D|\mathbf{x}_D)$ where

$$\begin{aligned} p(\theta_D|\mathbf{x}_D) &\propto p(\mathbf{x}_D|\theta_D)p(\theta_D) \\ &\propto \theta_D^{n(\mathbf{x}_D)} \exp(-|W|\theta_D)\, \theta_D^{(\alpha_D-1)} \exp(-\beta_D\theta_D) \\ &\propto \theta_D^{(n(\mathbf{x}_D)+\alpha_D-1)} \exp\left(-(|W|+\beta_D)\theta_D\right) \\ &= \Gamma(n(\mathbf{x}_D)+\alpha_D, \beta_D+|W|), \end{aligned}$$

where in the second step for $p(\mathbf{x}_D|\theta_D)$ the density of an independent Poisson process on W with intensity θ_D (equation 6.2 in Møller and Waagepetersen, 2003) was inserted, as well as the density of the gamma distribution for $p(\theta_D)$. So, samples for θ_D are drawn by sampling from the gamma distribution $\Gamma(n(\mathbf{x}_D)+\alpha_D, \beta_D+|W|)$ with α_D, β_D the shape and rate (hyper) parameters defining the prior gamma distribution of θ_D. Equivalent, samples for θ_A are drawn by sampling from $\Gamma(n(\mathbf{x}_A)+\alpha_A, \beta_A+|W|)$ with α_A, β_A the shape and rate (hyper) parameters defining the prior gamma distribution of θ_A. When updating θ_D or θ_A the steps 10 and 11 are skipped

and the program continues to make an update for the next parameter in line.

Step 10: Calculate MH ratio. The MH-ratio related to a proposal $\omega_j^{\mathbf{p}} \sim q_j(\cdot|\omega_j^{\mathbf{c}})$ for the j-th microscopic parameter is

$$\left(\frac{p(y|\psi^{\mathbf{p}},\mathbf{x}^{\mathbf{c}})}{p(y|\psi^{\mathbf{c}},\mathbf{x}^{\mathbf{c}})}\right)\left(\frac{p(\psi_j^{\mathbf{p}})}{p(\psi_j^{\mathbf{c}})}\right)\left(\frac{q(\psi_j^{\mathbf{c}}|\psi_j^{\mathbf{p}})}{q(\psi_j^{\mathbf{p}}|\psi_j^{\mathbf{c}})}\right),$$

where $\psi^{\mathbf{c}} = (\omega_1^{\mathbf{c}},\cdots,\omega_6^{\mathbf{c}})$ contains the current values for the microscope parameters and $\psi^{\mathbf{p}}$ contains the elements

$$\psi_k^{\mathbf{p}} = \begin{cases} \omega_k^{\mathbf{c}} & k \neq j \\ \omega_j^{\mathbf{p}} & k = j. \end{cases}$$

The logarithm of each of the terms in the MH-ratio are now specified by the

1. Likelihood ratio term[5]

$$\begin{aligned} &\ln\left[\frac{p(y|\psi^{\mathbf{p}},\mathbf{x}^{\mathbf{c}})}{p(y|\psi^{\mathbf{c}},\mathbf{x}^{\mathbf{c}})}\right] = \\ &\frac{1}{2}\sum_{i=1}^{n}\Big(\ln[\frac{G_D^{\mathbf{c}}\mu_{DD}^{i\mathbf{c}}+\sigma^{2,\mathbf{c}}}{G_D^{\mathbf{p}}\mu_{DD}^{i\mathbf{p}}+\sigma^{2,\mathbf{p}}}] + \frac{(y_{DD}^i-\mu_{DD}^{i\mathbf{c}})^2}{G_D^{\mathbf{c}}\mu_{DD}^{i\mathbf{c}}+\sigma^{2,\mathbf{c}}} - \frac{(y_{DD}^i-\mu_{DD}^{i\mathbf{p}})^2}{G_D^{\mathbf{p}}\mu_{DD}^{i\mathbf{p}}+\sigma^{2,\mathbf{p}}} \\ &\quad + \ln[\frac{G_A^{\mathbf{c}}\mu_{DA}^{i\mathbf{c}}+\sigma^{2,\mathbf{c}}}{G_A^{\mathbf{p}}\mu_{DA}^{i\mathbf{p}}+\sigma^{2,\mathbf{p}}}] + \frac{(y_{DA}^i-\mu_{DA}^{i\mathbf{c}})^2}{G_A^{\mathbf{c}}\mu_{DA}^{i\mathbf{c}}+\sigma^{2,\mathbf{c}}} - \frac{(y_{DA}^i-\mu_{DA}^{i\mathbf{p}})^2}{G_A^{\mathbf{p}}\mu_{DA}^{i\mathbf{p}}+\sigma^{2,\mathbf{p}}} \\ &\quad + \ln[\frac{G_A^{\mathbf{c}}\mu_{AA}^{i\mathbf{c}}+\sigma^{2,\mathbf{c}}}{G_A^{\mathbf{p}}\mu_{AA}^{i\mathbf{p}}+\sigma^{2,\mathbf{p}}}] + \frac{(y_{AA}^i-\mu_{AA}^{i\mathbf{c}})^2}{G_A^{\mathbf{c}}\mu_{AA}^{i\mathbf{c}}+\sigma^{2,\mathbf{c}}} - \frac{(y_{AA}^i-\mu_{AA}^{i\mathbf{p}})^2}{G_A^{\mathbf{p}}\mu_{AA}^{i\mathbf{p}}+\sigma^{2,\mathbf{p}}}\Big). \end{aligned} \tag{6}$$

2. Prior ratio term

$$\ln\left[\frac{p(\psi^{\mathbf{p}})}{p(\psi^{\mathbf{c}})}\right] = (\alpha-1)\ln[\frac{\omega_j^{\mathbf{p}}}{\omega_j^{\mathbf{c}}}] + \beta(\omega_j^{\mathbf{c}}-\omega_j^{\mathbf{p}}),$$

   with α and β the hyper parameters specifying the (prior) gamma distribution of parameter ω_j.
3. Proposal ratio term

$$\ln\left[\frac{q(\psi_j^{\mathbf{c}}|\psi_j^{\mathbf{p}})}{q(\psi_j^{\mathbf{p}}|\psi_j^{\mathbf{c}})}\right] = \ln[\frac{\omega_j^{\mathbf{p}}}{\omega_j^{\mathbf{c}}}].$$

[5]The previous footnote on page 10 also applies to this likelihood expression.

Step 11: Accept proposal?
With the value of the logarithm of the appropriate MH-ratio available as MHR, we draw a uniform random number u between 0 and 1. If $u < \exp(\text{MHR})$ then the proposal is accepted and otherwise it is declined. If the proposal is accepted, the current value of the microscope parameter is changed to the proposed value.

Step 12: mod(step,Tstep=0?)
If the current step number is equal to the total number of MCMC steps to be made (Tstep) the program stops, otherwise it continues.

2. Truncated likelihood expressions. For the typical case that the channel data does not allow for negative intensity values (i.e. non-synthetic data) we have defined the probability that a zero channel intensity value occurs as the probability mass in the left tail of the normal distribution over the negative intensity values. Therefore, in the case empirical three-cube FRET pixel intensities $y_k^i = 0$ are observed for pixel i and channel $k = DD, DA$ or AA, a truncated likelihood expression has to be used. In that case the three terms related to an observation $y_k^i = 0, k = DD$ in, for instance, (6) have to be replaced by

$$\ln\left[\frac{p(y_{DD}^i = 0|\psi^{\mathbf{p}}, \mathbf{x}^{\mathbf{c}})}{p(y_{DD}^i = 0|\psi^{\mathbf{c}}, \mathbf{x}^{\mathbf{c}})}\right] = \ln\left[\frac{\int_{-\infty}^0 \varphi(z|\mu_{DD}^{i\mathbf{p}}, G_D^{\mathbf{p}}\mu_{DD}^{i\mathbf{p}} + \sigma^{2,\mathbf{p}})\mathrm{d}z}{\int_{-\infty}^0 \varphi(z|\mu_{DD}^{i\mathbf{c}}, G_D^{\mathbf{c}}\mu_{DD}^{i\mathbf{c}} + \sigma^{2,\mathbf{c}})\mathrm{d}z}\right]$$

and when $k = DA$ or $k = AA$ by

$$\ln\left[\frac{p(y_k^i = 0|\psi^{\mathbf{p}}, \mathbf{x}^{\mathbf{c}})}{p(y_k^i = 0|\psi^{\mathbf{c}}, \mathbf{x}^{\mathbf{c}})}\right] = \ln\left[\frac{\int_{-\infty}^0 \varphi(z|\mu_k^{i\mathbf{p}}, G_A^{\mathbf{p}}\mu_k^{i\mathbf{p}} + \sigma^{2,\mathbf{p}})\mathrm{d}z}{\int_{-\infty}^0 \varphi(z|\mu_k^{i\mathbf{p}}, G_A^{\mathbf{c}}\mu_k^{i\mathbf{c}} + \sigma^{2,\mathbf{c}})\mathrm{d}z}\right]$$

where $\varphi(\cdot|\mu, \sigma^2)$ is the density of the normal distribution with mean μ and variance σ^2. The likelihood equations (5) are also updated accordingly to the strategy as described here above, in case the channel can not consist of negative intensity values.

References.

Baddeley, A. and Turner, R. (2005). spatstat: An R package for analyzing spatial point patterns. *Journal of Statistical Software* **12** 1–42.

R Core Team (2014). R: A language and environment for statistical computing R Foundation for Statistical Computing, Vienna, Austria.

Jan-Otto Hooghoudt
Aalborg University
Department of Civil Engineering
Aalborg, Denmark
E-mail: joh@civil.aau.dk

Rasmus Waagepetersen
Aalborg University
Department of Mathematics
Aalborg, Denmark
E-mail: rw@math.aau.dk