Bayesian filtering for spatial estimation of photo-switching fluorophores imaged in Super-resolution fluorescence microscopy

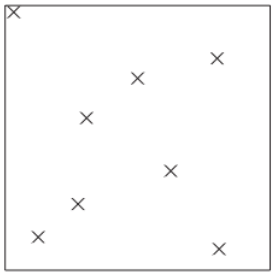
Lekha Patel* and Edward A.K. Cohen*
*Department of Mathematics,
Imperial College London, London, United Kingdom,
{lp1611,e.cohen}@imperial.ac.uk

***Abstract*—The success of many Super-resolution fluorescence microscopy methods lie in the exploitation of the inherent stochasticity of a light emitting molecule's photon emission state, allowing sparse subsets of molecules to be spatially detected with high precision. This photo-switching behavior, however, induces multiple localizations from each molecule during an imaging experiment, which therefore gives rise to misleading representations of their true spatial locations. By formulating a state-space model relating true molecular positions with observation sets collected across time, we show that the full Bayes filter for this problem can be derived and positions recovered via a Markov Chain Monte Carlo sampler.**

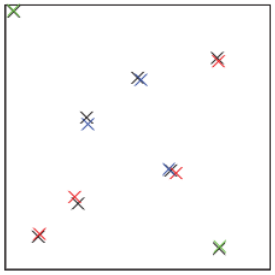
***Index Terms*—Super-resolution imaging, fluorescence microscopy, photo-switching, spatial-temporal point process, random set, target detection, state-space modeling, Bayes filter, reversible jump Markov Chain Monte Carlo, detection probability.**

I. Introduction

Many fluorescence microscopy techniques such as photo-activated localization microscopy (PALM) [1], [2] or stochastic optical reconstruction microscopy (STORM) [3], [4] are able to use optical imaging to produce high resolution images of biological organisms by utilizing a property of fluorescing molecules called *photo-switching*. A light emitting molecule or *fluorophore* can be made to stochastically photo-switch between a photon emission On state and several dark states [5], [6], enabling only a sparse subset of molecules to be detected by the camera at any one time. When a dense set of fluorophores (of unknown cardinality) are filmed over a number of frames, this enables accurate localizations to be made when molecules are detectable. However, their resulting non-stationary photo-switching behaviors invoke non constant detection rates per frame, which in turn generate a random number of *offspring* coordinates per molecule during an experiment [7]. Such measurements determined by fitting point spread functions (PSF) [8], [9] to high photon intensity spots, will differ between frames as the same molecules are localized about their true spatial locations. Figures 1a and 1b provide a visualization as to how these fluorophores are localized in space-time. When multiple localizations are made, the superimposed image (see Figure 1b) does not allow for individual resolvent. The inference problem therefore lies in identifying both the *number* of fluorophores and their true spatial positions.



(a) True spatial coordinates of fluorophores.



(b) Time aggregated localizations of fluorophores.

Fig. 1: Figure 1a shows the true spatial locations of a set of fluorophores from one fluorescence microscopy experiment. Figure 1b shows the aggregated measurements of localized fluorophores detected from the first 3 frames; objects detected from frame 1 (blue), frame 2 (red) and frame 3 (green) are shown with respect to the true locations (black).

The idea of modeling the clusters obtained in the superimposed image [10], [11] or the spatial distributions of molecules themselves [12] via spatial point patterns or *random finite sets* (RFS) [13]–[15], is familiar. While modeling clusters and sampling true locations (via e.g. MCMC methods [16]) is well studied, using the superimposed image will in general lead to spatial biases through false positive localizations and cases where molecules are spatially close. Nevertheless, inferring

upon point process cardinalities and positions via Bayesian filtering methods has been used for many radar/sonar tracking, navigation and computer vision applications [15], [17], [18]. Such methods are attractive due to their temporal incorporation and we thus look to formulate a model for this problem which rests on similar ideology. To the best of our knowledge, inference of detected points which so heavily rely on the temporal nature at which they occur has not been studied for this application.

II. Modeling Fluorophore Photo-Switching

The experimental set-up gives rise to an unknown initial probability mass over the hidden photon emission states of each fluorophore. This implies that in general, there is a non-zero amount of time taken for each molecule to first reach its photon-emission state, which therefore induces a pure spatial-temporal hidden birth process as new fluorophores are detected in each frame during an experiment. Due to this set-up, throughout this paper we adopt the discrete time indexing $n \in \mathbb{Z}_{\geq 0}$ to reference each frame of an imaging experiment.

Suppose at time n there are $K(n) := |X_n|$ fluorophores that have *already* reached their photon-emission states, where $X_n = \{\mathbf{x}_{n,1}, \ldots, \mathbf{x}_{n,K(n)}\}$ and each *parent* $\mathbf{x} \in S$ denotes the true position of a fluorophore that is present. Here, we assume that imaging occurs on some bounded window $S \subset \mathbb{R}^d$, whereby $d \in \{2, 3\}$ for 2D/3D applications. When N_F frames are recorded during an experiment, at every time $n = 0, 1, ..., N_F - 1$, a set of *offspring* measurements $Z_n = \{\mathbf{z}_{n,1}, \ldots, \mathbf{z}_{n,M(n)}\}$ is recorded. We define Z_n to be the *multi-target* measurement formed by the $K(n)$ present fluorophores and background noise.

A suitable model we choose to place is given by

$$\begin{aligned} X_0 &= B_0 \\ X_n &= X_{n-1} \cup B_n \qquad n \geq 1 \\ Z_n &= \Phi(X_n) \cup A_n \quad n \geq 0, \end{aligned}$$

where X_n denotes the *hidden* **parent** RFS and Z_n denotes the *observed* RFS. In particular, B_n denotes the independent birth RFS of fluorophores "born" in frame n, $\Phi(X_n)$ denotes the RFS of primary target generated measurements and A_n independently denotes the RFS of false positive measurements. An illustration of this model over three frames is shown in Figure 2, which compares the hidden RFS X_0, X_1 and X_2 with their observation counterparts Z_0, Z_1 and Z_2.

With an unknown number of molecules, both the birth RFS B_n and RFS for false positives are assumed to be a priori Poisson, with respective *spatial* densities $f_{B_n}(X)$ and $f_{A_n}(X)$ defined for any $X \subseteq S$. B_n is modeled to have cardinality distribution $N_{B_n} \sim \text{Poi}(\lambda p_{B,n})$, where λ denotes the number of molecules expected to be imaged prior to the experiment and $p_{B,n}$ denotes the probability of a fluorophore birth at time step n, assumed to be known given its photo-switching behavior. Moreover, A_n is modeled to have cardinality distribution $N_{A_n} \sim \text{Poi}(\alpha)$, where α is the unknown average number of false positive observations produced in a frame. In this manner, the spatial densities (with respect to the unit rate Poisson process), take the form [15]

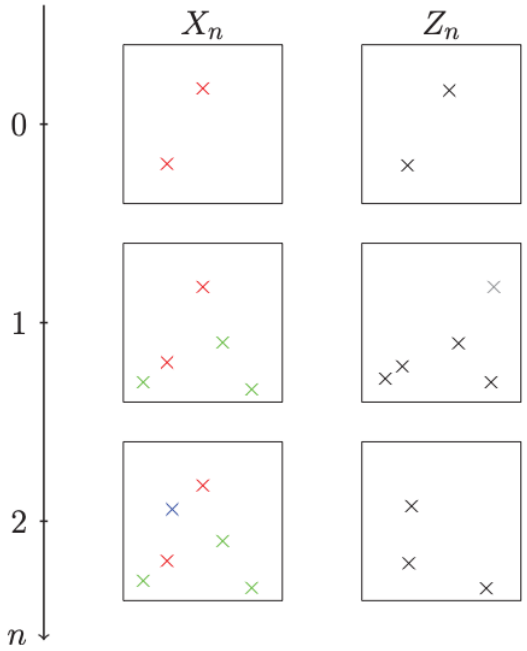


Fig. 2: An illustration of the RFS X_n compared with the RFS Z_n against time n. Left: $X_0 = B_0$ is plotted in red at time $n = 0$, $X_1 = X_0 \cup B_1$ is plotted with B_1 in green at time $n = 1$ and $X_2 = X_1 \cup B_2$ is plotted with B_2 in blue at time $n = 2$. Right: Observation sets Z_0, Z_1, Z_2 are shown, with false positive observations plotted in gray.

$$\begin{aligned} f_{B_n}(X) &= \mathrm{e}^{\mu(S)-\lambda p_{B_n}} \prod_{\mathbf{x}\in X} \lambda p_{B_n} b(\mathbf{x}) \\ f_{A_n}(X) &= \mathrm{e}^{\mu(S)-\alpha} \prod_{\mathbf{x}\in X} \alpha a(\mathbf{x}) \qquad n \geq 0, \end{aligned}$$

where $\mu(S)$ denotes the Lebesgue measure of S and $b(\mathbf{x})$, $a(\mathbf{x})$ denote the spatial distributions of births and false positives, respectively. In particular, these densities must satisfy $\int_S b(\mathbf{x})\ \mathrm{d}\mathbf{x} = \int_S a(\mathbf{x})\ \mathrm{d}\mathbf{x} = 1$. In this paper, we will only be dealing with the case where births and false positives are uniformly distributed over S and therefore when $b(\mathbf{x}) = a(\mathbf{x}) = \frac{1}{\mu(S)}$.

Each fluorophore or parent $\mathbf{x} \in X_n$ generates an offspring $\mathbf{z} \in \Phi(X_n)$ such that $\mathbf{z} = \emptyset$ with probability $1 - p_{D,n}$ or is a variate from the density $f(\mathbf{z}^*|\mathbf{x})$ otherwise. Here, $p_{D,n}$ denotes the detection probability of a parent in frame n (also assumed to be known) and $f(\mathbf{z}^*|\mathbf{x}) \propto \exp(-\frac{1}{2\sigma^2}(\mathbf{z}^* - \mathbf{x})^\top(\mathbf{z}^* - \mathbf{x}))$ denotes the likelihood that offspring point $\mathbf{z}^*$ is generated from parent $\mathbf{x}$. The localization standard deviation σ is assumed to be known.

A. Derivation of the Bayes Filter

By defining the limiting RFS $X^* := \lim_{n\to\infty} X_n$, which is interpreted as the *implicit* point pattern configuration of the parents, the true configuration of X^* can thereby be updated through time. Specifically, for any $n \in \mathbb{Z}_{>0}$, we wish to infer upon the parent RFS X_n (with its cardinality), and posterior estimates for α given all observations $Z^{(n)} := \cup_{i=0}^{n} Z_i$ through the Bayes-updated density

$$f_{X,\alpha|Z^{(n)}}(X_n, \alpha|Z^{(n)}) \propto f_{Z_n|X,\alpha}(Z_n|X_n,\alpha) f_{X|Z^{(n-1)},\alpha}(X_n|Z^{(n-1)},\alpha)\pi(\alpha), \tag{1}$$

where $\pi(\alpha)$ denotes a suitable prior probability density for α.

Equation (1) can be computed by first using the *multi-target* likelihood as described in [15] by

$$f_{Z_n|X,\alpha}(Z_n|X,\alpha) = (1-p_{D,n})^{|X|} f_{A_n}(Z_n) \sum_{\theta} \prod_{i:\theta(i)>0} \frac{p_{D,n} f(\mathbf{z}_{\theta(i)}|\mathbf{x}_i)\mu(S)}{\alpha(1-p_{D,n})}, \tag{2}$$

where the sum is taken over all mappings $\theta : \{1,\dots,|X|\} \to \{0,1,\dots,|Z_n|\}$, whereby $\theta(i) = \theta(i') > 0 \implies i = i'$. It should be noted here that when $|Z_n| = 0$, no observations are collected and the likelihood is therefore reduced to $f_{Z_n|X,\alpha}(\emptyset|X,\alpha) = \mathrm{e}^{\mu(S)-\alpha}(1-p_{D,n})^{|X|}$.

We can additionally show that up to a proportionality constant, the density $f_{X_n|Z^{(n-1)},\alpha}(X_n|Z^{(n-1)},\alpha)$ takes the form in (3) as is stated in Proposition 1.

Proposition 1. *For any $n \in \mathbb{Z}_{>0}$, the density $f_{X|Z^{(n-1)},\alpha}(X_n|Z^{(n-1)},\alpha)$ is proportional to the function*

$$\sum_{W_n\subseteq X_n} \sum_{W_{n-1}\subseteq W_n} \cdots \sum_{W_1\subseteq W_2} f_{B_n}(X_n \setminus W_n) \times \left[\prod_{i=0}^{n-1} f_{B_i}(W_{i+1}\setminus W_i) f_{Z_i|X,\alpha}(Z_i|W_{i+1},\alpha)\right], \tag{3}$$

with $W_0 = \emptyset$.

Proof. See Proof VI-A of the Appendix. □

When the number of parents is large, the posterior density in (1) will become computationally intractable. While difficulties lie both in the evaluations of (3) and the likelihood function (2), we implement a suitable approximation of (2) to improve speed and tractability in the evaluation of the posterior density.

Specifically, for whenever $|Z_n| > 0$, define the mappings $\phi_j : \{1,\dots,|X|\} \to \{0,1,\dots,|Z_n|\}\setminus\{j\}$ for $j = 1,\dots,|Z_n|+1$, such that $\phi_j(i) = \phi_j(i') > 0 \implies i = i'$. For each j, we can use a nearest neighbor search to identify the mapping $\hat{\phi}_j$ which gives $\prod_{i:\phi_j(i)>0} \frac{p_{D,n} f(\mathbf{z}_{\phi_j(i)}|\mathbf{x}_i)\mu(S)}{\lambda\alpha(1-p_{D,n})}$ its highest value. We then approximate (2) via

$$(1-p_{D,n})^{|X|} f_{A_n}(Z_n) \sum_{j=1}^{|Z_n|+1} \prod_{i:\hat{\phi}_j(i)>0} \frac{p_{D,n} f(\mathbf{z}_{\hat{\phi}_j(i)}|\mathbf{x}_i)\mu(S)}{\lambda\alpha(1-p_{D,n})}, \tag{4}$$

which is a valid approximation provided that the probability of more than one false positive observation appearing in each frame is negligible and that a sparse subset of the total number of molecules is detected in each frame.

This approximation to computing the density in (3) enables fast and accurate inference of the spatial coordinates of X_n and its cardinality via a Markov Chain Monte Carlo (MCMC) sampler.

III. Inference

Using the outlined decomposition of the Bayes filter, updating the parent locations X_n through its conditional density in (1) is done via a Birth-Death-Shift (Reversible-Jump) MCMC algorithm described in [16], [19]. This algorithm is designed to sample from posterior distributions with varying parameter spaces, as is the case for spatial point patterns of unknown cardinalities. Furthermore, updating the hyper-parameter α is done via the Metropolis-Hastings algorithm. ,

A. Birth-Death-Shift Moves

Let $\bar{X}_i = \{\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \dots, \bar{\mathbf{x}}_{k_i}\}$ be the configuration of the hidden parent process, and α_i the value of α at iteration i of the sampler. An initial choice between either a shift move or a birth/death move is made with probability q and $1-q$ respectively [16], [19]. If a birth/death move is chosen, then a component is added to the configuration with probability $p(k_i)$, otherwise a point $\mathbf{x}^* \in \bar{X}_i$ is randomly chosen to be deleted. When a birth choice is made, a component $\mathbf{x}'_{k_i+1}$ sampled uniformly over S is proposed. Then, $X' = \bar{X}_i \cup \mathbf{x}'_{k_i+1}$ is accepted with probability $\min(1, h_b)$, where the *birth hastings ratio* h_b can be shown [16] to take the form

$$h_b = PR \times \frac{1-f_K(k_i+1)}{f_K(k_i)} \times \frac{\mu(S)}{k_i+1}, \tag{5}$$

with $f_K(k) = \mathbb{P}(|X^*| = k)$. Similarly if a death move is chosen, then X' is accepted with probability $\min(1, h_d)$ where the *death hastings ratio* h_d is

$$h_d = PR \times \frac{f_K(k_i-1)}{1-f_K(k_i)} \times \frac{k_i}{\mu(S)}, \tag{6}$$

with $X' = \bar{X}_i \setminus \mathbf{x}^*$.

When a shift move is made, an index I is chosen uniformly from the set $\{1,2,\dots,k_i\}$ and $\mathbf{x}' \sim \mathcal{N}_d(\bar{\mathbf{x}}_I, \sigma_p^2 I_d)$ is sampled. Here, I_d denotes the d-dimensional identity matrix and σ_p^2 is a pre-specified fixed proposal variance. The proposal $X' = \{\bar{\mathbf{x}}_1,\dots,\bar{\mathbf{x}}_{I-1},\mathbf{x}',\bar{\mathbf{x}}_{I+1},\dots,\bar{\mathbf{x}}_{k_i}\}$ is then accepted with probability $min(1, h_s)$, where the *shift hastings ratio* $h_s = PR$. In all instances, the posterior ratio (PR) is given by

$$PR = \frac{f_{Z_n|X,\alpha}(Z_{N_F}|X',\alpha_i) f_{X|Z^{(n-1)},\alpha}(X'|Z^{(N_F-1)},\alpha_i)}{f_{Z_n|X,\alpha}(Z_{N_F}|\bar{X}_i,\alpha_i) f_{X|Z^{(n-1)},\alpha}(\bar{X}_i|Z^{(N_F-1)},\alpha_i)},$$

where $f_{Z_n|X,\alpha}(Z_{N_F}|\cdot,\cdot)$ is computed using (4) and $f_{X|Z^{(n-1)},\alpha}(\cdot|Z^{(N_F-1)},\cdot)$ from (3).

To avoid moves which sample over parameter spaces that are either too small or too large, we take

$$f_K(k) \propto \begin{cases} \frac{\lambda^k}{k!} & \text{if } k \in \{k_{\min}, k_{\min}+1, \ldots, k_{\max}-1, k_{\max}\} \\ 0 & \text{otherwise,} \end{cases}$$

and choose $p(k) = f_K(k)$ for all k. Here, $k_{\min}$ and $k_{\max}$ can either be user specified, or set to capture the majority of the variance of $f_K(k)$ over $k \geq 0$.

Final estimates of parent locations can be obtained via kernel smoothing methods applied on the samples obtained from these moves.

B. Update moves

As previously stated, α is updated via the standard Metropolis-Hastings algorithm. We assume that $\alpha \in (0,1)$; a constraint that aligns with standard Super-resolution imaging conditions, whereby α is observed to be small but non-zero. In doing so, this condition also validates the negligibility of more than one false positive observation appearing in a frame. As such, we chose $\pi(\alpha) \propto (1-\alpha)^{\beta^*-1}$, with $\beta^* > 0$ fixed. This choice ensures that α has a greater mass around zero, with the additional $\beta^* = \frac{1}{\mathbb{E}(\alpha)} - 1$ term used to characterize apriori knowledge of the average number of false positive observations produced in a frame.

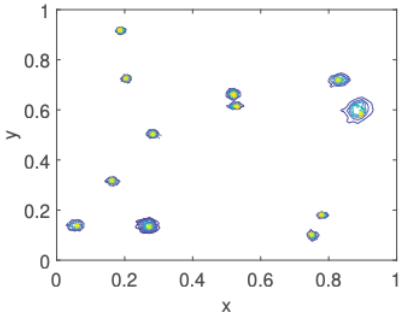
Given α_i, α' is sampled from a Beta$(1, \beta_1)$ distribution, whereby we choose $\beta_1 = \frac{1-\alpha_i}{\alpha_i}$ to ensure that the mean of this proposal is α_i. The proposed α' is then accepted with probability $\min(1, h_u)$, where the *update hastings ratio* h_u is given by

$$h_u = \frac{f_{Z_n|X,\alpha}(Z_{N_F}|\bar{X}_i, \alpha') f_{X|Z^{(n-1)},\alpha}(\bar{X}_i|Z^{(N_F-1)}, \alpha')}{f_{Z_n|X,\alpha}(Z_{N_F}|\bar{X}_i, \alpha_i) f_{X|Z^{(n-1)},\alpha}(\bar{X}_i|Z^{(N_F-1)}, \alpha_i)} \times \left(\frac{1-\alpha'}{1-\alpha_i}\right)^{\beta^*-1} \times \left(\frac{1-\alpha_i}{1-\alpha'}\right)^{\beta_1-1}.$$

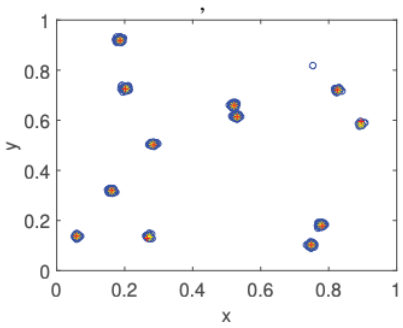
We note here that this methodology can be extended to sample from other unknown parameters of the model. For example, if the localization standard deviation σ is *unknown*, then the joint posterior distribution of unknown model parameters and X_{N_F} can also be sampled in this step.

IV. Simulations

We present results from two simulated fluorescence experiments. In the first experiment, $S = [0,1] \times [0,1]$, $\lambda = 10$, $|X^*| = 12$, $\alpha = 10^{-2}$ and $N_F = 100$ whereas in the second experiment, $\lambda = 20$, $|X^*| = 22$, $\alpha = 10^{-1}$ and $N_F = 200$. In all experiments, $\sigma = 0.005$, $\sigma_p = 0.01$, $q = 0.8$ and $\beta^* = 49$. The parameters $p_{B,n}$ and $p_{D,n}$ are computed using inferred photo-switching rates derived from a temporal distribution as is described in [7]. Results of coordinate estimates from the derived Bayes filter, are given in Figures 3 and 4 for the first and second experiments respectively. The posterior intensity maps of samples in Figures 3a and 4a show the distribution of samples gained from $f_{X|Z^{(n)}}(X_{N_F}|Z^{(N_F)})$ over S after a burn-in period of 10^3 samples. The coordinate estimates are then plotted against true positions in Figures 3b and 4b. The true cardinalities of the implicit point processes X^* were correctly identified in both experiments. From these plots, it is evident that the sampler is able to successfully identify the number of hidden parents, accurately estimate their true spatial positions and finally, able to identify observations generated directly by X_{N_F} with those generated by false positives measurements.



(a) Posterior intensity map.



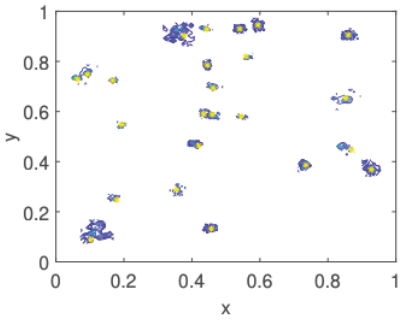
(b) Coordinate estimates against true positions.

Fig. 3: 3a: Posterior intensity map of MCMC samples against true positions (yellow stars). 3b: Coordinate estimates from the Bayes filter (red crosses) plotted with the true positions (yellow stars) and the superimposition of offspring observation sets and false positive observations (blue circles) collected over $N_F = 100$ frames. Simulation conducted with $S = [0,1] \times [0,1]$, $\lambda = 10$, $\sigma = 0.005$ and $\alpha = 10^{-2}$.

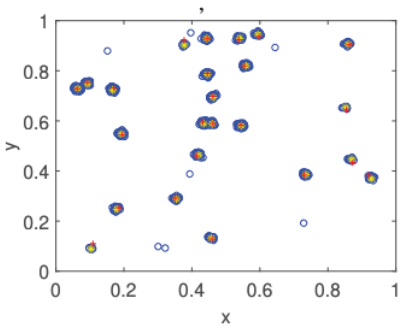
V. Conclusion

In this work, we have proposed a model to describe the time varying spatial distribution of fluorescing molecules imaged in Super-resolution fluorescence microscopy. By using a pure birth process to model when fluorophores first reach their photo-emission states, we have formulated a state-space model which links the hidden parent RFS with observation sets collected at each time frame n. This model, furthermore, accounts for the common scenario in which observation sets are corrupted by false positive observations. Using this, we have carefully derived the Bayes filter which is used to update the true point pattern of molecules across time. When computing the Bayes-updated density, we have also provided

a suitable approximation to increase computational feasibility and have used this when sampling from the joint posterior distribution of the parent process via a reversible jump MCMC algorithm. We have lastly shown, through simulations, that the true locations of parents or molecules and their cardinalities can be recovered using this method.



(a) Posterior intensity map.



(b) Coordinate estimates against true positions.

Fig. 4: 4a: Posterior intensity map of MCMC samples against true positions (yellow stars). 4b: Coordinate estimates from the Bayes filter (red crosses) plotted with the true positions (yellow stars) and the superimposition of offspring observation sets and false positive observations (blue circles) collected over $N_F = 200$ frames. Simulation conducted with $S = [0,1] \times [0,1]$, $\lambda = 20$, $\sigma = 0.005$ and $\alpha = 10^{-1}$.

VI. Appendix

A. Proof of Proposition 1

Proof. For every $n \in \mathbb{Z}_{>0}$, we can write $\{X_n|Z^{(n-1)}\} = \{X_{n-1}|Z^{(n-1)}\} \cup \{B_n\}$, where the point processes $\{X_{n\ 1}|Z^{(n-1)}\}$ and $\{B_n\}$ are independent by model construction.

By the fundamental theorem of convolution [15], we have for all $n \geq 1$ that

$$f_{X|Z^{(n\ 1)},\alpha}(X_n|Z^{(n-1)},\alpha) \propto \sum_{W \subseteq X_n} f_{X|Z^{(n-1)},\alpha}(W|Z^{(n-1)},\alpha) f_{B_n}(X_n \setminus W), \tag{7}$$

where the sum is taken over all subsets W of X_n.

Since $f_{B_0}(X_0)$ characterizes the distribution of X_0, we have $f_{X|Z^{(0)},\alpha}(X_0|Z_0,\alpha) \propto f_{Z_0|X,\alpha}(Z_0|X_0,\alpha) f_{B_0}(X_0)$. Using this as an initialization for (7), coupled with the Bayes update rule given in (1), yields the desired result. □

References

[1] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging Intracellular Fluorescent Proteins at Nanometer Resolution," *Science*, vol. 313(5793), pp. 1642–1645, 2006.

[2] S. T. Hess, T. P. K. Girirajan, and M. D. Mason, "Ultra-high resolution imaging by fluorescence photoactivation localization microscopy," *Biophysical journal*, vol. 91(11), pp. 4258–4272, 2006.

[3] M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nature methods*, pp. 793 795, 2006.

[4] M. Heilemann, S. Van de Linde, M. Schüttpelz, R. Kasper, B. Seefeldt, A. Mukherjee, P. Tinnefeld, and M. Sauer, "Subdiffraction - Resolution Fluorescence Imaging with Conventional Fluorescent Probes," *Angewandte Chemie International Edition*, vol. 47(33), pp. 6172–6176, 2008.

[5] S. Van de Linde and M. Sauer, "How to switch a fluorophore: from undesired blinking to controlled photoswitching," *Chemical Society reviews*, vol. 43(4), pp. 1076–1087, 2014.

[6] T. Ha and P. Tinnefeld, "Photophysics of Fluorescent Probes for Single-Molecule Biophysics and Super-Resolution Imaging," *Annual Review of Physical Chemistry*, vol. 63(1), pp. 595–617, 2012.

[7] L. Patel, N. Gustafsson, Y. Lin, R. Ober, R. Henriques, and E. Cohen, "A hidden Markov model approach to characterizing the photo-switching behavior of fluorophores," *Preprint: bioRxiv*, 2017.

[8] R. Ober, A. Tahmasbi, S. Ram, Z. Lin, and E. Ward, "Quantitative Aspects of Single-Molecule Microscopy: Information-theoretic analysis of single-molecule data," *IEEE Signal Processing Magazine*, vol. 32, no. 1, pp. 58–69, 2015.

[9] D. Sage, H. Kirshner, T. Pengo, N. Stuurman, J. Min, S. Manley, and M. Usher, "Quantitative evaluation of software packages for single-molecule localization microscopy," *Nature Methods*, vol. 12(8), pp. 717–724, 2015.

[10] C. Hsu and T. Baumgart, "Spatial Association of Signaling Proteins and F-Actin Effects on Cluster Assembly Analyzed via Photoactivation Localization Microscopy in T Cells," *PloS one*, vol. 6, pp. 1–13, 2011.

[11] M. Wiemhöfer, C. Thiel, J. Klingauf, and R. Chow, "Probing Protein-Protein Interactions on the NM Scale using TIRF-PALM," *Biophysical Journal*, vol. 102, 2012.

[12] D. Owen, C. Rentero, J. Rossy, A. Magenau, D. Williamson, M. Rodriguez, and K. Gaus, "PALM imaging and cluster analysis of protein heterogeneity at the cell surface," *Journal of Biophotonics*, vol. 3, no. 7, 2010.

[13] R. Mahler, "Multitarget Bayes filtering via first-order multitarget moments," *IEEE Trans. Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1152 1178, 2003.

[14] ——, "PHD filters of higher order in target number," *IEEE Trans. Aerospace and Electronic Systems*, vol. 43, no. 4, pp. 1523–1543, 2007.

[15] ——, *Statistical Multisource-Multitarget information fusion*. Arctech house, 2007.

[16] J. Møller and R. Waagepetersen, *Statistical Inference and Simulation for Spatial Point Processes*, ser. Chapman & Hall/CRC Monographs on Statistics & Applied Probability, 2003.

[17] Y. Bar-Shalom, T. Kirubarajan, and X. R. Li, *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. Wiley, 2001.

[18] E. Maggio, M. Taj, and A. Cavallaro, "Efficient multi-target visual tracking using random finite sets," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 8, pp. 1016 1027, 2008.

[19] C. J. Geyer and J. Møller, "Simulation procedures and likelihood inference for spatial point processes," *Scandinavian Journal of Statistics*, vol. 21, no. 4, pp. 359–373, 1994.