

A BAYESIAN APPROACH FOR SUPER-RESOLUTION MICROSCOPY IMAGE RECONSTRUCTION

Master thesis

Student: Cärolin Grahv

Supervisors: Martin Laasmaa, Assistant Professor, Tallinn University of Technology

Marko Vendelin, Professor, Tallinn University of Technology

Study program : Applied physics

Tallinn 2024

Declaration

Hereby I declare that I have compiled the paper independently and all works, important standpoints and data by other authors have been properly referenced and the same paper has not previously been presented for grading

Author: Cärolin Grahv

[*Signature*, 16.05.2024]

The paper conforms to requirements in force.

Supervisor: Martin Laasmaa, Marko Vendelin

[*Signature*, 16.05.2024]

Table of contents

1	Introduction	3
2	Overview of literature	4
2.1	Image formation in microscopy	4
2.2	Super-resolution microscopy	5
2.3	Super-resolution methods	5
2.3.1	Stimulated emission depletion - STED	5
2.3.2	Structured illumination microscopy - SIM	6
2.3.3	Airyscan	6
2.3.4	Single molecule localization microscopy - SMLM	7
2.4	Localization of molecules	8
3	Methods	10
3.1	Bayesian modelling in Python	10
3.2	Generation of synthetic data	10
3.3	Bayesian model	11
3.4	Frequentist approach	11
3.5	Evaluation of models	12
3.5.1	The Bayesian inference model	12
3.5.2	Comparison of Bayesian inference libraries	12
3.5.3	The frequentist model	13
3.5.4	Error and Kernel Density Estimates, visualisation of results	13
3.6	Model usage	14
3.7	Software	14
3.7.1	Packages used for Bayesian inference in Stan	14
3.7.2	Packages used for Bayesian inference in PyMC	14
3.7.3	Packages used for the frequentist approach	15
3.7.4	Packages used for visualisation, statistical analysis	15
4	Results	16
4.1	Generation of images	16
4.2	Comparison of Bayesian inference libraries	16
4.3	Results of Bayesian inference and comparison with frequentist approach	17
5	Discussion	26
6	Acknowledgements	27
7	Abstract	31
8	Annotatsioon	32

A Appendices	33
A.1 Non-exclusive licence	33

1 Introduction

The enhancement of microscope images is increasingly important in many fields. High resolution images are required for distinguishing intricacies of small cells in a sample. For example, Laasmaa et al. (2023) used high-resolution imaging techniques to study mouse cardiomyocytes. Microscopes are useful tools in multiple different fields, not just biology. Microscopes are used in fields like forensic research, medical fields and industrial fields.

Conventional light microscopes have a hard limit on their usefulness in many fields. The resolution acquired using a conventional light microscope is restricted by the resolution limit described by Ernst Abbe in 1873 (Abbe, 1873). Super-resolution techniques pass this resolution limit and allow for high resolution microscope images. The first concepts of techniques to surpass the Abbe resolution limit for far-field microscopy surfaced in the early 1990-s (Prakash et al., 2022). Ever since, numerous different techniques have been developed to resolve images in the nanoscale, these methods are called super-resolution methods. Super-resolution techniques use the independent behavior of fluorescent probes to overcome the diffraction limit (Fazel et al., 2022). Single-molecule localization techniques are one of the most popular super-resolution techniques. This is because they can be used with low-cost setups (Schermelleh et al., 2019). These techniques all use the stochastic "ON" - "OFF" switching of fluorescent molecules (Fazel et al., 2022).

In this work, a Bayesian approach is developed to localize fluorescent molecules (emitters) on synthetic single-molecule localization microscopy technique images. The Bayesian approach involves using Bayesian statistics to infer the emitter positions. The Bayesian approach differs from the "traditional" frequentist approach by the interpretation of uncertainty. The Bayesian approach uses prior information for the analysis, while the frequentist approach works only with the information from the data (Fornacon-Wood et al., 2022). Previous works have used a Bayesian approach in a method to localize emitters using grouping of localizations (Fazel et al., 2022).

The aim of this work is to develop a Bayesian approach to solve the problem of localizing emitter positions on microscope images and determine, whether the developed algorithm can resolve emitter positions better than the frequentist approach. For the comparison of frequentist inference and Bayesian inference, the Bayesian inference and frequentist inference are implemented. The models are tested on synthetic images generated using a stochastic image model. The images contain photons of a randomly placed emitter, the number of emitted photons is varied and images with noise are also included in the test data.

2 Overview of literature

2.1 Image formation in microscopy

In microscopy, light diffracts from the observed object and interferes with the undisturbed direct light. The resulting diffracted light is then captured by the objective lens of the microscope and forms an image (Rudi Rottenfusser et al., nd). The diffraction pattern emitted by a single point emitter is called the point spread function (PSF). The PSF is seen as blurry diffraction patterns that describe the blur of the optical system (Cheng, 2006). The formation of an image from a microscope can be defined using the PSF. Mathematically, the observed image could be viewed as a convolution operation between the viewed object and the PSF, with the existing noise (Cheng, 2006). The model for image formulation is the following:

$$i = \mathcal{P}(o \otimes h), \quad (1)$$

where \mathcal{P} is Poisson noise originating from photon counting, o is the observed object and h is the PSF. \otimes marks the convolution operation.

Images can be improved mathematically or experimentally using different imaging techniques. The mathematical approach is called deconvolution. Deconvolution aims to enhance image resolution by de-noising the convolved image (equation 1), using knowledge on the nature of noise (Laasmaa et al., 2011). The most substantial aberrations or noise result from the imaging system itself. This includes the optics of the microscope (Dey et al., 2006). As previously mentioned, noise also results from photon counting, which occurs when a detector is used to image the object using a microscope (Laasmaa et al., 2011). Deconvolution is able to greatly enhance microscope images, but is still restricted by the resolution limit, due to it being a post-processing method. This is because the optics used image the object are diffraction-limited (Dey et al., 2006). On the other hand, experimental methods, such as super-resolution methods are able to overcome the diffraction limit.

The resolution limit for light microscopy or optical microscopy is said to be constrained by the Abbe resolution limit. This spatial resolution is around half the wavelength of visible light ~ 200 nanometers. Super-resolution microscopy techniques get past this limit and allow us to accomplish spatial resolutions in the range of 1 to 100 nanometers. (Nienhaus and Nienhaus, 2016). This means that it is possible to potentially improve light microscopy images 200 times. The techniques used in super-resolution microscopy allow us to image cellular structures with the same level of detail as with electron microscopy (Schermelleh et al., 2019). The development of super-resolution techniques was also awarded with a Nobel Prize in Chemistry in 2014 to Eric Betzig, Stefan Hell, and William E. Moerner (Nienhaus and Nienhaus, 2016; Moerner, 2015; Betzig, 2015; Hell, 2015; Nienhaus, 2008). As deconvolution aims to reverse the effects of blur on the image, super-resolution methods attempt to construct a higher resolution image from observations with low resolution (Laasmaa et al., 2011; Maral, 2022).

2.2 Super-resolution microscopy

Super-resolution microscopy is able to overcome the Abbe resolution limit of conventional light microscopy by using techniques that allow higher resolution. Though the topic of super-resolution is still quite new, many techniques have been developed. Prakash et al. (2022) has brought up three major super-resolution microscopy techniques. The techniques have been applied to wide-field, total internal reflection and confocal microscope setups (Schermelleh et al., 2019). These techniques are stimulated emission depletion (STED), structured illumination microscopy (SIM) and single-molecule localization microscopy (SMLM) (Prakash et al., 2022). Furthermore, other methods exist – for example ground state depletion microscopy (GSD) and the Zeiss Airyscan microscope. (Dixon et al., 2017; Wu and Hammer, 2021)

The advantages of super-resolution microscopy originate from the nature of optical fluorescence microscopy. For one, a clear advantage of fluorescence microscopy is that it allows us to perform experiments on live specimens for an extended amount of time. This, with the added resolution lets us to look at dynamics and structures on the molecular level even when using live specimens (Nienhaus and Nienhaus, 2016). Another advantage is that super-resolution microscopy techniques maintain the advantage of sample preservation, similarly to optical fluorescence microscopy. It also allows to have target specificity, which means that we are able to distinguish between different targets (Schermelleh et al., 2019).

One of the more important disadvantages of super-resolution microscopy is difficulty of labelling of the structures in the specimen. It is important to carefully choose the correct markers for labelling of the molecules of interest for a successful experiment. Moreover, modern microscopes generate enormous amounts of data that, when analysed, require a lot of computational power as well as storage space (Nienhaus and Nienhaus, 2016).

Even considering the mentioned downsides to super-resolution techniques, the techniques could provide effective and easy acquisition of data with high resolution. It is likely that many of the limitations of the mentioned techniques will be overcome in the near future.

2.3 Super-resolution methods

2.3.1 Stimulated emission depletion - STED

STED relies on the general principle of light-matter interaction. This means that for this technique any fluorescent marker, using a suitable light wavelength, can be used (Nienhaus and Nienhaus, 2016). This technique utilizes non-linear excitation/emission, at the price of inaccurately counting fluorescent signals in different parts of the sample, to become diffraction-unlimited (Prakash et al., 2022). This means that it overcomes the diffraction limit of light. In practice this technique can be applied by overlaying the excitation beam with a depletion laser beam (Schermelleh et al., 2019). Using the two beams, the fluorescent probes are first excited, then de-excited via stimulated emission. This creates a doughnut-shaped focal intensity distribution, which is diffraction limited. But with high in-

tensities the stimulated emission is saturated and all the fluorophores that are not in the center of the doughnut-shape are in the off-state. The size of the effective excited area which in the center, however, reaches subdiffraction values. The size also decreases with the intensity of the depletion beam (Vicidomini et al., 2018).

Prakash et al. (2022) have composed the pros and cons of this technique. Firstly, the benefits include lateral resolution increase for sparse or isolated structures. Another benefit is that no post-processing is not required. It also allows for variable resolution increase and high penetration depth. The main limitations include low dynamic range, sensitivity to out-of-focus background, or noise, which is not great for dense 3D features. Another limitation is high power density, which leads to phototoxicity, which means that the specimen gets damaged by the light. Another important limitation is that this technique has low imaging speed per field-of-view (Prakash et al., 2022).

2.3.2 Structured illumination microscopy - SIM

Instead of using a point-like scanning pattern like STED microscopy, this technique uses periodic lines of excitation (Nienhaus and Nienhaus, 2016). The technique of SIM uses the widely known moiré effect. This effect occurs when two fine patterns are laid on one another in a multiplicative manner. As a result, moiré fringes or a beat pattern of periodic variation in amplitude occur. In SIM the two patterns are unknown spatial distribution of the fluorescent dye and the purposely structured excitation light (Gustafsson, 2000). Spatial resolution can be improved by creating a stack of images that have been algorithmically decoded and reassembled in frequency space (Schermelleh et al., 2019). Structured illumination microscopy techniques at best, double the spatial resolution in lateral and axial directions, due to it being still fundamentally bound by laws of diffraction. This method efficiently allows for photon detection because it relies on sensitive camera detection. This method is also suitable for volumetric live-cell imaging (Schermelleh et al., 2019).

In practice, SIM could be applied by passing the illumination light of a microscope through a line-patterned grating which is located at the secondary image plane of the microscope. This line grating could be, for example, sinusoidal illumination patterns. The resulting patterns contain information about the sample structure that is unobservable due to the diffraction limit (Gustafsson, 2000; Chen et al., 2023).

2.3.3 Airyscan

Airyscan is a technique that uses a special detector array. The pinhole of the confocal microscope, which is generally set to be the size of 1 Airy Unit, is replaced with a 32-channel Gallium Arsenide Phosphide detector or GaAsP detector. The individual channels of the detector are arranged in a circular disk. This array of detectors allows for optical sectioning in all 3 dimensions with 1 Airy Unit resolution (Wu and Hammer, 2021). The Airy Unit is equivalent to the lateral point spread function of a point emitter (Weisshart, 2014). The detector allows for a two fold improvement after a pro-

cessing step in which each 32 individual elements are linearly deconvolved. This further enhances the resolution in the axial and lateral directions (Wu and Hammer, 2021).

The detector works as follows. An image the size of 1.25 Airy units is projected onto the detector using zoom optics. Each detector element acts as a small 0.2 Airy unit pinhole. With this setup, the collection efficiency of a large pinhole is maintained and the resolution benefits of a small pinhole, from using the multiple small detector elements, are acquired (Huff, 2015).

2.3.4 Single molecule localization microscopy - SMLM

SMLM is a term used for multiple super-resolution methods. These methods achieve super-resolution by individually localizing molecules (Lelek et al., 2021). The methods that fall under this term are stochastic optical reconstruction microscopy (STORM), direct stochastic reconstruction microscopy (dSTORM), photoactivation localization microscopy (PALM) as well as fluorescence photoactivation localization microscopy (fPALM) and DNA-points accumulation for imaging in nanoscale topography (DNA-PAINT) (Schermele et al., 2019).

Single molecule localization microscopy is based on the fact that single emitters can be localized with high precision if emission light from different emitters do not overlap. The overlapping of emitter PSF-s is prevented by separating the distinct fluorescent emissions in time. This could be done by employing the phenomenon of photoswitching. Photoswitching means that the molecules can switch between an excited "ON"-state and an inactive "OFF"-state. The switching is most commonly achieved by using lasers or by controlling the chemical environment, both can change the probabilities of the stochastic switching events (Lelek et al., 2021). The key requirements for applying the SMLM techniques are a powerful laser to activate only the molecules of interest, while photobleaching or altering the fluorescence of other molecules, a high numerical aperture lens to collect photons efficiently and high quantum efficiency and low noise detector (Ma et al., 2017).

The principle for single-molecule localization microscopy techniques is wide-field excitation of fluorescent probes with temporal separation of stochastic emission (Prakash et al., 2022). The detection is most commonly done using electron-multiplying charge coupled devices (EMCCD-s) and scientific complementary metal-oxide-semiconductors (sCMOS). This is due to the cameras having relatively high quantum efficiency and low noise, which are necessary for effective photon collection (Ma et al., 2017). The advantages of single-molecule localization microscopy are that it provides single-molecule sensitivity, which means that the technique is able to detect single molecules in a sample, and another advantage is that this technique most likely provides the highest potential resolution. The more prominent disadvantages are that it is very slow and that it over-represents sparse features in an image and under-represents the very dense ones (Prakash et al., 2022).

The SMLM methods differ from each other by how the photoswitching is done and by the chosen fluorophores (Lelek et al., 2021). The aforementioned *STORM* is a single molecule-localization tech-

nique that requires the properties of the emission of the fluorescent labels to be controlled by the light (Nienhaus and Nienhaus, 2016). The on-off switching is done here by the use of photoswitching of activator and reporter dye-pairs. This technique requires special buffer conditions to induce the blinking of conventional dyes (Schermelleh et al., 2019). On the other hand, *dSTORM* uses conventional fluorescent probes to transfer dyes to long-lived off-states, which are important for allowing individual molecules to be imaged with high precision. This photoswitching is done using thiols (Schermelleh et al., 2019). In simpler words, STORM achieves photoswitching with the use of an activator-dye, and *dSTORM* without an activator-dye (Lelek et al., 2021).

PALM and *fPALM* both utilize labels or fluorescent dyes that can be photoswitched or activated (Schermelleh et al., 2019). These methods differ from STORM and *dSTORM* because it does not use activator-dyes or buffers. Instead these methods use different wavelengths of light, most commonly UV-light, to excite fluorescent proteins (Lelek et al., 2021).

DNA-PAINT differs from STORM and PALM because it does not rely on the photoswitching of the dyes. Instead, it uses dyes that can freely diffuse until they interact with the targets that we are interested in. This could happen by either transient or temporary binding and permanent binding (Lelek et al., 2021).

The *GSD* technique uses a different method to get past the diffraction limit of light. The method works by reducing the number of fluorophores, that are excited at the same time (Dixon et al., 2017). This is achieved by exciting the labelled samples with a high energy laser, this way the probability of the fluorophores entering a dark or "OFF"-state is greatly increased. This state is called the triplet state (Dixon et al., 2017; Hell and Kroug, 1995). When many fluorescent labels enter the dark state, the ground state or the bright state, where the emitters are "ON" gets depleted. Then it is possible to localize single molecules (Dixon et al., 2017).

2.4 Localization of molecules

In this work, the main focus is on SMLM methods. Over the years, many different methods or algorithms have been developed for localizing the emitters for SMLM techniques. The most widely known are the nonlinear least squares estimation and the maximum likelihood estimation. Abraham et al. (2009) has evaluated these methods. Both of these methods aim to solve parametric estimation problems. The least squares method works by finding the model parameter values that produce the least difference between the model and the observed data. The maximum likelihood estimation on the other hand finds values for the model parameters that are most likely to produce the observed data (Abraham et al., 2009).

A recently developed algorithm is the SNSMIL or Shot Noise based Single Molecule Identification and Localization. This algorithm demands the user to have a graphics processing unit (GPU), for real-time molecule identification and localization. The algorithm itself is based on the intrinsic nature of noise (Tang et al., 2015). The authors of the algorithm also introduced a new quality metric, for which

the threshold can be chosen by the user. The authors concluded that their algorithm provides good results even when the noise level of the image is very high (Tang et al., 2015).

Recently, Bayesian methods have been researched to precisely estimate emitter positions on super-resolution images. These methods use Bayesian statistics to design models, which can accurately localize the positions of the pointillistic emitters in SMLM. Fazel et al. (2022) introduced the Bayesian Grouping of Localizations (BaGoL) algorithm, which uses the Reversible Jump Markov Chain Monte Carlo (RJMCMC) methodology. This MCMC method allows the algorithm to estimate the number of parameters of the model, as well as the parameters themselves, from the data. This is important since their algorithm has a variable number of parameters, which is due to the number of emitters being unknown. The first step to applying their model is to split the area of interest into subregions with small overlaps, which is done to speed up the process of calculations. Secondly, the outliers or emitters that are spaced too close to each other are removed. The third step involves the application of the RJMCMC on the subregions, as a result, a single chain of data is acquired. The fourth step in the algorithm is to generate the model with the most probable number of emitters and the posterior probability results. The final step is to combine the results of the subregions to form one high resolution image. The authors applied their algorithm on images created by using DNA-PAINT and dSTORM methods as well as synthetic data and DNA origami constructs, and were able to produce sub-nanometer precision with their approach (Fazel et al., 2022). A Bayesian method is also used in (Jazani et al., 2019). The approach to which they have designed an alternative framework for is fluorescence correlation spectroscopy (FCS). FCS is a similar approach to SMLM in the way that they both can benefit from Bayesian statistics to deduce better results. In FCS the Bayesian approach allows to achieve results fast without losing accuracy compared to other methods. Their approach aims to capture processes faster and reduce photo-toxicity or cell death in live samples, which happens due to long exposure duration. It is technique which is used to deduce physical quantities of interest, such as diffusion (Jazani et al., 2019).

Overall, the Bayesian methods for molecule position determination is a potential candidate for further development due to showing good results in fluorescence microscopy.

3 Methods

3.1 Bayesian modelling in Python

In this work, two different approaches to localizing single molecules in synthetic images are presented and evaluated.

Bayesian inference could be done in Python using a few different libraries. One of the most well-known libraries is PyMC. This library allows users to easily build their Bayesian models and fit them using Markov Chain Monte Carlo (MCMC) methods. PyMC is a probabilistic programming library with tools to specify Bayesian models and inference (Abril-Pla et al., 2023). PyMC uses MCMC algorithms like the No-U-Turn-Sampler (NUTS), which is an extension to the Hamiltonian Monte Carlo method, that does not require the number of steps to be specified. The Hamiltonian Monte Carlo method itself is a MCMC method, that circumvents the random walk behavior (Hoffman et al., 2014; Abril-Pla et al., 2023).

Another library that can be used for Bayesian model in Python is Stan. The specific library that could be used in Python is PyStan. Similarly to PyMC, the NUTS sampler is available for use in Stan. The syntax of Stan is similar to that of C++ (Carpenter et al., 2017).

The developed Bayesian inference BayesM was compared to a frequentist approach FreqM. FreqM was also developed in Python.

3.2 Generation of synthetic data

The images used to test the models in this work are all generated using a stochastic image model implemented in Python.

The PSF is defined with the following function:

$$F(X, Y) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot \exp\left(-\frac{(X - X_0)^2 + (Y - Y_0)^2}{2 \cdot \sigma^2}\right). \quad (2)$$

The emitters are placed on the images in a random manner. The locations for the emitters are generated using a random uniform distribution, which is bound by the boundary size of 5 times the width of the PSF (equation 2). The emitter emits photons based on a random normal distribution. The distribution is centered at the generated location of the emitter and has a spread equivalent to the width of the PSF (equation 2). The background of the image is added using a random uniform distribution. The background is spread over the entire image. The information of the generated image frame is saved, so are the original emitter locations.

3.3 Bayesian model

The likelihood function of the BayesM model is defined as:

$$image \sim Poisson(i, \mu(X, Y)). \quad (3)$$

The total intensity for an image can be found in the following way:

$$\mu(X, Y) = (\mu_{mol} \cdot F(X, Y) + \mu_{bg}) \Delta t, \quad (4)$$

where μ_{mol} is the molecular brightness, X and Y are emitter center coordinates, $F(X, Y)$ is the PSF, μ_{bg} is the brightness of the background, and Δt is the time step or frame time.

The function defined in equation 3 is used to model the likelihood of the observed data i given the parameters of the model $\mu(X, Y)$ (equation 4). In the case of BayesM the observed data is the intensity values of the image and $\mu(X, Y)$ is the total intensity found using the parameters that are being optimized.

The priors are defined in the following manner. The prior for molecular brightness:

$$\mu_{mol} \sim Uniform(\alpha_{min}, \alpha_{max}). \quad (5)$$

The priors for the emitter center coordinates are defined as:

$$X \sim Normal(\mu_x, \sigma_x), \quad (6)$$

$$Y \sim Normal(\mu_y, \sigma_y), \quad (7)$$

The prior for background brightness is defined as:

$$\mu_{bg} \sim Uniform(\beta_{min}, \beta_{max}). \quad (8)$$

The prior for the width of the PSF is the following:

$$sigma \sim LogNormal(\mu_s, \sigma_s). \quad (9)$$

The parameter values for the priors in equations 5, 6, 7, 8, and 9 are the following. α_{min} is set to 1×10^3 . α_{max} is set to 1×10^7 . μ_x and μ_y are obtained by finding the center of mass of the image. $sigma_x$ and $sigma_y$ are equal to 20. β_{min} and β_{max} are 0 and 1×10^5 respectively. μ_s and σ_s are 3 and 1 respectively.

3.4 Frequentist approach

The FreqM used is a least squares model with a Gaussian function. The model aims to minimize the following sum of residuals:

$$R = \sum_{i=0}^n ((\mu_{mol} \cdot F(X, Y) + \mu_{bg}) \cdot \Delta t - intensity)_i^2, \quad (10)$$

where F is the PSF, defined by equation 2, and X and Y are the emitter center coordinates. Index n is the number of emitters, μ_{mol} is the molecular brightness, μ_{bg} is the background brightness, and Δt is the frame duration.

The initial conditions used for optimize the FreqM model, that minimizes the sum of residuals (equation 10) are the following. The initial condition for the amplitude or the molecular brightness is found by dividing the maximum value of the image intensity by the time step or frame time. The initial conditions for both coordinates are found using the same center of mass values used for the BayesM model priors. The initial value for the width of the PSF is 20 pixels. Finally the initial value for background brightness is estimated as the median value of the image.

3.5 Evaluation of models

For evaluating the models performance and comparing the results, different methods are used. The error between original emitter locations and those found using a model could be found by:

$$E = \sqrt{(orig - model)^2}, \quad (11)$$

where *orig* are the original parameter values we would like to compare to the model results *model*.

3.5.1 The Bayesian inference model

The BayesM approach was implemented as a set of scripts. The organization of files and data import was performed by a single script. This script sorted the files in a natural order and presented the executable file with the correct list of files to be sent to the model. The executable script loaded the image data from the file and the information of the image was extracted using a function from the same script that organizes the file order. The information of the current file is then sent by the executable file to the PyMC model. The model performed 10000 warm-up steps and 4000 sampling steps. The model had 4 chains. The summary of the execution of the model as well as the inference data was acquired and saved. The original emitter center coordinates were saved. Altogether, the BayesM implementation uses 3 different scripts written in the Python programming language: The file organization script, the executable script and the model script.

3.5.2 Comparison of Bayesian inference libraries

For comparing the different libraries in Python, two different scripts were created and their execution duration for performing Bayesian inference was timed. The model implemented in both libraries is visualised in figure 1. Both the script using the Stan library and the script using the PyMC library were executed at the same time. This was done five times and the results were saved. For comparison of the libraries, two convergence diagnostics were taken into account. The first is the Rhat, which is the most commonly used convergence diagnostic. The Rhat is the standard deviation found from all the chains of the MCMC model included together, divided by the root mean square of the separate standard deviations of the values within each separate chain. If the Rhat is one, the chains have all converged. Another diagnostic is the effective sample size (ESS). ESS shows how many independent

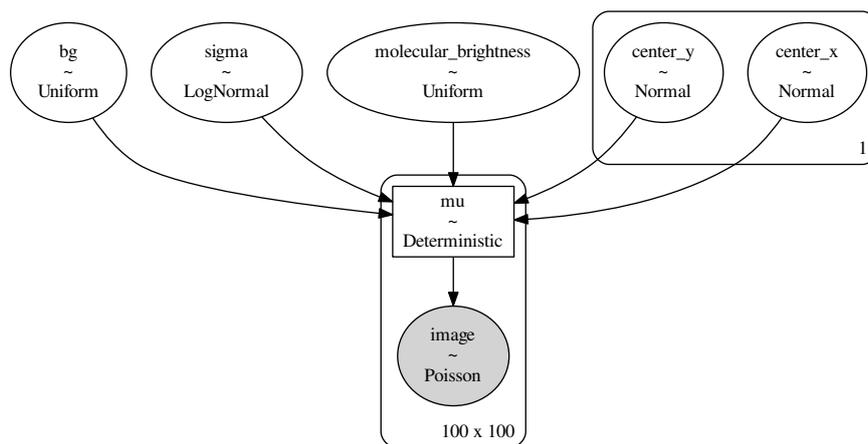


Figure 1: Model scheme for the BayesM. The top row of ovals indicate the priors for the parameters of the model. Bg signifies the prior for background, sigma the prior for the width of the PSF, molecular_brightness the molecular brightness and center_x and center_y signify the priors for the coordinates of the emitter center position. The rectangle around the coordinates shows with a marked number how many emitters are on the image. Mu is the total intensity and it is determined from all the parameters of upmost row. The image is created from the total intensity using the Poisson distribution, image is the observed data that the model is trying to explain. The rectangle around mu and image signifies that they have the same dimensions, with the dimensions shown on the bottom right of the rectangle.

draws contain the same amount of information as the dependent sample acquired by the MCMC algorithm. The larger the ESS, the better (Vehtari et al., 2021). Both the Rhat and the effective sample size (ESS) were noted and compared between the different libraries.

3.5.3 The frequentist model

The FreqM implementation uses the same file organization script used for the BayesM implementation to sort the input files and extract the information from the read files. The vector of residuals is found using a Gaussian function (equations 2 and 10) and the intensity of the image. The parameters that were being optimized by the model were molecular brightness, background brightness, the emitter center coordinates and the width of the PSF. Furthermore the modified Jacobian matrix is extracted from the results to find the covariance matrix. This is done by approximating the Hessian of the cost function.

3.5.4 Error and Kernel Density Estimates, visualisation of results

For the BayesM model results figures were created showing the results of the 4 different chains for the emitter coordinate results. Additionally, the chains were visualised as a coordinate versus iteration plot.

Student's T-test and ANOVA was performed on the data. ANOVA was performed only on images where the molecule had emitted 25 or more photons. For finding the errors for the BayesM method,

the mean values of all 4 chains were used. The found errors between the model results and the original values were visualised. The T-test results were visualised on the error figures using the following levels: if the p-value is less than 0.001, the label is "****", if the p-value is less than 0.01, the label is "***", the corresponding p-value to label "**" is less than 0.05. If p-value is greater than or equal to 0.05, the label is "ns" or not significant. The significance level for ANOVA was 0.05.

The Kernel density estimate (KDE) was found for both BayesM and FreqM. For FreqM, finding the KDE was less trivial. Since the results from the model were singular values it was necessary to make a few additional steps. The "chains" for the data were simulated using a multivariate normal distribution. Finally, the KDE was found using a Gaussian KDE function from. The results were visualised along with the original emitter location for both sets of images.

3.6 Model usage

The BayesM and FreqM model scripts were both developed, tested and executed in the TalTech HPC Center cluster. For running parallel jobs, slurm job array scripts were used. Using the cluster it was possible to efficiently acquire the results for the images. 350 jobs were run using the slurm job array and for each job (each image), 2 CPU-s were employed along with 6GB of memory.

3.7 Software

3.7.1 Packages used for Bayesian inference in Stan

The packages used for Bayesian inference in Stan include NumPy, which was used for finding the maximum intensity coordinates of the image for the model priors. NumPy was also used to make compatible data arrays from the input data for the model. PyStan was used to build and sample the model. Arviz was used to create the summary of the model as well as create inference data of the results, which could then be saved using the Python pickle module. Lastly, time module was used to time the duration of the model for library comparison.

3.7.2 Packages used for Bayesian inference in PyMC

The BayesM consisted of three separate scripts. The file organization and data import script used NumPy for array-related problems and for finding the maximum value of the image. The data import part of the script used pickle for handling the files. Natsorted package was used to sort the filenames in natural order. The centre of mass was found using SciPy *center_of_mass* function. Os module was used to access files in the cluster directories.

The model script uses two Python packages: NumPy and PyMC. PyMC is used to visualise the model using the *model_to_graphviz* function and to define the model priors. NumPy was used to perform the mathematical operations in the model script.

The executable script was more complicated and used many different modules and packages. Date-time and time were used to perform the naming of the data and the model execution time measure-

ments respectively. *Argparse* and *os* were used in file name related problems and *pickle* was used for saving the results of Bayesian inference. *PyMC* was used to perform Bayesian inference on the data and *Arviz* was used to create a summary of the model as well as create the inference data of the model results. *Pytensor* was used in slurm job array related problems.

3.7.3 Packages used for the frequentist approach

The *FreqM* script uses *NumPy*, *pickle*, *datetime*, *os*, *argparse* and *center_of_mass* similarly to the *BayesM* scripts. Least squares is performed on the data using the *SciPy* function *least_squares*.

3.7.4 Packages used for visualisation, statistical analysis

T-test was performed on the data using the *ttest_ind* function from *Scipy* package. ANOVA was performed using the *statsmodels* package. Visualisation of errors was done using the *Seaborn* package. Additionally, *Matplotlib* was used to plot the T-test results and for changing the figure parameters.

4 Results

4.1 Generation of images

350 images were generated using the image model. Each image had the dimensions of 100 x 100 pixels, the pixel size was 100 nanometers. The frame time was 0.001 seconds. Each image contained a single emitter in a random location. The emitters were located at least 5 times the width of the PSF (equation 2) away from the edges of the image. 50 images were generated for each number of photons per emitter. 7 different number of photons per emitter were chosen: 10, 25, 50, 100, 500, 650, 1000.

Additionally, 350 images were generated which also contained photons in the background. For each of these images, the number of photons in the background was 100 and they were uniformly distributed over the entire image. Figure 2 shows two examples of the generated images. Both images on the figure have 100 emitted photons per emitter, Figure 2 B. contains noise – 100 uniformly distributed background photons. Altogether, 800 images were generated for BayesM and FreqM model analysis.

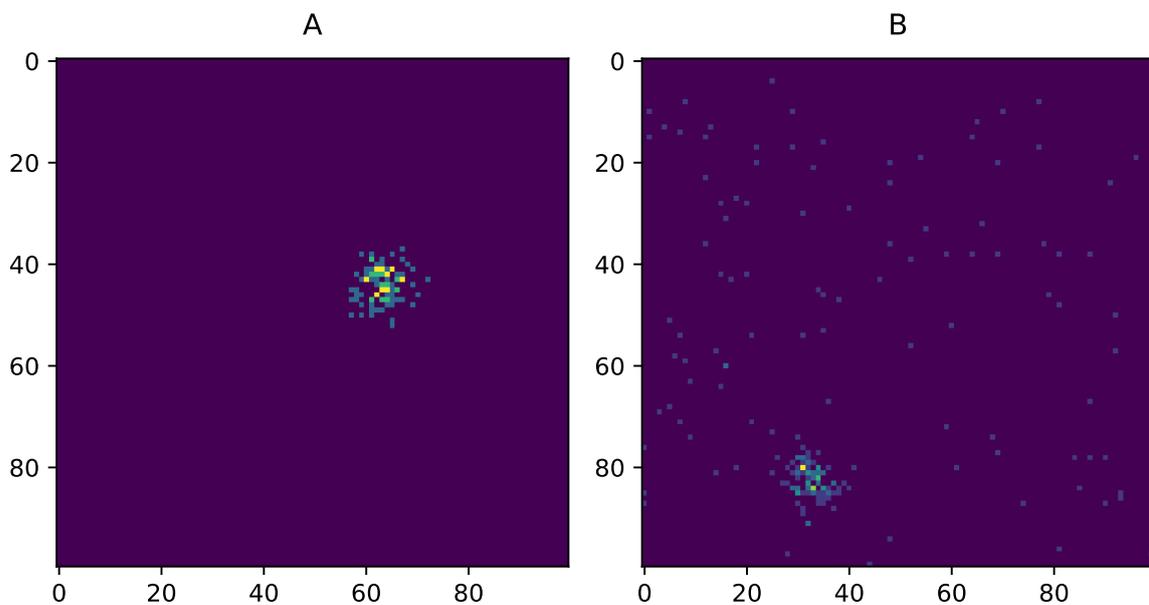


Figure 2: Synthetic images generated using the image model. A. An image with a single emitter that has emitted 100 photons. B. A single emitter with 100 emitted photons with added noise of 100 photons.

4.2 Comparison of Bayesian inference libraries

Two Python libraries, where Bayesian inference was possible to perform were tested. The two libraries were Stan and PyMC. From five runs of each Bayesian inference libraries: Stan and PyMC, resulted, that the runtime for performing Bayesian inference in Stan was 266 ± 18 seconds on average, whereas when performing Bayesian inference in PyMC the average runtime was 94 ± 8 seconds.

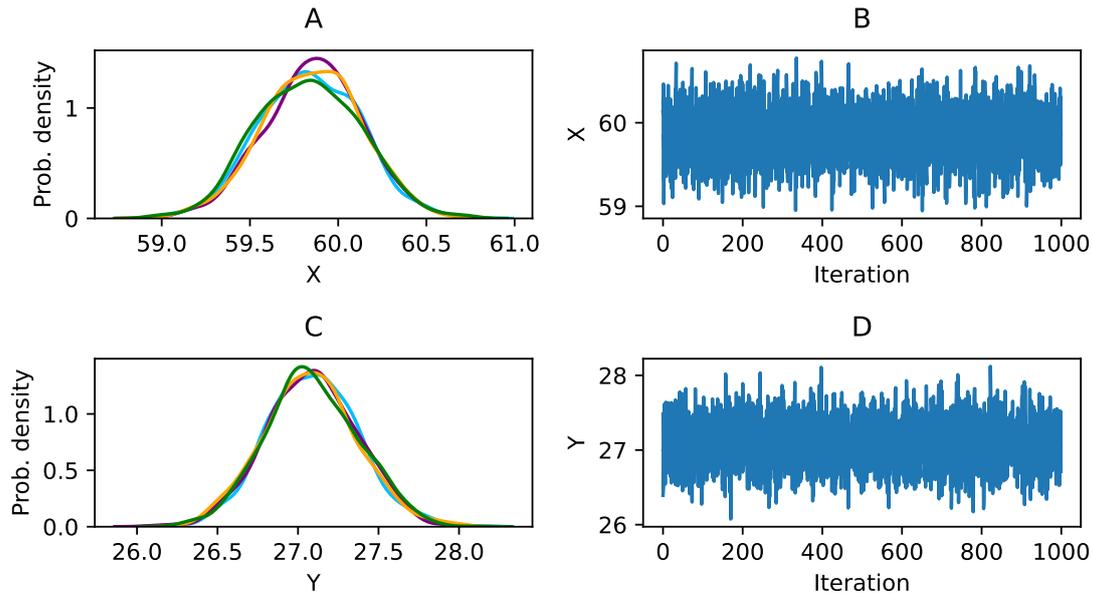


Figure 3: Trace plot for all chains for an image without added noise for BayesM results. A. and C. are plots for probability density versus image coordinates. The colors on subfigures A and B help to differentiate the 4 chains of the model. Note how the chains all follow the same curve approximately. B. and C. are plots for coordinate versus iterations or samples. The values of coordinates are in pixels for all subfigures.

That makes the PyMC library about two times faster than the Stan library.

4.3 Results of Bayesian inference and comparison with frequentist approach

The accuracy of BayesM is inspected visually by viewing the probability density plots for all 4 chains. Figures 3 and 4 show the probability density plots and the coordinate values versus iterations for an image with 100 emitted photons per emitter and an image with 100 emitted photons per emitter with an added uniformly distributed background of 100 photons respectively. From visual inspection, it is clear, that the densities of the chains are similar. The deviation of the calculations as seen on figures 3 B and D and 4 B and D is approximately 1 pixel.

Figures 5 and 6 display the original image, the emitter position found by the model and the reconstructed image side-by-side for BayesM. The same results for FreqM can be seen on figures 7 and 8. As seen from these figures, both models are able to determine the location of the emitter with reasonable accuracy. The reconstructed images portray the emitter correctly, the shape is uniform and round, when considering the two-dimensional images. The center appears the brightest and intensity fades gradually towards the outside of the emitter.

Errors found using equation 11 in the case of emitter position coordinates are displayed on figures 9, 10. Additionally these results were composed into tables 1. and 2. The results indicate that BayesM gradually surpasses the performance of FreqM, when images contain an emitter with 100 photons per emitter or more. The FreqM approach shows better results when the number of photons per emitter is small, this result is especially evident for 10 photons per emitter. In the case of images

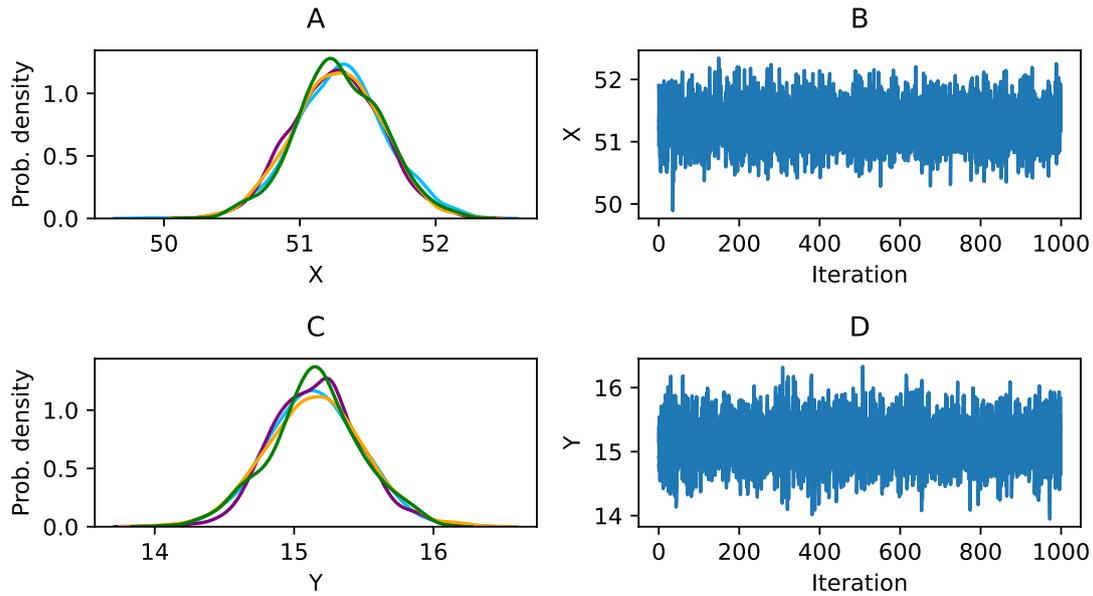


Figure 4: Trace plot for all chains for image with noise for BayesM results. The contents of the subfigures are as in Figure 3.

with no added background photons, the mean values for error are lower for both coordinate values for the cases with 25 and more photons per emitter for BayesM method compared to FreqM method. For images with 100 added background photons, the mean values of error of BayesM method are lower than those of the FreqM method when considering images with 500 and more photons per emitter. Both methods remain quite even in performance when the emitter has emitted 25, 50 and 100 photons for this case. These results are also indicated by the T-test results on the data. Though there are some discrepancies between the two coordinate results, it is clear, that there are significant differences between the means of the results from FreqM and BayesM in the case of 10 emitted photons per emitter. The results from ANOVA indicate, that the effect of number of photons on the error shows significance for the generated images (p -value 0.004 for images with no noise, p -value < 0.001 for images with noise). The effect of method on the error was found to be significant in the case of no background noise (p -values < 0.001), when the image contained noise, the effect of method on error was not significant (p -value 0.006). The interaction effect of method and number of photons showed significance when the image had no noise and this was the opposite for the case when the image contained background noise (p -value 0.005 for images with no noise and p -value 0.9 for images with noise).

The resulting errors of determining the width of the PSF with BayesM and FreqM are shown figures 11 and 12. It is apparent from both figures, that the BayesM method finds the width of the PSF more accurately for both, images with no noise and images with 100 added background photons. There is a noticeable difference between the results for images with no added background noise and images with added noise. The errors for the BayesM method in the case of determining the PSF width are much lower than those of the FreqM method. The results of T-test indicate significant differences between the methods from 10 to 500 photons per emitter for images with no noise. When the emit-

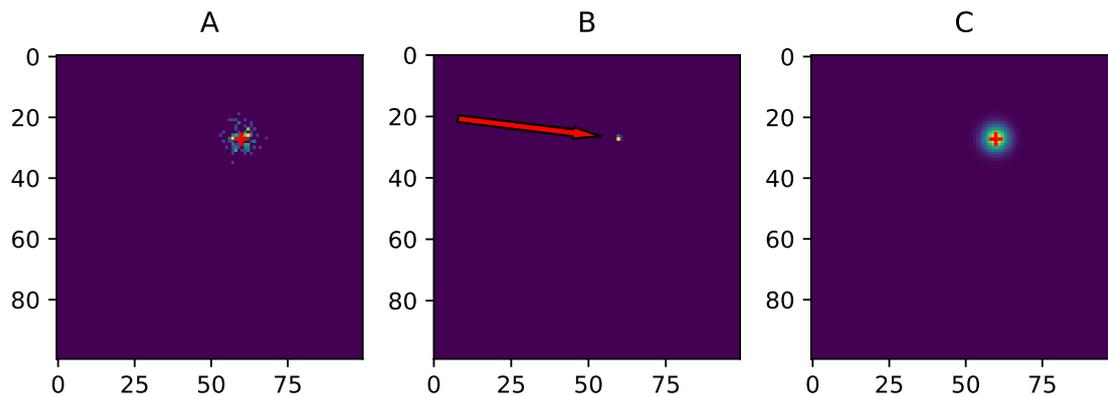


Figure 5: Localization accuracy for image with no noise (BayesM). Subfigure A displays the original generated image. The red cross marks the original emitter location. B. Detected location of the emitter. C. Reconstructed image, red cross marks the original emitter location. Subfigure B features a red arrow instead of a red cross for figure readability. The coordinate values are in pixels.

ter has emitted 650 to 1000 photons, the significance is not that prominent, for 650 photons, the difference is not significant. When the image contains background photons, there is a significance of differences for the cases when the emitter has emitted 10, 25, 100, 1000 emitters. The significance is more prominent for the case of 10 photons emitted per emitter. The ANOVA results indicated a significance of the effect of method and number of photons on the error for the case of no background noise and images with background noise (p -values < 0.001). Similarly, the interaction of method and number of photons showed a significance for both types of images (p -value < 0.001).

The KDE results for BayesM and FreqM can be found on figures 13 and 14. From visual inspection of the two figures, it is apparent, that for the selected images, BayesM more accurately estimates the emitter location, compared to FreqM. BayesM appears to have a good performance in the terms of bias and variance. The FreqM estimate is still quite near the original value, but is not as accurate as BayesM.

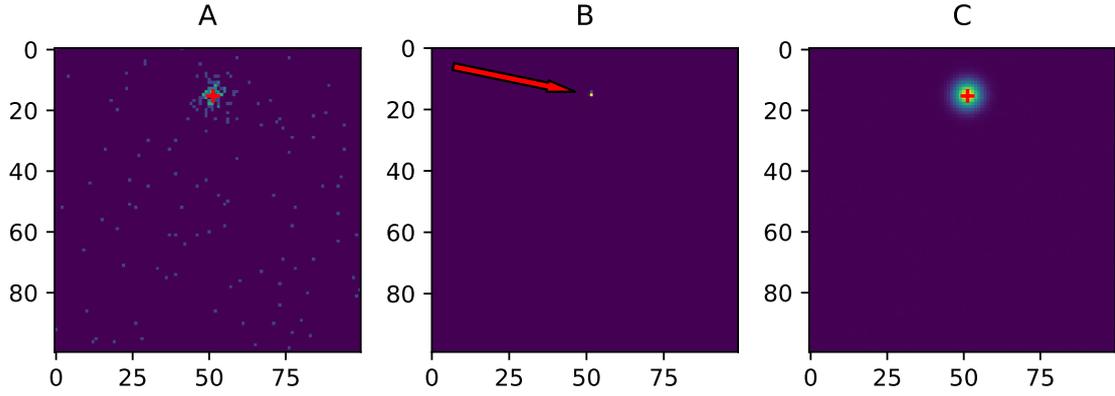


Figure 6: Localization accuracy for image with noise (BayesM). Same notation as in Figure 5 is used.

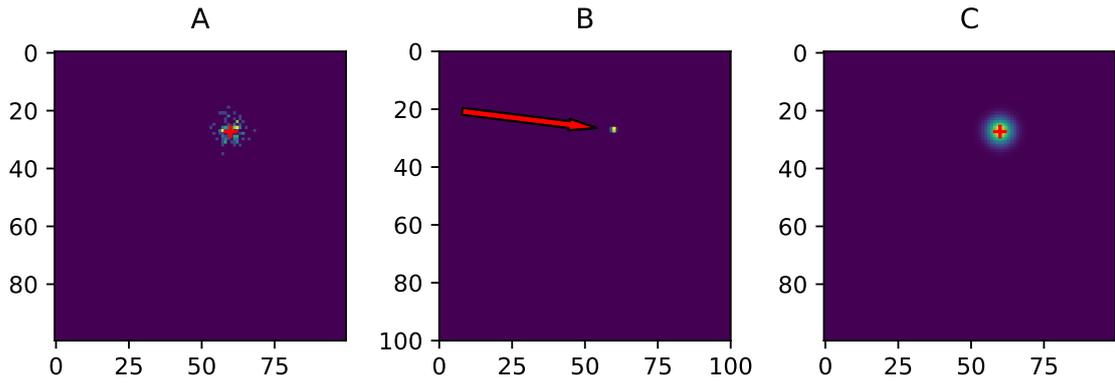


Figure 7: Localization accuracy for image without added noise (FreqM). Same notation as in Figure 5 is used.

Table 1: Mean errors for images without noise of Figure 9. All values in micrometers. X and Y signify the x- and y-coordinates of the emitter center coordinates.

Photons	Bayesian X		Bayesian Y		Gaussian X		Gaussian Y	
10	2.6	± 1.9	3.5	± 1.6	0.08	± 0.06	0.1	± 0.08
25	0.05	± 0.04	0.05	± 0.03	0.07	± 0.05	0.07	± 0.07
50	0.03	± 0.02	0.03	± 0.03	0.05	± 0.04	0.05	± 0.04
100	0.02	± 0.02	0.02	± 0.02	0.03	± 0.02	0.03	± 0.02
500	0.009	± 0.007	0.011	± 0.009	0.013	± 0.01	0.01	± 0.01
650	0.008	± 0.007	0.011	± 0.006	0.01	± 0.01	0.02	± 0.01
1000	0.009	± 0.006	0.008	± 0.006	0.011	± 0.008	0.01	± 0.007

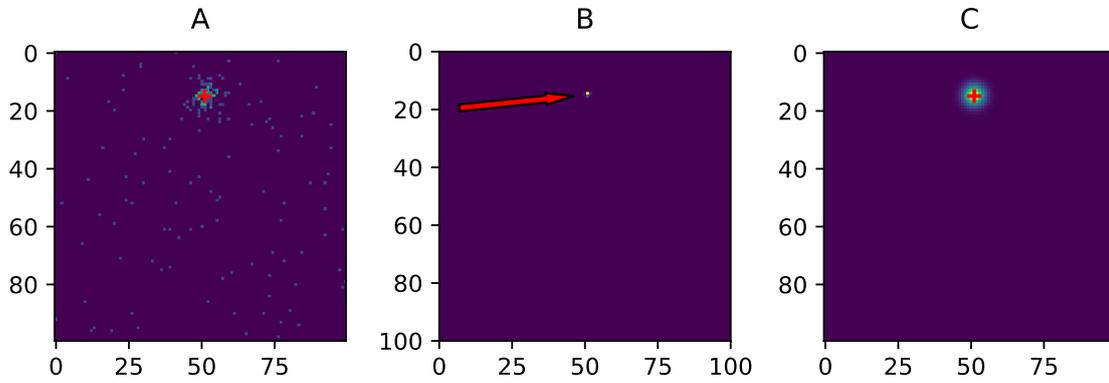


Figure 8: Localization accuracy for image with noise (FreqM). Same notation as in Figure 5 is used.

Table 2: Mean errors for images with noise of Figure 10. All values are in micrometers. X and Y signify the x- and y-coordinates of the emitter center coordinates.

Photons	Bayesian X		Bayesian Y		Gaussian X		Gaussian Y	
10	1.8	± 1.3	2	± 1.4	0.1	± 0.08	0.12	± 0.09
25	0.06	± 0.05	0.05	± 0.04	0.06	± 0.06	0.06	± 0.07
50	0.04	± 0.03	0.03	± 0.02	0.04	± 0.04	0.04	± 0.03
100	0.03	± 0.02	0.03	± 0.02	0.03	± 0.02	0.03	± 0.03
500	0.008	± 0.007	0.009	± 0.007	0.013	± 0.008	0.013	± 0.009
650	0.009	± 0.007	0.008	± 0.007	0.012	± 0.009	0.012	± 0.01
1000	0.008	± 0.006	0.007	± 0.006	0.009	± 0.007	0.011	± 0.008

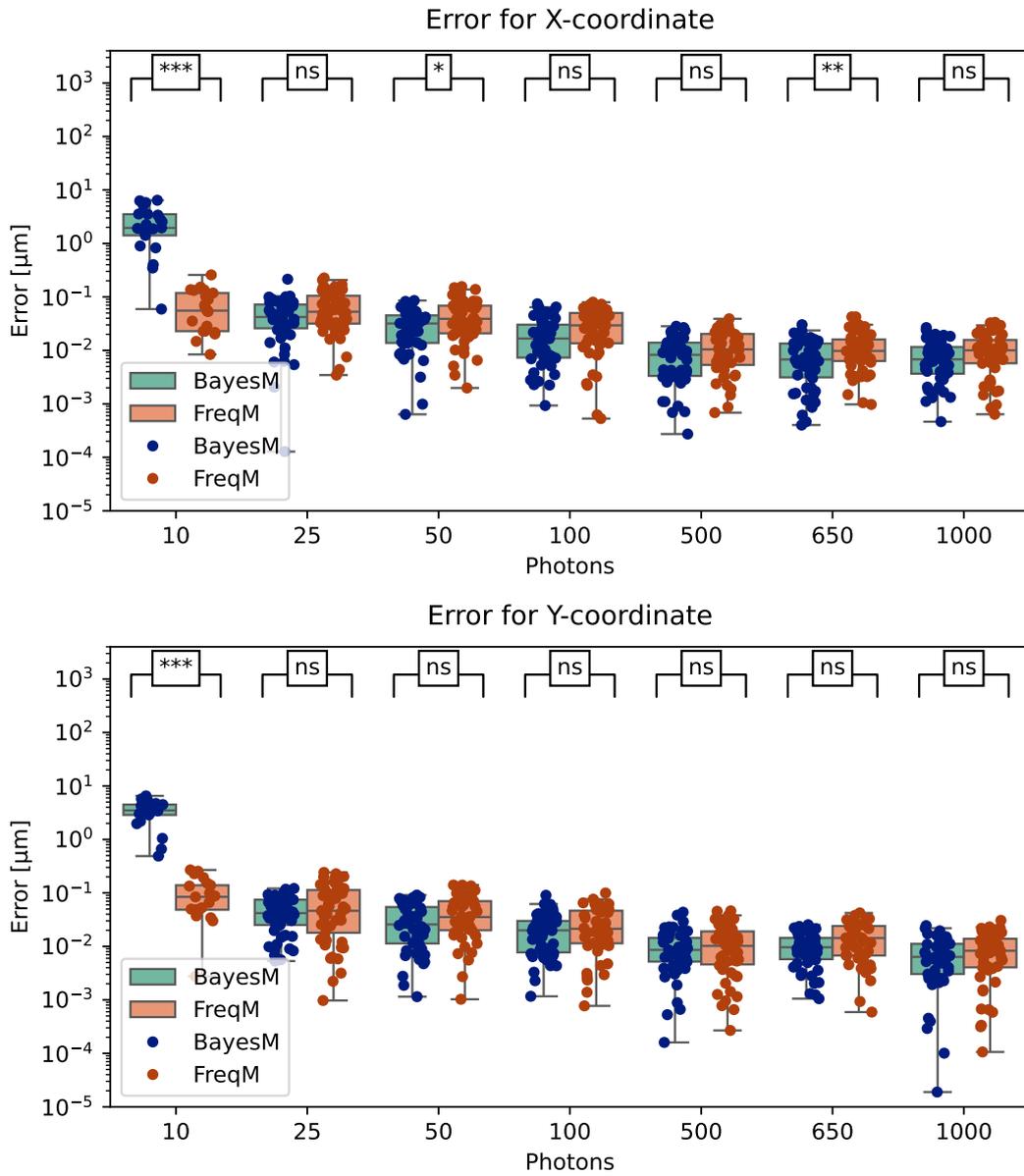


Figure 9: Errors for emitter center coordinates for results of image without added background noise. Displayed on the figure are T-test results performed to observe the significance of method on the error. See the results subsection for the significance levels. Note how the error values decrease as the number of photons per emitter is increased.

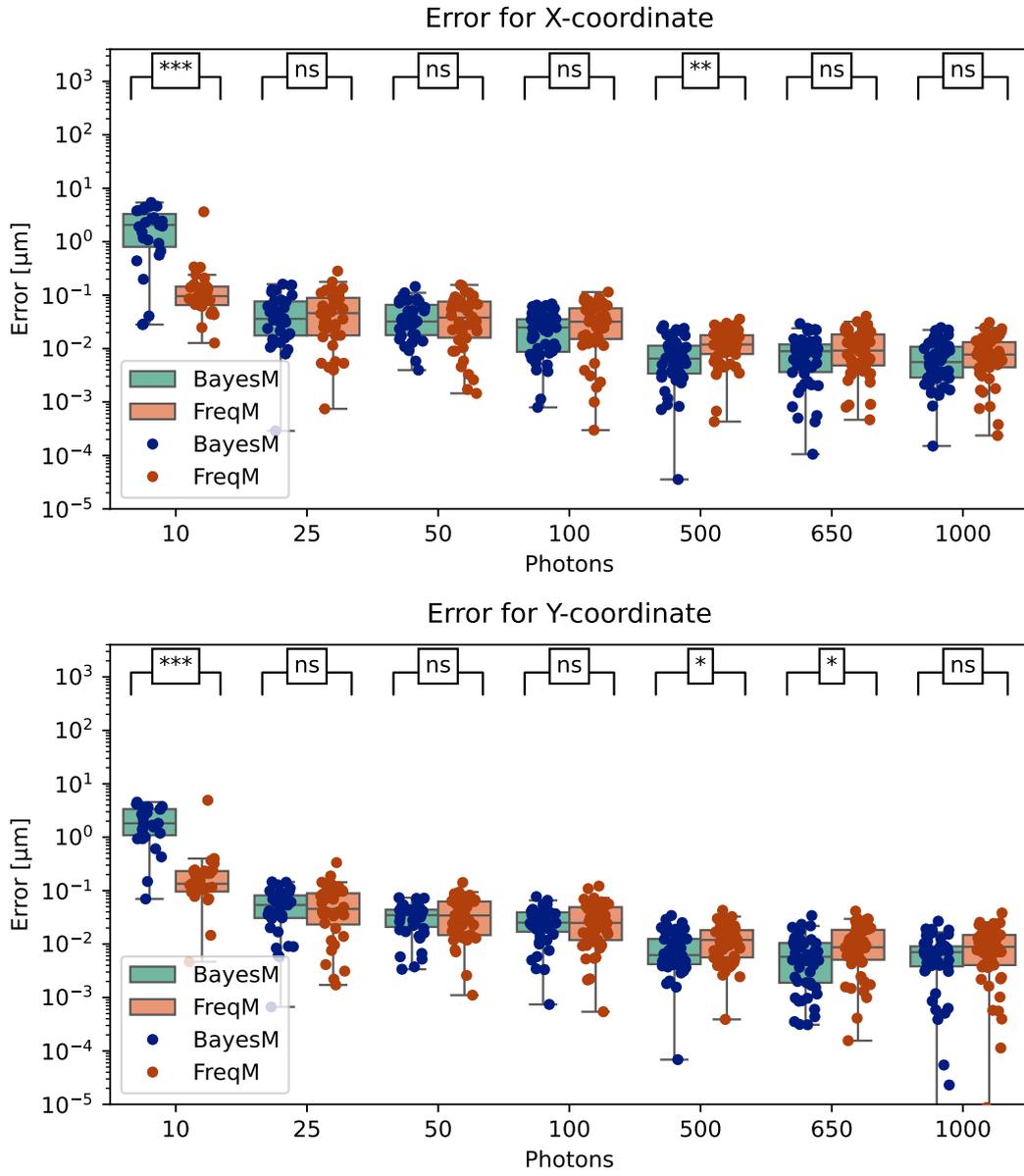


Figure 10: Errors for coordinates results of image with noise. Same notation as in Figure 9 is used.

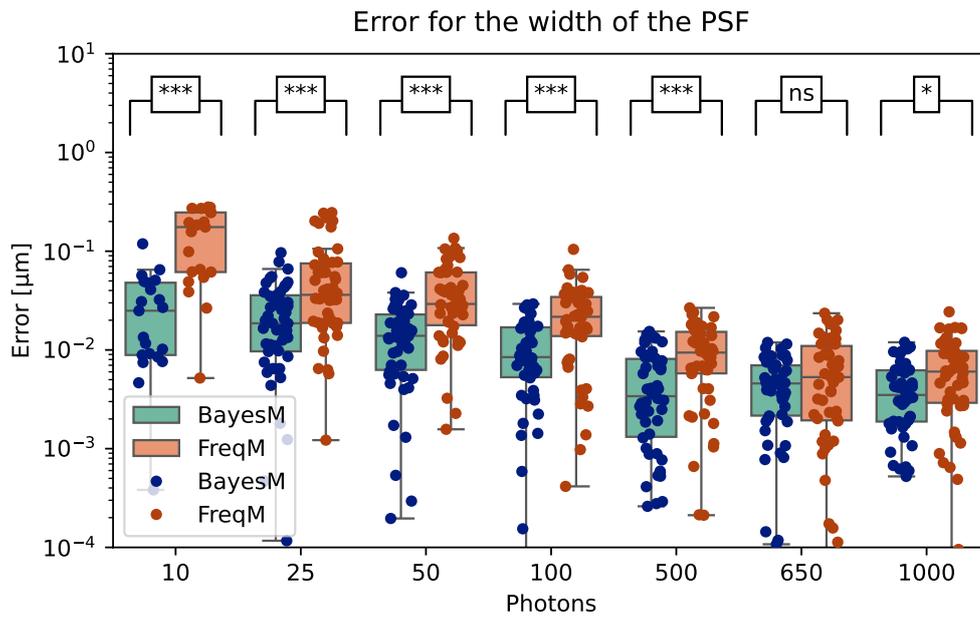


Figure 11: Errors for the width of the PSF for results of image without noise. T-test results are displayed at the top of this figure. Note how the difference of errors between the methods has a slight decrease as the number of photons per emitter is increased.

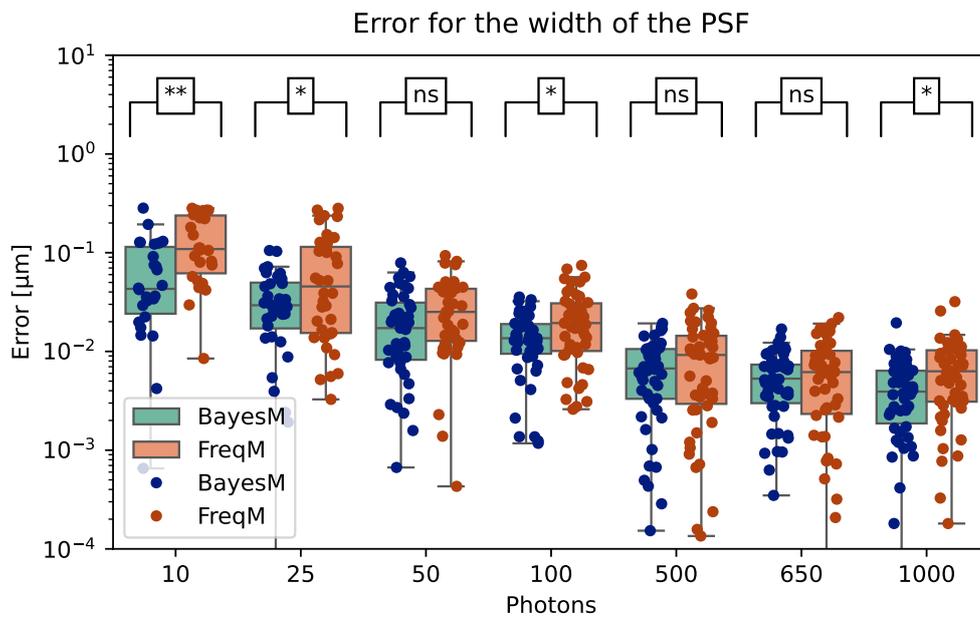


Figure 12: Errors for the width of the PSF for results of image with noise. Same notation as in Figure 11 is used.

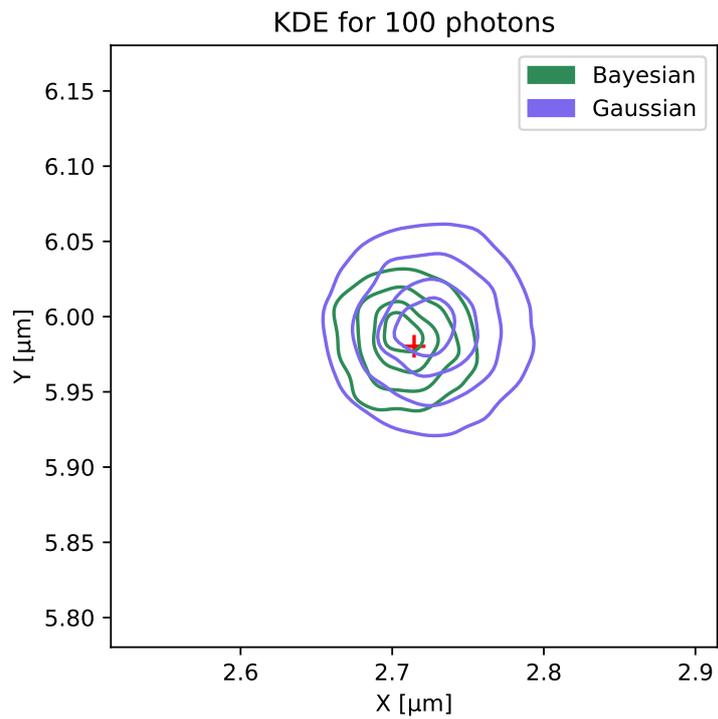


Figure 13: Figure showing the KDEs of BayesM and FreqM emitter detection of image without added noise. The red cross signifies the original emitter location. The levels represented correspond to 25, 50, 75 and 90 percentile.

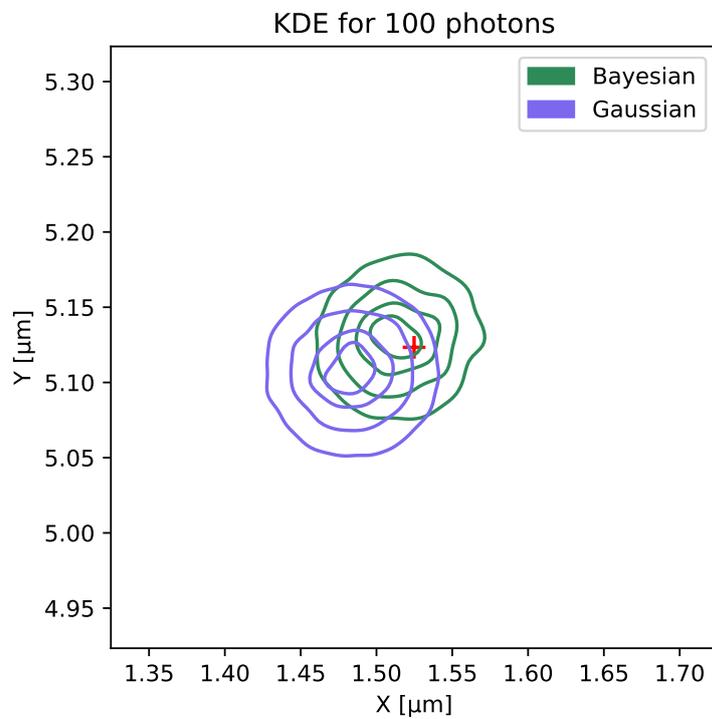


Figure 14: Figure showing the KDEs of BayesM and FreqM emitter detection of image with added noise. Same notation as in Figure 13 is used.

5 Discussion

Analysis of the BayesM model and the comparison with FreqM indicated that BayesM outperforms the FreqM model when the emitter has emitted around 25 or more photons. There is a clear difference in localization precision when the molecule has emitted 100 or more photons and in that case, BayesM gives better results. Considering this, the BayesM could be used in applications where the number of photons emitted by the molecule is moderate to large.

For using the BayesM in future applications, such as applying it to real microscope images, the model can be adapted to be suitable for the occasion. This involves changing the prior distributions based on information about the microscopy setup as well as the special information about the emitters when dealing with multiple frames.

When the image of interest contains multiple frames, a new problem arises for the BayesM. The emitter could be shown in a bright state on multiple frames, this causes the model to think that there are multiple emitters in the same spot instead of one. This problem could be investigated in the future and possibly solved using knowledge of the excitation probabilities of fluorophores. For even further investigations, the multiple emitter problem could be investigated. In this case, multiple emitters have been placed on one frame. In this case, a method needs to be devised to correctly label the emitters in the MCMC chains. Assume, we have emitter labelled as 1 and an emitter labelled as 2. The model does not recognize them as a scientist would naturally label them, that means that the coordinates could be swapped in different chains of the calculations leading to the complete image of just one emitter instead of two due to the chains being averaged to complete the final result. As an alternative, it is possible to correct the chain orders as a post-processing step. This can be done by arranging the chains according to the order of the first calculations.

Compared to previous works regarding Bayesian methods for localizing emitter positions, the developed and applied in this work uses 4 chains for MCMC, Fazel et al. (2022) only had one chain for the MCMC algorithm. The authors also did not perform convergence checks like R_{hat} due to the fact, that their algorithm made use of only one chain. They also did not perform the ESS convergence check. However, their algorithm is a promising method to more precisely localize and group multiple blinking events. This hints to the BayesM method possibly providing good results when extended to solve similar problems.

The results of BayesM show that it is not advised to apply for images where there are few emitted emitter photons. It is not quite clear why this is, thus the problem should be investigated further. Currently it is recommended to use the FreqM to localize emitters for the cases with few emitted photons emitted per fluorescent probe.

In conclusion, the BayesM shows a great potential to localizing emitters on super-resolution images. Future developments are required to make BayesM compatible to use for images with a small amount of photons emitted per emitter, as well as images with multiple frames and multiemitter cases.

6 Acknowledgements

The author would like to thank the supervisors of this work, Marko Vendelin and Martin Laasmaa, for skillful supervising. The author would also like to thank Diana Wide and the author's loved ones for support throughout the process of composing this work.

References

- Abbe, E. (1873). Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv für Mikroskopische Anatomie*, 9(1):413–468.
- Abraham, A. V., Ram, S., Chao, J., Ward, E. S., and Ober, R. J. (2009). Quantitative study of single molecule location estimation techniques. *Opt. Express*, 17(26):23352–23373.
- Abril-Pla, O., Andreani, V., Carroll, C., Dong, L., Fannesbeck, C. J., Kochurov, M., Kumar, R., Lao, J., Luhmann, C. C., Martin, O. A., et al. (2023). Pymc: a modern, and comprehensive probabilistic programming framework in python. *PeerJ Computer Science*, 9:e1516.
- Betzig, E. (2015). Single molecules, cells, and super-resolution optics (nobel lecture). *Angewandte Chemie International Edition*, 54(28):8034–8053.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, 76.
- Chen, X., Zhong, S., Hou, Y., Cao, R., Wang, W., Li, D., Dai, Q., Kim, D., and Xi, P. (2023). Superresolution structured illumination microscopy reconstruction algorithms: A review. *Light: Science & Applications*, 12(1):172.
- Cheng, P.-C. (2006). The contrast formation in optical microscopy. In Pawley, J. B., editor, *Handbook of Biological Confocal Microscopy*, pages 162–206. Springer US, Boston, MA.
- Dey, N., Blanc-Feraud, L., Zimmer, C., Roux, P., Kam, Z., Olivo-Marin, J.-C., and Zerubia, J. (2006). Richardson–Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution. *Microscopy Research and Technique*, 69(4):260–266.
- Dixon, R. E., Vivas, O., Hannigan, K. I., and Dickson, E. J. (2017). Ground state depletion super-resolution imaging in mammalian cells. *JoVE (Journal of Visualized Experiments)*, (129):e56239.
- Fazel, M., Wester, M. J., Schodt, D. J., Cruz, S. R., Strauss, S., Schueder, F., Schlichthaerle, T., Gillette, J. M., Lidke, D. S., Rieger, B., Jungmann, R., and Lidke, K. A. (2022). High-precision estimation of emitter positions using Bayesian grouping of localizations. *Nature Communications*, 13(1):7152.
- Fornaçon-Wood, I., Mistry, H., Johnson-Hart, C., Faivre-Finn, C., O’Connor, J. P., and Price, G. J. (2022). Understanding the differences between bayesian and frequentist statistics. *International journal of radiation oncology, biology, physics*, 112(5):1076–1082.
- Gustafsson, M. G. (2000). Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of microscopy*, 198(2):82–87.
- Hell, S. W. (2015). Nanoscopy with focused light (nobel lecture). *Angewandte Chemie International Edition*, 54(28):8054–8066.
- Hell, S. W. and Kroug, M. (1995). Ground-state-depletion fluorescence microscopy: A concept for breaking the diffraction resolution limit. *Applied Physics B*, 60(5):495–497.

- Hoffman, M. D., Gelman, A., et al. (2014). The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *J. Mach. Learn. Res.*, 15(1):1593–1623.
- Huff, J. (2015). The Airyscan detector from ZEISS: Confocal imaging with improved signal-to-noise ratio and super-resolution. *Nature Methods*, 12(12):i–ii.
- Jazani, S., Sgouralis, I., Shafraz, O. M., Levitus, M., Sivasankar, S., and Pressé, S. (2019). An alternative framework for fluorescence correlation spectroscopy. *Nature Communications*, 10(1):3662.
- Laasmaa, M., Branovets, J., Stolova, J., Shen, X., Rätsepso, T., Balodis, M. J., Grahv, C., Hendrikson, E., Louch, W. E., Birkedal, R., and Vendelin, M. (2023). Cardiomyocytes from female compared to male mice have larger ryanodine receptor clusters and higher calcium spark frequency. *The Journal of Physiology*, 601(18):4033–4052.
- Laasmaa, M., Vendelin, M., and Peterson, P. (2011). Application of regularized Richardson–Lucy algorithm for deconvolution of confocal microscopy images. *Journal of Microscopy*, 243(2):124–140.
- Lelek, M., Gyparaki, M. T., Beliu, G., Schueder, F., Griffié, J., Manley, S., Jungmann, R., Sauer, M., Lakadamyali, M., and Zimmer, C. (2021). Single-molecule localization microscopy. *Nature Reviews Methods Primers*, 1(1):39.
- Ma, H., Fu, R., Xu, J., and Liu, Y. (2017). A simple and cost-effective setup for super-resolution localization microscopy. *Scientific Reports*, 7(1):1542.
- Maral, B. C. (2022). Single image super-resolution methods: A survey. *arXiv preprint arXiv:2202.11763*.
- Moerner, W. E. W. E. (2015). Single-molecule spectroscopy, imaging, and photocontrol: Foundations for super-resolution microscopy (nobel lecture). *Angewandte Chemie International Edition*, 54(28):8067–8093.
- Nienhaus, G. (2008). The green fluorescent protein: A key tool to study chemical processes in living cells. *Angewandte Chemie International Edition*, 47(47):8992–8994.
- Nienhaus, K. and Nienhaus, G. U. (2016). Where Do We Stand with Super-Resolution Optical Microscopy? *Study of biomolecules and biological systems: Proteins*, 428(2, Part A):308–322.
- Prakash, K., Diederich, B., Heintzmann, R., and Schermelleh, L. (2022). Super-resolution microscopy: A brief history and new avenues. *Philosophical Transactions of the Royal Society A*, 380(2220):20210110.
- Rudi Rottenfusser, Erin E. Wilson, and Michael W. Davidson (n.d.). Image Formation in the Microscope. Accessed on: May 12, 2024.
- Schermelleh, L., Ferrand, A., Huser, T., Eggeling, C., Sauer, M., Biehlmaier, O., and Drummen, G. P. C. (2019). Super-resolution microscopy demystified. *Nature Cell Biology*, 21(1):72–84.

- Tang, Y., Dai, L., Zhang, X., Li, J., Hendriks, J., Fan, X., Gruteser, N., Meisenberg, A., Baumann, A., Katranidis, A., et al. (2015). Snsmil, a real-time single molecule identification and localization algorithm for super-resolution fluorescence microscopy. *Scientific reports*, 5(1):11073.
- Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., and Bürkner, P.-C. (2021). Rank-normalization, folding, and localization: An improved \hat{r} for assessing convergence of mcmc (with discussion). *Bayesian analysis*, 16(2):667–718.
- Vicidomini, G., Bianchini, P., and Diaspro, A. (2018). STED super-resolved microscopy. *Nature Methods*, 15(3):173–182.
- Weisshart, K. (2014). The basic principle of airyscanning. *Zeiss Technology Note*, 22.
- Wu, X. and Hammer, J. A. (2021). ZEISS airyscan: Optimizing usage for fast, gentle, super-resolution imaging. In *Confocal Microscopy: Methods and Protocols*, pages 111–130. Springer.

7 Abstract

The purpose of this thesis is to develop and evaluate the performance of a Bayesian inference model used for single-molecule localization microscopy. The Bayesian inference model aims to localize emitters on images with high precision. Another objective is to compare the Bayesian inference model to a frequentist inference approach.

In the main part of this thesis, the nature of super-resolution techniques are explained. The different techniques of super-resolution are described, mostly focusing on single-molecule localization microscopy techniques. Lastly the developed models and evaluation methods are described and results are presented.

In the practical part of the thesis a Bayesian inference model was developed to localize emitters on synthetically generated images. Two different Bayesian inference libraries are tested to determine which is faster to use for the images used in this work. The image data for this work, which was generated using a stochastic image model is used to test the Bayesian inference model and compare it to a frequentist approach. The models were implemented in the Python programming language.

As a result of applying the model on synthetic images, it was found that the Bayesian inference model works well to localize emitters on synthetic super-resolution images, where the emitter has emitted 25 or more photons. It was also concluded, that the Bayesian inference achieves smaller errors compared to the frequentist approach when the emitter has emitted many photons. As a result of using the models on images with emitters that have emitted few photons it was found, that the frequentist approach achieves smaller error values. Additionally, future directions for the Bayesian approach are given in the present work.

Key words: super-resolution, Bayesian inference, frequentist inference, single-molecule localization microscopy

8 Annotatsioon

Antud töö eesmärgiks on arendada Bayesi otsustusmudel ühe molekuli lokaliseerimise mikroskoopia jaoks ning selle jõudlust hinnata. Bayesi otsustusmudeli eesmärk on lokaliseerida kõrge täpsusega fluorestsete molekulide asukohti mikroskoopia piltidel. Teine eesmärk on võrrelda Bayesi otsustusmudelit sagedusotsustuse lähenemisega.

Töö põhiosas selgitatakse super-resolutsiooni tehnikate olemust. Kirjeldatakse erinevaid super-resolutsiooni tehnikaid, keskendudes ühe molekuli lokaliseerimise mikroskoopia tehnikatele. Viimasena kirjeldatakse arendatud mudeleid ja hinnangute meetodeid ning esitatakse saadud tulemusi.

Töö praktilises osas arendati Bayesi otsustusmudel molekulide lokaliseerimiseks tehislikel mikroskoopia piltidel. Testiti kahte erinevat teeki Bayesi otsustusmudel, et võrrelda kumb on kiirem antud töös kasutatud piltide parendamiseks. Töös kasutatud pildid genereeriti kasutades stohhastilist pilditekkemudelit, ning neid kasutati Bayesi otsustusmudeli ning sagedusotsustusmudeli võrdluseks. Mudeleid rakendati programmeerimiskeeles *Python*.

Mudeli rakendamisel tehislikel piltidel leiti, et Bayesi otsustusmudel saavutab häid tulemusi fluorestsete molekulide lokaliseerimisel tehislikel super-resolutsiooni piltidel, kus molekul on kiiranud 25 või rohkem footonit. Töö tulemusena järelitati, et Bayesi otsustusmudel saavutab väiksemaid vigu võrreldes sagedusotsustusmudel kui molekul on kiiranud palju footoneid. Mudelite rakendamisel piltidel, kus molekul on kiiranud vähe footoneid leiti, et sagedusotsustusmudel saavutab väiksemaid veaväärtusi. Lisaks antakse töö lõpus tulevikusuundi Bayesi lähenemisviisi kohta.

Võtmesõnad: super-resolutsioon, Bayesi otsustus, sagedusotsustus, ühe molekuli lokaliseerimise mikroskoopia

A Appendices

A.1 Non-exclusive licence

Annex
to Rector's directive No 1-8/17 of 7 April 2020

Non-exclusive licence for reproduction and publication of a graduation thesis¹

I _____Cärolin Grahv_____ (author's name)

1. grant Tallinn University of Technology free licence (non-exclusive licence) for my thesis
____A Bayesian approach for super-resolution microscopy image
reconstruction_____

(title of the graduation thesis)

supervised by _____Martin Laasmaa, Marko Vendelin_____,
(supervisor's name)

1.1 to be reproduced for the purposes of preservation and electronic publication of the graduation thesis, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright;

1.2 to be published via the web of Tallinn University of Technology, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright.

2. I am aware that the author also retains the rights specified in clause 1 of the non-exclusive licence.

3. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights, the rights arising from the Personal Data Protection Act or rights arising from other legislation.

____15.05.2024____ (date)

¹ The non-exclusive licence is not valid during the validity of access restriction indicated in the student's application for restriction on access to the graduation thesis that has been signed by the school's dean, except in case of the university's right to reproduce the thesis for preservation purposes only. If a graduation thesis is based on the joint creative activity of two or more persons and the co-author(s) has/have not granted, by the set deadline, the student defending his/her graduation thesis consent to reproduce and publish the graduation thesis in compliance with clauses 1.1 and 1.2 of the non-exclusive licence, the non-exclusive license shall not be valid for the period.