ADVANCING SUPER RESOLUTION MICROSCOPY FOR QUANTITATIVE IN-VIVO IMAGING OF CHROMATIN NANODOMAINS

by

Clayton Seitz

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Physics and Astronomy

West Lafayette, Indiana

December 2024

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Gautam Vemuri, Chair

Department of Physics

Dr. Jing Liu

Department of Physics

Dr. Ruihua Cheng

Department of Physics

Dr. Stephen Wassall

Department of Physics

Dr. Horia Petrache

Department of Physics

Approved by:

Dr. Jing Liu

I dedicate this thesis to those who have encouraged my pursuit of a doctoral degree.

To deal with a 14-dimensional space, visualize a 3-dimensional space and say 'fourteen' to yourself very loudly. Everyone does it

Geoffrey Hinton

Information is the resolution of uncertainty

Claude Shannon

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES 6

LIST OF FIGURES 7

LIST OF SYMBOLS 8

ABBREVIATIONS 9

ABSTRACT 10

1 Single molecule localization microscopy 11

1.1 Introduction 11

1.1.1 Elementary mathematical theory of SMLM 12

1.1.2 The definition of resolution in SMLM 15

1.1.3 The Cramer-Rao lower bound 18

2 Bromodomain protein 4 and chromatin organization 20

3 Uncertainty-aware localization microscopy by variational diffusion 21

A Appendix A 23

A.0.1 Interpolation by second order coherence 23

A.0.2 Generalization to nonzero background 25

A.0.3 Generalization to multi-level systems 25

A.0.4 Details of the Gaussian PSF 26

A.0.5 Fisher information for 2D integrated gaussian 27

B CITATIONS AND REFERENCES 29

B.1 Citations 29

B.2 References 31

VITA 33

LIST OF TABLES

LIST OF FIGURES

1.1 **Stochastic optical reconstruction microscopy (STORM)**. (A) Single molecules are resolved by separating their fluorescent emission in time, using fluorophores with multiple photophysical states (B) Example super-resolution image of H2B protein in a living Hela cell nucleus at 37C, 5 percent CO2. Image reconstructed from 10^3 10ms frames. Scalebar 5um. 14

1.2 **Noise model for CMOS cameras used for MLE**. (left)) CMOS offset for zero incident photons (middle) CMOS variance for zero incident photons (upper right) Cumulative mass function for the convolution distribution and its Poisson approximation for rate parameter $\mu_k = 500$ counts (lower right) Komogonov distance measured as a function of rate parameter μ_k 15

1.3 **Computing epistemic uncertanties with Metropolis-Hastings**. (top left) Simulated point spread function for $N_0 = 10^3$ photons with a red x at x_{MLE} and y_{MLE} (bottom left) Acceptance rate of Markov chain (middle) Markov chains sampling from the posterior distribution on molecule coordinates in 3D, with the maximum likelihood estimation in dashed blue (right) Estimated posterior marginals on the localization parameters with their respective uncertanties . . . 16

1.4 **Noise model for CMOS cameras used for MLE**. (left)) CMOS offset for zero incident photons (middle) CMOS variance for zero incident photons (upper right) Cumulative mass function for the convolution distribution and its Poisson approximation for rate parameter $\mu_k = 500$ counts (lower right) Komogonov distance measured as a function of rate parameter μ_k 18

LIST OF SYMBOLS

m mass

v velocity

ABBREVIATIONS

abbr	abbreviation
bcf	billion cubic feet
BMOC	big man on campus

ABSTRACT

This dissertation introduces single molecule localization microscopy and covers work discussed in the following papers: BRD4 phosphorylation state regulates structure of chromatin nanodomains [1] describes the role of the BRD4 phosphoswitch in the maintenance of chromatin nanodomains via super resolution microscopy and molecular dynamics simulation. We build on the notion that chromatin binding activity of BRD4 is regulated by phosphorylation by demonstrating that BRD4 phosphorylation regulated chromatin packing and mobility in mammalian nuclei.

Denoising diffusion probabilistic models for blind deconvolution in single molecule localization microscopy [2] describes an algorithm that leverages a novel paradigm for deep generative modeling using Gaussian diffusions in order to enhance the resolution of localization microscopy images.

Single-molecule localization microscopy (SMLM) techniques, such as direct stochastic optical reconstruction microscopy (dSTORM), can be used to produce a pointillist representation of fluorescently-labeled biological structures at diffraction-unlimited precision. Direct STORM approaches leverage the deactivation of standard fluorescent tags, followed by spontaneous or photoinduced reactivation, allowing resolution of fluorophores at distances below the diffraction limit. This basic principle remains one of the method's primary limitations - standard SMLM fitting routines require tight control of activation and reactivation to maintain sparse emitters, presenting a tradeoff between imaging speed and labeling density. Here, I present two parallel projects, which aim to push the current state of the art in SMLM and apply SMLM to the study of gene regulation. The former represents a novel localization technique for dense SMLM, based on deep probabilistic modeling and photon statistics. In the latter, conventional dSTORM is adapted for live cell imaging of chromatin nanodomains, demonstrating that BRD4 protein concentrates in nucleosome depleted regions.

1. Single molecule localization microscopy

1.1 Introduction

In the quest to understand cellular function, biologists aim to directly observe the processes enabling cells to maintain homeostasis and respond dynamically to internal and environmental cues at the molecular level. Super-resolution (SR) microscopy techniques have emerged as a pathway to this aim, surpassing the classical Abbe diffraction limit of optical resolution: $\lambda/2\text{NA}$ where λ is the emission wavelength and NA is the numerical aperture of an objective lens. Fluorescence microscopy techniques continually push the resolution boundary towards nanometer scales, facilitating imaging of cellular structures with a level of detail previously achievable only with electron microscopy (EM). Concurrently, SR techniques retain optical microscopy advantages in biological experiments, including sample preservation, imaging flexibility, and target specificity. SR enables extraction of quantitative information on spatial distributions and often absolute numbers of proteins, nucleic acids, or other macromolecules within subcellular compartments.

Many SR methods are based on wide-field (WF), total internal reflection fluorescence (TIRF) or confocal microscope setups and fundamentally differ in how fluorescently labeled samples are excited and how the emitted photons are detected. Here, I focus on single-molecule localization microscopy (SMLM) techniques a class of SR diffraction-unlimited SR methods which leverage fluorescence intermittency to resolve fluorophores in the sample whos spatially overlapping point spread functions would otherwise render them unresolvable at the detector. SMLM approaches, such as direct-STORM (dSTORM) have become quite popular because they can be implemented at low cost on conventional, camera-based, wide-field setups, shifting the complexity to biological sample preparation and image post processing. Common strategies for the temporal separation of molecules involve transient intramolecular rearrangements to switch from dark to fluorescent states or the exploitation of non-emitting molecular radicals. For example, in dSTORM, rhodamine derivatives can undergo intersystem crossing to a triplet state, which can be reduced by thiols to form a dark radical species. The dark state can then be quenched by oxidative processes, driving the fluorophore back to its ground state.

In SMLM applications, we seek the position and intensity of isolated fluorophores as well as to estimate the accuracy and precision of these parameters. Accuracy is a measure of the systematic error or bias, and precision is a measure of the statistical error of an estimator. To generate super-resolution images using SMLM, single emitters are located, and using the mosaic of their found positions, we produce a kernel density estimate (KDE). Such KDEs are often Gaussian, and are used to generate the final super-resolution images. The width of one such placed Gaussian function, σ is given by the precision of the fluorophore position localization. Therefore, in SMLM, it is necessary to both find the parameters and estimate their precision. Reported values are in the range of 2070 nanometers. In the following section, we derive a fundamental statistical description of fluorophore detection in SMLM, which is compatible with a coherent state of the quantized electromagnetic field. This description is necessarily simplified - background rates of light detection may vary across the field of view, and the fluorophore emission rate of chemically identical fluorophores can vary owing to effects such as uneven illumination profile, dipole orientation or different optical path lengths.

1.1.1 Elementary mathematical theory of SMLM

$$\omega = \mathrm{i}_0 \int O(u)du \int O(v)dv \tag{1.1}$$

where $\mathrm{i}_0 = \eta N_0 \Delta$. The optical impulse response $O(u, v)$ is often approximated as a 2D isotropic Gaussian with standard deviation σ (Zhang 2007). The parameter η is the photon detection probability of the sensor and Δ is the exposure time. N_0 represents the number of photons emitted.

Using the common definition $\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^t \mathrm{e}^{-t^2} dt$,

$$\int O(u)du = \frac{1}{2}\left(\mathrm{erf}\left(\frac{u_k + \frac{1}{2} - u_0}{\sqrt{2}\sigma}\right) - \mathrm{erf}\left(\frac{u_k - \frac{1}{2} - u_0}{\sqrt{2}\sigma}\right)\right) \tag{1.2}$$

For the sake of generality, the number of photoelectrons at a pixel k, $\mathbf{s}_k$, is multiplied by a gain factor g_k [ADU/e$^-$], which is often unity. The readout noise per pixel ζ_k can be Gaussian with some pixel-specific offset o_k and variance σ_k^2. Ultimately, we have a Poisson

component of the signal, which scales with N_0 and may have Gaussian component, which does not. Therefore, in a single exposure, we measure:

$$\mathbf{x}_t = \mathbf{s}_t + \zeta \tag{1.3}$$

What we are after is the likelihood $p(\mathbf{x}_t|\theta)$ where θ are the molecular coordinates. Fundamental probability theory states that the distribution of $\mathbf{x}_k$ is the convolution of the distributions of $\mathbf{s}_k$ and ζ_k,

$$p(\mathbf{x}_t|\theta) = A\sum_{q=0}^{\infty}\frac{1}{q!}\mathrm{e}^{-\omega_k}\omega_k^q\frac{1}{\sqrt{2\pi}\sigma_k}\mathrm{e}^{-\frac{(\mathbf{x}_k-g_kq-o_k)}{2\sigma_k^2}} \tag{1.4}$$

where $P(\zeta_k) = \mathcal{N}(o_k, \sigma_k^2)$ and $P(S_k) = \text{Poisson}(g_k\omega_k)$, A is some normalization constant. In practice, (4) is difficult to work with, so we look for an approximation. We will use a Poisson-Normal approximation for simplification. Consider,

$$\zeta_k - o_k + \sigma_k^2 \sim \mathcal{N}(\sigma_k^2, \sigma_k^2) \approx \text{Poisson}(\sigma_k^2) \tag{1.5}$$

Since $\mathbf{x}_k = \mathbf{s}_k + \zeta_k$, we transform $\mathbf{x}'_k = \mathbf{x}_k - o_k + \sigma_k^2$, which is distributed according to

$$\mathbf{x}'_k \sim \text{Poisson}(\omega'_k) \tag{1.6}$$

where $\omega'_k = g_k\omega_k + \sigma_k^2$. This result can be seen from the fact the the convolution of two Poisson distributions is also Poisson. The quality of this approximation will degrade with decreasing signal level, since the Poisson distribution does not retain its Gaussian shape at low expected counts. Nevertheless, the quality of the approximation can be predicted by the Komogonov distance between the convolution distribution (4).

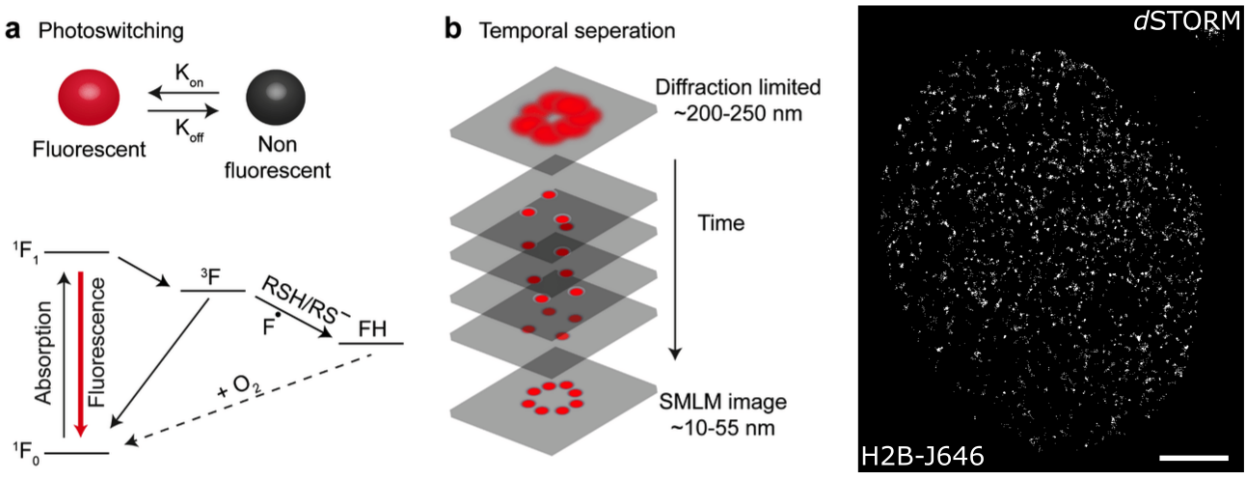

Figure 1.1. Stochastic optical reconstruction microscopy (STORM). (A) Single molecules are resolved by separating their fluorescent emission in time, using fluorophores with multiple photophysical states (B) Example super-resolution image of H2B protein in a living Hela cell nucleus at 37C, 5 percent CO2. Image reconstructed from 10^3 10ms frames. Scalebar 5um.

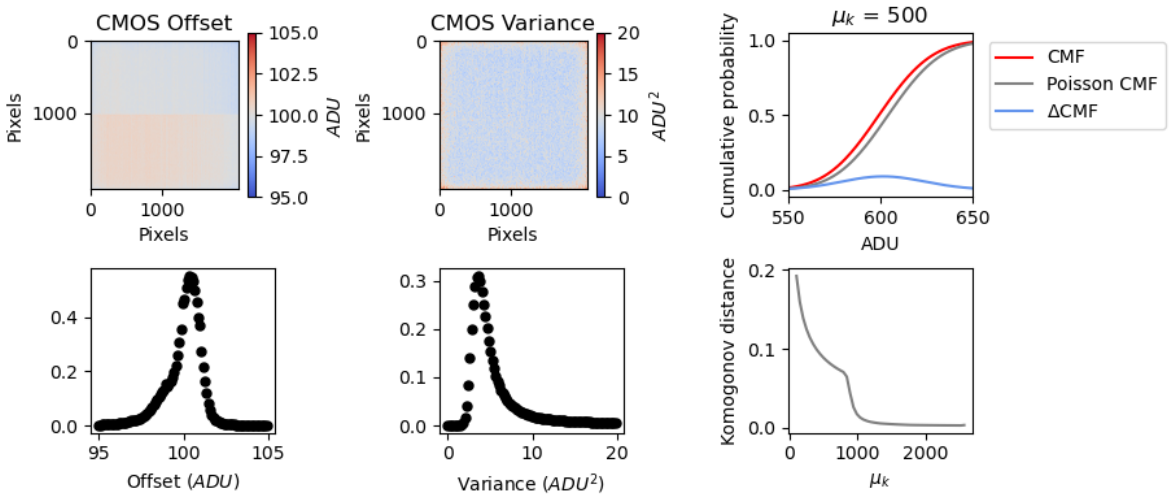

Figure 1.2. Noise model for CMOS cameras used for MLE. (left)) CMOS offset for zero incident photons (middle) CMOS variance for zero incident photons (upper right) Cumulative mass function for the convolution distribution and its Poisson approximation for rate parameter $\mu_k = 500$ counts (lower right) Komogonov distance measured as a function of rate parameter μ_k

1.1.2 The definition of resolution in SMLM

The distribution of a particular biomolecule in the cell can be described as a probability density over a two-dimensional space, casting super-resolution as a density estimation problem. Intuitively, the spatial resolution of SMLM images then increases as we draw more samples from this density - a concept which is made mathematically precise by the so-called Fourier ring correlation or FRC. Using FRC, one can compute image resolution as the spatial frequency at which a correlation function in the frequency domain drops below a threshold, typically taken to be 1/7 (See Supplement). According to this theory, reducing localization uncertainty while increasing the number of samples, results in an increase in image resolution (Nieuwenhuizen 2013). However, there remains a fundamental limit to the the minimal localization uncertainty which can be obtained.

$$\mathrm{FRC}(q) = \frac{\sum_{\vec{q}\in\mathrm{circle}} \tilde{f}_1(\vec{q})\tilde{f}_2(\vec{q})^*}{\sqrt{\sum_{\vec{q}\in\mathrm{circle}}|f_1(\vec{q})|^2}\sqrt{\sum_{\vec{q}\in\mathrm{circle}}|f_2(\vec{q})|^2}}$$

Localization uncertainty, typically the RMSE of a maximum likelihood or similar statistical estimator, is bounded from below by the inverse of the Fisher information matrix, known as the Cramer-Rao lower bound (Chao 2016). Localization uncertainties in sparse conditions are often tens of nanometers, although recent work on integration of Bayesian priors with modulation enhanced SMLM (meSMLM) or structured illumination with MINFLUX, has reduced spatial resolution below to a few nanometers (Kalisvaart 2022, Gwosh 2020). Nevertheless, managing the increase in localization uncertainty at high labeling density remains a major bottleneck to SMLM. Static uncertainty due to molecular crowding can be partially amelioriated by using pairwise or higher-order temporal correlations within a pixel neighborhood, known as stochastic optical fluctuation imaging or SOFI (Dertinger 2009). Other approaches such as stimulated emission and depletion (STED) imaging bring control over the photophysical state of a chosen subset of the sample, yet the need for laser scanning prevents widespread application in live-cell studies. The spatial resolution and relative simplicity of SMLM techniques remains unmatched, inciting an effort to increase the resolution of SMLM techniques and explore avenues towards time resolved SMLM.

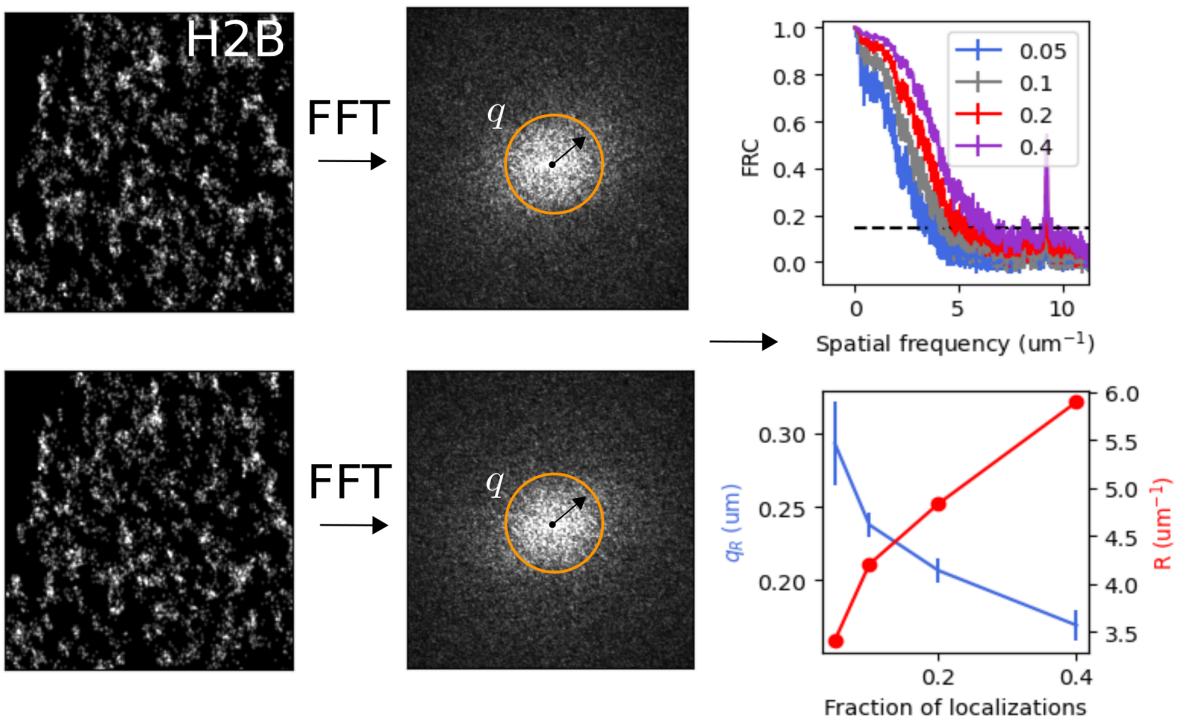

Figure 1.3. Noise model for CMOS cameras used for MLE. (left)) CMOS offset for zero incident photons (middle) CMOS variance for zero incident photons (upper right) Cumulative mass function for the convolution distribution and its Poisson approximation for rate parameter $\mu_k = 500$ counts (lower right) Komogonov distance measured as a function of rate parameter μ_k

1.1.3 The Cramer-Rao lower bound

The Poisson approximation is also convenient for computing the Fisher information matrix for θ_{MLE} and thus the Cramer-Rao lower bound, which bounds the variance of a statistical estimator of θ_{MLE}, from below (Chao 2016). The Fisher information is

$$I_{\text{ij}}(\theta) = \mathop{\mathbb{E}}_{\theta}\left(\frac{\partial \ell}{\partial \theta_{\text{i}}}\frac{\partial \ell}{\partial \theta_{\text{j}}}\right) \tag{1.7}$$

Let $\mu_k' = g_k\mu_k + \sigma_k^2$. For an arbitrary parameter,

$$\begin{aligned}\frac{\partial \ell}{\partial \theta_{\mathrm{i}}} &= \frac{\partial}{\partial \theta_{\mathrm{i}}} \sum_k x_k \log x_k + \mu'_k - x_k \log\left(\mu'_k\right) \\ &= \sum_k \frac{\partial \mu'_k}{\partial \theta_{\mathrm{i}}} \left(\frac{\mu'_k - x_k}{\mu'_k}\right)\end{aligned}$$

$$I_{\mathrm{ij}}(\theta) = \mathop{\mathbb{E}}_{\theta}\left(\sum_k \frac{\partial \mu'_k}{\partial \theta_{\mathrm{i}}} \frac{\partial \mu'_k}{\partial \theta_{\mathrm{j}}} \left(\frac{\mu'_k - x_k}{\mu'_k}\right)^2\right) = \sum_k \frac{1}{\mu'_k} \frac{\partial \mu'_k}{\partial \theta_{\mathrm{i}}} \frac{\partial \mu'_k}{\partial \theta_{\mathrm{j}}}$$

2. Bromodomain protein 4 and chromatin organization

3. Uncertainty-aware localization microscopy by variational diffusion

In single molecule localization, uncertainty quantification is a critical aspect of parameter estimation. Uncertainty can be defined in either the frequentist or the Bayesian statistical paradigm. These primarily differ in their fundamental definition of probability. The frequentist defines probability in terms of the frequency of observed events and examines the empirical distribution of these samples to express uncertainties. The Bayesian would express uncertainty *apriori* based on, say, domain knowledge. Frequentist methods typically estimate uncertainties under a *fixed* model $\mathcal{M}$ using techniques such as a Laplace approximation around the maximum likelihood estimate (MLE). For example, a frequentist might compute

$$\mathcal{L}(D|\mathcal{M}) = \prod_{\mathrm{i}=1}^{N} P(D_{\mathrm{i}}|\mathcal{M})$$

for a dataset of size N, where the model M is held fixed. A frequentist might also make use of the Cramér-Rao lower bound (CRLB), which provides theoretical bounds on the uncertainty achievable by any unbiased estimator.

In contrast, Bayesian approaches to uncertainty quantification in single molecule localization often involve parameter exploration techniques such as Markov Chain Monte Carlo (MCMC). MCMC methods sample from the posterior distribution over the parameters, allowing for comprehensive exploration of parameter space and estimation of uncertainty.

$$P(\mathcal{M}|D) \propto P(D|\mathcal{M})P(\mathcal{M})$$

which can also be optimized by maximum aposteriori estimation (MAP). Additionally, Bayesians may utilize inequalities analagous to the CRLB such as the van Trees inequality, which provides bounds on the expected error in parameter estimation. In the context of single molecule localization, where only a single measurement is often feasible, the treatment of uncertainty demands a Bayesian perspective. We are unable to consider the data that we could have gotten; rather, we must express our uncertainties conditioned on the data that we do have. However, Bayesian approaches are often avoided in this domain, due to limited

computational resources. Diffusion models are a promising direction to speed up and simplify Bayesian inference.

A. Appendix A

A.0.1 Interpolation by second order coherence

Photoswitching fluorescent molecules are described in the density matrix formalism

$$\rho = \sum_k \xi_k \left|\alpha_k\right\rangle \left\langle\alpha_k\right| \quad \sum_k \xi_k = 1$$

where $|\alpha_k\rangle$ is a coherent state with amplitude α_k i.e., $\langle n\rangle = \langle\alpha_k| \, n \, |\alpha_k\rangle = |\alpha_k^2|$. Typically ξ_k and $\langle n_k\rangle$ are heterogeneous. We consider a simplified model consisting of a single mode field

$$E^+(r_\mathrm{i}) = h(r_\mathrm{i} - s_0)\hat{a}_n$$

$$g_{\mathrm{ij}}^{(2)}(0) = \frac{\langle E^-(r_\mathrm{i})E^-(r_\mathrm{j})E^+(r_\mathrm{i})E^+(r_\mathrm{j})\rangle}{\langle E^-(r_\mathrm{i})E^+(r_\mathrm{i})\rangle\langle E^-(r_\mathrm{j})E^+(r_\mathrm{j})\rangle} = \frac{\mathrm{Tr}(E^-(r_\mathrm{i})E^-(r_\mathrm{j})E^+(r_\mathrm{i})E^+(r_\mathrm{j})\rho)}{\mathrm{Tr}(E^-(r_\mathrm{i})E^+(r_\mathrm{i})\rho)\mathrm{Tr}(E^-(r_\mathrm{j})E^+(r_\mathrm{j})\rho)}$$

Terms related to point spread function will cancel. It is instructive to compute

$$\begin{aligned}
\mathrm{Tr}(a^\dagger a^\dagger aa \left(\xi_k \left|\alpha_k\right\rangle \left\langle\alpha_k\right|\right) &= \mathrm{Tr}\left(\xi_k \mathrm{e}^{-|\alpha|^2} \sum_{n,m}^{\infty} \frac{\alpha^n}{n!} \left|n\right\rangle \left\langle m\right|\right) \\
&= \mathrm{Tr}\left(\xi_k \mathrm{e}^{-|\alpha|^2} \sum_{n}^{\infty} \frac{|\alpha|^{2n}}{n!} n(n-1)\right) \\
&= \mathrm{Tr}\left(\xi_k \mathrm{e}^{-|\alpha|^2} \sum_{n=2}^{\infty} \frac{|\alpha|^{2n}}{(n-2)!}\right) \\
&= \xi_k |\alpha_k|^4
\end{aligned}$$

Similarly,

$$
\begin{aligned}
\mathrm{Tr}(a^\dagger a\left(\xi\left|\alpha\right\rangle\left\langle\alpha\right|\right)) &= \mathrm{Tr}\left(\xi \mathrm{e}^{-|\alpha|^2}\sum_{n,m}^{\infty}\frac{\alpha^n(\alpha^m)^*}{\sqrt{n!}\sqrt{m!}}a^\dagger a\left|n\right\rangle\left\langle m\right|\right)\\
&= \xi \mathrm{e}^{-|\alpha|^2}\sum_{n=0}^{\infty}\frac{(|\alpha|^2)^n}{n!}n\\
&= \xi \mathrm{e}^{-|\alpha|^2}\sum_{n=1}^{\infty}\frac{(|\alpha|^2)^n}{(n-1)!}\\
&= \xi \mathrm{e}^{-|\alpha|^2}\left(|\alpha|^2+\frac{|\alpha|^4}{1!}+\frac{|\alpha|^6}{2!}+...\right)\\
&= \xi \mathrm{e}^{-|\alpha|^2}|\alpha|^2\left(1+\frac{|\alpha|^2}{1!}+\frac{|\alpha|^3}{2!}+...\right)\\
&= \xi \mathrm{e}^{-|\alpha|^2}\mathrm{e}^{|\alpha|^2}|\alpha|^2 = \xi|\alpha|^2
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Tr}(aa^\dagger\left(\xi\left|\alpha\right\rangle\left\langle\alpha\right|\right)) &= \mathrm{Tr}\left(\xi \mathrm{e}^{-|\alpha|^2}\sum_{n,m}^{\infty}\frac{\alpha^n(\alpha^m)^*}{\sqrt{n!}\sqrt{m!}}aa^\dagger\left|n\right\rangle\left\langle m\right|\right)\\
&= \xi \mathrm{e}^{-|\alpha|^2}\sum_{n=0}^{\infty}\frac{(|\alpha|^2)^n}{n!}(n+1)\\
&= \xi \mathrm{e}^{-|\alpha|^2}\left(\sum_{n=1}^{\infty}\frac{(|\alpha|^2)^n}{(n-1)!}+\mathrm{e}^{|\alpha|^2}\right)\\
&= \xi \mathrm{e}^{-|\alpha|^2}\left(|\alpha|^2\mathrm{e}^{|\alpha|^2}+\mathrm{e}^{|\alpha|^2}\right) = \xi(|\alpha|^2+1)
\end{aligned}
$$

Putting it all together yields a simple expression for the two-point coherence function

$$
g^{(2)}_{\mathrm{ij}}(0) = \frac{\sum_k \xi_k|\alpha_k|^4}{\left(\sum_k \xi_k|\alpha_k|^2\right)\left(\sum_k \xi_k|\alpha_k|^2\right)}
$$

For example, if we have a two-level system consisting of a fluorescent state with amplitude α and the vacuum state, this becomes

$$
g^{(2)}_{\mathrm{ij}}(0) = \frac{\xi|\alpha|^4}{\xi^2|\alpha|^4} = \frac{1}{\xi}
$$

As $\xi \to 1$ (always on) we recover a coherent state. As $\xi \to 0$ we observe $g^{(2)}_{\mathrm{ij}}(0) > 1$ i.e., bunching.

A.0.2 Generalization to nonzero background

$$E_0^+ \sim \sum_{\mathrm{j}=1}^{M} \delta(s - s_\mathrm{j}) a_\mathrm{j} \quad E^+(r_\mathrm{i}) = \int d^2 s E_0^+ = \sum_n h(r_\mathrm{i} - s_n) a_n$$

$$\rho_S = \xi \left|\alpha\right\rangle \left\langle\alpha\right| + (1-\xi) \left|0\right\rangle \left\langle 0\right| \quad \rho_B = \left|\beta\right\rangle \left\langle\beta\right| \quad \rho = \rho_S \otimes \rho_B$$

$$E(r_\mathrm{i})^+ = E_S(r_\mathrm{i})^+ + E_B(r_\mathrm{i})^+ = h(r_\mathrm{i} - s_n) a_S + a_B$$

$$\begin{aligned}
G^2_{\mathrm{ij}}(0) &= \langle (E_S^\dagger + E_B^\dagger)(E_S^\dagger + E_B^\dagger)(E_S + E_B)(E_S + E_B) \rangle \\
&= h_\mathrm{i}^2 h_\mathrm{j}^2 \langle a_S^\dagger a_S^\dagger a_S a_S \rangle + h_\mathrm{i}^2 \langle a_S^\dagger a_B^\dagger a_S a_B \rangle + h_\mathrm{j}^2 \langle a_B^\dagger a_S^\dagger a_B a_S \rangle + \langle a_B^\dagger a_B^\dagger a_B a_B \rangle \\
&= \xi (h_\mathrm{i}^2 h_\mathrm{j}^2 |\alpha|^4 + h_\mathrm{i}^2 |\alpha|^2 |\beta|^2 + h_\mathrm{j}^2 |\alpha|^2 |\beta|^2 \rangle + |\beta|^4) \\
&= \xi (h_\mathrm{i}^2 h_\mathrm{j}^2 |\alpha|^4 + |\alpha|^2 |\beta|^2 (h_\mathrm{i}^2 + h_\mathrm{j}^2) + |\beta|^4)
\end{aligned}$$

The normalized second order coherence function then reads

$$g^2_{\mathrm{ij}}(0) = \frac{\xi h_\mathrm{i}^2 h_\mathrm{j}^2 N_0^2 + \xi N_0 B_0 (h_\mathrm{i}^2 + h_\mathrm{j}^2) + B_0^2}{\xi^2 h_\mathrm{i}^2 h_\mathrm{j}^2 N_0^2 + \xi N_0 B_0 (h_\mathrm{i}^2 + h_\mathrm{j}^2) + B_0^2}$$

Notice the PSF factor h_i appears squared. This squared value can be seen as the probability of photon detection at a point s_i, while h_i is the amplitude of the electric field.

A.0.3 Generalization to multi-level systems

$$G^2_{\mathrm{ij}}(0) = \frac{h_\mathrm{i}^2 h_\mathrm{j}^2 \sum_k \xi_k |\alpha_k|^4 + |\beta|^2 (h_\mathrm{i}^2 + h_\mathrm{j}^2) \sum_k \xi_k |\alpha_k|^2 + |\beta|^4}{(h_\mathrm{i}^2 \sum_k \xi_k |\alpha_k|^2 + |\beta|^2) \left(h_\mathrm{j}^2 \sum_k \xi_k |\alpha_k|^2 + |\beta|^2 \right)}$$

A.0.4 Details of the Gaussian PSF

We will derive the gradients for the integrated astigmatic Gaussian, since it is the more general case. As before, define $\mathrm{i}_0 = g_k \gamma \Delta t N_0$ such that $\mu'_k = \mathrm{i}_0 \lambda_k$

$$J_{x_0} = \beta_k \lambda_y \frac{\partial \lambda_x}{\partial x_0} \quad J_{y_0} = \beta_k \lambda_x \frac{\partial \lambda_y}{\partial y_0} \quad J_{z_0} = \frac{\partial \mu'_k}{\partial \sigma_x}\frac{\partial \sigma_x}{\partial z_0} + \frac{\partial \mu'_k}{\partial \sigma_y}\frac{\partial \sigma_y}{\partial z_0}$$

$$\begin{aligned} J_{x_0} &= \beta_k \lambda_y \frac{\partial \lambda_x}{\partial x_0} \\ &= \frac{\beta_k \lambda_y}{2} \frac{\partial}{\partial x_0} \left(\operatorname{erf}\left(\frac{x_k + \frac{1}{2} - x_0}{\sqrt{2}\sigma_x} \right) - \operatorname{erf}\left(\frac{x_k - \frac{1}{2} - x_0}{\sqrt{2}\sigma_x} \right) \right) \\ &= \frac{\beta_k \lambda_y}{\sqrt{2\pi}\sigma_x} \left(\exp\left(\frac{(x_k - \frac{1}{2} - x_0)^2}{2\sigma_x^2} \right) - \exp\left(\frac{(x_k + \frac{1}{2} - x_0)^2}{2\sigma_x^2} \right) \right) \end{aligned}$$

$$\begin{aligned} J_{y_0} &= \beta_k \lambda_x \frac{\partial \lambda_y}{\partial y_0} \\ &= \frac{\beta_k \lambda_x}{2} \frac{\partial}{\partial y_0} \left(\operatorname{erf}\left(\frac{y_k + \frac{1}{2} - y_0}{\sqrt{2}\sigma_y} \right) - \operatorname{erf}\left(\frac{y_k - \frac{1}{2} - y_0}{\sqrt{2}\sigma_y} \right) \right) \\ &= \frac{\beta_k \lambda_x}{\sqrt{2\pi}\sigma_y} \left(\exp\left(\frac{(y_k - \frac{1}{2} - y_0)^2}{2\sigma_y^2} \right) - \exp\left(\frac{(y_k + \frac{1}{2} - y_0)^2}{2\sigma_y^2} \right) \right) \end{aligned}$$

$$\begin{aligned} J_{\sigma_x} &= \beta_k \lambda_y \frac{\partial \lambda_x}{\partial \sigma_x} \\ &= \frac{\beta_k \lambda_y}{2} \frac{\partial}{\partial \sigma_x} \left(\operatorname{erf}\left(\frac{x_k + \frac{1}{2} - x_0}{\sqrt{2}\sigma_x} \right) - \operatorname{erf}\left(\frac{x_k - \frac{1}{2} - x_0}{\sqrt{2}\sigma_x} \right) \right) \\ &= \frac{\beta_k \lambda_y}{\sqrt{2\pi}} \left(\frac{\left(x - x_0 - \frac{1}{2}\right) \mathrm{e}^{-\frac{\left(x - x_0 - \frac{1}{2}\right)^2}{2\sigma_x^2}}}{\sigma_x^2} - \frac{\left(x - x_0 + \frac{1}{2}\right) \mathrm{e}^{-\frac{\left(x - x_0 + \frac{1}{2}\right)^2}{2\sigma_x^2}}}{\sigma_x^2} \right) \end{aligned}$$

$$
\begin{aligned}
J_{\sigma_y} &= \beta_k \lambda_x \frac{\partial \lambda_y}{\partial \sigma_y} \\
&= \frac{\beta_k \lambda_x}{2} \frac{\partial}{\partial \sigma_y} \left(\operatorname{erf}\left(\frac{y_k + \frac{1}{2} - y_0}{\sqrt{2}\sigma_y} \right) - \operatorname{erf}\left(\frac{y_k - \frac{1}{2} - y_0}{\sqrt{2}\sigma_y} \right) \right) \\
&= \frac{\beta_k \lambda_x}{\sqrt{2\pi}} \left(\frac{\left(y - y_0 - \frac{1}{2}\right) \mathrm{e}^{-\frac{\left(y - y_0 - \frac{1}{2}\right)^2}{2\sigma_y^2}}}{\sigma_y^2} - \frac{\left(y - y_0 + \frac{1}{2}\right) \mathrm{e}^{-\frac{\left(y - y_0 + \frac{1}{2}\right)^2}{2\sigma_y^2}}}{\sigma_y^2} \right)
\end{aligned}
$$

Luckily, computing the Hessian matrix for (2.9) is tractable, and is actually quite simple when one takes advantage of the chain rule for Hessian matrices. Looking at (2.9), the likelihood is a hierarchical function that maps a vector space Θ to a vector space Λ to a scalar value. Formally, we define $T : \Theta \to \Lambda$ and $W : \Lambda \to \mathbb{R}$. The parameter vector $(x_0, y_0, z_0, \sigma_0, N_0) \in \Theta$, the Poisson rate vector $\vec{\lambda} \in \Lambda$ and $\ell \in \mathbb{R}$. Note that we choose to optimize σ_x and σ_y directly and compute z_0 to simplify the computation of the Hessian. To get the Hessian, we need the chain-rule for Hessian matrices, which can be quickly computed in terms of the jacobian and hessian of T and W.

$$
H_\ell = J_\mu^T H_\ell J_\mu + (J_\ell \otimes I_n) H_\mu
$$

where we have used J_μ to represent the jacobian of T and J_ℓ for the jacobian of W. Similar notation is used for the corresponding Hessian matrices. In the 3D case, the Hessian matrix is not directly separable since $\mu \propto \lambda_x(x_0, \sigma_0, \sigma_x)\lambda_y(y_0, \sigma_0, \sigma_y)$. To see this, an abstract representation of the Hessian reads

A.0.5 Fisher information for 2D integrated gaussian

For the 2D integrated gaussian point spread function, the Hessian only contains separable second order derivatives, so the Fisher information matrix takes on a convenient form

$$
I_{\mathrm{ij}}(\theta) = \mathop{\mathbb{E}}_{\theta} \left(\frac{\partial \ell}{\partial \theta_{\mathrm{i}}} \frac{\partial \ell}{\partial \theta_{\mathrm{j}}} \right) \tag{A.1}
$$

For an arbitrary parameter then we have

$$\begin{aligned}\frac{\partial \ell}{\partial \theta_{\mathrm{i}}} &= \frac{\partial}{\partial \theta_{\mathrm{i}}} \sum_k x_k \log x_k + \mu'_k - x_k \log\left(\mu'_k\right) \\ &= \sum_k \frac{\partial \mu'_k}{\partial \theta_{\mathrm{i}}} \left(\frac{\mu'_k - x_k}{\mu'_k}\right)\end{aligned}$$

$$I_{\mathrm{ij}}(\theta) = \mathop{\mathbb{E}}_{\theta}\left(\sum_k \frac{\partial \mu'_k}{\partial \theta_{\mathrm{i}}} \frac{\partial \mu'_k}{\partial \theta_{\mathrm{j}}} \left(\frac{\mu'_k - x_k}{\mu'_k}\right)^2\right) = \sum_k \frac{1}{\mu'_k} \frac{\partial \mu'_k}{\partial \theta_{\mathrm{i}}} \frac{\partial \mu'_k}{\partial \theta_{\mathrm{j}}}$$

To compute the bound, it turns out all we need is the jacobian $\frac{\partial \mu'_k}{\partial \theta_{\mathrm{j}}}$.

B. CITATIONS AND REFERENCES

```
\chapter{CITATIONS AND REFERENCES}
```

This chapter contains information about citations and references—how to cite a reference in the text and the fine points of defining a bibliography (also called "References") entry.

```


This chapter contains information about citations
and references---how to cite a reference in the text
and the fine points of defining a bibliography
(also called ``References'')
entry.
```

B.1 Citations

```


\section{Citations}
```

For LaTeX answers I refer to **lamport1994** and then to **goossens1994** or **kopka1999**. **kopka1999** is an update to **kopka1995**.

```
For \LaTeX\ answers I refer to
\cite{lamport1994}
and then to
\cite{goossens1994}
or
\cite{kopka1999}.
\cite{kopka1999}
is an update to
\cite{kopka1995}.
```

Here is an example .bib file entry:

```
@misc{example2020,
  address   = {Imaginaryville, Indiana},
  author    = {Andrew Anteater and Bertha Bear and Charles Cheetah and Davida Deer
               and Ethan Eagle},
  date      = {2020-10-27},
  doi       = {00.0000/000-0-000-00000-0},
  editor    = {Mark Senn},
  edition   = {2},
  isbn      = {{000\FigureDash 0\FigureDash 000\FigureDash 00000\FigureDash 0}},
  publisher = {Bogus International Publishing Company},
  title     = {An Imaginary Document Not About {Mark Senn} or {NASA}},
  url       = {https://bogus.com/bogus.html},
  urldate   = {2020-10-27},
  version   = {1.0},
}
```

```

Here is an example .bib file entry:

{\footnotesize
\begin{verbatim}
@misc{example2020,
  address   = {Imaginaryville, Indiana},
  author    = {Andrew Anteater and Bertha Bear and Charles Cheetah and Davida Deer
               and Ethan Eagle},
  date      = {2020-10-27},
  doi       = {00.0000/000-0-000-00000-0},
  editor    = {Mark Senn},
  edition   = {2},
  isbn      = {{000\FigureDash 0\FigureDash 000\FigureDash 00000\FigureDash 0}},
  publisher = {Bogus International Publishing Company},
  title     = {An Imaginary Document Not About {Mark Senn} or {NASA}},
  url       = {https://bogus.com/bogus.html},
  urldate   = {2020-10-27},
  version   = {1.0},
}
\end{verbatim}
}
```

PurdueThesis only uses BibLaTeX. Here are some example BibLaTeX citations for your document.

Input	**Output**
\cite{example2020}	**example2020**
\cite*{example2020}	**example2020**
\citeauthor{example2020}	**example2020**
\citeauthor*{example2020}	**example2020**
\citedate{example2020}	**example2020**
\citetitle{example2020}	**example2020**
\citetitle*{example2020}	**example2020**
\citeurl{example2020}	**example2020**
\citeyear{example2020}	**example2020**
\parencite{example2020}	[**example2020**]
\textcite{example2020}	**example2020**

```
\PurdueThesisLogo\ only uses \BibLaTeXLogo.
Here are some example \BibLaTeXLogo\ citations for your document.

\begin{tabular}{@{}ll@{}}
  \bf Input&                          \bf Output\\
  \verb+\cite{example2020}+&          \cite{example2020}\\
  \verb+\cite*{example2020}+&         \cite*{example2020}\\
  \verb+\citeauthor{example2020}+&    \citeauthor{example2020}\\
  \verb+\citeauthor*{example2020}+&   \citeauthor*{example2020}\\
```

```
    \verb+\citedate{example2020}+&    \citedate{example2020}\\
    \verb+\citetitle{example2020}+&   \citetitle{example2020}\\
    \verb+\citetitle*{example2020}+&  \citetitle*{example2020}\\
    \verb+\citeurl{example2020}+&     \citeurl{example2020}\\
    \verb+\citeyear{example2020}+&    \citeyear{example2020}\\
    \verb+\parencite{example2020}+&   \parencite{example2020}\\
    \verb+\textcite{example2020}+&    \textcite{example2020}\\
  \end{tabular}
```

B.2 References

```

\section{References}
```

Emily Spreen wrote that the following URLs are invisible in the PDF file. They worked fine for me on 2021-04-08. See **hambleton**, **gerstenmaier**, and **gerstenmaier2** in the REFERENCES.

```
@misc{hambleton,
  key = {Deep Space Gateway},
  title = {{Deep Space Gateway to Open Opportunities for Distant Destinations}},
  note = {Editor: Kathryn Hambleton},
  year = {2018},
  month = {August 24,},
  howpublished = {\url{https://www.nasa.gov/feature/deep-space-gateway-to-open-...}},
  organization = {NASA},
}

@misc{gerstenmaier,
  author = {William H. Gerstenmaier},
  title = {{Progress in Defining the Deep Space Gateway and Transport Plan}},
  month = {March},
  year = {2017},
  howpublished = {\url{https://www.nasa.gov/sites/default/files/atoms/files/...}},
  organization = {NASA},
}
```

I suggest using the following (added a '2' to the key so they'd have separate entries in the references.).

```
@misc{gerstenmaier2,
  author = {William H. Gerstenmaier},
  date = {2017-03},
  title = {{Progress in Defining the Deep Space Gateway and Transport Plan}},
  url = {https://www.nasa.gov/sites/default/files/atoms/files/nss_chart_v23.pdf},
  organization = {NASA},
}
```

```

Emily Spreen wrote that the following URLs are invisible in the PDF file.
They worked fine for me on 2021-04-08.
```

```
See \cite{hambleton}, \cite{gerstenmaier}, and \cite{gerstenmaier2} in the REFERENCES.

{\footnotesize
\begin{verbatim}
@misc{hambleton,
  key = {Deep Space Gateway},
  title = {{Deep Space Gateway to Open Opportunities for Distant Destinations}},
  note = {Editor: Kathryn Hambleton},
  year = {2018},
  month = {August 24,},
  howpublished = {\url{https://www.nasa.gov/feature/deep-space-gateway-to-open-...}},
  organization = {NASA},
}
\end{verbatim}
}

{\footnotesize
\begin{verbatim}
@misc{gerstenmaier,
  author = {William H. Gerstenmaier},
  title = {{Progress in Defining the Deep Space Gateway and Transport Plan}},
  month = {March},
  year = {2017},
  howpublished = {\url{https://www.nasa.gov/sites/default/files/atoms/files/...}},
  organization = {NASA},
}
\end{verbatim}
}

I suggest using the following
(added a '2' to the key so they'd have separate entries in the references.).
{\footnotesize
\begin{verbatim}
@misc{gerstenmaier2,
  author = {William H. Gerstenmaier},
  date = {2017-03},
  title = {{Progress in Defining the Deep Space Gateway and Transport Plan}},
  url = {https://www.nasa.gov/sites/default/files/atoms/files/nss_chart_v23.pdf},
  organization = {NASA},
}
\end{verbatim}
}
```

VITA

[Put a brief autobiographical sketch here.]

Temporary page!

LaTeX was unable to guess the total number of pages correctly. As there was some unprocessed data that should have been added to the final page this extra page has been added to receive it.

If you rerun the document (without altering it) this surplus page will go away, because LaTeX now knows how many pages to expect for this document.