

Visualizing nucleosome cluster dynamics with dense single molecule localization microscopy

Clayton W. Seitz

July 20, 2023

Introduction

TO DO

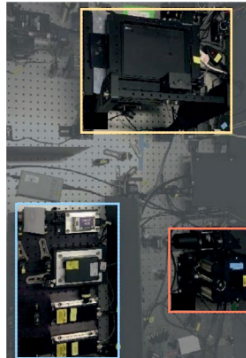
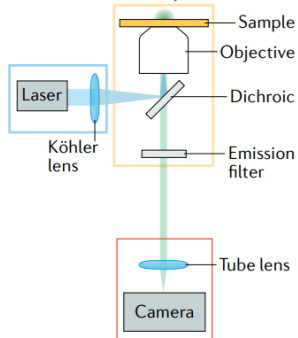
- ▶ Figure 1b: k_1, k_2 inference (Monday)
- ▶ Figure 1c,d: Performance comparison of CNN with LoG algorithm (Tuesday-Thursday)
- ▶ Implement the Dataset object - a universal data object for simulated and experimental data. It can have an arbitrary shape, with axes for different conditions, for example. Ground truth information is attached to the object
- ▶ Hardware autofocus (focus drifts after 5-10 seconds)
- ▶ Precise exposure control (using an Arduino)
- ▶ Optimizing two-color labeling
- ▶ Setting up CO₂
- ▶ Later: Caged diffusion dynamics (Singh 2018, Ashwin 2019)

Methods

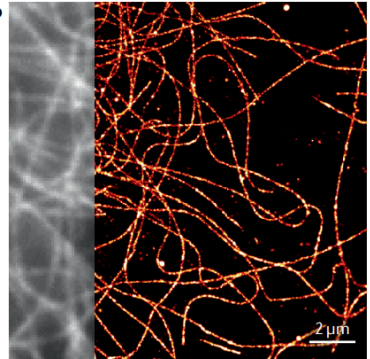
Direct stochastic optical reconstruction microscopy

Direct stochastic optical reconstruction microscopy

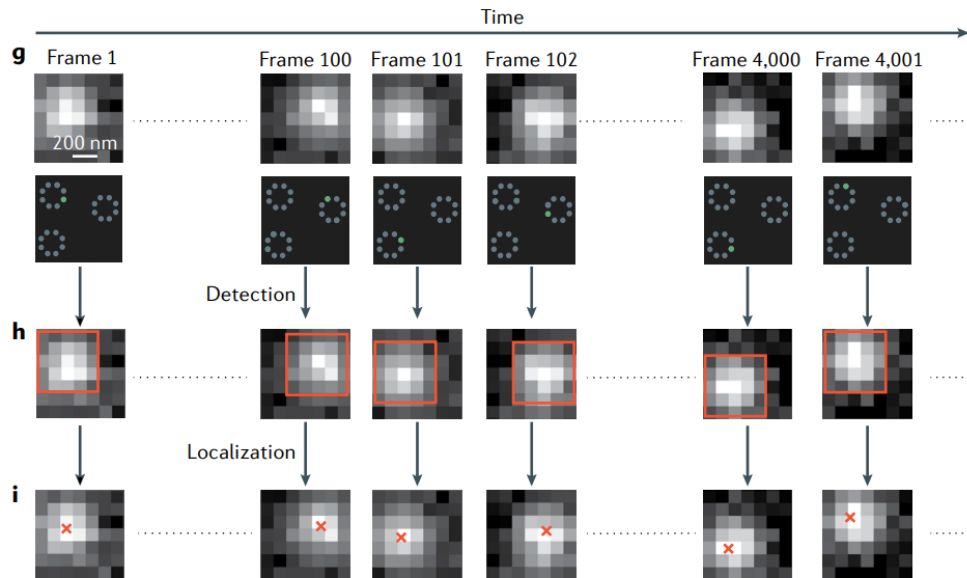
a Basic SMLM set-up



b

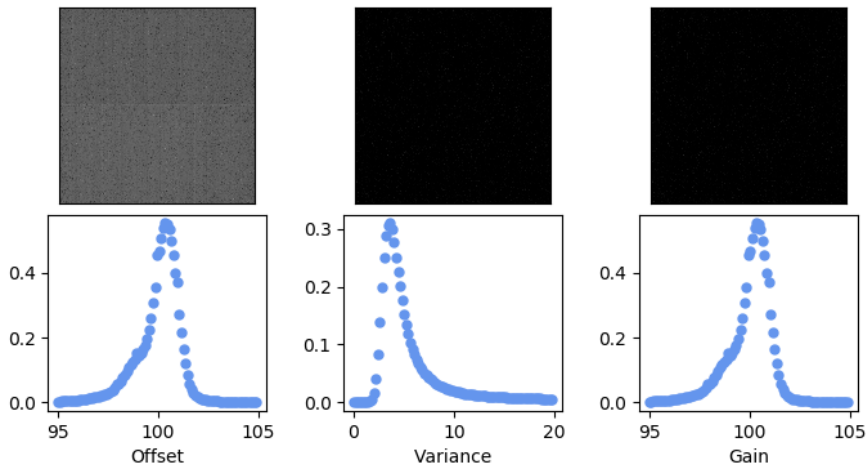


Direct stochastic optical reconstruction microscopy



Readout noise of sCMOS cameras

Hamamatsu ORCA v3 CMOS, air cooled to -10C



Measured signal: $H_k = S_k + \xi_k$, $S_k \sim \text{Poisson}(\mu_k)$, $\xi_k \sim \mathcal{N}(o_k, \sigma_k^2)$

Maximum likelihood localization of an isolated fluorescent emitter

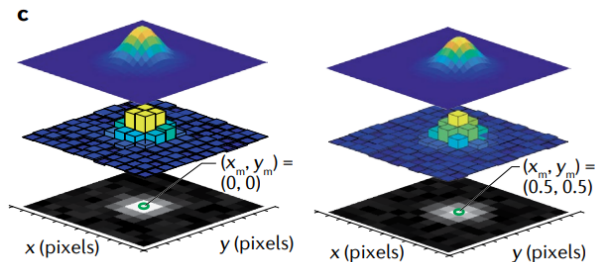
Localization: $\theta^* = \underset{\theta}{\operatorname{argmax}} \prod_k P(H_k|\theta) = \underset{\theta}{\operatorname{argmin}} - \sum_k \log P(H_k|\theta)$

$$\mu_k = g_k \eta N_0 \Delta \int_{\text{pixel}} G(x, y) dA$$

η – quantum efficiency

N_0 – emission rate

Δ – exposure time

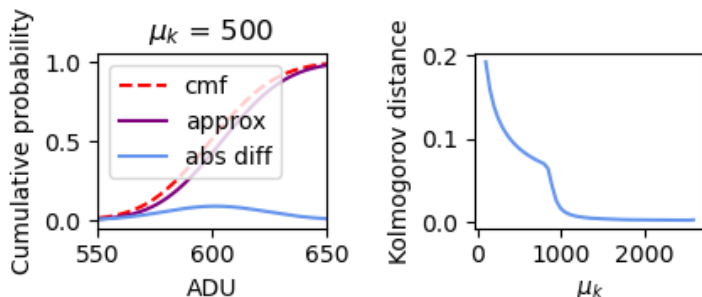


$$P(H_k|\theta) = A \sum_{q=0}^{\infty} \frac{1}{q!} e^{-\mu_k} \mu_k^q \frac{1}{\sqrt{2\pi\sigma_k}} e^{-\frac{(H_k - g_k q - o_k)^2}{2\sigma_k^2}}$$

$P(H_k|\theta)$ can be approximated as Poisson at high signal-to-noise (SNR)

Quality of the Poisson approximation depends on SNR

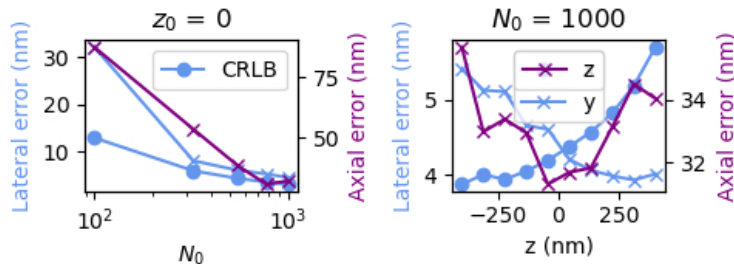
$P(H_k|\theta) \approx \text{Poisson}(\mu_k + \sigma_k^2)$ for $N_0 > 500$ assuming $\Delta = 100\text{ms}$



Using the approximation we can write

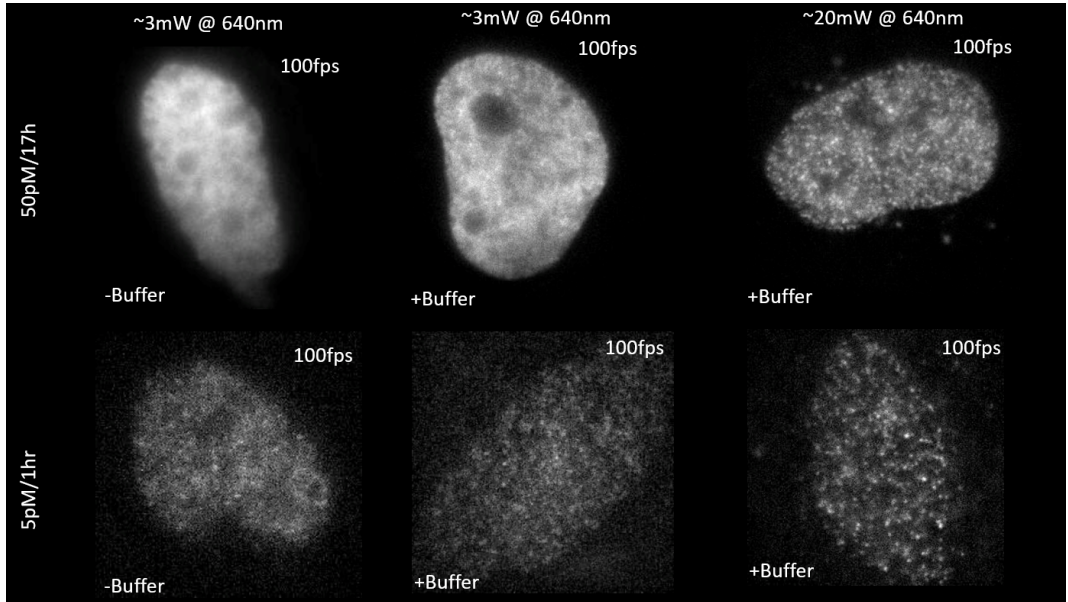
$$\ell(\vec{H}|\theta) = -\log \prod_k \frac{e^{-(\mu'_k)} (\mu'_k)^{n_k}}{n_k!} = \sum_k \log n_k! + \mu'_k - n_k \log (\mu'_k)$$

Estimator precision sets the resolution limit in localization microscopy

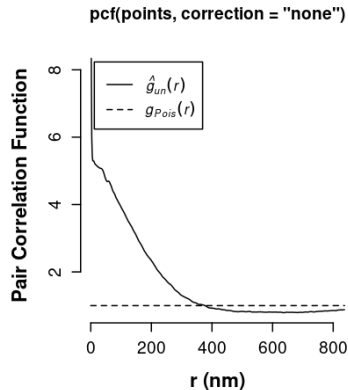
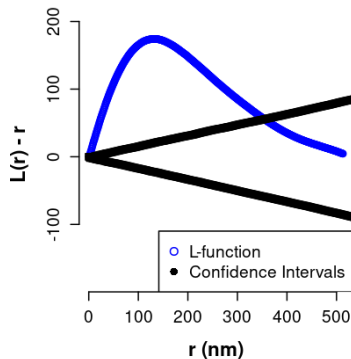
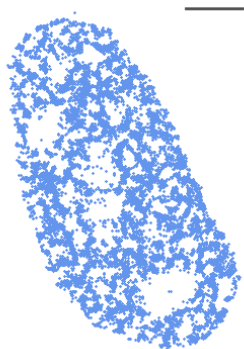


- MLE can approach the CRLB on simulated isolated emitter data

The metastable OFF state can be maintained with high laser power



Resolution is dependent on photoswitching kinetics



- MLE can approach the CRLB on simulated isolated emitter data

Resolution is dependent on photoswitching kinetics

A molecule is considered "detected" in principle if the measured ADU signal satisfies $\tilde{s} = \mu\tau \geq \delta$ where δ is a number of photons which satisfy a criterion on localization accuracy.

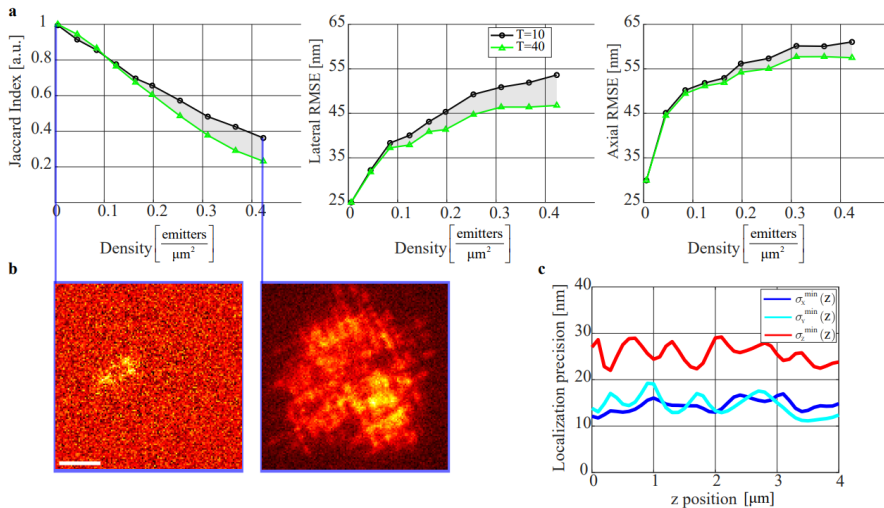
$$\alpha = \int_{\delta}^{\Delta} \left(\sum_{n=0}^{\infty} Q(N=n) \psi(\tau|n; \vec{k}) \right) d\tau \approx \mathbb{E}_{\tau \sim P(\tau)} (\mathbb{I}[\tau > \delta])$$

$P(\tau)$ is usually obtained by Monte Carlo simulation. This is useful for computing density measures and the total acquisition time:

$$D = \alpha K \left(\frac{\lambda}{2NA} \right) \quad T = \left(\Delta_{SR} + \frac{2N}{\log(1 - \alpha)} \right)^2$$

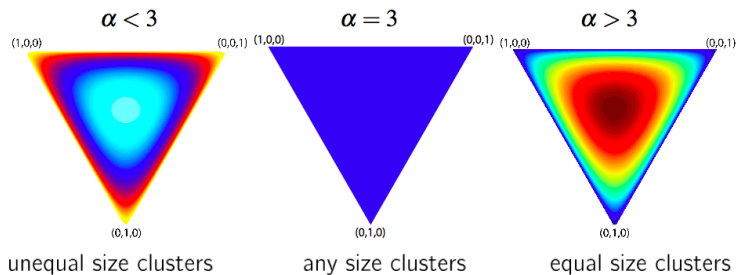
For actually inferring k_1, k_2 , we need a measure of distance between $P(\tilde{s})$ and $P(s|k_1, k_2)$ for many k_1, k_2 pairs. Luckily we only need to compute $P(s|k_1, k_2)$ once, and we can then perform a grid search

Resolution is dependent on photoswitching kinetics

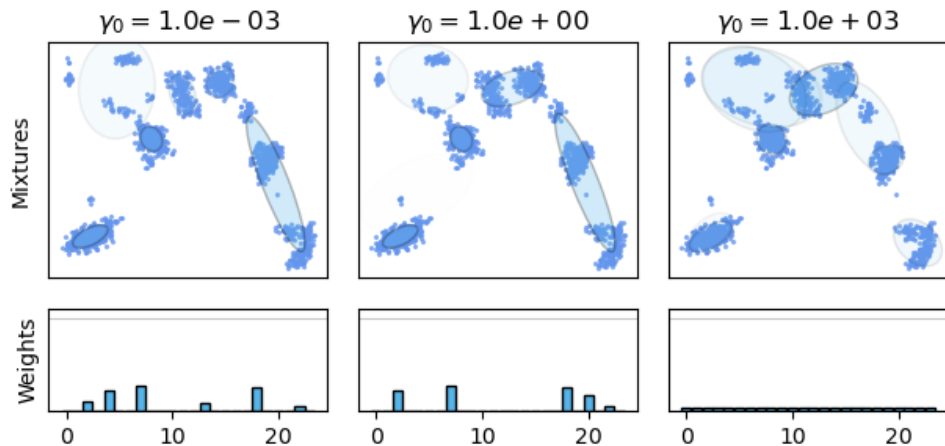


Deep learning enables accurate 3D localization and single molecule tracking

Dirichlet process Gaussian mixture model (DPGMM)



Dirichlet process Gaussian mixture model (DPGMM)



Results

GMM cluster analysis of H2B

Number of clusters is unknown apriori - Bayesian nonparametrics

Mesoscale nucleosome organization and dynamics

Mesoscale nucleosome organization and dynamics

BRD4 associates with the small clusters

Besag's L-Function

Besag's L-Function and 3D diffusion