

NORTHWESTERN UNIVERSITY

Quantitative Analytics for Spectroscopic Single-Molecule Localization Microscopy

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Biomedical Engineering

By

Janel Lesedi Davis

EVANSTON, ILLINOIS

June 2020

Copyright by Janel Lesedi Davis 2020
All Rights Reserved

ABSTRACT

Quantitative Analytics for Spectroscopic Single-Molecule Localization Microscopy

Janel L. Davis

Single-molecule localization microscopy (SMLM) has significantly stimulated the development of methods to quantitatively visualize and characterize biomolecules *in vitro* and *in situ*. SMLM is a class of super-resolution microscopy (SRM) techniques, which exploits the “on-off” switching of individual fluorescent molecules to estimate their location with nanometer precision and offers spatial resolution up to 10 nm. However, the invaluable spectroscopic information of fluorescent labels has previously been overlooked in conventional SMLM. Recently developed spectroscopic SMLM (sSMLM), integrates a dispersive component into the traditional SMLM system to concurrently capture the spatial and spectral information of each single-molecule emission event. Thus far, the emission spectra have been used to expand the multiplexing capabilities of SMLM and develop functional SRM. While traditional SMLM has widely been adopted by the scientific community, the limited availability of software tools to process sSMLM data has hindered the routine use of sSMLM for SRM studies. Additionally, the benefits of using spectroscopic information to improve quantitative SRM have remained unexplored.

This dissertation aims to build an image processing platform for sSMLM and develop three quantitative spectroscopic analysis methods for the characterization of nanostructures. First, we present RainbowSTORM, a freely available ImageJ plug-in, which includes functions to calibrate sSMLM systems, process two-dimensional and three-dimensional sSMLM data, and generate

pseudo-colored sSMLM image reconstructions. Second, we develop a regression method to evaluate spectroscopic information and reject signals from fluorescent impurities which can lead to molecular misidentification and degraded spatial resolution in SMLM. Using this method, we quantify immobilized nanorulers and improve sample identification of surface-combed DNA fibers. Third, we develop quantitative spectroscopic analysis for cluster extraction (qSPACE), a post-processing method for the sSMLM variant referred to as spectroscopic point accumulation for imaging in nanoscale topography (sPAINT), which captures the fluorescence induced by transient interactions between Nile Red (NR) dye and polymersomes. We use qSPACE to accurately size and count the polymersomes, while rejecting non-specific interactions between NR and the poly-L-lysine surface. Finally, we develop photon-accumulation enhanced reconstruction (PACER) which numerically increases the photon budget of fluorescent labels and estimates the emitter's spatial location with improved localization precision. Using PACER, we image quantum dots and Alexa Fluor 647 labeled DNA nanostructures with a minimum spacing of 6 nm to uncover individual molecules that would be unresolvable using conventional SMLM.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my Ph.D. advisor Prof. Hao F. Zhang, who has played an instrumental role in my development as a scientist. Throughout my experience Prof. Zhang, has fostered my growth by encouraging me to apply to different fellowship programs, attend conferences, and try new applications of our technology. Prof. Zhang has been a constant source of advice both for advancing my research and my future career. Overall, he has done his best to help me to develop my technical writing skills and shown me what it takes to manage an academic lab. I will always be grateful for the time he has invested in guiding me through my Ph.D.

I would also like to thank my co-advisor Prof. Matthew Glucksberg. Prof. Glucksberg played an important role in helping me to lay a strong foundation for my Ph.D. Through my involvement with The Center for Innovation in Global Health Technologies (CIGHT) directed by Prof. Glucksberg, I was able learn more about Global Health research which was the reason I chose to pursue my Ph.D. at Northwestern. Although, my goals ultimately changed my experience in CIGHT was a significant learning experience for me.

Thank you to my committee members Prof. Sun, Dr. McFall, and Prof. Backman. With his deep technical understanding of optics, physics, materials science, and research, Prof. Sun has been an exceptional resource throughout my Ph.D. Prof. Sun has always provided helpful feedback and our many meetings to discuss strategies for presenting my research findings have helped me make progress in my projects and grow as a researcher. Dr. McFall been an essential resource for planning experimental studies and preparing DNA and protein samples. I really have appreciated our many chats which ranged from current events to the technical details of

designing DNA probes. Prof. Backman provided very helpful feedback during my qualifying exam and committee meetings. Additionally, as a student in Prof. Backman's Advanced Optics course during my second year, I learnt the fundamental theories which govern electromagnetism and light scattering.

I would also like to thank all my lab mates from the Functional Optical Imaging Laboratory (FOIL) for their continued friendship and support throughout my Ph.D. I really value the time we've spent discussing research, life, and future plans. I would especially like to thank Dr. Biqin Dong, for his time and patience as he helped me learn the basics of super-resolution microscopy. I would also like to thank Dr. Yang Zhang, whose solid knowledge of chemistry has played a vital role in helping me to develop strategies for labeling cells and preparing samples. I also really appreciate the time Yang has spent discussing strategies for solving my experimental challenges, and for all the invaluable recommendations for improving my articles and presentations. I would also like to thank my colleague Ki-Hee Song for the time and effort he devoted to helping me to improve my optical alignment. Ki-Hee has also for been a great resource for discussing sSMLM data analysis and application development. I would also like to thank Dr. Brian Soetikno for his initial work on the RainbowSTORM project. Thank you also to Yang, Ki-Hee, and Lisa Beckmann for their help testing RainbowSTORM.

I would like to thank my colleagues in the CIGHT lab for their help laying a strong foundation during the beginning of my Ph.D. I would especially like to thank Jen Reed and Dr. Debby Basu for helping me to learn new sample preparation methods and providing advice for preparing samples. I would also like to thank Debby and Dr. David Pinelli who have opened

their hearts and home to me during my last year of my Ph.D. I have been extremely fortunate to have such generous friends to help guide me through this process.

I would also like to thank my collaborators in the Scott lab, especially Dr. Sijia Yi and Dr. FanFan Du for their help preparing polymersome samples. Thank you to my collaborators in the McNally lab, especially Dr. Ellis Kim for her help preparing myotonic dystrophy samples. Thank you to the National Science Foundation for supporting my research through the Graduate Research Fellowship Program and the National Institutes of Health for their support through the Biotechnology Training Program (BTP). My experience in BTP was particularly helpful for gaining a broader perspective on biomedical and biotechnology research.

I am also eternally grateful, to my family and friends for their love, and support. I would especially like to thank my parents who have instilled in me the value of education and cleared the way for me to pursue any area of study I desired. I cannot begin to express what a source of inspiration you have been to me throughout this journey. You've both encouraged me to do what's best for me and have helped me to make tough decisions.

Finally, but certainly not least, to my husband Vladimir Coxall- thank you for being with me every step of the way, through the long-distance, the really hard times, and the happy times. I really would not have been able to complete this journey without your love and emotional support. Thank you for listening to every presentation and reading every article I have ever asked you to proofread. I look forward to celebrating the end of this chapter and the beginning of the next leg of our journey together.

LIST OF ABBREVIATIONS

2D	Two-dimensional
3D	Three-dimensional
λ_{\max}	Emission Maximum
λ_{SC}	Spectral Centroid
AF	Alexa Fluor
BCP	Block copolymer
BSA	Bovine Serum Albumin
CR	Cluster Radius
CryoTEM	Cryogenic transmission electron microscopy
CTRL	Control
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DNA	Deoxyribonucleic Acid
EMCCD	Electron Multiplying Charge-Coupled Device
FOV	Field of View
FRC	Fourier Ring Correlation
FRET	Förster resonance energy transfer
FWHM	Full Width Half Maximum
GLOX	Glucose Oxidase
HC	High Concentration
KOH	Potassium Hydroxide
LC	Low Concentration
MINFLUX	Minimum photon Fluxes
Min-NN	Minimum Number of Neighbors
MIP	Maximum Intensity Projection
NEU	Neutravidin
NN	Nearest Neighbor

NR	Nile Red
NTA	Nanoparticle Tracking Analysis
PALM	Photoactivatable Localization Microscopy
PACER	Photon-ACcumulation Enhanced Reconstruction
PAINT	Point Accumulation for Imaging in Nanoscale Topography
PLL	Poly-L-Lysine
PS	Polymersome
PSF	Point Spread Function
QD	Quantum Dots
qSPACE	Quantitative SPectroscopic Analysis for Cluster Extraction
R^2	Coefficient of Determination
RNA	Ribonucleic Acid
ROI	Region of Interest
SH	Spectral Heterogeneity
SMLM	Single-Molecule Localization Microscopy
sSMLM	Spectroscopic Single-Molecule Localization Microscopy
SNR	Signal to Noise Ratio
sPAINT	Spectroscopic Point Accumulation for Imaging in Nanoscale Topography
SRM	Super-Resolution Microscopy
STORM	Stochastic Optical Reconstruction Microscopy
SW	Spectral Window
TIRF	Total Internal Reflection
UV	Ultraviolet Light

DEDICATION

*To my love,
you kept me going even when it felt as though there was nowhere to go.
Thank you for being my guiding light.*

TABLE OF CONTENTS

ABSTRACT	3
ACKNOWLEDGEMENTS	5
LIST OF ABBREVIATIONS	8
DEDICATION	10
TABLE OF CONTENTS	11
LIST OF FIGURES	13
LIST OF TABLES	23
CHAPTER 1: Introduction	24
1.1 Background and Motivation	24
1.1.1 Fluorescence Microscopy	25
1.1.2 Quantitative Fluorescence Microscopy	29
1.1.3 Super-Resolution Microscopy	32
1.1.4 Quantitative Single-Molecule Localization Microscopy (SMLM)	36
1.3 Dissertation Outline	42
CHAPTER 2: Spectroscopic Single-Molecule Localization Microscopy (sSMLM)	45
2.1 Introduction	45
2.2 sSMLM System Design	46
2.3 Multicolor Imaging	47
2.4 Functional Imaging	50
2.5 High-throughput spectroscopy	52
2.6 Intrinsic Fluorescence	52
2.7 Summary	53
CHAPTER 3: Developing Software Tools for Spectroscopic Single-Molecule Localization	
Microscopy Analysis	54
3.1 Introduction	54

	12
3.2 Features and Methods	56
3.3 Algorithms and Equations	58
3.5 Methods	64
3.4 Conclusions	67
CHAPTER 4: Quantitative sSMLM #1: Removing Artifacts from SMLM Image	
Reconstructions	68
4.1 Introduction	68
4.2 Methods and Materials	70
4.3 Results and Discussion	85
4.5 Conclusions	104
CHAPTER 5: Quantitative sSMLM #2: Counting and Sizing Self-Assembled Nanocarriers	
using Quantitative Spectroscopic Analysis for Cluster Extraction	106
5.1 Introduction	107
5.2 Results and Discussion	109
5.3 Conclusions	132
5.4 Methods	132
CHAPTER 6: Quantitative sSMLM #3: Improving image resolution using Photon-	
Accumulation Enhanced Reconstruction	141
6.1 Introduction	141
6.2 Results and Discussion	150
6.3 Conclusions	160
6.4 Materials and Methods	161
CHAPTER 7: Conclusions and Future Perspectives	
	165
REFERENCES	167
VITA	184

LIST OF FIGURES

- Figure 1-1:** Simplified Jablonski diagram showing various fluorescent and non-fluorescent states. During traditional FM experiments, a fluorophore absorbs a photon from the excitation beam causing an electron in the ground state (S_0) to be promoted to the excited state (S_1). Fluorescence occurs when the excited electron undergoes internal conversion and vibrational relaxation resulting in a release of energy in the form of a photon with less energy or a longer wavelength. In SMLM electrons in the S_1 can undergo intersystem crossing (ISC) to a triplet state or weakly fluorescent state. In this state many fluorophores can be reduced using chemical reagents (e.g. β -mercaptoethanol (BME) or mercaptoethylamine (MEA)) which cause transitions to more stable non-fluorescent or dark states via ISC. The electrons will then undergo additional ISC or react with oxygen (O_2) and undergo non-radiative relaxation (NRR) returning to S_0 releasing fluorescence in the process.
- Figure 1-2:** Illustration of the Rayleigh Criterion showing the pairs of (a) resolvable, (b) barely resolvable and (c) unresolvable molecules. An example, showing a (d) 200 nm fluorescent sample (ground truth), imaged using conventional microscopy. (e) The image appears blurred due to the diffraction limit of light. However, using (f) SMLM where photo-switchable or photoactivatable dyes and proteins are used to separate emission events in time and space. (g) The locations of single-molecules after Gaussian fitting and the (h) resulting super-resolution reconstruction.
- Figure 1-3:** (A and B) Using probes specific to the RNA repeat expansions, FISH was used to monitor RNA foci and their proximity to MBNL using total internal reflection fluorescence. The cells were co-labeled with an antibody to MBNL1 (green). Control iPSC-CMs were labeled with a repeat probe for DM1 (CAG)₅ in A and for DM2 (CAGG)₁₀ in B. The distance between RNA foci and MBNL1 foci was quantified. Because MBNL1 foci were only readily detected in DM1 cardiomyocytes, the data from control and DM2 cells represent background signal. Thus, this method was used to measure the distance between RNA repeat expansions and MBNL1 in DM1. (A) RNA foci colocalized with MBNL1 foci in DM1 cardiomyocytes, and the distance between RNA foci and MBNL1 averaged 200 nm, consistent with a very close physical association between RNA repeat expansions and MBNL. Control cardiomyocytes displayed a random distribution of distances (top panel), reflecting the absence of RNA foci and MBNL foci. (B) In DM2 cardiomyocytes, there was no colocalization of RNA foci with MBNL foci, reflecting the absence of MBNL clusters and a pattern similar to cells from healthy controls. Correspondingly, the distances were randomly distributed, similar to control iPSC-CMs (similar distribution between top and bottom panels). Scale bar: 5 μ m (left); 1 μ m (right).

- Figure 1-4:** SMLM image reconstructions of (a) a representative undifferentiated (UD) cell and (b) a representative differentiated (DF) cell (Scale bars: 5 μm). Selected 10 $\mu\text{m} \times 10 \mu\text{m}$ ROIs within the nucleus of the (c) UD and (d) DF cells. (Scale bars: 1 μm) (e) Comparison of the number of NPCs in the ROIs from 7 UD and 7 DF cells. (f) Size histograms of the sizes of the NPCs found in both cell types. (g) The density of the NPC clusters within the 10 $\mu\text{m} \times 10 \mu\text{m}$ ROIs.
- Figure 2-1:** The principle of sSMLM. (a) An example of a sSMLM detection channel which uses a diffraction grating to separate the spatial image (zeroth order) and spectral image (first order) simultaneously. For example, when a (b) 200 nm fluorescent sample (ground truth), is imaged using conventional microscopy the (c) images appears blurred due to the diffraction limit of light. However, using (d) SMLM where photo-switchable or photoactivatable dyes and proteins are used to separate emission events in time and space a (e) super-resolution image can be reconstructed by recording the location of each emission event. However, the identity of each emitter remains unknown. sSMLM can be used to recover this information because (f) the spectra of each emission event can be identified and used to reconstruct a (g) color-coded super-resolution image which better represents the true sample.
- Figure 3-1:** (a) General sSMLM system schematic. (b) sSMLM images with the spatial and spectral images simultaneously captured on different parts of a detector. (c) RainbowSTORM workflow showing how the system calibration module interacts with the analysis module (d) SMLM reconstruction (e) Pseudo-colored sSMLM reconstruction. Images of the three separate channels showing (f) mitochondria labeled with AF647, (g) microtubules labeled with CF660, (h) peroxisomes labeled with CF680, and (i) the overlay image of the three channels.
- Figure 3-2:** Flowchart of the algorithm to generate the background image
- Figure 3-3:** Flowchart of the algorithm to remove overlapping spectra
- Figure 3-4:** Comparison of the expected spectral precision and the average spectral precision estimated by RainbowSTORM at different spectral photon counts [180 to 3800]. Inset shows the zoomed in comparison for the spectral photon count from 1000 to 3800.
- Figure 3-5:** General sSMLM system schematic for 3D imaging using the astigmatism method.
- Figure 4-1:** Schematics of SMLM and sSMLM experimental systems. (a) Excitation optics and instrumentation; (b) SMLM detection channel used to capture images of

cleaned and functionalized surfaces; (c) sSMLM detection channel used to capture spatial and spectral images simultaneously.

Figure 4-2: Flowchart of the algorithm used to compare intensity thresholding and spectral fitting filtering methods.

Figure 4-3: The origin of fluorescent impurities. MIP images (scalebars 5 μm) of unlabeled glass surface (a) before cleaning, (b) after plasma cleaning, and (c) after PLL functionalization. Comparison of the density of fluorescent impurities from 5 different FOVs (d) before surface cleaning (BC) and after cleaning via the piranha solution (Pir), rinsing with potassium hydroxide and sterilization using UV light (KOH+UV), rinsing with Hydrochloric acid and Prop-2-anol (Acid+Alcohol), exposure to UV activated ozone (UV-zone) and exposure to argon and oxygen plasma (Plasma). Comparison of the density of fluorescent impurities for 5 different FOVs on surfaces (e) before and after plasma cleaning (as a reference) and plasma-cleaned surfaces after functionalization via PLL coating, silanization with a final wash of chloroform (Sil+C), BSA and NEU (BBS) functionalization with glucose oxidase buffer (BBS+G) and BBS water as the buffer (BBS+W).

Figure 4-4: Representative MIP images of a bare Fisherbrand™ coverslip (a) before cleaning, (b) after cleaning using the piranha solution, (c) after sonication in 1 M KOH and sterilization using UV illumination, (d) after rinsing with HCl and prop-2-anol, (e) after cleaning with UV-activated ozone, and (f) after exposure to a mixture of oxygen and argon plasma. All images were captured using 532 nm illumination at a power density of 3 kW/cm². Scale bars are 5 μm .

Figure 4-5: Variability in surface cleanliness and uniformity. (a) Comparison of standard chemicals previously purchased for general lab use to ultrapure chemicals purchased specifically for cleaning. Images of a regular coverslip cleaned with standard chemicals (b) and cleaned with ultrapure chemicals (c). Variability of uniformity using ultrapure chemicals. (d-e) fluorescence images at different areas on the same coverslip cleaned with the piranha solution. (f-g) fluorescence images at different areas on the same coverslip cleaned with KOH then sterilized with UV.

Figure 4-6: Representative MIP images of plasma cleaned Fisherbrand™ coverslips functionalized with (a) PLL, (b) silane with chloroform as the final rinse, (c) biotinylated BSA and NEU with glucose oxidase (GLOX) buffer, and (d) biotinylated BSA and NEU with water buffer. All images were captured using 532 nm illumination at a power density of 3 kW/cm².

- Figure 4-7:** (a) Representative spectra from three fluorescent impurities on a Fisherbrand™ coverslips before cleaning. (b) Representative spectra from three fluorescent impurities associated with PLL functionalization.
- Figure 4-8:** Spectroscopic analysis of functionalized coverslips. (a) Schematic of sSMLM detection channel with the slit closed to acquire average images and line. (b) Image of coverslip functionalized with biotinylated BSA and NEU with GLOX buffer (BBS+GLOX). (c) Line spectra of spectra from BBS+GLOX coverslip. (d) Average line spectra from PLL, silane with chloroform rinse, silane with water rinse, BBS+GLOX, and BBS+Water.
- Figure 4-9:** Excitation dependent emission of fluorescent impurities of PLL coated plasma cleaned coverslips under illuminations at wavelengths of (a) 445 nm, (b) 532 nm, and (c) 645 nm. (d) Average line spectra under illuminations at wavelengths of 532 nm and 645 nm.
- Figure 4-10:** Identifying fluorescent impurities during SMLM. (a) Average spatial and spectral image of DNA origami nanorulers, containing two emitting points labeled with single AF532 and AF568 molecules 10 nm apart, immobilized on a PLL coated surface. Images were acquired under illuminations with power densities associated with conventional fluorescence imaging (0.5 kW/cm^2). (b) Stack of 1500 frames of the spatial and spectral images of the nanoruler sample for sSMLM (3 kW/cm^2) using the same FOV. (c) MIP images of the spatial and spectral of the same FOV. (d) Photon count versus time of two nanorulers (1,4) and two fluorescent impurities (2,3) highlighted in average and MIP of SSMLM images. (e) Corresponding spectra of the point sources identified in the average and SSMLM images representing true positive, false positive, true negative and false negative cases for the spectral fitting method. (f) Sensitivity and (g) Specificity comparison for 9 datasets using an emission intensity a threshold of 180 and a spectral fitting filter adjusted R^2 threshold of 0.84.
- Figure 4-11:** (a) Cartoon of the 10 nm AF532 and AF568 nanoruler. (b) Representative spatial image and (c) extracted spectra of nanorulers detected using LPD illumination using the field of view highlighted in Figures 4-11a-c. (d) Representative spatial image and (e) extracted spectra of fluorescent impurities detected using LPD illumination using the field of view highlighted in Figures 4-11a-c.
- Figure 4-12:** Histograms of adjusted R^2 values of average spectra from 174 locations in the LPD image after spectral fitting to the reference spectrum of (a) AF532 and (b) both AF532 and AF568 with the adjusted R^2 threshold of 0.89 highlighted.

Figure 4-13: The number of emitters detected using HPD excitation, the number of nanorulers, and the number of fluorescent impurities detected under HPD excitation for all 9 field of views. The median is highlighted for each box plot

Figure 4-14 Comparing results in minimizing artifacts induced by fluorescent impurities using intensity filtering and our spectral fitting methods. (a) sSMLM spatial and spectral MIP images of nanorulers with fluorescent impurities. (b) Reconstructed super-resolution images without rejecting fluorescent impurities; (c) result after emission intensity filtering; (d) result after spectral fitting. ROI 1 highlights the localized fluorescent impurities that are eliminated by our spectral fitting method but are misidentified by intensity filtering method. ROI2 highlights the case of spatial overlapping of fluorescent impurities and nanorulers results in higher localization uncertainty. The resulting super-resolution images of ROI2 are further magnified in (e) before filtering (standard deviation (S.D.) 52.9 nm), (f) after intensity filtering (S.D. 40.1 nm), and (g) after spectral fitting (S.D. 22.5 nm). (h) Averaged spectra of fluorescent impurities and nanoruler emission. (i) Reconstructed color-coded super-resolution image of stretched lambda phage DNA labeled with YOYO-1 dye on a silane functionalized surface before rejecting emission unrelated to the DNA-YOYO sample (resulting artifacts highlighted by white triangles); (j) result after intensity filtering contains artifacts from unwanted fluorescence; (k) result after spectral fitting specifically removed artifacts induced by unwanted fluorescence.

Figure 5-1: (a) Chemical structure of NR; (b) Chemical structure and illustration of the BCP for PS assembly; (c) Illustration of the assembled PS (the green color represents the polar end of the BCP and the blue color represents the non-polar end) and the difference in the emission spectra of NR when transiently bound to the PS (yellow) and the PLL-coated glass substrate (red). Free non-fluorescent NR is shown in gray; (d) Schematic of our sPAINT experimental setup. TL: tube lens; S: slit; G: transmission grating; L: lens; EMCCD: electron multiplying charge-coupled device.

Figure 5-2: (a) Comparison of the average emission spectra of NR in solutions containing PS and sample preparation reagents (PLL and PBS) measured using a fluorimeter and sPAINT. (b) Representative super-resolution reconstructions of the PLL coated glass used as the control sample. (Scale bar: 1 μm). (c-e) Histograms of the emission maxima (λ_{max}) of NR interactions in ROIs containing PS from three different images (six clusters were selected in each image). The selected spectral window (SW, 595 – 625 nm) is highlighted in red. (f) Reproducibility test showing the histogram of λ_{max} from PS acquired on a different day with similar experimental conditions. (g) Comparison of the identification percent of localizations from the PS λ_{max} histogram and

identification percent of localizations from the control λ_{\max} histogram as the window size of the SW increases (peak position of SW = 610 nm).

Figure 5-3: (a) A representative reconstructed super-resolution image of the immobilized PS sample and NR interactions (Scale bar: 1 μm); (b) Histogram of the λ_{\max} of NR interactions in three ROIs containing PS as highlighted by the yellow squares numbered 1-3; (c) Histogram of the λ_{\max} of non-specific NR interactions in the control sample. The SW used for detecting NR interactions with PS is highlighted in red

Figure 5-4: sPAINT reconstructions before applying the spectral window for selecting PS for (a) example 1 and (b) example 2. Overlay of the reconstructions with the localizations designated as PS in cyan and localizations designated from non-specific binding in red from (c) example 1 and (d) example 2. sPAINT reconstructions of the PS channel for (e) example 1 and (f) example 2. sPAINT reconstructions of the non-specific binding channels for (g) example 1 and (h) example 2 (Scale bar: 500 nm).

Figure 5-5: (a) Representative super-resolution reconstruction showing individual clusters with varying sizes identified by the optimized density-based spatial clustering for applications with noise (DBSCAN) algorithm marked by the red plus signs (Scale bar 500 nm). (b) Comparison of the differences in the cluster counting accuracy using DBSCAN for trials using data with different cluster densities (number of nearest neighbors or number of NN), the red bar shows data selected based on the spectroscopic and spatial information while the blue bar shows data selected for analysis based on spatial information alone.

Figure 5-6: Histograms of the (a) spatial photon counts, (b) spectral photon counts and (c) spectral precision used for sSMLM analysis from a representative image.

Figure 5-7: Histograms showing the comparison between the number of NN of localizations with 450 photons in the spatial domain before ($n \sim 10^5$) and after application of the selected spectral window ($n \sim 10^4$) of (a-c) 3 control images and (d-f) 3 PS images. (g) Plots showing the number of clusters using localizations within the SW as the min-NN increases in representative PS and control images. The black dashed line shows the threshold used in this study.

Figure 5-8: Plots showing the number of clusters with localizations with at least 450 photons in the spatial domain as the min-NN increases in the (a) PS and control images. The black dashed line shows the threshold used to compare spatial clustering without considering the PS window. Representative super-resolution reconstructions of the (b) PS sample (white arrows indicate changes in PS morphology) and (c) the control sample after application of the clustering thresholds (min- $\text{NN}_c = 30$ and Int=450). (Scale bar 500 nm).

- Figure 5-9:** Clusters extracted from (a) the PS sample and (b) control based on DBSCAN alone; Clusters extracted from (c) the PS sample and (d) control using sPAINT (Scale bar: 500 nm).
- Figure 5-10:** Localizations excluded from spectroscopic analysis based on photon number (Average localization uncertainty ~ 25 nm, Photon budget: 300-450 in the spatial domain).
- Figure 5-11:** (a) The qSPACE workflow shows the (i) detected spatial localizations with the location of the sample highlighted by the green circles. A subset of localizations containing spectroscopic information is used to create a (ii) validation map that shows clusters with the selected spectra. All detected localizations are used for (iii) spatial clustering without considering spectroscopic information. (iv) Localizations from spatial clusters, which are spatially correlated with the validation map, are recovered for further analysis while artifacts are rejected. (v) The number, size, and morphology of the extracted sample can be further analyzed. (b) Detailed qSPACE flowchart outlines how spectroscopic information is used to recover localizations from clusters from the PS sample while rejecting artifacts from non-specific binding.
- Figure 5-12:** The average percent variance in size measurements ($P_{\Delta\text{SIZE}}$) and localization density for clusters as the number of localizations per cluster increases. The black line shows where $P_{\Delta\text{SIZE}} = 5\%$. The dashed blue and red lines indicate the min-NN of 35 and LD of $3.5 \times 10^{-3} \text{ nm}^{-2}$ required for accurate sizing.
- Figure 5-13:** Plots showing the number of clusters at least 300 photons in the spatial domain as the minimum number of NN increases in the (a) PS and control images. The dashed black line represents the threshold used for detecting potential clusters for qSPACE. Representative super-resolution reconstructions of the (b) PS sample and (c) the control sample after application of the clustering thresholds (min- $\text{NN}_C=35$ and Int=300), (d) PS sample and (e) the control sample after application of the clustering thresholds (min- $\text{NN}_C=45$ and Int=300). The white arrows indicate morphology changes in d. (Scale bar: 500 nm).
- Figure 5-14:** Plots showing the tuning of the min- NN_v for noise removal in the validation channel using qSPACE with a SW window of 595-625 nm and min- NN_C of 35 for potential cluster assignment. The dashed black line shows the selected min- NN_v .
- Figure 5-15:** (a) Histogram of the localization densities (black dashed lines indicate the $3.5 \times 10^{-3} \text{ nm}^{-2}$ LD threshold) for qSPACE and (b) sPAINT; (c) Comparison of the global FRC curves for sPAINT (red) and qSPACE (blue). The dashed lines are the corresponding FRC resolution for each method at the 1/7 FRC

threshold; (d) Size distribution of PS (0.02 mg/mL) measured by NTA and qSPACE.

Figure 5-16: CryoTEM image of PS sample.

Figure 5-17: Representative super-resolution reconstructions of the (a) CTRL sample with artifacts pseudo-colored in red and misidentified PS pseudo-colored in cyan. The three white arrows highlight three examples of sample misidentification; (b) LC sample (0.02 mg/mL) with artifacts pseudo-colored in red and validated PS pseudo-colored in cyan; and (c) HC sample (0.2 mg/mL) with artifacts pseudo-colored in red and validated PS pseudo-colored in cyan (Scale bar: 1 μ m). The size distributions for the (d) misidentified PS in the CTRL samples; (e) validated PS in the LC samples; and (f) validated PS in the HC samples. Comparison between the total of number of PC found and the qSPACE VC in the (g) CTRL samples, (h) LC samples, and (i) HC samples.

Figure 5-18: (a) Histograms showing the size distribution of artifacts identified from 4 reconstructions of the control sample (809 clusters with average size 120 ± 4 nm) and clusters identified as artifacts in the LC (504 clusters with average size 111 ± 5 nm) and HC (501 clusters with average size 90 ± 7 nm) PS samples by qSPACE. (b) Histograms showing the size distribution of artifacts identified in 4 reconstructions of the control sample (229 clusters with average size 199 ± 2 nm) and clusters identified as artifacts in the LC (273 clusters with average size 220 ± 11 nm) and HC (186 clusters with average size 211 ± 8 nm) PS samples by sPAINT (clusters formed from localizations with λ_{\max} outside the SW filter).

Figure 5-19: The average emission spectra of NR in solutions containing PS and sample preparation reagents (BSA and NEU) measured using a fluorimeter.

Figure 6-1: Schematic showing how image resolution is improved by molecular discrimination. (a) Due to the wave nature of light, when light comes from a point emitter focused by an optical imaging system, the interference can result in a blurred distribution of light called a PSF. (b-c) The size of the PSF sets up the fundamental resolution limit of an optical imaging system, namely Abbe diffraction limit. (d-f) In SMLM, the probable location of a single fluorescent molecule can be estimated from the centroid of the fluorescence diffraction pattern produced on a camera. The localization precision is determined by the photon count collected in each captured frame. The scatter plot and rendered image of 100 blinking events indicate a localization precision of ~ 30 nm. (g-i) Through molecular discrimination, blinking events can be combined, resulting in an improved localization precision of sub-3 nm. (j) Improved resolution using photon accumulation with respect to the number of blinking events (NB). White crosses denoted positions of all localizations in each test and red crosses denoted their centroids (upper panel). Reconstructed images only using

centroids (middle panel). Localization precision is calculated along the x-axis and its intensity was then normalized for comparison, which are 28.57 nm, 13.24 nm, 9.02 nm, 6.40 nm, 3.89 nm, and 2.85 nm for NB=1, 5, 10, 20, 50, and 100, respectively (lower panel). (k) Simulated localization precision with respect to NB. Error bars are from 100 independent simulations. (l) The black line is the curve calculated by σ_0/\sqrt{NB} , where σ_0 is the localization precision at NB=1.

Figure 6-2: Illustration of two different paths to combine all photons from the same molecule. Simulated stochastic emission events from the same single molecule (a) Simulated diffraction-limited spatial images and their respective reconstructed locations. (b) Sum of the photons in the diffraction-limited image (c) Location of the emitter after combining all photons from different blinking events. (d) All single molecule localizations are combined after reconstruction. (e) The centroid of all localizations from the same molecule. (f) The location of the emission events after the reconstruction based on option 1 overlaps with the centroid of the combined localizations after taking option 2.

Figure 6-3: Experimental demonstration of PACER in achieving a 1.7-nm localization precision. (a) The schematic of sSMLM. (b) One frame of the spatial images and (c) one frame of the simultaneously acquired spectral images of the QD sample. (d) The scatter plot of localization events in the red dashed box in (b). (e) Histogram of the SC distribution. (f) Fluorescence spectra of three QDs after classifying by SCs using spectral intensity threshold of 300, 460, and 480 photons, respectively, and spectral windows of 575-585 nm, 585-600 nm, and 615-625 nm, respectively, as filtering criteria. The corresponding SC of each fluorescence spectrum is noted as an open circle in the plot. (g) Rendered sSMLM images after combining multiple emission events with NB of 1, 10, and 100, respectively. (h) Line profiles across two QDs in (g) with NB of 1, 10, 100.

Figure 6-4: Identification of individual QDs. (a-c) Single frames with emission from QD1, QD2 and QD3 respectively occurring from the same location highlighted in the red dashed box. (d) Normalized emission spectra of QD1, QD2 and QD3 extracted from spectral images in (a-c).

Figure 6-5: Illustration of resolution improvement using QDs through PACER. (a) Rendered sSMLM images after combining multiple emission events with NB of 1, 5, 10, 50, and 100, respectively. (b) The corresponding line profiles across two QDs.

Figure 6-6: Experimental validation of PACER using DNA nanoruler samples. (a) Schematic illustration of a DNA nanoruler labeled with a pair of AF647 featuring a predefined mark-to-mark distance. (b) Representative emission

spectra of two molecules on one nanoruler. (c) Histogram of SCs indicates the existence of two molecules with distinct spectral signatures on one nanoruler. (d) The average spectra of the two molecules separated by SC at the wavelength of 669 nm. (e) Schematic of a DNA nanoruler featuring mark-to-mark distance of 23 nm. (f) The scatter plot and (g) the rendered sSMLM image of localization events with colors indicating distinct spectral signatures after stage drift correction. (h) Comparison of molecule location in SMLM image (the dashed black line) and sSMLM images without (the dashed colored lines) and with (the solid colored lines) PACER. (i) Calculated location of molecules through PACER. (j) Histogram of mark-to-mark distance measured from 57 nanorulers. The representative results of a DNA nanoruler featuring mark-to-mark distance of (k-n) 11 nm and (o-r) 6 nm.

Figure 6-7: PACER imaging DNA origami nanogrids consisting of a 3×3 array of AF647 fluorescent molecules with 11-nm inter-molecular spacing. (a) Schematic illustration of the DNA origami nanogrid labeled with 3×3 array of AF647 with 11-nm inter-molecular spacing. (b) Conventional SMLM reconstruction of all stochastic blinking events. (c) All events can be separated to seven clusters based on their spatial locations and SCs in the extended spatial-spectral domain. (d) The corresponding sSMLM reconstruction with pseudo-colors assigned to seven clusters. (e) Locations of fluorophores calculated after PACER. Black circles represent the localization precision after PACER. The gray dashed circles represent the best guess of the location and orientation of the nanogrid.

LIST OF TABLES

- Table 4-1:** Summary of chemical reagents used in this study
- Table 4-2:** Summary of proteins used in this study
- Table 4-3:** Sensitivity and Specificity comparison between single molecule filtering based on emission intensity (threshold 180) and spectral fitting (threshold 0.84).

CHAPTER 1: Introduction

1.1 Background and Motivation

Since its invention, optical imaging has enabled discoveries which have improved our understanding of the world around us and beyond. Fluorescence microscopy (FM) in particular, has been an essential tool for visualizing cells, biomolecules (e.g. proteins, nucleic acids and cellular vesicles), and synthetic materials (e.g. synthetic vesicles, polymer fibers, and quantum dots) (1, 2). As a result, many quantitative methods have been developed to maximize the amount of information which can be extracted from imaging experiments. For example, quantitative image analysis for FM has been used to quantify the number of particles within the cell (3-6), measure the uptake of nanoparticles by cells (7), and characterize immobilized biomolecules (8-12). However, due to the diffraction limit of light (13), FM cannot be used to directly visualize molecules and structures smaller than ~250 nm. To address the limited resolution of FM, many studies have been combined with electron microscopy (EM) (14) and atomic force microscopy (AFM) (15) which can achieve nanometer scale resolutions. However, invasive sample preparation steps for both EM and AFM can alter the native environment of the cells and biomolecules making these methods prone to artifacts which can obscure the true structure of the sample (16).

Scientists have developed super-resolution techniques such as single-molecule localization microscopy (SMLM) (17, 18), stimulated emission depletion (STED) microscopy (19), and structured illumination microscopy (SIM) (20) to overcome the diffraction limit of light. The discovery of SMLM and STED led to the award of the Nobel Prize in Chemistry to Profs. Eric Betzig, William E. Moerner and Stefan Hell in 2014. The advancement of super-resolution

microscopy (SRM) has also led to the development of new quantitative analyses particularly using SMLM. Quantitative SMLM, has allowed scientists to count the number of subunits which make up a complex structure (8, 21-23), quantify the number of nanostructures in a region of interest (24-26), and characterize biomolecules within cells (27-29).

This dissertation will focus on the development of quantitative methods using a recently developed derivative of SMLM called spectroscopic single-molecule localization microscopy (sSMLM) which simultaneously captures the full emission spectrum and location of stochastically emitting molecules (30-32). This chapter will provide the relevant background to understand FM and SRM with a focus on SMLM. Additionally, this chapter will discuss the challenges of developing quantitative methods using SMLM. Finally, a brief summary of the remaining chapters of this dissertation will be provided.

1.1.1 Fluorescence Microscopy

FM is enabled by fluorophores or fluorochromes which absorb photons of a specific energy level and emit photons with less energy (red-shifted or with a longer wavelength with respect to the absorbed light) (33, 34). Figure 1-1 features a simplified Jablonksi diagram which describes the physical and molecular basis for the theory of fluorescence excitation and emission (35). Using this basic principle, scientists and engineers have developed a range of optical system, fluorescent molecules, labeling strategies, and experimental techniques to maximize the high contrast achieved using FM.

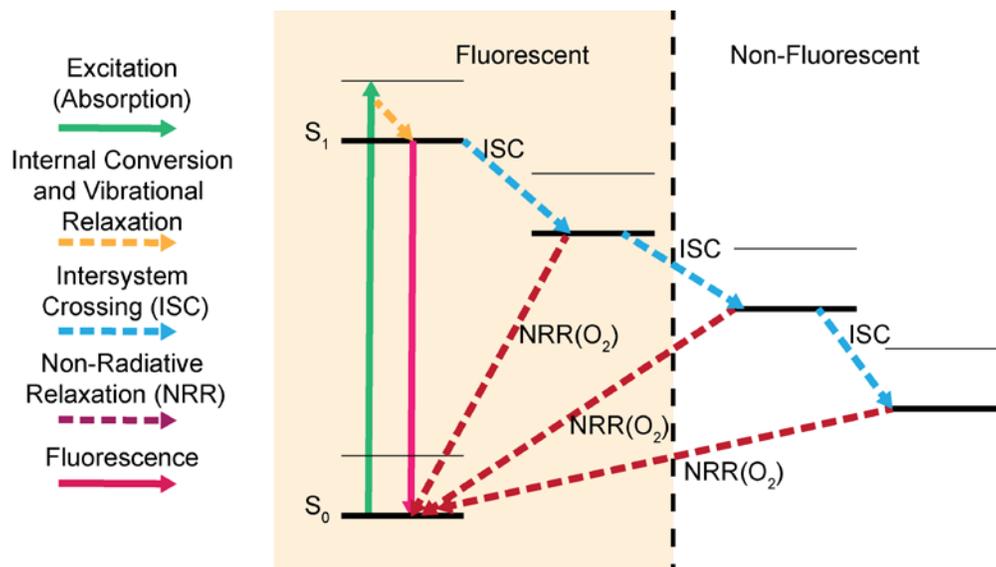


Figure 1-1: Simplified Jablonski diagram showing various fluorescent and non-fluorescent states. During traditional FM experiments, a fluorophore absorbs a photon from the excitation beam causing an electron in the ground state (S_0) to be promoted to the excited state (S_1). Fluorescence occurs when the excited electron undergoes internal conversion and vibrational relaxation resulting in a release of energy in the form of a photon with less energy or a longer wavelength. In SMLM electrons in the S_1 can undergo intersystem crossing (ISC) to a triplet state or weakly fluorescent state. In this state many fluorophores can be reduced using chemical reagents (e.g. β -mercaptoethanol (BME) or mercaptoethylamine (MEA)) which cause transitions to more stable non-fluorescent or dark states via to ISC. The electrons will then undergo additional ISC or react with oxygen (O_2) and undergo non-radiative relaxation (NRR) returning to S_0 releasing fluorescence in the process.

In general, standard FM systems utilize a widefield or epifluorescence setup where microscopes are equipped with filter cubes which contain two filters and a dichroic mirror placed at a 45-degree angle (2). First, the excitation filter allows light within the excitation band to be transmitted to the dichroic mirror which reflects the light onto the sample labeled with fluorophores. When the excitation light interacts with the sample, the fluorophores absorb the light and emit red-shifted light which is transmitted through the dichroic mirror and emission filter to the detector (2). Widefield systems are the most common type of FM optical system due to its straightforward and relatively inexpensive design (2). Additionally, widefield FM is ideal for

dynamic or time-sensitive measurements. However, widefield systems have several limitations which include the fading of fluorophores and the detection of unwanted fluorescence from out-of-focus fluorophores (2).

Fading can occur in two ways, first there is photobleaching where the fluorophores become unable to absorb light due to prolonged exposure to the illumination source (2). Fluorophores can also be quenched, meaning photons are transferred to a dark state where fluorescence is suppressed for a longtime period as described in Figure 1-1. Fading is particularly challenging for thick samples and three-dimensional (3D) imaging (1). To reduce the impact of fading, antifade buffers and controlling the illumination intensity have been used, however, these approaches only have a limited overall impact. Additionally, antifade buffers cannot always be used when imaging biological samples. Typically, to reduce out-of-focus background signals, total internal reflection fluorescence (TIRF) has been used to image molecules at the glass interface with significantly higher signal-to-noise ratio (SNR) (1). In a widefield system, light propagation at the interface between two media with different refractive indices n_1 and n_2 is described by Snell's Law

$$n_1 \sin \theta_{in} = n_2 \sin \theta_{ref} , \quad (1-1)$$

where θ_{in} and θ_{ref} are the angle of incidence and the angle of refraction respectively (1). When illuminating the coverslip using a collimated beam, light propagates through the first medium (glass: $n_1 = n_{glass}$) to the second medium (the aqueous environment of the sample: $n_2 = n_{sample}$) with $n_{sample} < n_{glass}$, the beam is then refracted through the glass-sample interface at an $\theta_{in} > \theta_{ref}$ (1).

By changing the θ_{in} to the critical angle (θ_c) which is described by $\sin \theta_c = \frac{n_{sample}}{n_{glass}}$, the refracted beam propagates parallel to the surface and when $\theta_{in} > \theta_c$ light is propagated back into the first medium, thus achieving TIRF (1). In the TIRF configuration, the excitation intensity decreases

exponentially with increasing distance from the glass surface along the axial direction (1). As a result, only fluorophores close to the surface are excited. Therefore, TIRF is only applicable for surface measurements, and unwanted fluorescence from out-of-focus fluorophores continues to be a challenge for 3D imaging and for imaging thick samples (1).

Alternatively, confocal and multi-photon optical systems have both been developed to resolve these challenges. A confocal microscope is equipped with a pinhole to prevent out-of-focus fluorescence from being detected (1, 2, 36). Additionally, since only fluorophores in the focal plane can be detected, confocal microscopes rely on laser scanning in order to capture the full sample (1). This makes confocal microscopy relatively slow and therefore ill-suited for imaging highly dynamic processes (1). Confocal microscopy is also still susceptible to fading since the out-of-focus fluorophores are still exposed to the excitation beam. Two-photon and multi-photon microscopy are based on non-linear optical processes which significantly reduce fading and limit the detection of out-of-focus light (1). Briefly, multi-photon imaging systems rely on a single fluorophore absorbing two or more photons at the same time (1). To achieve this, multi-photon microscopes rely on high-power femtosecond mode locked pulsed lasers with a longer wavelength typical used for a given fluorophore (1). Since multi-photon systems confine fluorescence to the focal plane, they are immune to fading and detection of fluorescence from out-of-focus fluorophores. This allows for improved resolution in the axial domain for 3D imaging and improves the image quality of thick samples (37). However, like confocal microscopes, multi-photon microscopes also rely on laser scanning, thus limiting its applicability for imaging dynamic processes (1). Even when combined with deconvolution methods which numerically improve the

image resolution, the achievable resolution using these systems is still limited to around 200 nm (1, 37).

1.1.2 Quantitative Fluorescence Microscopy

In quantitative FM studies, the intensity of the detected fluorescence is used as a proxy for the local concentration of labeled molecules (38, 39). This is very challenging since many factors can affect intensity measurements, including but not limited to the brightness of the fluorescent signal, noise from experimental components, background signals, and the uniformity of the illumination beam (38). Therefore, much care and attention to detail must be taken when designing experimental and analysis protocols for quantitative studies (38, 39). One typical application of quantitative FM is to characterize the dimensions of structures and measure the distances between molecules. However, due to the fundamental diffraction limit of light measurements can only be directly measured for molecules ~250 nm apart (1, 37, 40).

This fundamental limit is due to wave nature of light which was first described by Ernst Abbe (13) as

$$R_{x,y} = \frac{\lambda}{2NA}, \quad (1-2)$$

where $R_{x,y}$ is the lateral resolution, λ is the wavelength of light and NA is the numerical aperture of the optical system. Therefore, even with an ideal optical system, the achievable resolution is limited to half the wavelength of visible light. Additionally, Abbe described the axial resolution (R_z) as

$$R_z = \frac{2\lambda}{NA^2}, \quad (1-3)$$

The diffraction limit can be further described using Rayleigh's criteria ($R_{x,y} = \frac{0.61\lambda}{NA}$) which captures the distance at which the images of two distinct point sources can no longer be distinguished (Figure 1-2a-c) (37). In addition to the optical resolution limit, the spatial density of the image, which is dependent on the number of pixels contained in the digital image, is also important for achieving the highest possible image resolution (1). The Nyquist-Shannon sampling theory requires the sampling rate to be at least twice the highest spatial frequency in the diffraction limited image (1). For example, when imaging a single molecule with a full width half maximum (FWHM) of 200 nm the maximum pixel size of the camera should be \sim FWHM/2 (1). Therefore, to ensure the desired image resolution is achieved, detectors with appropriate pixel sizes must be selected. Figure 1-2d shows a nanostructure and Figure 1-2e shows the resulting from conventional microscopy.

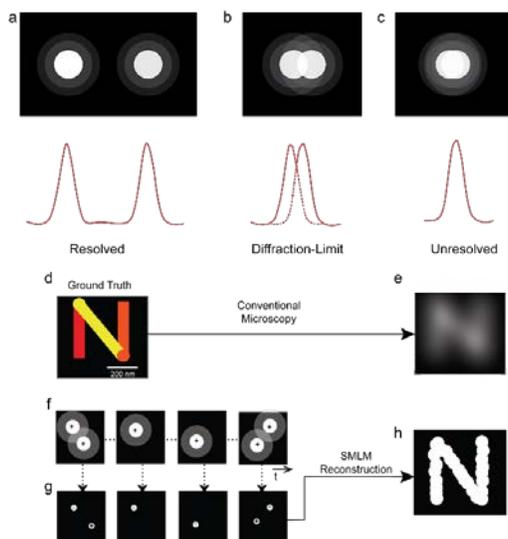


Figure 1-2: Illustration of the Rayleigh Criterion showing the pairs of (a) resolvable, (b) barely resolvable and (c) unresolvable molecules. An example, showing a (d) 200 nm fluorescent sample (ground truth), imaged using conventional microscopy. (e) The image appears blurred due to the diffraction limit of light. However, using (f) SMLM where photo-switchable or photoactivatable dyes and proteins are used to separate emission events in time and space. (g) The locations of single-molecules after Gaussian fitting and the (h) resulting super-resolution reconstruction.

Single-molecule imaging and spectroscopy has been achieved by imaging well separated fluorescent molecules (3, 41-43). The utility of this technique for performing quantitative imaging was first demonstrated by tracking of myosin V molecules movement along actin filaments (44). However, because single-molecule imaging techniques require the sparse distribution of molecules for direct quantitative measurements, alternative methods have been developed to characterize dense clusters of molecules (40, 45). For example, sequential photobleaching experiments have been developed to estimate the number of fluorescent proteins or dye molecules in a dense cluster in order to identify subunits of proteins or quantify dyes encapsulated in vesicles (8, 12). Additionally, Förster resonance energy transfer (FRET) has been developed to measure distances below the diffraction limit (46, 47). FRET experiments measure the intensity change between two

appropriate fluorophores which are 10 nm or less apart. The FRET efficiency (46) can be calculated by

$$E_{FRET} = \frac{1}{1 + \left(\frac{r}{R_0}\right)^6}, \quad (1-4)$$

where r is the distance between two molecules and R_0 is the characteristic distance where the FRET efficiency is 50 percent (46). FRET probes also need to be appropriately selected so the emission spectrum of the donor probe (e.g. Cyanine 3) overlaps with the absorption spectrum of the reporter probe (e.g. Cyanine 5) (46, 47). However, FRET measurements are limited to short length scales (<10 nm) and do not enable direct visualization of the individual molecules (46, 47).

1.1.3 Super-Resolution Microscopy

SRM was developed to overcome the diffraction limit, thus enabling new discoveries. SIM uses patterned illumination, typically a grating pattern, projected onto the sample and recorded at a range of orientations (1, 20, 48). Frequency information is determined using Fourier analysis to calculate the correct spatial positions of the signals (20, 48). The resolution improvement is constrained by the periodicity of the grating pattern; however, other optical constraints limit the lateral and axial resolutions to about ~125 nm and ~300 nm respectively (20, 48-50). Though the resolution improvements using SIM are limited, the use of relatively low power illumination, high speed interference pattern generation along with standard dyes and labeling protocols has enabled live cell imaging with a temporal resolution on the order of milliseconds to seconds (51, 52). To further improve the technique, non-linear SIM which uses photo-switchable probes has been developed to improve the lateral resolution to 50 nm but at reduced temporal resolution on the order of seconds to minutes (20).

STED uses a donut-shaped illumination pattern where a low power illumination beam is used to restrict fluorescence to the central focal point with a diameter well below the diffraction limit (19). Using high-power illumination of a longer wavelength than the excitation beam, fluorescent emission from fluorophores surrounding the central focal point are depleted (19). This allows only fluorophores in the central point to be recorded and enables images with ~50 nm lateral resolution (40, 48). Additionally, because STED uses laser scanning it has been combined with optical sectioning to achieve 3D imaging of both thin and thick samples with typical axial resolutions of ~150 nm (53). STED has also been demonstrated in live cell (54), live tissue (55), and live mouse (56) studies with a temporal resolution on the order of seconds. While the high-power depletion beam is necessary for achieving the improved resolution, it can result in photodamage to the samples (57). To address this challenge, reversible saturated optical fluorescent translations (RESOLFT) a variation of STED which takes advantage of photo-switchable fluorophores to enable reduction of emission from fluorophores surrounding the center point of the illumination pattern using lower power illumination was developed (57). Further, the development of minimal photon fluxes (MINIFLUX) which uses single-molecule localization techniques combined with STED/RESOLFT to more precisely determine the true location of the detected fluorophores has allowed for sub-10 nm imaging resolutions to be achieved (58-60). Since STED's illumination pattern only allows a small FOV to be excited at a time, a scanning laser is used to image larger FOVs. Additionally, arrays of STED illumination beams have been used to image dynamic processes (59). While, this eliminates the time-scale constraints of STED it makes large FOV imaging expensive and challenging to implement.

SMLM relies on the detection of “on-off” emission events from fluorescent molecules which allows the position of individual molecules to be estimated with high precision (17, 18). In order for SMLM to be achieved, multiple frames of well-spaced emission events from molecules at random positions in the sample are recorded. Next, the blurred distribution of light or point spread function (PSF) (Figure 1-2f) of each emission event is fitted with a Gaussian and the location of the emitter is estimated with a localization precision (σ) (Figure 1-2g) described by (61, 62)

$$\sigma = \sqrt{\left(\frac{s_i^2 + a^2/12}{N}\right)\left(\frac{16}{9} + 4\tau\right)}, \quad (1-5)$$

where s_i is the standard deviation of the Gaussian fit in the x and y directions, a is the pixel size of the electron multiplying charge-coupled device (EMCCD), N is the number of detected photons, $\tau = \frac{2\pi b^2(s_i^2 + a^2/12)}{Na^2}$, and b is the standard deviation of the camera background. Finally, after locating each emission event, a super-resolution image can be reconstructed as shown in Figure 1-3h. To enable the “on-off” emission required for SMLM, a variety of strategies have been developed.

The most common SMLM strategies are photoactivation localization microscopy (PALM) (17), stochastic optical reconstruction microscopy (STORM) (18, 63, 64), and point accumulation for imaging in nanoscale topography (PAINT) (65). PALM traditionally relies on the photoactivation and subsequent photobleaching of photoactivatable proteins produced by genetically modified cells (17, 27). Repeated photoactivation and photobleaching steps are used to control the random photoactivation of the proteins (17). PALM is ideal for live cell imaging since the cell itself produces the fluorescent proteins, however, the use of genetically modified

cells has made sample preparation challenging. Additionally, the limited number of fluorescent proteins available reduces its applicability for multicolor studies. However, these challenges are being overcome by the production of new photoactivable dyes and proteins for imaging live and fixed cells (66, 67). Meanwhile, STORM relies on the photo-switching of specially designed FRET probes or dyes which react with reducing buffers (e.g. BME and MEA as described in Figure 1-1) (18, 63, 68, 69). The photo-switching process reduces the probability of multiple fluorophores simultaneously emitting within the same spatial region (18, 63, 68, 69). While STORM is easier to implement and enables the use of a wide range of off the shelf fluorophores (70), various labeling strategies are required to attach the dyes to the biomolecules and cellular features being imaged. For example, immuno-labeling which uses antibodies chemically functionalized with STORM dyes requires the cell membrane to be permeabilized in order for the labels to be attached to the cellular targets (69). These additional sample preparation steps can introduce fluorescent impurities or cause artifacts which may obscure experimental results (71). PAINT relies on dyes which have an affinity for specific conditions (65, 72-74). For example, Nile Red (NR) is a fluorescent molecule which transiently binds to non-polar (hydrophobic or water repelling) environments and only fluoresces when bound to the target molecule (65). The transient binding of the freely diffusing probes, therefore, creates the “on-off” signals required for SMLM (65, 72-74). DNA-PAINT is a variation of PAINT which uses labeled DNA molecules which transiently bind to complementary DNA strands (75, 76). One significant challenge for PAINT and its variants, is the high background fluorescence due to the freely diffusing probes (73, 76, 77). To address this, TIRF systems are typically used to improve the SNR. However, this limits PAINT to use in thin samples or surface immobilized molecules. Recently, FRET-based DNA-PAINT has

been developed to address this issue, however, photobleaching of the acceptor probes has limited its utility (78).

Since SMLM, uses widefield imaging, a variety of optical designs have been developed to achieve 3D imaging. These methods include the astigmatism method which uses a cylindrical lens to modify the PSF in the axial domain (64). Additionally, biplane optical designs split the single molecule emission events into two separate channels and use the spatial information from each channel to estimate the axial positions of the emission events (79, 80). Additionally, 4Pi interferometric methods which capture the self-interference of each emission event to estimate the emitter's axial location have also been developed (81, 82).

While all three super-resolution imaging modalities allow for visualization of molecules with high resolutions, SMLM has unlocked the potential for a range of new quantitative analyses. Unlike SIM, SMLM enables much higher image resolutions (SIM: 100 nm vs SMLM: 10 nm). Further, SMLM captures widefield images and unlike STED variants, SMLM doesn't require laser scanning or laser arrays to capture large FOVs. Although the temporal resolution is limited to the order of seconds to minutes and the imaging depth is limited to a few micrometers, the high resolution achieved by SMLM makes it useful for many quantitative applications (48, 83).

1.1.4 Quantitative Single-Molecule Localization Microscopy (SMLM)

Quantitative SMLM has facilitated our improved understanding of biomolecules both within and outside of cells. By taking advantage of the localization of individual molecules, many studies have been developed to quantify the number of biomolecules (5, 25, 84), characterize them in terms of their subunits (12, 21, 22, 24), size (85, 86), and volume (86). Additionally, the

colocalization of biomolecules have been determined by measuring the distance between two molecules (87). The stoichiometry of densely packed protein clusters has also been measured using SMLM (21, 22, 88). One requirement for quantitative SMLM is the identification of emission events from one or more fluorophores attached to a single biomolecule. Nearest neighbor analysis (89), Ripley's K (90), pair correlation (91), and density-based clustering (92-94) are all techniques used to identify spatial patterns in SMLM data.

Here we give two examples of quantitative SMLM, the first details the colocalization measurements of muscleblind-like (MBNL) ribonucleic acid (RNA) splicing protein and RNA molecules using nearest neighbor analysis from a study published in *JCI Insight* (87). The second example shows the use of cluster analysis to count and size nuclear pore clusters (NPCs) in the cell nucleus. We first estimate the distance between MBNL and RNA within fixed induced pluripotent stem cells (iPSC) from patients with Myotonic Dystrophy (DM) and healthy controls (87). DM is a genetic disease which is caused by repeating sequences in the autosomal genes resulting in prolonged skeletal muscle contraction, muscle weakness, and cardiac complications (87). Figure 1-3 shows the measurement of the distance between MBNL protein immuno-labeled with Alexa Fluor 488 (AF488) antibodies and RNA sequences labeled with Alexa Fluor 647 (AF647) using RNA fluorescent in situ hybridization (FISH) probes and comparing DM Type 1 (DM1) cells relative to control (healthy) cells (Figure 1-3a) and DM Type 2 (DM2) cells relative to control cells (Figure 1-3b) (87). Using nearest-neighbor analysis, the center location or foci for labels associated with individual RNA and MBNL molecules was determined (87). The distances between RNA foci and the nearest MBNL foci in 14-16 cells for each type were then estimated using pair correlation functions (91) and the histogram plotted for the cells (Figure 1-3) (87). These

single cell measurements supported the hypothesis that patients with DM1 have MBNL proteins which are colocalized with DM1 RNA (87).

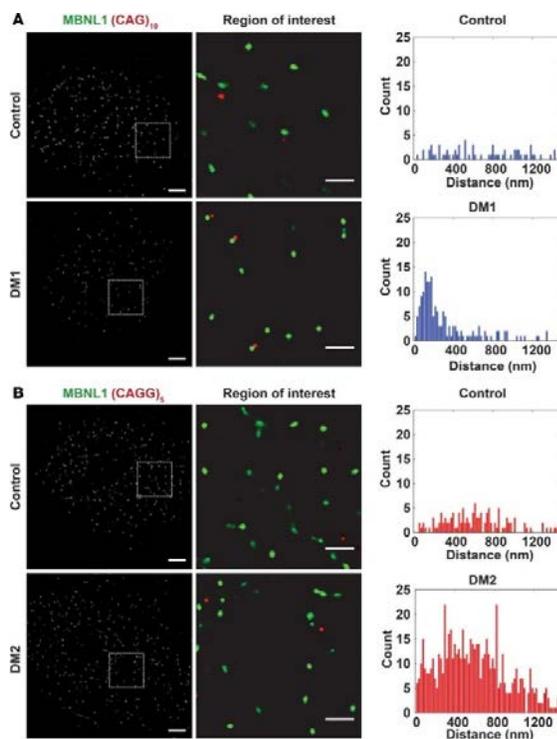


Figure 1-3: (A and B) Using probes specific to the RNA repeat expansions, FISH was used to monitor RNA foci and their proximity to MBNL using total internal reflection fluorescence. The cells were colabeled with an antibody to MBNL1 (green). Control iPSC-CMs were labeled with a repeat probe for DM1 (CAG)₅ in A and for DM2 (CAGG)₁₀ in B. The distance between RNA foci and MBNL1 foci was quantified. Because MBNL1 foci were only readily detected in DM1 cardiomyocytes, the data from control and DM2 cells represent background signal. Thus, this method was used to measure the distance between RNA repeat expansions and MBNL1 in DM1. (A) RNA foci colocalized with MBNL1 foci in DM1 cardiomyocytes, and the distance between RNA foci and MBNL1 averaged 200 nm, consistent with a very close physical association between RNA repeat expansions and MBNL. Control cardiomyocytes displayed a random distribution of distances (top panel), reflecting the absence of RNA foci and MBNL foci. (B) In DM2 cardiomyocytes, there was no colocalization of RNA foci with MBNL foci, reflecting the absence of MBNL clusters and a pattern similar to cells from healthy controls. Correspondingly, the distances were randomly distributed, similar to control iPSC-CMs (similar distribution between top and bottom panels). Scale bar: 5 μ m (left); 1 μ m (right).

While this approach allows for successful identification of colocalization between MBNL and RNA, there are several key experimental steps which need to be considered. Since the proteins are labeled with fluorophores which do not absorb the same wavelength of light, sequential acquisition steps using different laser lines and filter sets to image the protein and RNA are required. This approach presents several challenges, the first is the potential cross talk between fluorophores in each channel (63). To prevent this, appropriate filter cubes, fluorophores, and excitation strategies must be selected when designing the experiment (63). The second challenge is presented by the system alignment since different lasers and filter sets must be used. In order to prevent uneven illumination of the sample during the experiment, both lasers must be carefully aligned (38). Additionally, a pixel registration process can be implemented with both channels in order to remove any aberrations induced by the differences in the optical components used for each channel (76). Sample drift between acquisition steps must also be minimized. Typically, landmarks in the image can be used to correct artifacts due to drifting, these landmarks can come from the sample itself or by adding multicolor fluorescent beads to the sample before beginning the experiment (76). While most of these challenges can be addressed by careful experiment design, the presence of any fluorescent impurities which may originate from sample preparation steps cannot be identified using this approach (71, 95). Fluorescence may also be generated from the cell itself (96, 97). For example, intrinsic fluorescence observed in DNA molecules has been used to reconstruct images of DNA oligonucleotides, chromatin within cells, and chromosomes separated from cells (98). The influence of the length and nucleotide composition of the DNA was also investigated in order to better understand this phenomenon (99). Additionally, non-specifically bound probes can make it challenging to identify the fluorophores specifically bound

to the sample (73, 100). Typically, photon thresholds or a threshold on the number of emission events from each molecule can be used to remove unwanted emission events, however, without understanding the origin of these emission events it is difficult to develop methods to avoid them (83).

In our second example, we immuno-labeled the NUP98 proteins in the NPCs (101) with AF647 antibodies and used density-based clustering analysis to identify emission events spatially related to individual NPCs in fixed undifferentiated (UD) (Figures 1-4a and c) and differentiated (DF) (Figures 1-4b and d) primary keratinocytes. Figure 1-4e shows the comparison between the number of NPCs in 7 UD and 7 DF cells. We also estimated the size of the individual clusters and show the histograms for clusters in both cell populations in Figure 1-4f. Finally, we show a change in the density of the NPCs between DF and UD cells in Figure 1-4g. This study indicates that NPCs are more numerous and more densely distributed in UD cells than DF cells.

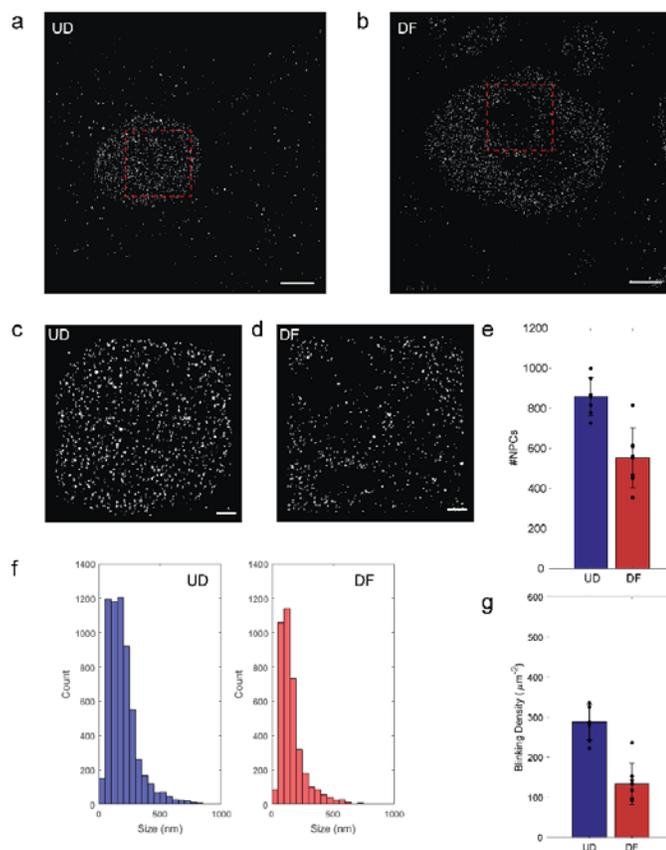


Figure 1-4: SMLM image reconstructions of (a) a representative undifferentiated (UD) cell and (b) a representative differentiated (DF) cell (Scale bars: 5 μm). Selected 10 $\mu\text{m} \times 10 \mu\text{m}$ ROIs within the nucleus of the (c) UD and (d) DF cells. (Scale bars: 1 μm) (e) Comparison of the number of NPCs in the ROIs from 7 UD and 7 DF cells. (f) Size histograms of the sizes of the NPCs found in both cell types. (g) The density of the NPC clusters within the 10 $\mu\text{m} \times 10 \mu\text{m}$ ROIs.

In this example, only a single channel is used for analysis. The major concerns for this type of analysis is establishing appropriate parameters for cluster analysis. In this case, fluorescent impurities (95) and non-specifically bound probes (73, 100) can make it challenging to select appropriate cluster analysis parameters specifically the minimum number of points within a cluster and the associated cluster radius (CR) for defining clusters (83, 92). Typically, the number of emission events or the photons from each emission event is used to set thresholds to exclude

emission associated with impurities and non-specific binding (83, 95). However, these thresholds may also remove many emission events associated with fluorophores labeled to the NPCs. This can result in an under estimation of the size of the NPC clusters due to poor molecular sampling (the portion of the molecule which is visualized) (102). Additionally, undercounting and overcounting artifacts can also occur due to the exclusion of the true sample or inclusion of impurities and non-specifically bound probes (83). Inaccurate merging of emission events can also lead to erroneous spatial and temporal measurements during quantitative SMLM.

While quantitative SMLM is a powerful tool for visualizing and characterizing biomolecules, there are several challenges which can result in misinterpretation of the results. These challenges include overcounting of blinking events, artifacts from fluorescent impurities and non-specific binding, undercounting due to over filtering (83), focal drift during acquisition, and registration of different colored imaging channels (76). In this dissertation we introduce sSMLM and demonstrate how it can be used to address several challenges of quantitative SMLM.

1.3 Dissertation Outline

This dissertation has three central objectives. The first will provide the necessary information for understanding the advancements in super-resolution imaging achieved by sSMLM. Second, efforts to make sSMLM more accessible by implementing an open-source ImageJ plug-in will be discussed. Finally, the use of spectroscopic analysis provided by sSMLM to improve quantitative measurements of nanostructures will be outlined using three examples. Chapter 2 of this dissertation will provide a review of sSMLM and its applications to date.

Chapter 3 will focus on the development of RainbowSTORM, a user-friendly ImageJ plugin, for calibrating sSMLM systems, analyzing sSMLM images, and visualizing sSMLM data. We also show the equations for calculating spectroscopic fields as well as the algorithms for performing background subtraction and identifying localizations with overlapping spectra. Additionally, we compare the performance of RainbowSTORM's estimation of the spectral precision to the theoretical spectral precision using simulated sSMLM images.

Chapters 4 to 6 will show examples of spectroscopic analysis for improving quantitative SMLM. Chapter 4 of this dissertation shows the existence and origins of fluorescent impurities in SMLM. We further demonstrate that these impurities are unavoidable since they originate from necessary sample preparation steps. To address this challenge, we developed a spectral fitting algorithm to identify and reject fluorescent impurities based on their spectroscopic signatures. We used immobilized nanorulers to establish a ground truth and define the sensitivity and specificity of the established method for identifying the true sample. We also show how sSMLM and its ability to reject fluorescent impurities can improve specificity in quantitative imaging and prevent misrepresentation of the structure of the true sample. We then apply our spectral fitting method to reject fluorescent impurities in images of immobilized DNA fibers.

In Chapter 5, we utilized spectroscopic PAINT (sPAINT) to quantitatively image immobilized nanocarriers or polymersomes (PS). We first show how sPAINT can be used to specifically identify the PS sample. We also show that sPAINT can reduce the number of localizations used for image reconstruction due to the restrictive photon budget required for spectroscopic analysis. To address this, we developed quantitative spectroscopic analysis for cluster extraction (qSPACE) which combines the specificity achieved by spectroscopic analysis

with density-based cluster analysis to enable accurate sizing of PS under aqueous conditions. We then employed qSPACE to accurately size and count immobilized PS from solution with different PS concentrations.

Chapter 6 of this dissertation will demonstrate how sSMLM and spectral heterogeneity (SH) of fluorescent molecules of the same species can be used to classify individual molecules. Using the molecular discrimination capabilities of sSMLM, we developed photon-accumulation enhanced resolution (PACER) to perform sub-2 nm precision imaging. We then show how PACER can be used to accurately measure the distance between fluorophores on nanorulers. We further show how PACER can be used to reconstruct the image of a densely packed nanogrid sample.

Finally, Chapter 7 will summarize the findings of this dissertation and present future perspectives for advancing quantitative sSMLM.

CHAPTER 2: Spectroscopic Single-Molecule Localization Microscopy (sSMLM)

2.1 Introduction

Spectroscopic single-molecule localization microscopy (sSMLM) is a correlative imaging technique which simultaneously records the location and the full emission spectrum of each single-molecule emission event (30-32). Figure 2-1 shows an example of a typical grating-based sSMLM system (31) which separates the fluorescent signals into the zeroth order containing the spatial information and the first order containing the spectral information. Both signals are then captured on different sections of an electron multiplying charge-coupled device (EMCCD). As shown in Figure 2-1f and 2-1g, the spectra can be used to classify molecules using a single detection channel. This capability allows for a wider range of dyes to be detected during one acquisition cycle. This method has enabled the generation of multicolor images (30, 31, 103) as well as the development of new functional imaging techniques which use variations in the spectral signals to understand the dyes' local environment (104-108). This chapter will focus on the different types of sSMLM systems and provide a brief review of the sSMLM applications to date.

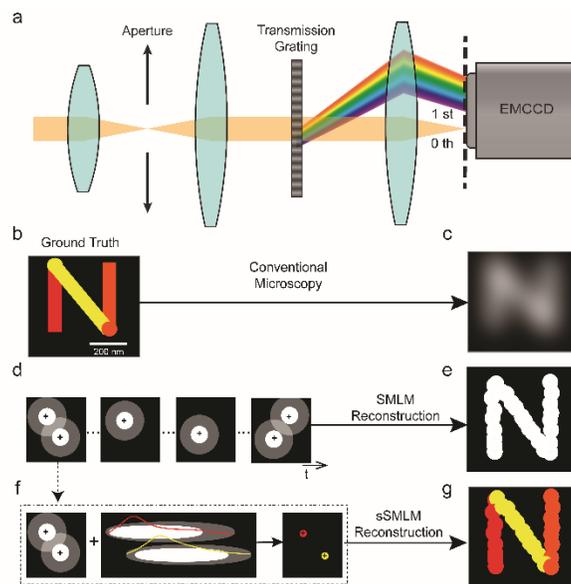


Figure 2-1: The principle of sSMLM. (a) An example of a sSMLM detection channel which uses a diffraction grating to separate the spatial image (zeroth order) and spectral image (first order) simultaneously. For example, when a (b) 200 nm fluorescent sample (ground truth), is imaged using conventional microscopy the (c) images appears blurred due to the diffraction limit of light. However, using (d) SMLM where photo-switchable or photoactivatable dyes and proteins are used to separate emission events in time and space a (e) super-resolution image can be reconstructed by recording the location of each emission event. However, the identity of each emitter remains unknown. sSMLM can be used to recover this information because (f) the spectra of each emission event can be identified and used to reconstruct a (g) color-coded super-resolution image which better represents the true sample.

2.2 sSMLM System Design

sSMLM imaging systems utilize a dispersive element, typically a prism or a grating to generate the spectrum of a fluorescent signal (30-32). For sSMLM, the most common designs utilize prisms, reflective gratings (31), or transmission gratings (109). Unlike prisms, which require the use of a beam splitter to split the incident beam into two paths (30, 32), transmission gratings directly split light at a fixed ratio without any additional optical components (109). This allows transmission gratings to be used for more compact sSMLM system designs (109). Meanwhile, reflective gratings also split light into two paths at a fixed ratio, but they require mirrors to redirect the light

towards the detector (31, 110, 111). In addition, gratings disperse light linearly which makes the spectral calibration process more straightforward than prisms which disperse light non-linearly (112). However, prisms have higher efficiency than both types of gratings allowing for better SNR of the spatial and spectroscopic signals (112). In comparison to the spatial signal captured by the PSF, the dispersed spectral signal typically requires a higher photon budget in order for useful information to be extracted (109, 112). In typical sSMLM studies, either the weighted spectral mean (or spectral centroid) (λ_{SC}) (30, 99) or the emission maximum (λ_{max}) (104) of each spectrum is calculated and used to color-code sSMLM image reconstructions. Additionally, depending on the number of spectral species which need to be detected, the spectral photons required for accurate identification of each species can be more restrictive (30, 109).

2.3 Multicolor Imaging

Multicolor sSMLM, use samples labeled with several dyes which can be excited by the same laser source and each emission event is mapped to its associated dye based on its λ_{SC} (30, 109). Thus far, four-color sSMLM has been used to image different cellular features (30). To achieve this, a dual-objective 4Pi prism-based system was used in order to maximize the photons from each single-molecule emission event (30). However, this design requires complicated alignment and is limited to thin samples which can be imaged using two objective lenses (30). For grating-based systems, it was found that tuning the level of spectral dispersion can be used to optimize the photon count in the spectral domain (109). Using this strategy, three different cellular features were imaged using sSMLM (109). Both methods require localizations with spectral photon counts below a specified level to be excluded which reduces the number of localizations which are utilized

for image reconstruction (113). This practice can have an impact on the image quality of the reconstructed images, since in some case the molecular sampling requirements for accurately reconstructing the labeled structures fail to be met (102). This is especially challenging for grating-based systems which have lower signal levels than prism-based systems (109, 112). Machine-learning (ML) has presented a promising solution to improving the utilization while maintaining the classification accuracy (113). To demonstrate this, the utilization ratio of two-color classification using a spectral photon cutoff criteria of 1000 photons was ~18% while the utilization ratio for ML was 42%, in both cases the classification accuracy was ~99% (113). Based on these results, new ML modules can be developed to classify the spectra from a wider range of dyes acquired using grating-based systems (113).

sSMLM has also enabled the identification of fluorescent impurities and the detection of variations in the emission spectra or spectral heterogeneity (SH) of dye molecules (30, 31, 109). When performing two-color sSMLM, spectrally distinct fluorescent signals not associated with the dyes used in the experiment were observed (31). These signals are believed to originate from fixation chemicals (71), residues associated with the fluorophores (95), or from intrinsic fluorescence from the cell itself (98, 99). This observation highlights the potential for the emission spectra to be used to identify and remove fluorescent impurities from image reconstructions. The removal of fluorescent impurities can have a significant impact on quantitative studies allowing for more accurate parameters for quantitative methods to be established (83, 92). We further explore this idea in Chapters 4 and 5. sSMLM can also be used to detect SH in multicolor imaging studies, by further studying this phenomenon the spectra can be used as a fingerprint for individual

dyes (31, 111). In Chapter 6, we will show how SH can be used to achieve sub-2 nm precision measurements using dyes of the same species.

Three-dimensional (3D) sSMLM has also been implemented using two strategies. The first strategy is the astigmatism method which is implemented by placing a cylindrical lens into the spatial imaging path to manipulate the PSF as the z position changes (30, 64, 114). A calibration process is used to map the axial position to the ellipticity of the PSF by imaging immobilized single-color fluorescent beads (64). This approach has allowed for four-color sSMLM imaging with 10 nm lateral localization precision and 20 nm axial localization precision (30). However, there are several challenges presented by the astigmatism approach. The first challenge is the further reduction in the photon levels and consequentially the localization precision of sSMLM due to the addition of the cylindrical lens into the spatial imaging path (115). Additionally, since the cylindrical lens is only placed in the path for the spatial domain, there are differences in the spherical aberrations experienced by the spatial and spectral paths making wavelength calibration of the sSMLM system challenging (115). Further, the astigmatism method is inherently susceptible to optical aberrations when imaging thick samples and the elliptical shape of the PSF results in inhomogeneous lateral precision along the axial direction (64, 115). The biplane 3D SMLM method uses two paths separated along two axial image planes to encode the in the z position of emission events and is robust against variations in lateral precision at different depths (79, 80, 115). By taking advantage of the spatial information encoded in the spectra, 3D biplane sSMLM imaging can be performed without further reduction to the SNR (115). This method has been demonstrated by imaging the mitochondria and microtubules in fixed cells with 20 nm lateral localization precision and 50 nm axial localization precision (115). In addition to multicolor

imaging, sSMLM also enabled multicolor single-molecule tracking. Using sSMLM, three freely diffusing fluorescent molecules were tracked in live cells with 20-40 nm localization precision (116).

2.4 Functional Imaging

The information rich spectroscopic signatures have also been used to identify variations in the local environment of fluorescent molecules. Functional SRM was most clearly demonstrated by sPAINT, where the spectrum of Nile Red (NR) dye varies in response to the polarity of its environment (104, 105, 107). Thus far, sPAINT has been used to measure the polarity in the membranes of different amphiphilic vesicles (104, 107). Using, model 1–2 Dioleoyl-sn-glycero-3-phosphocholine (DOPC), sphingomyelin (SM), and sphingomyelin mixed with cholesterol (SM+CL) vesicles, NR's spectral response to changes in polarity were identified (104, 107). The DOPC bilayer was the least ordered membrane due to the DOPC molecules which were highly polar and enable water molecules to move across their membranes. The SM+CL vesicles had the most ordered membrane due to SM+CL molecules which were highly non-polar and prevented the transport of water across the membrane (104). SM vesicles had membranes with intermediate ordering relative to DOPC and SM+CL (104). An increase in polarity resulted in a red-shift of NR's fluorescence emission spectra as observed when comparing the spectra of NR in DOPC vesicles with λ_{max} of 634 nm to NR in SM with λ_{max} of 619 nm and SM+CL vesicles with λ_{max} of 587 nm (104). In addition, sPAINT was used to investigate changes in aggregates of proteins associated with Parkinson's and Alzheimer's (α -synuclein and amyloid- β) (104, 105). sPAINT was also used to map the surface polarity of aggregates and fibrils formed by α -synuclein proteins

(105). From this study it was found that NR in association with small soluble α -synuclein aggregates had spectral centroids λ_{SC} of 612 nm, indicating low surface polarity while large insoluble fibrils had λ_{SC} of 634 nm indicating high surface polarity (105). NR has also been used to investigate the membrane polarity of the plasma membrane and other organelle membranes in living cells (107). From this study, it was found that the NR in plasma membranes was \sim 30 nm blue-shifted relative to NR in the membranes of the endoplasmic reticulum and mitochondria (107). The impact of cholesterol levels on the membrane polarity in living cells was also investigated. By depleting cholesterol levels in the plasma membrane using methyl-beta-cyclodextrin (M β CD), a strong red-shift of NR spectra was measured in the plasma membrane. In contrast, M β CD had very little impact on NR's spectra when it was associated with organelle membranes (107). After adding water-soluble cholesterol (cholesterol-M β CD), the organelle membranes underwent a significant blue-shift and the NR spectra became closer to that of NR in the plasma membrane (107). These results indicate that the cellular cholesterol level regulates differences in the polarity of the plasma and organelle membranes (107). Additionally, NR has been used to image the formation of surface adlayers due to the mixture of liquids with different polarities with \sim 30 nm localization precision (108). It is important to note that though NR serves as an excellent polarity sensor, NR binding is non-specific which can make it challenging to work with. Additionally, the high background of PAINT experiments due to freely the presence of diffusing NR can contribute to reduced spectral and localization precisions. To address these challenges, new NR derivatives have been developed to specifically target the plasma membrane allowing for new discoveries of lipid organization within the cell (117). Functional imaging has also enabled, 3D sSMLM using the astigmatism method combined with simultaneous polarization

sensing to track the trajectories of diffusing quantum rods (QR) (114). By detecting variations in the polarization and spectral signals, individual QR were identified and tracked with 8 nm lateral localization precision and 35 nm axial localization precision (114).

2.5 High-throughput spectroscopy

Previously, single-molecule spectroscopy was used to study individual fluorescent molecules (3, 41, 43). However, since single-molecule analysis is limited to sparse distributions of molecules, it was challenging to identify rare sub-populations of fluorescent molecules. sSMLM has overcome this challenge since it enables the spectroscopic analysis of thousands of single-molecules in a single experiment. For example, using sSMLM rare conformational isomers have been discovered due to differences in their spectra (118). Additionally, sSMLM has led to the discovery of multiple reaction pathways for spectrally distinct fluorescent molecules (106).

2.6 Intrinsic Fluorescence

sSMLM was also used to identify heterogeneous spectra associated with the intrinsic fluorescence from the polymer polymethyl methacrylate (PMMA) (110). Stochastic emission events were also observed in polystyrene (PS) and SU-8 (SU-8 2005) indicating that polymers used in sample preparation could generate unwanted fluorescence during SMLM (110). Similarly, studies of DNA oligonucleotides with 20 base pairs of each type of nucleotide had distinct spectral variations ranging from 520 nm to 580 nm (111). Based on the distinct spectral signatures of each type of nucleotide, a method was developed to enable sub-10 nm imaging resolution of stretched DNA fibers without the use of any dyes (111).

2.7 Summary

sSMLM combines sub-diffraction imaging capability with spectroscopy, unlocking new multiplexing, functional super-resolution imaging, and high-throughput spectroscopy capabilities. Recently, a number of advancements in 3D, live cell, and single-molecule tracking using sSMLM have been made enabling the discovery of new isomer conformations (118), the measurement of the polarity of cellular membranes (106), and the tracking of proteins in living cells (116) to name a few examples. The following chapters of this dissertation will demonstrate the development of software tools for advancing sSMLM and the development of quantitative analyses using the spectroscopic signatures from individual emission events to improve characterization of biomolecules.

CHAPTER 3: Developing Software Tools for Spectroscopic Single-Molecule Localization Microscopy Analysis

Spectroscopic single-molecule localization microscopy (sSMLM) simultaneously captures the spatial locations and full spectra of stochastically emitting fluorescent single molecules. It provides an optical platform to develop new multi-molecular and functional imaging capabilities. While several open-source software suites provide sub-diffraction localization of fluorescent molecules, software suites for comprehensive spectroscopic analysis of sSMLM data remain unavailable. RainbowSTORM is an open-source, user-friendly ImageJ/FIJI plug-in for end-to-end spectroscopic analysis and visualization for sSMLM images. RainbowSTORM allows users to calibrate, preview, and quantitatively analyze emission spectra acquired using different reported sSMLM system designs and fluorescent labels. RainbowSTORM is a java plug-in for ImageJ (<https://imagej.net>)/FIJI (<http://fiji.sc>) freely available through: <https://github.com/FOIL-NU/RainbowSTORM>. RainbowSTORM has been tested with Windows and Mac operating systems and ImageJ/FIJI version 1.52. This chapter is adapted from a preprint currently available on *BioRxiv* (119) prior to submission to *Bioinformatics*.

3.1 Introduction

Single-molecule localization microscopy (SMLM) (17, 18, 65) overcomes the optical diffraction limit by localizing stochastically emitting fluorescent molecules with high localization precision (typically 10-20 nm). Recently, spectroscopic single-molecule localization microscopy (sSMLM) (30, 31, 104), which simultaneously detects the location and full emission spectra of each emission

event was reported. Thus far, sSMLM has enabled multi-color imaging (30) and tracking (116) of as many as four different fluorescent species using a single excitation source. sSMLM has also led to new functional imaging capabilities through analyzing variations in the spectra of individual molecules. For example, sSMLM detected the polarity of the environment surrounding dye molecules (104) and enabled the discovery of previously undetected molecular conformations of dyes (106). Overall, sSMLM shows great promise to further extend existing SMLM. While a variety of software algorithms and packages are currently available for processing and analyzing traditional SMLM images (120), software tools for comprehensive spectroscopic analysis of sSMLM images remain unavailable.

Here, we present RainbowSTORM, an open-source spectroscopic analysis plug-in for ImageJ/FIJI. RainbowSTORM leverages the functionality of the existing SMLM processing tool ThunderSTORM (121) to attain spatial information while providing crucial spectroscopic tools for system calibration as well as spectral identification and classification. RainbowSTORM uses the spectral centroids (or intensity-weighted spectral means) of each localized stochastic event to define a range of spectral colors and render pseudo-colored sSMLM super-resolution images (30, 31, 104). Multicolor images can be generated by setting different user-defined spectral centroid ranges for channels with predefined colors. We provide test calibration and sSMLM images along with a testing protocol, and a detailed user guide which includes descriptions and workflows for the processes implemented in RainbowSTORM. Derivations for spectroscopic analysis (Song, et al., 2018) and flowcharts of the algorithms used in RainbowSTORM are included in the supplementary information.

3.2 Features and Methods

System calibration

RainbowSTORM calibrates sSMLM images acquired using systems, where the dispersive element (Figure 3-1a) can be either a grating or a prism. While grating-based systems are calibrated by linearly fitting pixel positions to known wavelengths (31), prism-based systems are calibrated using second-order (116) or third-order (30) polynomial fittings. Calibration in RainbowSTORM can be performed using both calibrated light sources (e.g. calibration lamps or multiple laser lines) and multicolor fluorescent beads.

sSMLM image processing

In addition to providing a flexible calibration tool, RainbowSTORM also includes a sSMLM analysis module for processing sSMLM images (Figure 3-1b). The general workflow for RainbowSTORM analysis is outlined in Figure 3-1c. RainbowSTORM first requires sSMLM images to be cropped for spatial and spectral analysis. Next two-dimensional (2D) spatial images and three-dimensional (3D) spatial images, captured using the astigmatism method (30, 64), can be processed using ThunderSTORM. Figure 3-1d shows the resulting SMLM reconstruction after spatial analysis. Next, RainbowSTORM removes background signals from the spectral images, automatically excludes emission events which spatially overlap, and previews results of spectroscopic analysis using the current processing parameters. Finally, RainbowSTORM identifies the full spectra and calculates the spectroscopic fields for all localizations.

Visualization and post-processing

After processing the spectral images, pseudo-colored super-resolution reconstructions (Figure 3-1e) are rendered using the spectral centroids and spatial coordinates of each localization. For 3D sSMLM images, a stack of pseudo-colored super-resolution reconstructions can be rendered, where images in the stack are separated by the axial position of each localization. The histograms of the calculated spatial and spectral fields for the processed localizations can be displayed and used to select subsets of the data for independent visualization. Additionally, localizations with large point spread function and spectrum widths as well as localizations with low photon counts and precisions in the spatial and spectral domains can be excluded. RainbowSTORM can also apply ThunderSTORM drift-correction files and assess the image quality of sSMLM images using Fourier Ring Correlation (FRC) analysis (102). In addition, the spectral centroid information can be assigned to multiple channels to create multicolor super-resolution images using the classification module. For example, Figures 3-1d-h shows images of the mitochondria, microtubules, and peroxisomes of COS-7 cells respectively labeled by Alexa Fluor 647 (AF647), CF660, and CF680. Figure 3-1i shows the overlay of the three images from the selected spectral centroid windows. After post-processing, sSMLM results can be saved. Previously saved sSMLM results can be loaded using the sSMLM import module for further analysis.

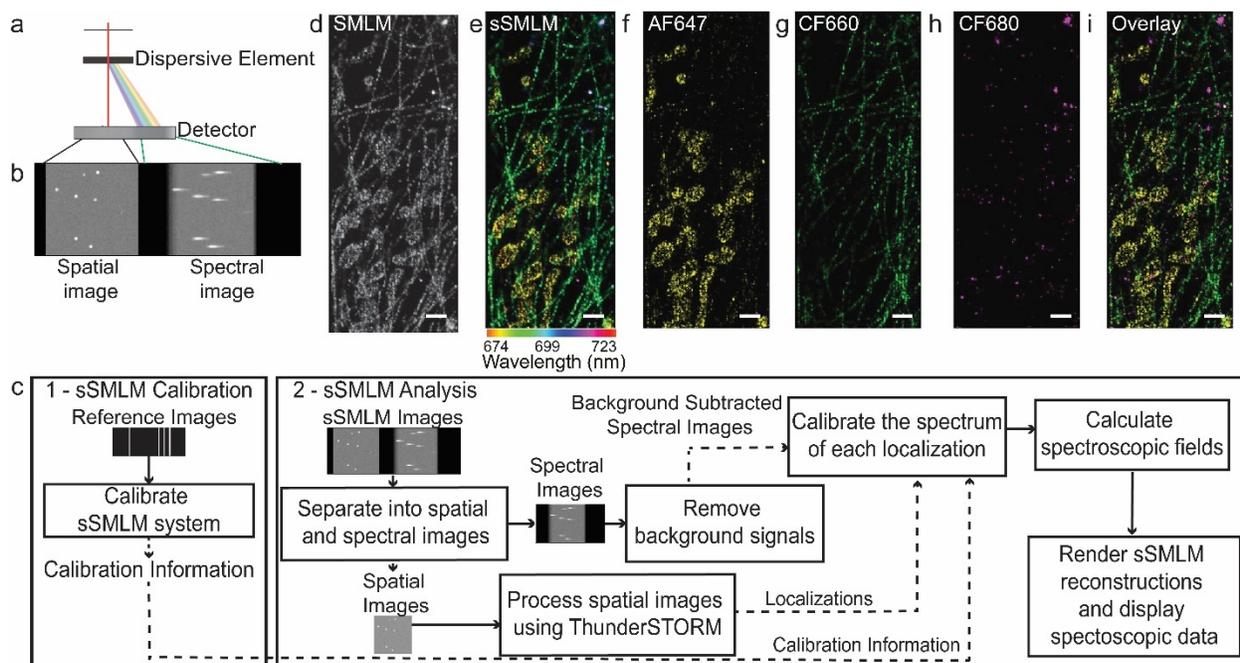


Figure 3-1: (a) General sSMLM system schematic. (b) sSMLM images with the spatial and spectral images simultaneously captured on different parts of a detector. (c) RainbowSTORM workflow showing how the system calibration module interacts with the analysis module (d) SMLM reconstruction (e) Pseudo-colored sSMLM reconstruction. Images of the three separate channels showing (f) mitochondria labeled with AF647, (g) microtubules labeled with CF660, (h) peroxisomes labeled with CF680, and (i) the overlay image of the three channels.

3.3 Algorithms and Equations

The algorithms and equations used for system calibration and spectral analysis of sSMLM images are detailed in the following sections.

System Calibration

The peak pixel positions of known wavelengths from a calibration light source or from multi-color emitter (e.g. Tetraspeck Microsphere, Thermofisher) are selected and fit using the least-squares method. Grating-based systems are calibrated by fitting the pixel positions to a straight line ($y = a + bx$). Meanwhile, prism-based systems are calibrated by fitting the pixel positions and

corresponding wavelength information using either a second-order polynomial ($y = a + bx + cx^2$) or a third-order polynomial ($y = a + bx + cx^2 + dx^3$). The coefficients a , b , c and d represent the offset and the coefficients of the first, second, and third orders. The resulting coefficients are used to calibrate the fluorescence emission spectra from single-molecule dyes or fluorescent proteins captured during sSMLM experiments.

Global background estimation

An averaging filter is used to estimate the background of the spectral images. Pixel values which are associated with the sample are identified by comparing each pixel value to a threshold. The threshold for this filter can be either user-defined or automatically generated. Figure 3-2 shows the flowchart of the algorithm used to generate background image (B) from the stack of input images (I). The background image is then subtracted from each frame of the input images to generate a stack of background-subtracted spectral images.

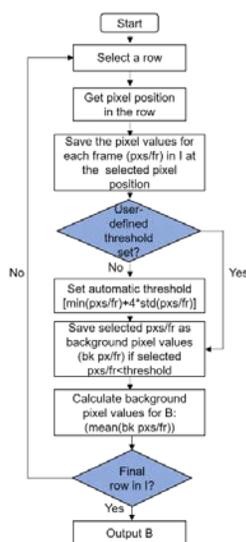


Figure 3-2: Flowchart of the algorithm to generate the background image

Overlapping spectra removal

Localizations with spectral images which overlap in space can be excluded from analysis by enabling the ‘Remove overlapping spectra’ checkbox on the sSMLM analysis module. The localization pixel shift (lps) and the spectral dispersion (sd) are used to remove the overlapping spectra. The lps controls the number of pixels used to capture the point spread function (PSF) of each localization. By default, lps is set to ± 1 for a total of three pixels to capture the PSF. The sd captures the number of pixels which cover the spectral image (spectra pixels) and is set based on the sSMLM system. Increasing either the lps or the sd will result in more localizations being removed as overlapping. The thresholds along the x axis is twice the number pixels to capture the spectra and the y axis is twice the number of pixels to capture the PSF. Figure 3-3 shows the flowchart of the algorithm used to determine whether the spectra are overlapping.

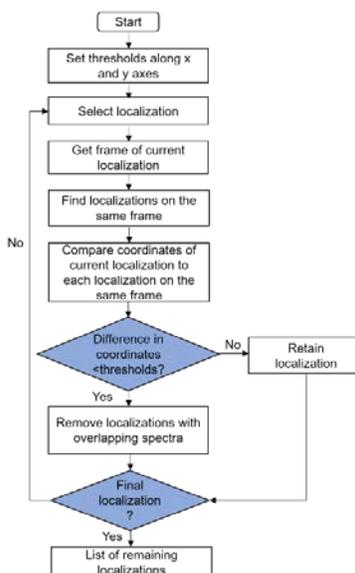


Figure 3-3: Flowchart of the algorithm to remove overlapping spectra

Spectral centroid calculation

The intensity-weighted spectral mean or the spectral centroid (λ_{SC}) of the line spectrum for each localization is calculated by the equation below

$$\lambda_{SC} = \frac{\sum_{\lambda=\lambda_1}^{\lambda=\lambda_n} I(\lambda) * \lambda}{\sum_{\lambda=\lambda_1}^{\lambda=\lambda_n} \lambda}, \quad (3-1)$$

where λ is the wavelength between (λ_l and λ_n) and I is the photon count at that wavelength.

Spectrum width estimation

The line spectrum for each blinking event is fit using the least-squares method to a Gaussian function:

$$y = a * e^{-\frac{(x-b)^2}{2c^2}}, \quad (3-2)$$

The coefficients a , b , and c represent the peak intensity, mean value (μ) and standard deviation (σ) respectively.

Spectral precision assessment

The spectral precision (σ_λ) of each spectrum is estimated as reported by Song et al (Song, et al., 2018) (122):

$$\sigma_\lambda = \sqrt{n_{bg}^2 + n_s^2 + n_{ro}^2 + \sigma_{sse}^2}, \quad (3-3)$$

where n_s , n_{bg} , n_{ro} , and σ_{sse} are the shot noise, background noise, readout noise, and uncertainty of the spectral-shift error respectively.

The background noise of the signal is:

$$n_{bg} = \sqrt{\frac{2\left(\frac{1024B_t\eta\Delta\lambda W_p}{64a\Delta\lambda} * a^3 * \sigma_{PSF}\right)}{3\Delta\lambda W_p(\eta I)^2}}, \quad (3-4)$$

where B_t is the total background of the spectral signal, η is the quantum efficiency, $\Delta\lambda$ is the spectral dispersion, W_p is the pixel size, $a = \sqrt{\sigma_{SPE}^2 + \frac{\Delta\lambda^2}{12}}$ where σ_{SPE} is the standard deviation of the fitted spectrum, σ_{PSF} is the standard deviation of the fitted PSF and I is the total photon count of the spectrum.

The shot noise of the signal is calculated by:

$$n_s = \sqrt{\frac{2\left(\sigma_{SPE}^2 + \frac{\Delta\lambda^2}{12}\right)}{\eta I}}, \quad (3-5)$$

The readout noise is calculated by:

$$n_{ro} = \sqrt{\frac{1024\sigma_{SPE}^3 \Delta\lambda^2 N_r^2}{3\Delta\lambda W_p(\eta I)^2}}, \quad (3-6)$$

where N_r is the readout noise per pixel.

The uncertainty of the spectral-shift error limits the spectral precision which can be achieved and can be quantitatively expressed as:

$$\sigma_{sse} = \sqrt{\frac{\Delta\lambda^2}{12}}, \quad (3-7)$$

Figure 3-4 shows the comparison of the expected spectral precision and the spectral precision estimated with RainbowSTORM using simulated sSMLM images of single-molecule emission events with increasing photon counts. For the estimation, we used $\Delta\lambda$ of 5.9 nm/pixel, σ_{SPE} of 22 nm, σ_{PSF} of 117 nm, N_r of 1 e-, W_p of 16 μm , and η of 1.

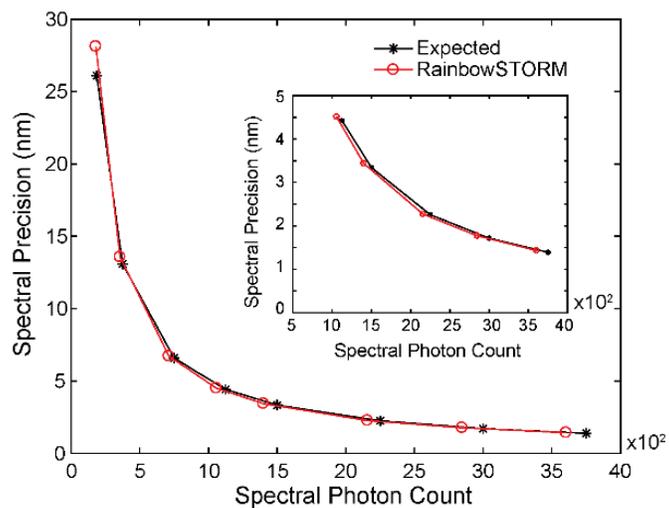


Figure 3-4: Comparison of the expected spectral precision and the average spectral precision estimated by RainbowSTORM at different spectral photon counts [180 to 3800]. Inset shows the zoomed in comparison for the spectral photon count from 1000 to 3800.

Three-dimensional sSMLM images

Three-dimensional sSMLM images can be acquired using the astigmatism method (64) by placing a cylindrical lens in the spatial imaging path (30) as shown in Figure 3-5.

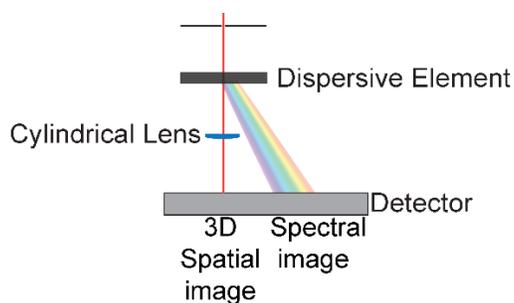


Figure 3-5: General sSMLM system schematic for 3D imaging using the astigmatism method.

3.5 Methods

sSMLM Optical Setup

The sSMLM system is setup as previously described (109) and briefly described as follows. A 642-nm CW laser was focused on the back focal plane of a Nikon Ti microscope body and illuminated through a 100× Total Internal Reflection Fluorescence (TIRF) objective lens and a numerical aperture (NA) of 1.49 (CFI Apochromat, Nikon). Fluorescence from the sample is collected by the objective lens and passed through a tube lens then directed to an entrance slit using a mirror. The entrance slit is used to confine the field of view (FOV) of the spatial image to one section of the electron multiplying charge-coupled device (EMCCD). The photons then pass through a transmission grating (100 grooves/mm, STAR100 Paton Hawksley Education Ltd) which splits the photons into spatial and spectral images with a ratio of ~1:3. The remaining photons are then passed through an imaging lens and then focused onto the EMCCD (ProEM, Princeton Instrument) by a second imaging lens.

Sample Preparation

COS-7 cells (ATCC) were grown in Dulbecco's Modified Eagle Media (Gibco/Life Technologies) supplemented with 2-mM L-glutamine (Gibco/Life Technologies), 10% fetal bovine serum (Gibco/Life Technologies), and 1% penicillin/streptomycin (10,000 U mL⁻¹, Gibco/Life Technologies) at 37°C with 5% CO₂. The cells were plated on No. 1 borosilicate bottom eight-well Lab-Tek Chambered coverglass with 30%-50% confluency. The cells were fixed after 48 h in pre-warmed 3% paraformaldehyde and 0.1% glutaraldehyde in phosphate-buffered saline (PBS) for 10 min, washed with PBS twice and quenched with freshly prepared 0.1% sodium borohydride in PBS for 7 mins. The cells were then rinsed three times in PBS at 25°C. The fixed cells were

then permeabilized with blocking buffer (3% bovine serum albumin (BSA), 0.5% Triton X-100 in PBS) for 20 min and then incubated with primary antibodies (sheep anti-tubulin, 10 ug mL^{-1} , rabbit anti-PMP70 and mouse anti-TOM-20, 5 ug mL^{-1}) in blocking buffer for 1 hr at room temperature and rinsed with washing buffer (0.2% BSA, 0.1% Triton X-100 in PBS) three times. The cells were then incubated in donkey secondary antibody conjugates (anti-sheep AF647, anti-mouse CF660C and anti-rabbit CF680) in blocking buffer for 40 min and then washed three times with PBS and stored at 4°C . The secondary antibodies were prepared as previously described and briefly described here. 0.5 uL of 5mM dye (NHS-ester functionalized AF647 (Thermofisher), CF660C and CF680 (Biotinum)) were combined at 25°C with 100 uL of 1 mg mL^{-1} IgG/IgY antibodies in PBS and sodium bicarbonate (10 uL of 1 M). The mixture was incubated overnight then purified by a Nap-5 size exclusion column. The concentrated sample was then extracted using an Amicon Ultra-0.5 Centrifugal Filter unit to give 1-2 dyes per antibody. Finally, the absorption and emission spectra of the dyes were tested using a NanoDrop Spectrophotometer and then stored at 4°C .

System Calibration

Using a calibration light source (Neon lamp, 6032 Newport) reference images were captured by using a narrow slit. The images in the spatial domain was a narrow straight line while the spectral images showed multiple spectral lines representing known emission maxima related (640.23 nm, 703.24 nm, 724.52 nm, and 743.89 nm) to the calibration light source.

Experimental sSMLM Data

Imaging buffer (50 mM Tris (pH=8.0), 10 mM NaCl, 0.5 mg mL⁻¹ glucose oxidase (Sigma, G2133), 2000 Uq mL⁻¹ catalase (Sigma, C30), 10% (w/v) D-glucose, and 100 mM cysteamine) was added to the COS-7 cells and the cells were imaged using the sSMLM system described above. The exposure time was set to 10 ms, pixel size of the system was 160 nm, the analog to digital unit (ADU) was 4.6 e-/analog to digital count, and the electron multiplying gain (EM Gain) was 100. For this experiment 30000 frames were recorded.

Simulated sSMLM Data

We simulated 500 frames of sSMLM images with a back-projected pixel size of 160 nm, camera base level of 200 digital counts, EM Gain of 100, and ADU of 4.6 e-/analog to digital count. For spectral precision testing, the simulated emitter was AF647 with total photon counts of 250, 500, 1000, 1500, 2000, 3000, 4000 and 5000. The photons from the simulated emitters were split into spatial and spectral images using a 1:3 ratio. The spatial image of each single-molecule event was modeled as a 2D gaussian with a mean of 0 and std of 0.73 pixels. A reference line spectrum for AF647 was convolved with the point spread function image to generate the spectral image. The shifted spectral image was set based on a calibration lamp with known emission maxima (485.5 nm, 546.5 nm, 611.6 nm, and 707 nm). The locations of each localization event and its associated spectrum were randomly assigned. For multicolor simulated data, reference spectra of CF660 and CF680 were used to generate spectral images.

3.4 Conclusions

RainbowSTORM provides a plug-in for performing spectroscopic analysis of 2D and 3D sSMLM images acquired using both grating-based and prism-based sSMLM implementations. RainbowSTORM fills the need for a spectroscopic analysis platform and provides spectral classification methods, spectral and spatial filtering methods, pseudo-colored visualization of sSMLM datasets, and built-in FRC analysis. Future updates will make RainbowSTORM compatible with a wider range of spatial analysis platforms. Additional spectroscopic analysis methods such as spectral unmixing (123), machine-learning based spectral classification (113), and cluster analysis (104) will be added to RainbowSTORM.

CHAPTER 4: Quantitative sSMLM #1: Removing Artifacts from SMLM Image Reconstructions

The existence of fluorescent impurities has been a long-standing obstacle in single-molecule imaging, which results in sample misidentification and higher localization uncertainty. Spectroscopic single-molecule localization microscopy (sSMLM) can record the full fluorescent spectrum of every stochastic single-molecule emission. This capability allows us to quantify the spatial and spectral characteristics of fluorescent impurities introduced by sample preparation steps, based on which we developed a method to effectively separate fluorescent impurities from target molecules. This chapter is adapted from a previously published article in the *Journal of Biomedical Optics* (123).

4.1 Introduction

The term fluorescent impurity usually refers to unintended fluorescence emission from unknown molecules or chemical complexes. The presence of fluorescent impurities represents a long-standing issue in single-molecule imaging and spectroscopy (3, 96, 97). To reduce the impact of these fluorescent impurities, stringent cleaning and sample preparation techniques need to be utilized (3, 96, 97). In recent years new imaging techniques, such as single-molecule localization microscopy (SMLM) (17, 18, 43, 63, 68, 69), emerged to offer super-resolution single-molecule imaging far beyond the diffraction limit of the light. However, the impact of fluorescent impurities on correctly interpreting single-molecule imaging results has not been thoroughly investigated (95, 124-126). In conventional fluorescence microscopy, fluorescent impurities are often negligible due

to their apparent lower absorption cross-sections and weak fluorescent emissions (98, 99, 127).

However, growing evidence has shown that fluorescent impurities significantly impacts SMLM by inducing imaging artifacts, which include sample misidentification and higher localization uncertainty in cases where fluorescent impurities overlap in space with target molecules (95, 125, 126). While SMLM accumulates the stochastic emissions from individual fluorophores and proteins to collectively render super-resolution images (17, 18, 43, 63, 69), the required high-power-density illumination to excite stochastic emissions also unfavorably intensifies emissions from fluorescent impurities (31, 110, 126). When a large number of photons are stochastically emitted from fluorescent impurities, they behave similarly to target molecules and are difficult to distinguish and remove (31, 95, 126). Preventing sample misidentification is a particularly significant challenge when imaging low number density ($<1 \mu\text{m}^{-2}$) single-molecules without distinct structural or morphological features (124, 125).

Currently, the reported methods to identify target molecules in reconstructed SMLM image mainly rely on spatial and temporal profiling of their stochastic emissions, such as width of the fitted point spread function (PSF) (128), repetition rate of blinking events (21), and emission intensity (126, 128). Emission intensity in particular is commonly compared against a user-defined intensity threshold and one can remove any emission with lower intensity than the threshold, hoping to exclude fluorescent impurities (125, 126, 128). However, due to their diverse origins, emissions from fluorescent impurities can often exceed the threshold value, resulting in low specificity (125, 126). A more specific criterion is needed to faithfully identify target molecules while rejecting fluorescent impurities. The spectra of all stochastic emissions can be such signatures; however, existing SMLM technologies are unable to measure these spectra. Recently

we and other groups reported spectroscopic single-molecule localization microscope (sSMLM) (30, 31, 110), which simultaneously detects the spatial and spectral information of each stochastic fluorescent emission event. Hence, we anticipate that sSMLM, by analyzing emission spectrum of every stochastic emission, will provide a highly specific criterion to identify target molecules and to reject fluorescent impurities. In this study, we seek to answer two questions: [1] is it possible to reduce or ultimately eliminate fluorescent impurities and [2] can we utilize the emission spectra to remove fluorescent impurities from all the detected stochastic emissions in a low number density sample.

4.2 Methods and Materials

Coverslip cleaning

Fisherbrand™ 22×22 mm #1.5 borosilicate coverslips (Fisher Scientific) and precleaned FisherFinest™ 22×22 mm #1 borosilicate coverslips (Fisher Scientific) were imaged using a 532-nm laser at four typical power densities (1.5-5.8 kW/cm²) used in SMLM. Before imaging, the coverslips were air blown to remove any large particles. Additional cleaning processes were performed on Fisherbrand™ coverslips as described below.

Piranha solution

A beaker was cleaned and placed in a fume hood. Sulfuric acid (H₂SO₄) (Sigma Aldrich) was added to hydrogen peroxide(H₂O₂) (Sigma Aldrich) at a ratio of 3:1 (90 mL to 30 mL) (129). The coverslips were submerged in the solution for 20 mins. The coverslips were then submerged in

distilled nuclease-free water (Ambion, ThermoFisher) and then dried by air blowing. The piranha solution was allowed to cool disposal in an appropriate waste container.

Potassium hydroxide (KOH) and ultra-violet (UV) light sterilization

The coverslips were sonicated in 1 M KOH (Sigma Aldrich) for 15 mins (18). The coverslips were then rinsed in Milli-Q water and dried using nitrogen (N₂) gas. The cleaned coverslips were placed in a petri dish and sterilized using UV light for 30 mins (18).

Hydrochloric acid (HCl) and prop-2-anol cleaning

Each coverslip was sequentially submerged for 30 secs in 36 % HCl (Sigma Aldrich), Milli-Q water, and then prop-2-anol (Sigma Aldrich) before drying with nitrogen (N₂) gas (130).

UV and ozone cleaning

Coverslips were placed in the ZoneSEM Cleaner (131) (Hitachi) and exposed to ozone activated by UV light for 2 mins per side.

Plasma cleaning

The operating conditions for the plasma cleaner (PC 2000, South Bay Technology) for a mixture of argon and oxygen gas was set to use a forward power of 20 W and a minimized reflection power. A cleaning time of 2 mins was selected and a precleaning step was performed to clean the chamber. The coverslips were placed in glass petri dishes and plasma cleaned uncovered for 2 mins (132, 133). Metal tweezers used for handling the coverslips were plasma cleaned during this

cycle. Using the cleaned tweezers, the coverslips were turned over and the exposed surface was cleaned using the same settings. Cleaned coverslips were stored in sealed glass petri dishes.

Coverslip functionalization

Plasma cleaned coverslips were functionalized via poly-L-lysine (PLL), (7-octen-1yl) trimethoxysilane (silane), and biotinylated bovine-serum albumin (BSA) and neutravidin (NEU).

Poly-L-Lysine

Coverslips were incubated in 1 ppm PLL (134) (Sigma Life Science) solution for 2 mins. The surface was then rinsed three times using nuclease free water (Ambion, ThermoFisher) before air blowing.

Silanization

A 250 mL Pyrex crystallizing dish was tripled rinsed using methanol (Sigma Aldrich) and then n-heptane (Sigma Aldrich). Working in a chemical hood, 100 mL of n-heptane was added to the dish and 100 μ L of silane (129, 130) (Sigma Aldrich). Coverslips were added to the silane treatment using tweezers and left overnight in a desiccator without a vacuum. The next day, the coverslips were sequentially sonicated for 5 mins in n-heptane, Milli-Q water, and finally chloroform (Sigma Aldrich) before drying using air.

BSA-biotin-neutravidin

Coverslips were rinsed 3 times with 500 μ L phosphate-buffered saline (PBS) (Gibco, Life

Technologies). The coverslips were then incubated for 5 mins in 200 μ L of 0.5 mg/mL BSA-biotin (11) (Sigma Aldrich) in PBS. The BSA-biotin solution was removed, and the coverslip was triple rinsed in 500 μ L PBS then incubated for 5 mins in 200 μ L of 0.5 mg/mL NEU (11) (Invitrogen, ThermoFisher) in PBS. The coverslips were then triple rinsed in 500 μ L immobilization buffer (PBS supplemented with 10 mM of magnesium chloride (MgCl_2) (Ambion, ThermoFisher). During imaging water was used to prevent the treatment from drying. A second surface with glucose-oxidase imaging buffer was also tested.

Immobilization buffer and oxygen scavenger system

Immobilization buffer containing 10 mM MgCl_2 in PBS (pH 7.4) was freshly prepared and added to the BSA-biotin-NEU sample. The immobilization buffer was supplemented with an oxygen scavenging system containing 0.5 mg/mL glucose oxidase (Sigma Aldrich), 40 μ g/mL catalase (Sigma Aldrich) and 10 % (w/v) glucose (Sigma Aldrich) and 143 mM 2-mercapethanol (Sigma-Aldrich).

Reagent purity

Purity information for the chemical reagents and proteins used in this study is detailed in Table s4-1 and 4-2 respectively.

Table 4-1: Summary of chemical reagents used in this study

Chemical	Supplier, Product Number	Purity	Notes
Ethyl alcohol- 200 proof	Sigma Aldrich, 459844	≥99.5 %	ACS reagent
2-Propanol	Sigma Aldrich, 650447	99.9 %	HPLC Plus
Potassium Hydroxide Pellets	Sigma Aldrich, 306568	99.99 %	Semiconductor grade
Hydrogen Peroxide Solution	Sigma Aldrich, 316989	99.999 %	Semiconductor grade
Sulfuric Acid	Sigma Aldrich, 258105	95 % - 98 %	ACS Reagent
α-D-Glucose, anhydrous	Sigma Aldrich, 158968	96 %	
2-Mercaptoethanol	Sigma Aldrich, 63689	≥ 99.0 %	BioUltra
Trimethoxy(7-octen-1-yl) silane	Sigma Aldrich, 452815	80 %	Technical grade
n-Heptane, anhydrous	Sigma Aldrich, 246654	99 %	
Chloroform	Sigma Aldrich, 650498	≥ 99.9 %	HPLC-Plus
Hydrochloric Acid	Sigma Aldrich, 339253	99.9%	

Table 4-2: Summary of proteins used in this study

Protein	Supplier, Product Number	Purity	Notes
Glucose oxidase aspergillus niger	Sigma Aldrich, G2133	≥60 % Protein	
Poly-L-Lysine	Sigma Aldrich, P4707	Lysine concentration ≥0.45mmol	Sterile filtered
Neutravidin, lyophilized powder	Thermo Scientific, 31000	14 ug/mg active protein	Salt Free
Albumin, Biotin labeled bovine, lyophilized powder	Sigma Aldrich, A8549	80% protein	
Catalase	Sigma Aldrich, C40	≥10,000 units/mg Protein	≤0.2 wt. % Thymol

SMLM and sSMLM experimental setup

In these experiments, a diode-pumped solid state 532 nm laser with a maximum output power of 300 mW was used to illuminate the sample. The laser output was filtered (LL01-532-12.5, Semrock) and passed through a half-wave plate and a linear polarizer to control the output power. The laser was then coupled to an inverted microscope body using a telescopic system and dichroic mirror to focus the light on the back focal plane of a Nikon CFI apochromat total internal reflection objective lens (100 \times , 1.49 numerical aperture) shown in Figure 4-1a. Adjusting the position of the beam path to the edge of the objective allowed for illumination at the critical angle at the water-coverslip interface, thus limiting the volume of material illuminated. A long-pass filter (BLP01-532R-25, Semrock) was used to reflect the 532 nm laser. SMLM was performed using only position data collected using an EMCCD (iXon 512B, Andor) as shown in Figure 4-1b. For sSMLM, light was guided through a home-made spectrometer equipped with a 100 lines/mm blazed transmission grating (STAR100, Panton Hawksley Education Ltd.), which separated the spatial and spectrally dispersed images. The spatial and the spectral information for each emission event was collected simultaneously on different regions of an EMCCD (ProEm HS 512X3, Princeton Instruments) as shown in Figure 4-1c.

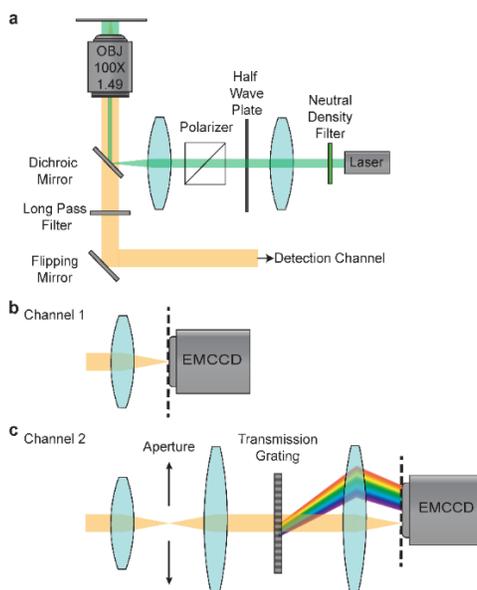


Figure 4-1. Schematics of SMLM and sSMLM systems. (a) Schematic of Excitation optics and instrumentation. (b) SMLM detection channel. This setup was used to capture images of cleaned and functionalized surfaces. (c) sSMLM detection channel. This setup was used to capture images and spectra simultaneously allowing for spectroscopic analysis of single molecules.

Optical power density measurements

We used a power meter (Newport 1918-R) with a high-power detector (Newport, 918D-SL-OD2R) to measure the power of the excitation laser after beam expansion and before entering the microscope. In comparing with the power measured right after the objective lens, we found a 76% transmission within the microscope body. For all experiments, the power was measured before entering the microscope and scaled by the transmission loss. Power density measurements of 1.5 kWcm^{-2} , 3.0 kWcm^{-2} , 4.4 kWcm^{-2} , and 5.8 kWcm^{-2} at the sample plane were calculated from power measurements at the microscope base (25 mW, 50 mW, 75 mW, and 100 mW) and an illumination radius of $20 \text{ }\mu\text{m}$. The power level was adjusted by changing the angle of the linear polarizer. To calibrate this process, corresponding angles for each power level was recorded and used for all experiments.

Imaging procedure for quantitatively assessing the origin of fluorescent impurities

One coverslip from each treatment was imaged under 532 nm illumination. We randomly selected 5 positions on the coverslip and recorded 1000 frames using an integration time of 10 ms. While imaging cleaned surfaces a 200×200-pixel FOV was used and a 256×256-pixel FOV was used for imaging functionalized surfaces. For comparison, the number of fluorescent impurities were normalized by the area of their respective FOVs.

To investigate the impact of excitation power density on the number of detectable fluorescent impurities, Fisherbrand™ (Fisher Scientific) and Fisherfinest™ (Fisher Scientific) coverslips were imaged at four different power density levels (1.5 – 5.8 kWcm⁻²). For each dataset, a maximum intensity projection (MIP) image was generated and the number of fluorescent impurities per FOV was determined using the ImageJ plug-in ThunderSTORM. There was an average of 2.0×10⁷ cm⁻² fluorescent impurities from Fisherbrand and 1.7×10⁷ cm⁻² fluorescent impurities from Fisherfinest™ coverslips before cleaning.

Since the tested power densities did not have a further impact on the number of fluorescent impurities, we used a typical SMLM power density of 3 kWcm⁻² in our investigations. Spectroscopic information from the surfaces was collected by randomly selecting multiple FOVs on a Fisherbrand™ coverslip before cleaning and a plasma cleaned coverslip functionalized with PLL. Each FOV was imaged until photobleaching occurred. We captured 1000 frames from the unprocessed coverslip and 3000 frames from the PLL coverslip under 532 nm at 3 kWcm⁻² with 20-ms integration time per frame.

Spectral fitting method

We used a nonlinear least-square fitting method to fit each recorded spectrum to a reference spectrum. Since the recorded emission events overlapped in space, the mixed spectrum S attributed to each point spread function can be expressed as

$$S = a_1 s_1(x_1 + d_1) + a_2 s_2(x_2 + d_2) + w, \quad (4-1)$$

where $s_i(x)$ is the emission spectrum for each type of molecule at position x ; a_i is the emission intensity of the molecule; d_i is the spectral shift due to conformation heterogeneity of each dye molecule; and w is the error term accounting for additive noise (98). Using this equation, parameters for the recorded intensity, spectral heterogeneity, and noise were used to fit experimentally recorded spectra to reference spectra of the dye being studied. The adjusted coefficient of determination (R^2) was calculated as

$$R_{\text{adj}}^2 = 1 - \left(\frac{n-1}{n-p} \right) \frac{\text{SSE}}{\text{SST}}, \quad (4-2)$$

where SSE is the sum of the squared residuals ($\text{SSE} = \sum_{i=1}^n (y_i - f(x_i))^2$), SST is the total sum of squares ($\text{SST} = \sum_{i=1}^n (y_i - \bar{y})^2$), n is the number of observations, and p is the number of regression coefficients, respectively. The adjusted R^2 was used to assess the goodness of fit of the recorded spectrum to the reference.

Establishing the ground truth within each FOV

We selected 10-nm DNA origami nanorulers (Gattaquant) labeled with Alexa Fluor 532 (AF532) and Alexa Fluor 568 (AF568) to test whether the spectrum could be used to separate target molecules from fluorescent impurities. The nanorulers were the ideal model system for this study since their spacing was unable to be resolved by the 20-nm spatial resolution of SMLM but their

spectral separation was greater than the 3-nm spectral dispersion of our sSMLM. Though the peaks of emission spectra of the dyes used were well separated, both dyes can be directly excited by 532 nm laser. The combined signal from a single resolvable pixel provided a unique spectral signature, which could be used to establish a faithful ground truth for the sample in the presence of fluorescent impurities under low power density (LPD) excitation of 0.5 kWcm^{-2} . We then tested using spectral fitting and intensity thresholding to categorize recorded emission events using high power density (HPD) excitation of 3 kWcm^{-2} .

We observed steady fluorescence emission with rather small temporal fluctuations from all fluorescent point emitters in the LPD condition, we used the average of the 300 frames to extract the spectra with high signal-to-noise ratio (SNR). The approximate location of the immobilized nanorulers in the sample were estimated using the average image of each FOV. Overlapping spectra in the average LPD images were removed from the LPD and HPD datasets.

Consequently, a total of 15 emitters were excluded from further analysis. Due to their high absorption cross-section and quantum yield compared with the fluorescent impurities, we anticipate that the observed fluorescent emissions mainly originated from nanorulers. The minority of fluorescent impurities excited were removed using the spectral fitting method. The extracted spectra were first normalized using the emission maximum of the record spectra then fit to the reference spectra. We attributed fluctuations in the position of the spectra to conformation heterogeneity of each dye and the influence of noise was ignored in this case. From the reference sample for both dyes, we found that full width half maximum (FWHM) of that emission centroids of AF532 was 20 nm and AF568 was 40 nm. We also observed spectral shift parameters of ± 10

and ± 20 for AF532 and AF568, respectively. Since 532 nm laser illumination could directly excite 100 % of AF532 and 42 % of AF568, each dye had to exceed the noise floor.

Therefore, the background should not exceed 10% of the peak intensities for both dyes. Because AF532 could be optimally excited using 532 nm laser illumination, the influence of AF568 was determined by first fitting all 174 points using only the reference spectra of AF532. The data was then fit using both spectra and the difference in the peak adjusted R^2 value was used to select a threshold of 0.89. Single molecules excited under LPD, which had an adjusted R^2 value of 0.89 after spectral fitting were considered to be true nanorulers. The determined spatial and spectral characteristics of the nanorulers established the ground truth for each FOV.

Preparation of nanoruler sample

Nanorulers (Gattaquant) DNA origami samples were prepared by adding 1 μL of the nanorulers to 200 μL nuclease free water (Ambion, ThermoFisher). The 10 μL of the nanoruler solution was deposited on a PLL coated surface via spin deposition (Laurell WS-650- 23) at 1200 rpm for 30 secs.

Imaging procedure for nanoruler samples

One coverslip containing immobilized nanorulers was imaged under 532-nm illumination. 9 positions on the coverslip were randomly selected and each FOV was imaged using the following procedure. The nanoruler sample was imaged for 4 secs (300 frames) at LPD (0.5 kWcm^{-2}). The observed fluorescence from the dye molecules was stable and non-blinking at this power density level. The power density was then increased by changing the polarizer position to

reach a HPD (3 kWcm^{-2}) to allow stochastic fluorescence emission of the dye molecules. Images were recorded for 30 secs (1500 frames). An integration time of 20 ms was used to record each FOV. This data was used for sample classification as detailed in the algorithm in Figure 4-2. The LPD frames were averaged, and the location and spectra used as references for the single molecule quantification experiments. The HPD frames were used to compare the performance of filters based on emission intensity and spectral fitting.

Sensitivity and specificity calculation

We tested the performance of filtering emission events using the emission intensity thresholding and our spectral fitting method. The sensitivity of each method to correctly identify emission events from nanorulers and the specificity of each method to correctly remove emission events from fluorescent impurities was determined by identifying true positives, false positives, true negatives and false negatives. Nanorulers, which were correctly included by the filtering method were marked as true positives, while any nanorulers which were excluded were marked as false negatives. True negatives were any fluorescent impurities which were correctly excluded by the filtering method while false positives were any fluorescent impurities incorrectly marked as nanorulers. These definitions were used to calculate the sensitivity and specificity of each filtering method using the following equations:

$$\begin{cases} \text{Sensitivity} = \text{True Positives} / (\text{True Positives} + \text{False Negatives}) \\ \text{Specificity} = \text{True Negatives} / (\text{True Negatives} + \text{False Positives}) \end{cases} \quad (4-3)$$

SMLM ground truth

To determine the locations and the number of true nanorulers and fluorescent impurities in each FOV under HPD excitation, incorrect localizations due to background noise were removed from 27396 recorded points from 9 FOVs using a simple density filter. To do this, the nearest neighbors within a 200-nm radius of a localization were identified. For clusters with more than 5 neighbors, the centroid was found and localizations within a 200-nm radius were assigned to that cluster. The average of the localizations was used to estimate the location of the detected emitter. The estimated locations were classified as nanorulers or fluorescent impurities by comparing the results to the ground truth established using the locations and spectra from the averaged image of the same FOV under LPD excitation. On average, we observed 6 ± 2 nanorulers and 35 ± 7 fluorescent impurities among all 9 FOVs being measured.

Threshold selection

For both emission intensity thresholding and spectral fitting, the generated histogram from 27396 emission events were used to select a range of possible thresholds. For intensity thresholding the background intensity range (120:400) was selected from the histogram of emission intensities to ensure an SNR of at least 6 dB. For spectral fitting the range (0.8:0.94) was selected from the histogram of adjusted R^2 values. This range was selected since it fell between two-peak adjusted R^2 values. Examples using an intensity threshold of 180 and a spectral fitting threshold of 0.84 were compared due to their similar high sensitivities (~90%).

Filtering SMLM data

For the intensity thresholding method, emission events with an average intensity greater than 180 were classified as fluorescence from nanorulers and all other events were classified as fluorescent impurities. For spectral fitting, the spectrum was first normalized using the maximum intensity of the signal. The accepted spectral shift parameter was ± 10 nm for AF532 and ± 20 nm for AF568. The spectrum from each emission event in the SMLM dataset was fit to the reference and the adjusted R^2 value determined. Emission events with an adjusted R^2 value greater than 0.84 were classified as fluorescence from nanorulers and all other events were classified as fluorescent impurities.

The localizations identified as emission events from nanorulers were then used to reconstruct SMLM images. For an emitter to be reconstructed, more than 5 emission events within a 200-nm radius of the centroid was required. The location of the emitters after each filtering method were compared to the known location of the nanorulers using the established ground truth. The sensitivity and specificity of each method was then calculated and compared. To estimate the size of each cluster, the standard deviation of emission events within each cluster was used (135).

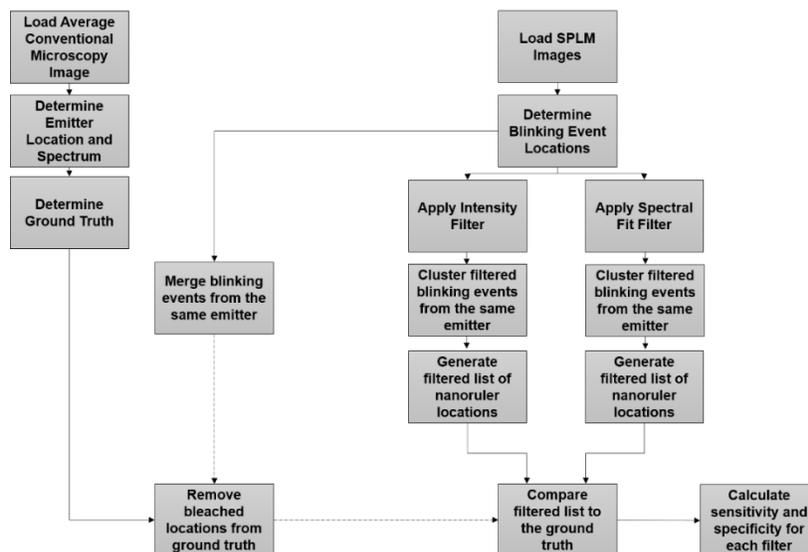


Figure 4-2: Flowchart of the algorithm used to compare intensity thresholding and spectral fitting filtering methods.

DNA sample preparation

Lambda phage DNA (Thermo Scientific) was diluted to 100 ng/ μ L in Tris EDTA (TE) buffer (10 mM Tris and 1 mM EDTA). YOYO-1 dye (Invitrogen) was diluted to 2 μ M in TE buffer. 32 μ L of DNA was mixed with 480 μ L of YOYO-1 for a base pair to dye labeling ratio of 5:1 (132). The mixture was incubated for 1 hr at room temp covered using aluminum foil. The sample was then heated to 65 $^{\circ}$ C for 10 mins (132). 50 μ L of the labeled DNA was spin stretched on silanized coverslips at 1200 rpm for 30 secs.

DNA imaging and analysis

A 488-nm laser was used to excite the DNA labeled with YOYO-1 dye and imaged using sSMLM. 940 frames of the stretched DNA were captured using at an integration time of 10 ms. The recorded spectrum of each localization was used to calculate the spectral centroid. Color-coded sSMLM

images were generated using the centroid for each localization. Intensity and adjusted R^2 values for each localization were used to generate histograms. An intensity threshold of 240 and an adjusted R^2 threshold of 0.78 were selected and used to remove localizations unrelated to the DNA-YOYO sample. For spectral fitting, the reference spectrum of YOYO-1 was fit to the normalized signal with the accepted intensities from 0 to 1.2 and the spectral shift parameter was ± 5 nm. The reference spectrum and selected spectral shift parameter was based on measurements of YOYO-1 bound to DNA immobilized on a glass surface.

4.3 Results and Discussion

To quantitatively understand the origin of fluorescent impurities, we first focused on the essential initial step in sample preparation: preparing optically transparent substrate via various established surface cleaning (18, 67, 129-132) and functionalization (11, 129, 130, 136) methods. We recorded SMLM images of the unlabeled glass substrates (Fisherbrand™, Fisher Scientific) (Figures. 4-3a-c). As shown in Figure 4-3a, the representative MIP of SMLM images from a non-processed glass substrate clearly shows the existence of stochastic fluorescent emission with an average number density of $2.0 \pm 0.3 \times 10^7 \text{ cm}^{-2}$ (Figure 4-3d). Without adding fluorescence dye, such observed stochastic emission can only be contributed by fluorescent impurities. These observed fluorescent impurities are likely caused by contaminants introduced during the manufacturing, packing, and transportation stages, which may potentially be removed by cleaning the substrate.

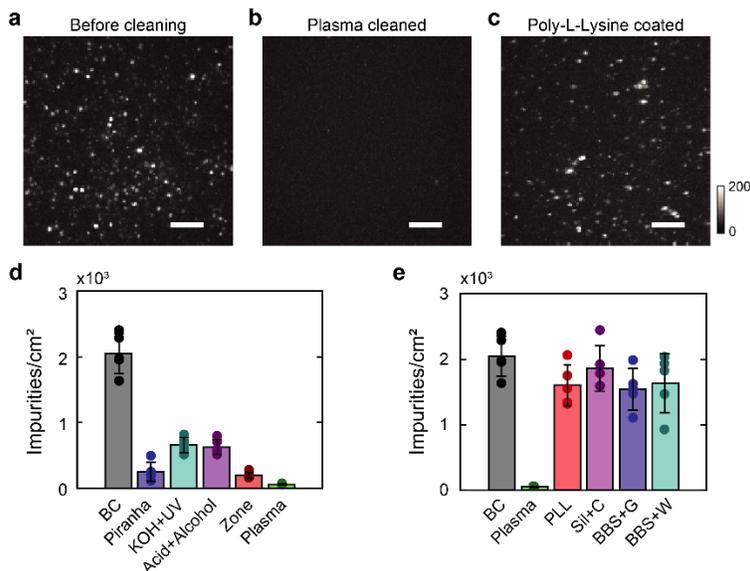


Figure 4-3: The origin of fluorescent impurities. MIP images (scalebars 5 μm) of unlabeled glass surface (a) before cleaning, (b) after plasma cleaning, and (c) after PLL functionalization. Comparison of the density of fluorescent impurities from 5 different FOVs (d) before surface cleaning (BC) and after cleaning via the piranha solution (Pir), rinsing with potassium hydroxide and sterilization using UV light (KOH+UV), rinsing with Hydrochloric acid and Prop-2-anol (Acid+Alcohol), exposure to UV activated ozone (UV-zone) and exposure to argon and oxygen plasma (Plasma). Comparison of the density of fluorescent impurities for 5 different FOVs on surfaces (e) before and after plasma cleaning (as a reference) and plasma-cleaned surfaces after functionalization via PLL coating, silanization with a final wash of chloroform (Sil+C), BSA and NEU (BBS) functionalization with glucose oxidase buffer (BBS+G) and BBS water as the buffer (BBS+W).

Second, we tested literature-reported cleaning methods, including three chemical methods (piranha solution (129), KOH solution (18), and HCl solution (130)) and two physical methods (UV-ozone (131) and plasma cleaning (67, 132)). The MIP of SMLM images of the substrate after each cleaning method is shown in Figure 4-4.

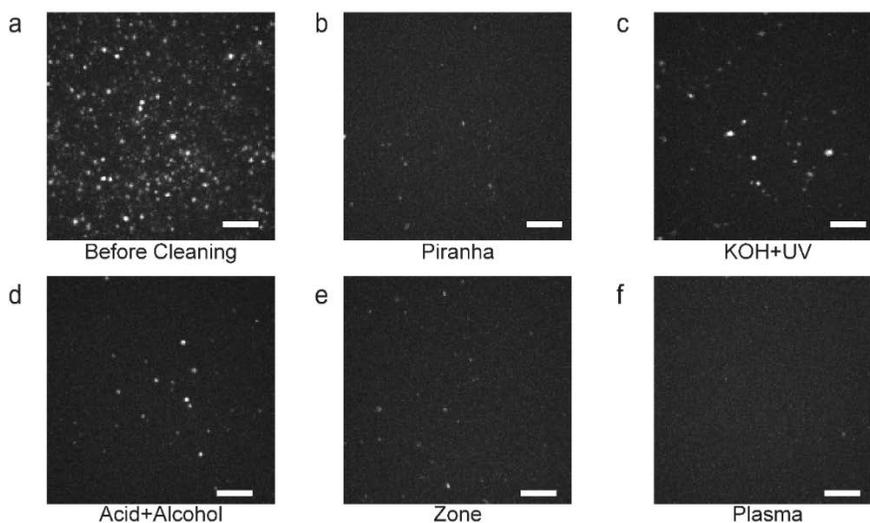


Figure 4-4: Representative MIP images of a bare Fisherbrand™ coverslip (a) before cleaning, (b) after cleaning using the piranha solution, (c) after sonication in 1 M KOH and sterilization using UV illumination, (d) after rinsing with HCl and prop-2-anol, (e) after cleaning with UV-activated ozone, and (f) after exposure to a mixture of oxygen and argon plasma. All images were captured using 532 nm illumination at a power density of 3 kW/cm². Scale bars are 5 μm.

As expected, we found that all tested surface cleaning methods effectively reduced the number of fluorescent impurities (Figure 4-3d). Using piranha solution, KOH solution, and HCl solution, the fluorescent impurity number density dropped to $2.5 \pm 1.4 \times 10^6 \text{ cm}^{-2}$, $6.4 \pm 1.1 \times 10^6 \text{ cm}^{-2}$, and $6.2 \pm 1.2 \times 10^6 \text{ cm}^{-2}$, respectively. Using physical cleaning methods, the fluorescent impurity number density respectively dropped to $1.7 \pm 0.1 \times 10^6 \text{ cm}^{-2}$ and $5.5 \pm 0.9 \times 10^5 \text{ cm}^{-2}$ after UV-ozone and plasma cleaning. The fluorescent impurity number density for each cleaning method was calculated using 1000 frames recorded using an integration time of 10 ms and a power density of 3 kW/cm². We hypothesize that while chemical cleaning methods can effectively remove the possible contaminants on the bare substrate, the chemical solution itself may contain new contaminants. Additionally, these methods require rinsing and drying, which could contribute to potential sources of fluorescent impurities. Further, chemical methods did not always uniformly

clean the surface. This was mostly due to variation in drying of the surface and the quality of the chemicals used in the cleaning (see demonstrations of its impact in Figure 4-5). Therefore, care should be taken when using chemical cleaning methods since sections of the coverslip may have an accumulation of chemical impurities along the direction the coverslip was rinsed. Consequently, these sources of fluorescent impurities, reduce the effectiveness of chemical cleaning. Figure 4-3d suggests that plasma cleaning is the most appropriate method in consistently minimizing the occurrence of the fluorescent impurities.

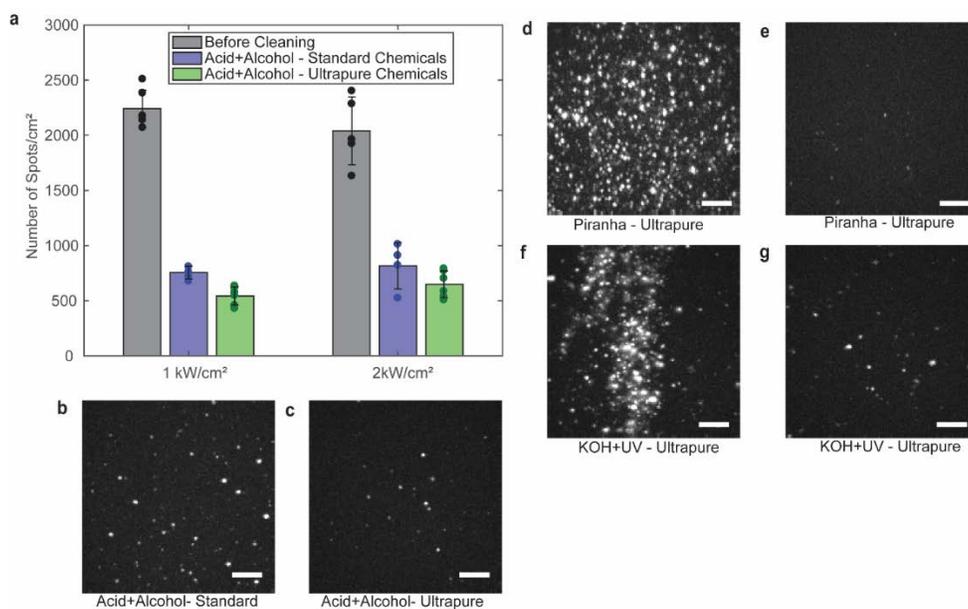


Figure 4-5: Variability in surface cleanliness and uniformity. (a) Comparison of standard chemicals previously purchased for general lab use to ultrapure chemicals purchased specifically for cleaning. Images of a regular coverslip cleaned with standard chemicals (b) and cleaned with ultrapure chemicals (c). Variability of uniformity using ultrapure chemicals. (d-e) fluorescence images at different areas on the same coverslip cleaned with the piranha solution. (f-g) fluorescence images at different areas on the same coverslip cleaned with KOH then sterilized with UV.

After cleaning, we examined fluorescent impurities introduced by other essential sample preparation steps, which requires a wide variety of chemical reagents and may introduce new sources of fluorescent impurities. To this end, we tested three commonly used surface functionalization methods (PLL (136), silane (129), and biotinylated BSA with NEU or BBS (11)) after plasma cleaning. We found a significant increase of the fluorescent impurities after the functionalization process (Figure 4-3e). Figure 4-3c shows a representative SMLM MIP image after surface functionalization using PLL (see Figure 4-6 for results of other functionalization methods). Although we used chemical reagents with the highest purity grade (see Tables 4-1 and 4-2 for purity information), we found that the trace amount of fluorescent impurities still imposed significant effects on the fluorescent impurities in SMLM.

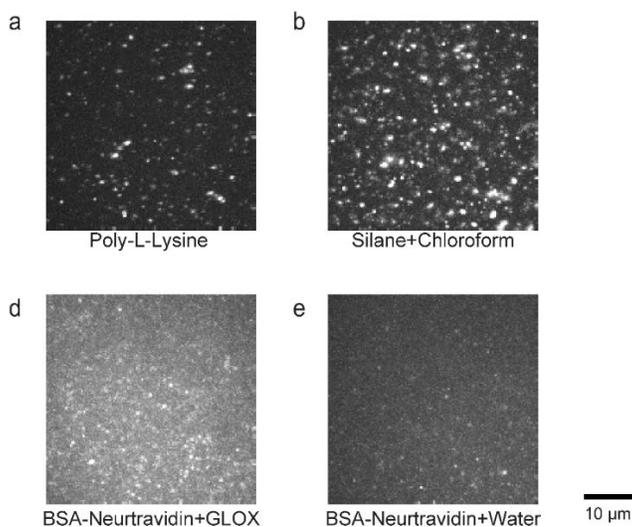


Figure 4-6: Representative MIP images of plasma cleaned Fisherbrand™ coverslips functionalized with (a) PLL, (b) silane with chloroform as the final rinse, (c) biotinylated BSA and NEU with glucose oxidase (GLOX) buffer, and (d) biotinylated BSA and NEU with water buffer. All images were captured using 532 nm illumination at a power density of 3 kW/cm².

As shown in Figure 4-3e, after treating with PLL, silane solution, and BBS, the observed fluorescent impurities number density increased to $1.6 \pm 0.3 \times 10^7 \text{ cm}^{-2}$, $1.9 \pm 0.351 \times 10^7 \text{ cm}^{-2}$, and $1.5 \pm 0.3 \times 10^7 \text{ cm}^{-2}$, respectively. Adding typical oxygen scavenging imaging buffer (containing glucose, glucose oxidase, catalase, and 2-mercapethanol in phosphate buffered saline (PBS) supplemented with 10 mM MgCl_2) to BBS functionalized surfaces further increased the fluorescent impurities number density to $1.6 \pm 0.5 \times 10^7 \text{ cm}^{-2}$.

Fluorescent impurity number densities were calculated using the same number of frames, integration time, and power density as aforementioned. Clearly, we observed a positive correlation between the fluorescent impurity number density and the use of chemicals, even at the highest available purity grade (3, 96). One common practice in single-molecule imaging and spectroscopy is to photobleach the prepared surface prior to sample introduction (96), however, any fluorescent impurities associated with the buffer for the sample would be ignored. Additionally, photobleaching could potentially damage or inactivate the functionalized surface if care is not taken to select the appropriate photobleaching power and wavelength (96, 97). Therefore, an alternative approach would be necessary to address these problems associated with the removal of all fluorescent impurities. In answering our first question, is it possible to reduce or ultimately eliminate fluorescent impurities, Figures 4-3d-e indicate that it is impractical to fully eliminate fluorescent impurities as long as any chemical reagent is used. These results further suggest that researchers should take precaution of the impact of fluorescent impurity in interpreting single-molecule imaging results and underscores the need for a strategy is to distinguish fluorescent impurities in SMLM.

We hypothesize that sSMLM is more effective to identify target molecules and reject

fluorescent impurities. To test this, we first recorded the spectra of fluorescent impurities associated with surfaces before cleaning and after functionalization. Figure 4-7a shows representative spectra of fluorescent impurities in Figure 4-3a. While fluorescent impurities 1 and 2 have spectra at 569 nm and 593 nm, respectively, the spectrum of impurity 3 ranges from 566 nm to 610 nm.

Figure 4-7b shows three representative spectra from fluorescent impurities associated with PLL functionalization. We found that these fluorescent impurities displayed a significant amount of inhomogeneity with the different fluorescent impurities having spectra at 562 nm, 623 nm, and 642 nm. These findings indicate that fluorescent impurities have diverse spectral characteristics and can emit a large number of photons when excited using high power densities. Though the nature of fluorescent impurities remains unknown, their spectral signatures can be used to guide experimental design and data analysis.

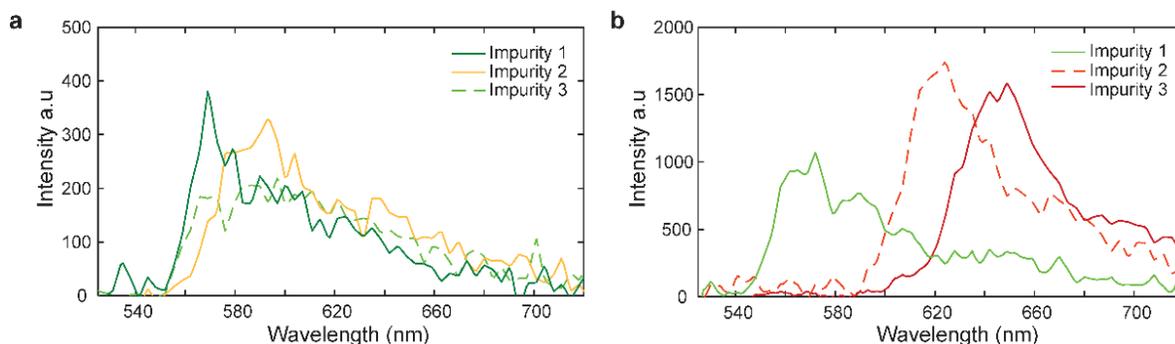


Figure 4-7: (a) Representative spectra from three fluorescent impurities on a Fisherbrand™ coverslips before cleaning. (b) Representative spectra from three fluorescent impurities associated with PLL functionalization.

Because surface functionalization is necessary for the proper deposition of biological molecules, it would be difficult to remove this step in the preparation of many samples. Therefore, to better understand the impurities associated with different surface functionalization methods, we measured the average fluorescence spectrum of coverslips treated with each method under 532-nm illumination using sSMLM with a narrow slit as shown in Figure 4-8a-c. The average spectrum for each functionalization method was determined by measuring the emission spectra of multiple impurities (see Figure 4-8d). This study showed that each functionalization method had a distinct associated spectrum. The spectrum from impurities associated with PLL was found to have main peaks around 575 nm and 625 nm, impurities associated with silane and chloroform have a single peak around 585 nm, while silane impurities have having a broader spectrum and a peak around 590 nm, BSA-biotin with NEU had the broadest spectrum with a peak around 610 nm in both buffer conditions as shown in Figure 4-8d.

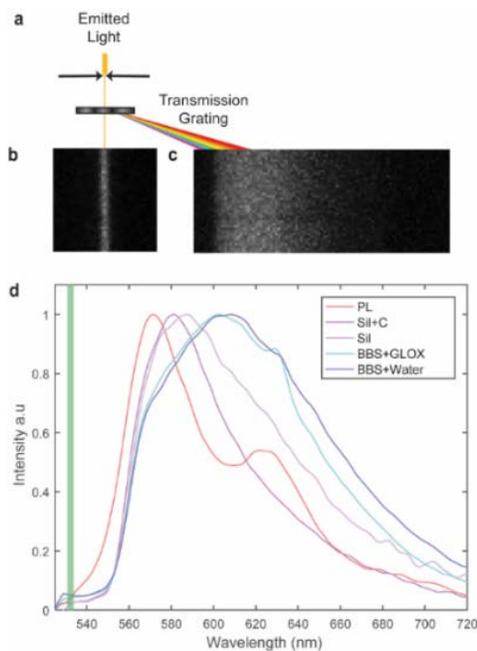


Figure 4-8: Spectroscopic analysis of functionalized coverslips. (a) Schematic of sSMLM detection channel with the slit closed to acquire average images and line. (b) Image of coverslip functionalized with biotinylated BSA and NEU with GLOX buffer (BBS+GLOX). (c) Line spectra of spectra from BBS+GLOX coverslip. (d) Average line spectra from PLL, silane with chloroform rinse, silane with water rinse, BBS+GLOX, and BBS+Water.

Additionally, the PLL impurities were found to be sensitive to specific wavelengths as shown in Figure 4-9, emission events could be detected under 532 nm and 647 nm excitation but not under 445 nm excitation. Using the same detection system shown in Figures 4-8a-c, the average spectrum from experiments using 532 nm and 647 nm are included in Figure 4-9.

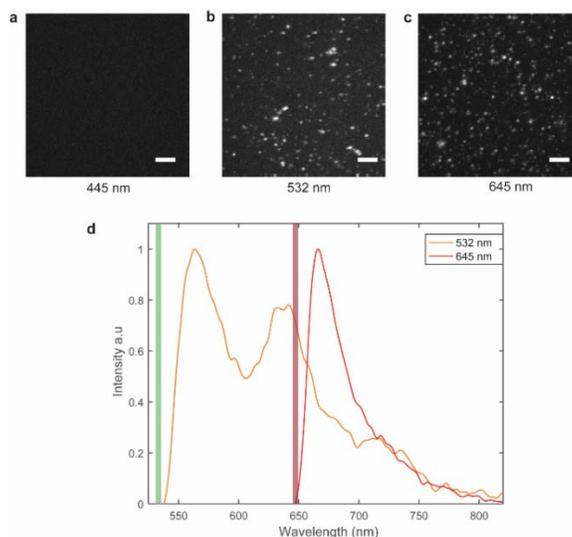


Figure 4-9: Excitation dependent emission of fluorescent impurities of PLL coated plasma cleaned coverslips under illuminations at wavelengths of (a) 445 nm, (b) 532 nm, and (c) 645 nm. (d) Average line spectra under illuminations at wavelengths of 532 nm and 645 nm.

Using sSMLM, we developed a spectral fitting method and compared it with the intensity thresholding method to experimentally evaluate their sensitivity in identifying target molecules and specificity in rejecting fluorescent impurities. We used DNA origami nanorulers (labeled with AF532 and AF568 with 10 nm spatial separation, Gattaquant) (137, 138) as the target molecules because the spacing of the dyes was beyond the spatial resolution of SMLM but their spectral separation was greater than the spectral dispersion of our sSMLM system. We spin-coated the nanorulers on PLL functionalized glass substrate. We acquired images within the same FOV using both low power density (LPD, 0.5 kWcm^{-2}) and high-power-density (HPD, 3 kWcm^{-2}) illuminations. LPD and HPD illuminations respectively represented the conditions of conventional fluorescent microscopy and SMLM (Figures 4-1a-c). Under LPD illumination, the observed fluorescent emissions are highly likely from the nanorulers (Figure 4-10a) (139). Additionally,

since photo-switching is suppressed under LPD illumination the average emission spectrum of the nanorulers and the minority of fluorescent impurities can be recorded.

Therefore, to establish the ground truth, we examined and fitted the emission spectra in the average LPD image with known nanoruler emission spectra. Overlapping spectra in the average LPD images were excluded from this analysis. Detected emissions that fit the spectra of AF532 and AF568 with an adjusted R^2 value greater than 0.89 after spectral fitting were considered to be true nanoruler emissions.

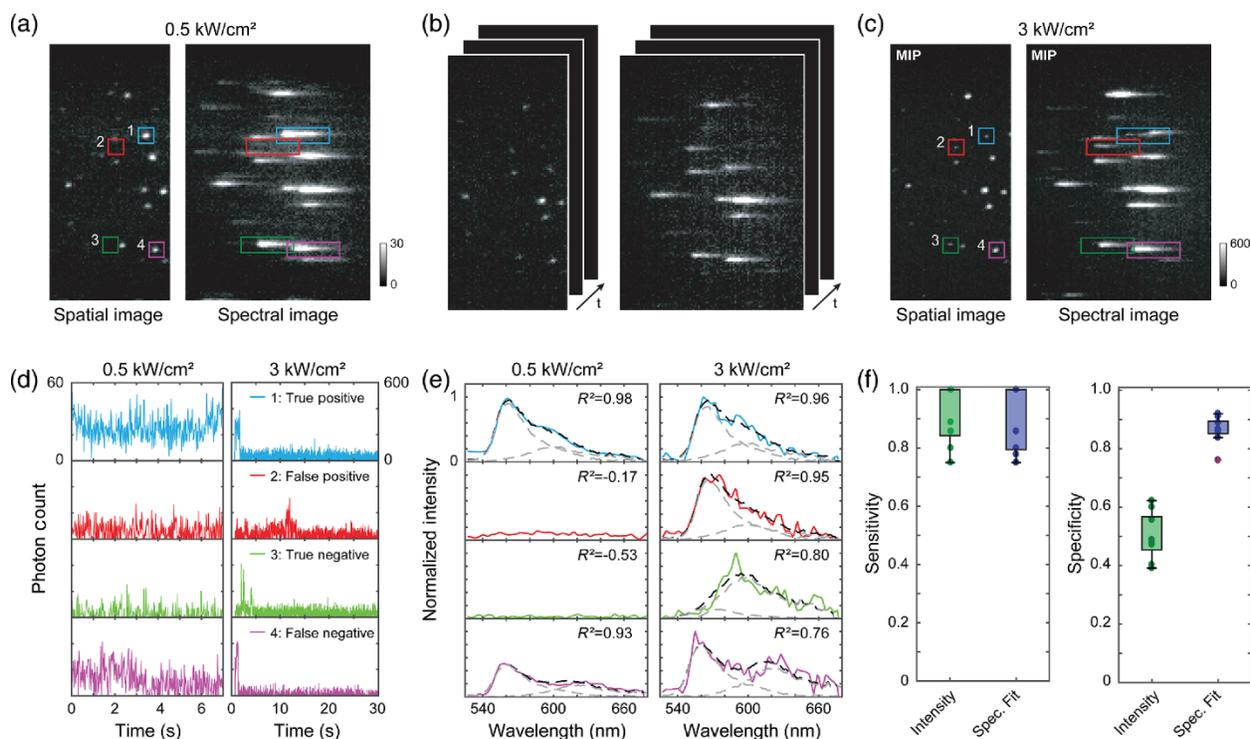


Figure 4-10: Identifying fluorescent impurities during SMLM. (a) Average spatial and spectral image of DNA origami nanorulers, containing two emitting points labeled with single AF532 and AF568 molecules 10 nm apart, immobilized on a PLL coated surface. Images were acquired under illuminations with power densities associated with conventional fluorescence imaging (0.5 kW/cm²). (b) Stack of 1500 frames of the spatial and spectral images of the nanoruler sample for sSMLM (3 kW/cm²) using the same FOV. (c) MIP images of the spatial and spectral of the same FOV. (d) Photon count versus time of two nanorulers (1,4) and two fluorescent impurities (2,3) highlighted in average and MIP of SSMLM images. (e) Corresponding spectra of the point sources identified in the average and SSMLM images representing true positive, false positive, true negative and false negative cases for the spectral fitting method. (f) Sensitivity and (g) Specificity comparison for 9 datasets using an emission intensity a threshold of 180 and a spectral fitting filter adjusted R^2 threshold of 0.84.

Figure 4-11 shows the spectra of nanorulers and fluorescent impurities. To select a R^2 threshold for ground truth analysis, the histograms of the R^2 fitting parameter before and after the influence of the Alexa 568 terms were assessed.

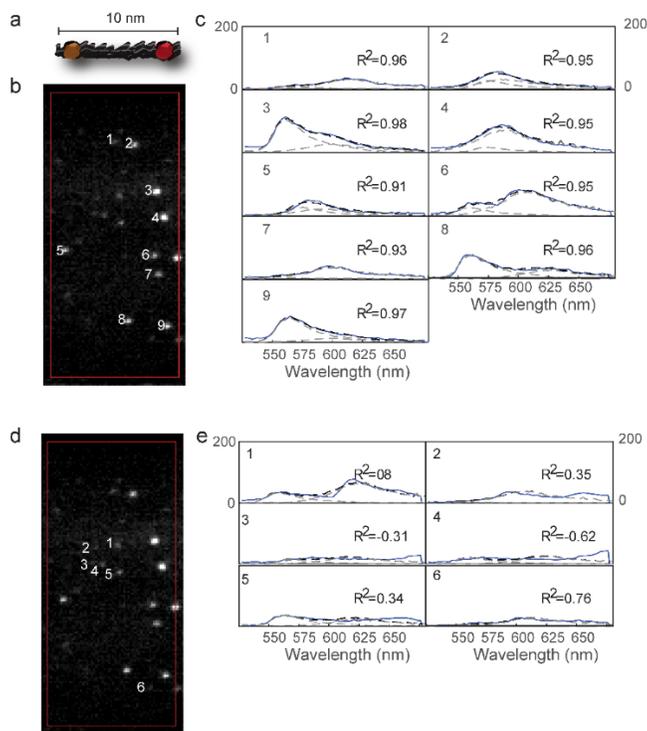


Figure 4-11: (a) Cartoon of the 10 nm AF532 and AF568 nanoruler. (b) Representative spatial image and (c) extracted spectra of nanorulers detected using LPD illumination using the field of view highlighted in Figures 4-11a-c. (d) Representative spatial image and (e) extracted spectra of fluorescent impurities detected using LPD illumination using the field of view highlighted in Figures 4-11a-c.

As shown in Figure 4-12, a threshold of 0.89 was selected to include only emitters whose spectra included both AF532 and AF568.

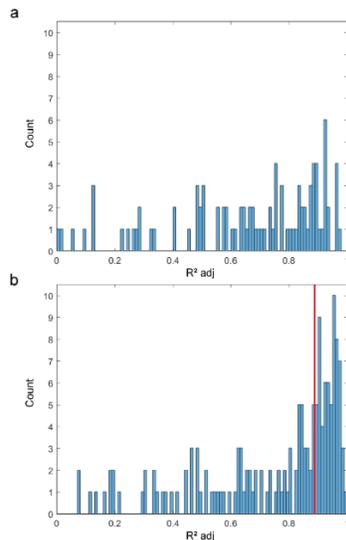


Figure 4-12: Histograms of adjusted R^2 values of average spectra from 174 locations in the LPD image after spectral fitting to the reference spectrum of (a) AF532 and (b) both AF532 and AF568 with the adjusted R^2 threshold of 0.89 highlighted.

We acquired 1500 sSMLM images from the same FOV under HPD illumination (Figure. 4-10b) and plotted both the spatial and spectral MIP images in Figure 4-10c. Since the nanorulers have already been identified in the LPD experiment, any additional fluorescent emission identified in HPD experiment can be treated as fluorescent impurities. Figure 4-13 shows the number of emitters for all 9 FOVs in the high-power density datasets (3 kWcm^{-2}) and their categorization as nanorulers or fluorescent impurities after comparison to the established ground truth.

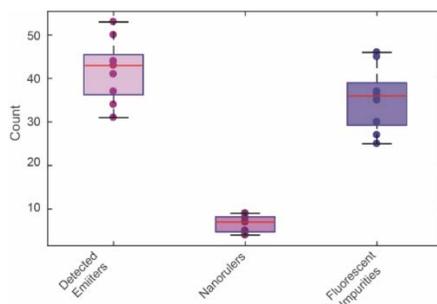


Figure 4-13: The number of emitters detected using HPD excitation, the number of nanorulers, and the number of fluorescent impurities detected under HPD excitation for all 9 field of views. The median is highlighted for each box plot.

We compared the sensitivities and specificities of our spectral-fitting method and the commonly used emission intensity thresholding method. We used the histograms for the adjusted R^2 and emission intensity of each emission event to select a range of possible thresholds. For the spectral fitting method, a range of 0.80 to 0.94 was tested for the adjusted R^2 values and for the emission intensity thresholding method a range from 120 to 400 was tested allowing the SNR to be at least 6 dB above the background. For fair comparison, we selected the case with ~90% sensitivities in both methods (Table 4-3). In this example, for the spectral fitting method emission spectra fitted with an adjusted R^2 value greater than 0.84 were considered as positive identification of nanorulers, while others were considered as negative identification. On the other hand, in the emission intensity thresholding method stochastic emission with the intensity above 180 will be recognized as a nanoruler while others were categorized as a fluorescent impurity.

We classified nanoruler identifications in the HPD experiments against the ground truth established in the LPD experiments into four categories: true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN). Representative intensities and spectra of the four

categories are shown in Figure 4-10d and Figure 4-10e, respectively. As shown in Figures 4-10d.2 and 4-10d.3, the emission intensity thresholding method would fail to remove both fluorescent impurities since their intensities exceed the established threshold.

We compared the sensitivities (Figure 4-10f) and specificities (Figure 4-10g) of both methods using the datasets collected from 9 FOVs (see Table 4-3 for actual values). The sensitivity and specificity for the emission intensity thresholding method are $91\pm 9\%$ and $50\pm 8\%$, respectively; the sensitivity and specificity for our spectral fitting method are $89\pm 10\%$ and $87\pm 4\%$, respectively.

While both methods showed comparable sensitivity in identifying nanorulers, the specificity of rejecting fluorescent impurities by our spectral fitting methods is close to two-fold higher than the emission intensity thresholding method. Though an 85% specificity for the emission intensity thresholding method can be achieved by increasing the threshold to 300, this will result in a 13% reduction in sensitivity. On the other hand, the threshold for the spectral fitting method can be increased up to 0.89 allowing for a specificity of 90% with only a 4% reduction in sensitivity. This study shows that the specificity of spectral fitting is less dependent on the user defined R^2 threshold than the threshold for emission intensity thresholding. However, due to diverse origins of fluorescent impurities, their spectra can overlap with nanorulers (as shown in Figure 4-10e.2), which contributed to 13% FP identification in spectral fitting method. Further reducing FP identification can be accomplished by incorporating additional signatures related to dye photophysics, such as switching time constant (83, 140, 141) or fluorescence lifetime (140, 141).

Table 4-3: Sensitivity and Specificity comparison between single molecule filtering based on emission intensity (threshold 180) and spectral fitting (threshold 0.84).

	Intensity Threshold	Spectral Fitting		Intensity Threshold	Spectral Fitting
Sensitivity	1.00	1.00	Specificity	0.62	0.91
	0.89	0.78		0.40	0.86
	0.80	0.80		0.47	0.89
	1.00	1.00		0.39	0.76
	0.86	0.86		0.47	0.86
	1.00	1.00		0.48	0.92
	0.86	0.86		0.60	0.87
	1.00	1.00		0.49	0.84
	0.75	0.75		0.56	0.89
Average	0.91	0.89	Average	0.50	0.87
Median	0.89	0.86	Median	0.48	0.87
STD	0.09	0.10	STD	0.08	0.04

Figure 4-14 demonstrates that our spectral-fitting method better identifies and minimizes artifacts caused by fluorescent impurities. Figure 4-14a shows the sSMLM spatial and spectral MIP images of the same nanoruler sample imaged in Figure 4-10, but from a different FOV. We highlighted two regions of interest (ROIs) that contain both nanorulers and fluorescent impurities. Figure 4-14b shows the reconstructed super-resolution image using ImageJ plug-in ThunderSTORM (121) without excluding fluorescent impurities. The results after emission intensity thresholding and spectral fitting are shown in Figures 4-14c and 4-14d, respectively. ROI1 is an example of a misidentified molecule. Within ROI1, among the 189 localized events being originally identified in Figure 4-14b, 114 events were treated by emission intensity thresholding method as nanorulers (Figure 4-14c). By comparing corresponding spectra of all the localized events (representative spectrum is shown as the black curve in Figure 4-14h) with the

spectroscopic signature of the nanoruler (Figure 4-10e), our spectral fitting method determined that none of the 189 events is from nanorulers (Figure 4-14d).

Using the spectral fitting method in ROI1 prevented sample misidentification. Figures 4-14e-g are the magnified view of the ROI2 shown in Figures 4-14b-d, respectively. ROI2 is an example of a fluorescent impurity which overlaps in space with a nanoruler. Within ROI2, among the localized events being originally identified in Figure 4-14e, which corresponds to a standard deviation (S.D.) of localizations of 52.9 nm (135). Among them, 269 events were treated by emission intensity thresholding method as nanorulers, which reduces the S.D. of localizations to 40.1 nm (Figure 4-14f). After spectral fitting, we identified 103 events from the nanoruler and determined that 389 of the originally identified events were fluorescent impurities. As shown in Figure 4-14h, the representative spectrum of nanoruler (red curve) shows distinct spectroscopic signatures in clear contrast with the spectrum from the fluorescent impurity (blue curve), which further validates the specificity of our spectral fitting method. We demonstrate here that our spectral fitting method can effectively reduce localization uncertainty of samples by removing localizations from fluorescent impurities, with approximately two-fold improved localization precision (S.D.: 22.5 nm) comparing with emission intensity thresholding method.

Finally, we compared the performance of emission intensity thresholding and spectral fitting in removing artifacts induced by unwanted fluorescence when imaging DNA samples. For this demonstration we stretched lambda phage DNA labeled with YOYO-1 on a silane treated coverslip. We imaged the sample using sSMLM and color-coded the reconstructed image using the spectral centroid for 831 localizations as shown in Figure 4-14i. After applying an intensity filter with an intensity threshold of 240, the reconstructed image contained 476 localizations as

shown in Figure 4-14j, however, localizations unrelated to the DNA-YOYO sample were not completely removed. We then applied our spectral fitting method with an adjusted R^2 threshold of 0.78 and found that only 221 localizations were more specifically associated with the DNA-YOYO sample as shown in Figure 4-14k. The successful removal of the unwanted SMLM imaging artifacts is highlighted as triangles in Figure 4-14i-k, which results in a clear image after applying our spectra fitting method.

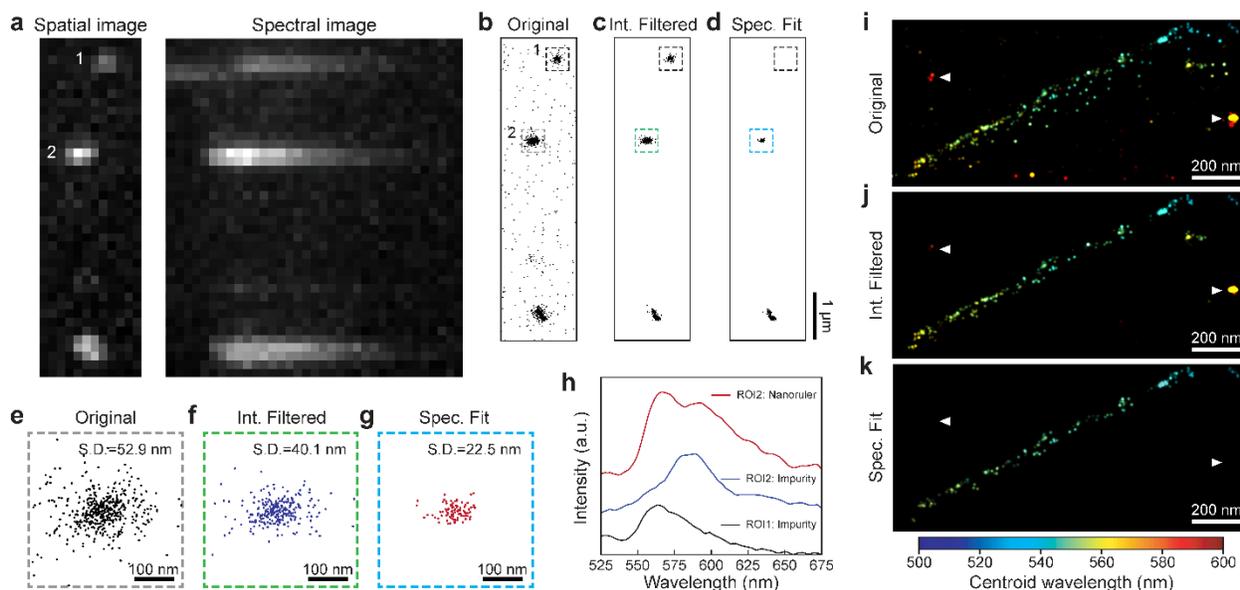


Figure 4-14: Comparing results in minimizing artifacts induced by fluorescent impurities using intensity filtering and our spectral fitting methods. (a) sSMLM spatial and spectral MIP images of nanorulers with fluorescent impurities. (b) Reconstructed super-resolution images without rejecting fluorescent impurities; (c) result after emission intensity filtering; (d) result after spectral fitting. ROI 1 highlights the localized fluorescent impurities that are eliminated by our spectral fitting method but are misidentified by intensity filtering method. ROI2 highlights the case of spatial overlapping of fluorescent impurities and nanorulers results in higher localization uncertainty. The resulting super-resolution images of ROI2 are further magnified in (e) before filtering (standard deviation (S.D.) 52.9 nm), (f) after intensity filtering (S.D. 40.1 nm), and (g) after spectral fitting (S.D. 22.5 nm). (h) Averaged spectra of fluorescent impurities and nanoruler emission. (i) Reconstructed color-coded super-resolution image of stretched lambda phage DNA labeled with YOYO-1 dye on a silane functionalized surface before rejecting emission unrelated to the DNA-YOYO sample (resulting artifacts highlighted by white triangles); (j) result after intensity filtering contains artifacts from unwanted fluorescence; (k) result after spectral fitting specifically removed artifacts induced by unwanted fluorescence.

4.5 Conclusions

We show that fluorescent impurities are unavoidable. Although thorough plasma cleaning significantly reduced the number of detectable fluorescent impurities, a large amount of fluorescent impurities can be introduced by required substrate treatments, such as surface functionalization. Although the true origins of fluorescent impurities remain unclear, using

sSMLM to perform spectral fitting can effectively improve the specificity of rejecting fluorescent impurities by nearly two-folds comparing with commonly used method while maintaining comparable sensitivity in identifying target molecules. Additionally, we found that the specificity of spectral fitting is less dependent on the user-defined R^2 threshold than the intensity threshold for intensity filtering. This study suggests that sSMLM, with newly added spectral analysis capability, is a powerful tool for single-molecule studies to guide sample preparation for better experimental design and analysis.

CHAPTER 5: Quantitative sSMLM #2: Counting and Sizing Self-Assembled Nanocarriers using Quantitative Spectroscopic Analysis for Cluster Extraction

Self-assembled nanocarriers have inspired a range of applications for bioimaging, diagnostics, and drug delivery. Non-invasive visualization and characterization of nanocarriers are important for understanding their structure to function relationship. However, quantitative visualization of nanocarriers in the sample's native environment remains challenging using existing technologies. Single-molecule localization microscopy (SMLM) has the potential to provide both high-resolution visualization and quantitative analysis of nanocarriers in their native environment. However, non-specific binding of fluorescent probes used in SMLM can introduce artifacts, which impose challenges in quantitative analysis of SMLM images. We showed the feasibility of using spectroscopic point accumulation for imaging in nanoscale topography (sPAINT) to visualize self-assembled polymersomes (PS) with molecular specificity. Furthermore, we analyzed the unique spectral signatures of Nile Red (NR) molecules bound to the PS to reject artifacts from non-specific NR bindings. We further developed quantitative spectroscopic analysis for cluster extraction (qSPACE) to increase the localization density by 4-fold compared to sPAINT; thus, reducing variations in PS size measurements to less than 5%. Finally, using qSPACE we quantitatively imaged PS at various concentrations in aqueous solutions with ~20-nm localization precision and 97% reduction in sample misidentification relative to conventional SMLM. This chapter is adapted from a previously published article in *Langmuir* (142).

5.1 Introduction

Nanocarriers are soft nanomaterials with particle size typically less than 500 nm in diameter and are frequently used as transport vehicles for other substances (143). The composition, size, morphology, and functionality of nanocarriers have been extensively investigated for a wide range of biomedical applications, including drug delivery, diagnostics and therapeutics (143). The visualization of nanocarriers with minimal perturbation of the nanomaterials and the surrounding environments are crucial steps for understanding the relationship between their structure, spatial distribution, and function (16, 144); however, this remains a challenge using existing imaging techniques. Conventional fluorescence microscopy is unable to fully resolve the nanocarriers because of the optical diffraction limit. Other non-optical techniques (14, 15) offer higher resolution (e.g. electron microscopy and atomic force microscopy); however, they often require restrictive sample preparations, which alter the sample's native environment. The limitations in the available technologies created the need for a better solution, which can provide both high resolution images of nanomaterials in their native environment while simultaneously quantifying their structural and spatial distributions.

Super-resolution optical microscopy allows visualization of nanomaterials with sub-diffraction resolution and molecular specificity (16, 145-149). In particular, single-molecule localization microscopy (SMLM) (17, 18, 65, 66) techniques detect the stochastic emissions from individual fluorescent molecules, approximate their spatial coordinates with nanometer precision and reconstruct the image after compiling all the molecular coordinates. For quantitative characterization, sophisticated image processing is required to identify individual nanocarriers and reject background noise coming from, for example, fluorescent impurities or non-specific labeling.

Spatial clustering techniques such as Ripley's K-function (90, 135), pair correlation (91), and density-based clustering algorithms (93, 94, 150) have been routinely applied to identify the boundaries of the samples in SMLM (83, 151). However, Ripley's K-function and pair correlation restrict sample identification to nanocarriers exhibiting homogenous size distributions. Meanwhile, density-based clustering algorithms allow for the identification of a range of samples with inhomogeneous size and shape; however, they are sensitive to background signals and variations in localization density (83, 92, 151).

Recently developed spectroscopic SMLM or sSMLM (30-32, 103, 104, 109, 110, 112, 115, 116, 122) techniques distinguish fluorescent signals from individual molecules according to their emission spectra. They provide new strategies for multicolor (30-32, 109, 116) and functional super-resolution imaging (104-108, 110, 123). Furthermore, the spectral signatures have been used to identify and remove the influence from fluorescent impurities (123). The removal of non-specific fluorescent signals, however, remains challenging since their spectral signatures are identical to spectral signatures from probes that are bound to the sample. In contrast, a variant of sSMLM referred to as spectroscopic point accumulation for imaging in nanoscale topography (sPAINT) (104, 105, 107, 108), has been developed to detect specific transient binding of probes, whose fluorescence emission spectra depends on their binding environment. For instance, sPAINT has been employed to image membrane hydrophobicity in liposomes and cells (104, 107), to observe heterogeneity of protein oligomers (105), and to characterize the formation of surface adlayers (108).

In this work, we demonstrate that sPAINT can be used to quantitatively image and characterize nanocarriers after rejecting non-specific bindings. We use sPAINT and Nile Red (NR)

dye (Figure 5-1a), to image, count, and measure the size distribution of nanocarriers in an aqueous environment, using polymersomes (PS) as a model material. We also report a quantitative Spectroscopic Analysis for Cluster Extraction (qSPACE) post-processing method to increase the sizing accuracy by optimizing the utilization of the spatial and spectral information obtained from sPAINT. Finally, we validate the accuracy of qSPACE in quantifying the size distribution of PS using nanoparticle tracking analysis (NTA). We also used cryogenic transmission electron microscopy (cryoTEM) to validate the size range and morphology of the PS in the qSPACE super-resolution image reconstructions.

5.2 Results and Discussion

PS imaging using sPAINT

The PS is assembled from block copolymers (BCP) with the non-polar end facing inwards and the polar tails exposed to the aqueous interface (Figures 5-1b & 5-1c) (7). NR is well-known for its polarity-dependent emission spectrum (152, 153) with unique spectral variations upon binding to PS. Particularly, the bulk emission measurement of NR in aqueous solutions with the presence of PS shows a 55-nm hypsochromic shift in the emission maximum (λ_{\max}) in comparison to solutions without PS (Figure 5-2a).

Figure 5-1c illustrates the sPAINT imaging principle of PS with NR. The PS sample was immobilized on poly-L-Lysine (PLL) coated glass substrate and imaged in phosphate buffer saline (PBS) solution containing 5 nM NR. The single-molecule blinking signals of free-diffusing NR molecules can only be detected upon transient binding to the PS or PLL substrate. The distinct spectral shifts of the two types of binding events are further distinguished by analyzing the single-molecule spectra using sPAINT. As shown in Figure 5-1d, the locations and emission spectra of

single-molecule signals are concurrently collected after passing through a transmission diffraction grating (see details in the methods section).

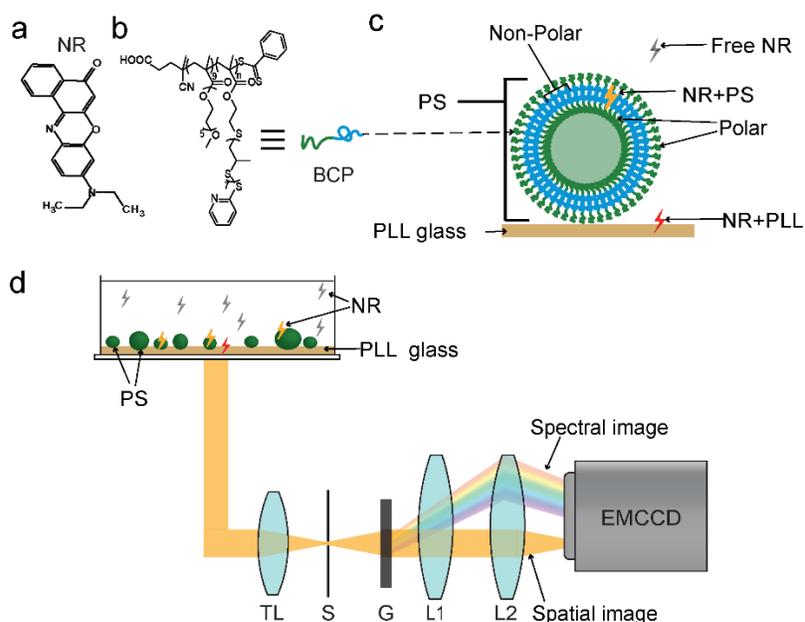


Figure 5-1: (a) Chemical structure of NR; (b) Chemical structure and illustration of the BCP for PS assembly; (c) Illustration of the assembled PS (the green color represents the polar end of the BCP and the blue color represents the non-polar end) and the difference in the emission spectra of NR when transiently bound to the PS (yellow) and the PLL-coated glass substrate (red). Free non-fluorescent NR is shown in gray; (d) Schematic of our sPAINT experimental setup. TL: tube lens; S: slit; G: transmission grating; L: lens; EMCCD: electron multiplying charge-coupled device.

A representative sPAINT reconstruction of PS immobilized on the PLL substrate (Figure 5-3a) shows distinct clustered features (as highlighted by 1, 2, and 3), ranging from tens to hundreds of nanometers and a relatively uniform background with hollow patterns. Using the same concentration of NR, we imaged a control sample (CTRL, PLL substrate without PS). A representative sPAINT reconstruction of the control sample (Figure 5-2b) shows consistent

background patterns without the clustered features, suggesting a large amount of non-specific bindings of NR to the PLL substrate.

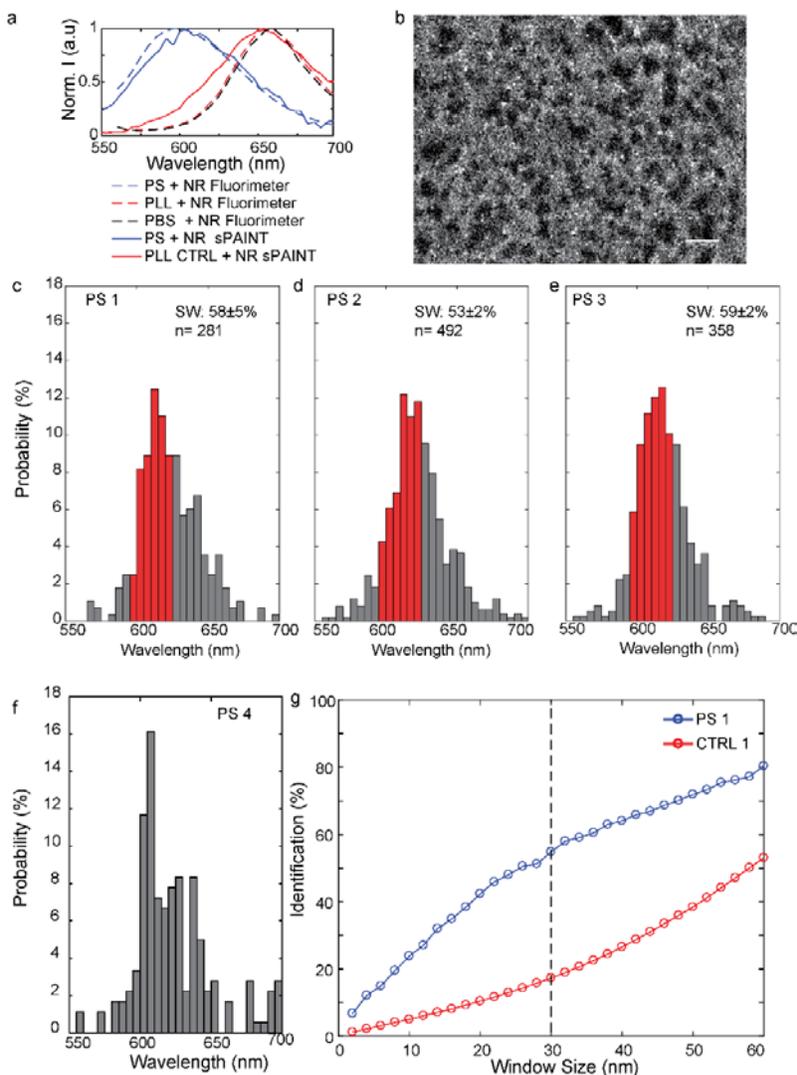


Figure 5-2: (a) Comparison of the average emission spectra of NR in solutions containing PS and sample preparation reagents (PLL and PBS)) measured using a fluorimeter and sPAINT. (b) Representative super-resolution reconstructions of the PLL coated glass used as the control sample. (Scale bar: 1 μm). (c-e) Histograms of the emission maxima (λ_{max}) of NR interactions in ROIs containing PS from three different images (six clusters were selected in each image). The selected spectral window (SW, 595 – 625 nm) is highlighted in red. (f) Reproducibility test showing the histogram of λ_{max} from PS acquired on a different day with similar experimental conditions. (g) Comparison of the identification percent of localizations from the PS λ_{max} histogram and identification percent of localizations from the control λ_{max} histogram as the window size of the SW increases (peak position of SW = 610 nm).

Statistical analysis of 1131 single-molecule spectra collected from 18 clustered features in three different images of the PS sample (Figure 5-2c-e) showed the peak of the histograms from the images had the mean λ_{\max} of 610 with a standard deviation of 3 nm (610 ± 3 nm). Three examples of the clustered features are highlighted in Figure 5-3a and their corresponding histograms of the λ_{\max} distribution are shown in Figure 5-3b. To identify non-specific binding events, we relied on the λ_{\max} values of individual single-molecule emissions from three representative images on the control sample each with $\sim 10^5$ localizations. Figure 5-3c shows a representative histogram of the λ_{\max} distributions from non-specific binding events. From the histograms from the three control images, we found that the control sample has two main populations with λ_{\max} of 585 nm ($26 \pm 2\%$) and 645 nm ($57 \pm 2\%$). Using the histograms of λ_{\max} from the PS, we set the peak wavelength for specifically identifying NR interactions with PS as 610 nm. Next, we tested various spectral window (SW) sizes and found that the most efficient SW for rejecting non-specific binding while specifically identifying NR interactions with PS was from 595 nm - 625 nm (see Figure 5-2g), as highlighted by the red color in Figures 5-3b and 5-3c. We show three typical single-molecule spectra of NR in Figure 5-3d. They represent NR specific binding to PS (black curve with λ_{\max} value around 607 nm) and non-specific binding to PLL substrate (blue and red curves with λ_{\max} values around 571 nm and 651 nm, respectively).

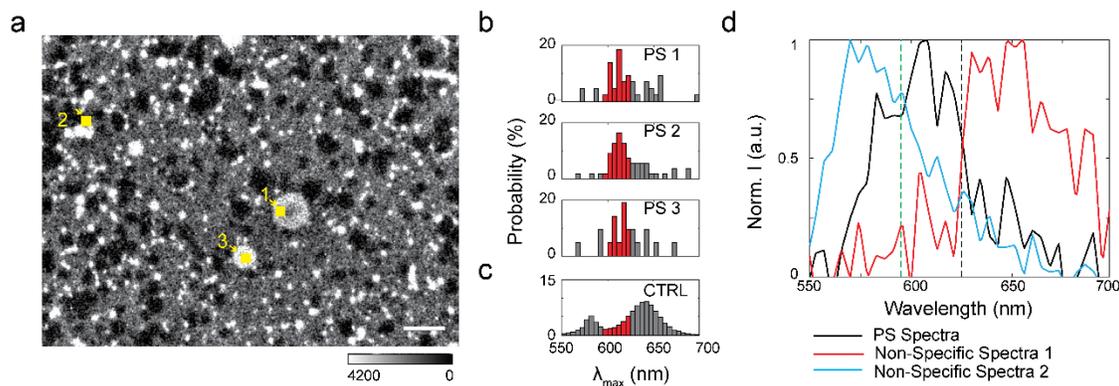


Figure 5-3: (a) A representative reconstructed super-resolution image of the immobilized the PS sample and NR interactions (Scale bar: 1 μ m); (b) Histogram of the λ_{\max} of NR interactions in three ROIs containing PS as highlighted by the yellow squares numbered 1-3; (c) Histogram of the λ_{\max} of non-specific NR interactions in the control sample. The SW used for detecting NR interactions with PS is highlighted in red

We found that the average single-molecule spectrum from the localizations associated with PS matched the bulk measurements of NR in a PBS solution containing PS (Figure 5-2a). Additionally, the average spectrum from localizations with $\lambda_{\max} > 645$ nm in the control sample resembled the bulk measurements of NR in solutions without PS (Figure 5-2a). For sPAINT measurements of the average spectrum, 300 localizations from a representative image were used (see methods section for details). The binding events with $\lambda_{\max} < 595$ nm in the control sample are likely from aggregates of NR in aqueous solutions that are not resolved in the fluorimeter measurements (154).

We selected SW centered around 610 ± 3 nm based on the histograms from the PS shown in Figures 5-3 and Figures 5-2c-e. We then selected a 30 nm SW where the non-specific binding to the PLL coated glass was minimized in the control sample (PLL coated glass with no PS) while maximizing the size of the window to allow as much NR interactions with PS to be included (Figure 5-2g). The selected SW rejected 83% of the localizations from the control sample while

60% of localizations associated with the PS were accepted. Changing the SW window size would affect the percentage localizations from both the non-specific binding and the PS sample. For example, using a window size of 10 nm, the percentage of localizations from non-specific binding would be reduced to 5% and the percentage of localizations from the PS sample would be reduced to 23%. This change would reduce the number of localizations per cluster in both cases causing inaccuracy in the size and morphology during subsequent cluster analysis steps.

Quantitative analysis of sPAINT images involves two steps. First, we selected localizations within SW (595-625 nm) for PS identification. We confirmed that the NR specific binding to PS can be separated from all other non-specific binding events using the selected SW (Figure 5-4). In the control samples, $17\pm 1\%$ of the localizations had λ_{\max} values within the SW. Meanwhile, we found that $58\pm 5\%$ of the localizations from the clusters in the PS sample were within this SW (Figure 5-2c). Figure 5-3b shows examples of the histograms of three regions of interest (ROIs) highlighted in Figure 5-3a with SW highlighted in red. Histograms of the λ_{\max} from ROIs in similar images (Figures 5-2d-e) show a similar distribution of localizations within SW.

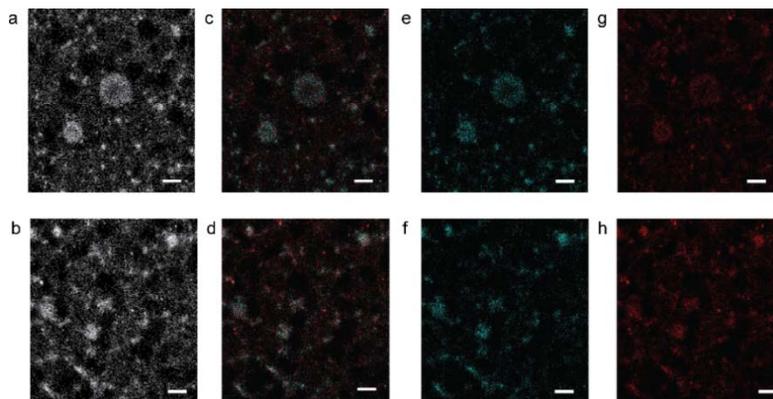


Figure 5-4: sPAINT reconstructions before applying the spectral window for selecting PS for (a) example 1 and (b) example 2. Overlay of the reconstructions with the localizations designated as PS in cyan and localizations designated from non-specific binding in red from (c) example 1 and (d) example 2. sPAINT reconstructions of the PS channel for (e) example 1 and (f) example 2. sPAINT reconstructions of the non-specific binding channels for (g) example 1 and (h) example 2 (Scale bar: 500 nm)

Second, we applied density-based spatial clustering of applications with noise (DBSCAN) optimized for clusters with heterogeneous cluster densities (93) to count and size PS. DBSCAN (94) and its variants (92, 93, 150) are supervised learning algorithms which are routinely used for SMLM post-processing. A cluster radius (CR) and a minimum number of localizations in the CR are user-defined to identify localizations belonging to particular clusters. We defined CR as twice the average localization precision (40 nm) (27, 92). Next, we identified the minimum number of nearest neighbors (min-NN) within the CR of each localization. The localizations were then ordered based on their spatial distribution and grouped into individual clusters (93). Finally, the number of clusters and the size of each cluster was estimated (see details of cluster analysis in the methods section and Figure 5-5).

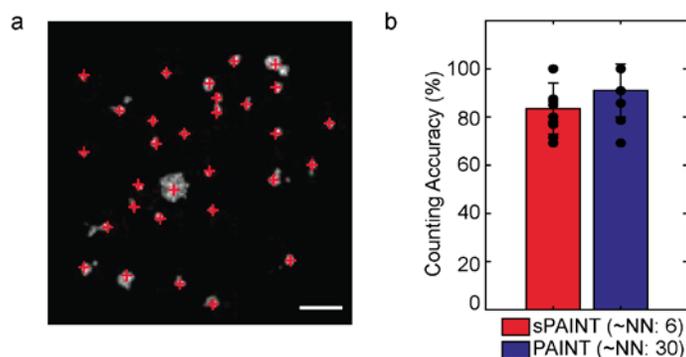


Figure 5-5: (a) Representative super-resolution reconstruction showing individual clusters with varying sizes identified by the optimized density-based spatial clustering for applications with noise (DBSCAN) algorithm marked by the red plus signs (Scale bar 500 nm). (b) Comparison of the differences in the cluster counting accuracy using DBSCAN for trials using data with different cluster densities (number of nearest neighbors or number of NN), the red bar shows data selected based on the spectroscopic and spatial information while the blue bar shows data selected for analysis based on spatial information alone.

We investigated the influence of non-specific binding on DBSCAN by comparing the number of clusters found by DBSCAN using localizations within SW (sPAINT) and DBSCAN without considering SW. For both cases, only bright localizations that contained high-precision spectral information (localizations at least 450 photons in the spatial domain and average spectral precision ~ 9.2 nm (122)). The photon distribution and related spectral precision information for the localizations used in the analysis are shown in Figure 5-6. The distribution of the spectral precision (Figure 5-6c) shows that the spectral precision of a typical localization used in this study would be sufficient to be separated from the non-specific binding with emission maxima populations ~ 30 nm away from the emission maximum of the sample. We plotted the histograms of the number of nearest neighbors (NN) with and without the SW filter for three different controls and three different PS images (Figure 5-7). We found that they followed similar distributions. We selected one representative control image and one representative PS image for setting clustering

thresholds. For sPAINT and DBSCAN, we selected min-NN thresholds by achieving the minimal number of clusters from the control sample.

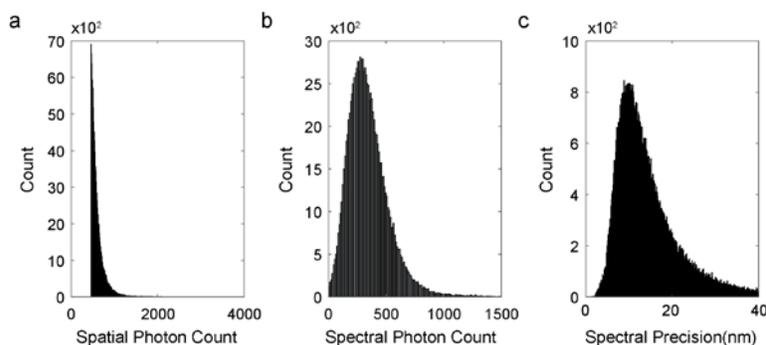


Figure 5-6: Histograms of the (a) spatial photon counts, (b) spectral photon counts and (c) spectral precision used for sSMLM analysis from a representative image.

To select min-NN thresholds we first confirmed the distributions of NN was the same for all datasets. For clarity, the min-NN threshold set based on DBSCAN alone is referenced by NN_C . Meanwhile, nearest neighbor thresholds based on sPAINT will be designated as NN_V . Figures 5-7a-c and 5-7d-f show the comparison between the histograms of NN for localizations in the control and PS samples with and without application of the SW. For the control and PS images the number of localizations was $\sim 10^5$ before the SW filter and $\sim 10^4$ the SW filter. By selecting localizations within the SW, the average number of NN reduced from 13 to 7 in the PS sample. Alternatively, in the control image the average number of NN reduced 5-fold from 10 to 2. This indicates that though the number of localizations was reduced for both groups, the application of the spectra-based criteria preferentially selected localizations which were members of clusters.

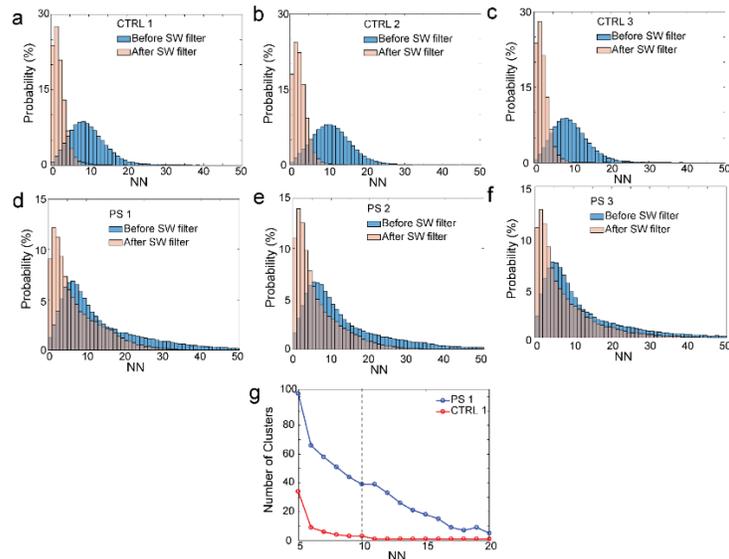


Figure 5-7: Histograms showing the comparison between the number of NN of localizations with 450 photons in the spatial domain before ($n \sim 10^5$) and after application of the selected spectral window ($n \sim 10^4$) of (a-c) 3 control images and (d-f) 3 PS images. (g) Plots showing the number of clusters using localizations within the SW as the min-NN increases in representative PS and control images. The black dashed line shows the threshold used in this study.

We then tested a range of thresholds in or to achieve the minimal number of clusters from the control sample. For sPAINT a range of min-NN_V from 5 to 20 were tested (Figure 5-7g). Meanwhile, for analysis based on DBSCAN alone, a range of min-NN_C from 15 to 50 was tested (Figure 5-8a). A threshold was selected to achieve the minimal number of clusters in the control without causing morphology changes in the PS image (Figure 5-8b). Using this threshold, localizations with too few neighbors were rejected as noise. For sPAINT, a min-NN_V of 10 was selected while a min-NN_C of 20 was selected for DBSCAN alone.

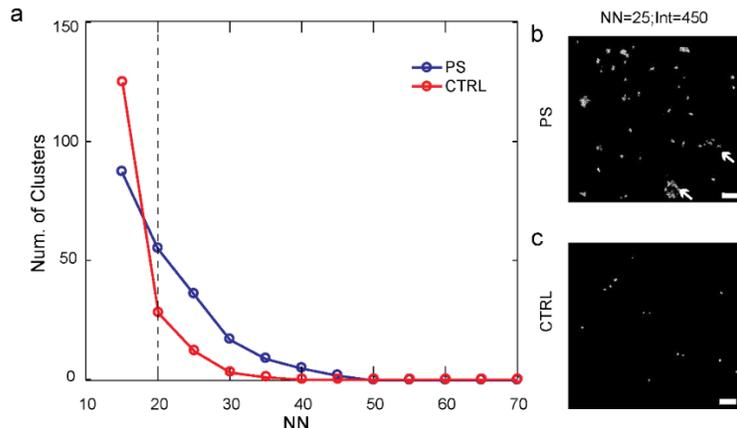


Figure 5-8: Plots showing the number of clusters with localizations with at least 450 photons in the spatial domain as the min-NN increases in the (a) PS and control images. The black dashed line shows the threshold used to compare spatial clustering without considering the PS window. Representative super-resolution reconstructions of the (b) PS sample (white arrows indicate changes in PS morphology) and (c) the control sample after application of the clustering thresholds (min- $NN_c = 30$ and Int=450). (Scale bar 500 nm)

Without using SW, we found that DBSCAN was sensitive to non-specific binding events. We found 55 and 28 clusters with averaged size of 126 nm and 90 nm in the PS and the control samples, respectively, when performing DBSCAN alone using a min-NN threshold of 20 (Figures 5-9a & 5-9b). In comparison, by first selecting signals within the SW, we found 39 and 3 clusters in the PS and the control samples respectively after DBSCAN using a min-NN threshold of 10 (Figures 5-9c & 5-9d). This shows an 89% reduction in the artifacts from non-specific binding in the control image. While stricter min-NN thresholds can be used to reduce the influence of non-specific binding using DBSCAN alone, we found that this often causes over-filtering (83, 135, 151) which presents an additional challenge since it can erroneously misrepresent the morphology of the sample (Figure 5-8b).

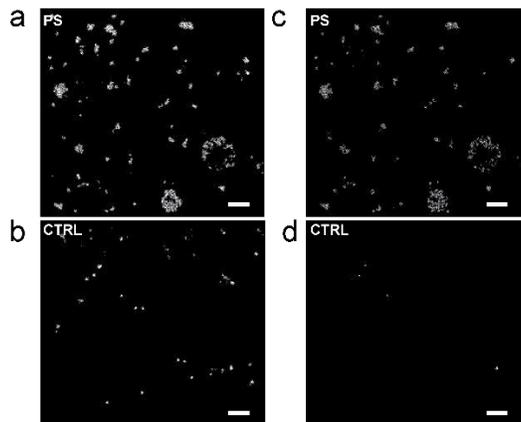


Figure 5-9: Clusters extracted from (a) the PS sample and (b) control based on DBSCAN alone; Clusters extracted from (c) the PS sample and (d) control using sPAINT (Scale bar: 500 nm).

sPAINT allowed us to image and quantify PS in aqueous environments. Particularly, it enabled the specific detection of NR to PS binding by analyzing the single-molecule emission spectra of each localization event. Meanwhile, sPAINT only used bright localization events (>450 photons in the spatial domain). As a result, clusters identified by sPAINT can have low localization densities (LD, number of localizations in a cluster within the cluster area), which might affect the visualization and size measurement for PS. Therefore, we further investigated the effect of LD on the size measurements and visualization of the PS in sPAINT imaging, and further developed a quantitative spectroscopic analysis for cluster extraction (qSPACE) post-imaging processing method to increase LD without compromising the specificity achieved by sPAINT in the next section.

Improving localization density using qSPACE

While sPAINT can readily image PS, we recognized that it rejected a large portion of the localization events (Figure 5-10). In fact, this is a common issue in all sSMLM because the limited

photon budget from single-molecule emission events is further split into the spatial and spectral channels (30, 31, 109). Although dim molecules (300-450 photons in the spatial domain) can still be localized at ~ 25 nm precision (Figure 5-10), they are discarded since λ_{\max} cannot be estimated precisely (estimated spectral precision worse than 15 nm (122)) due to their low photon budget.

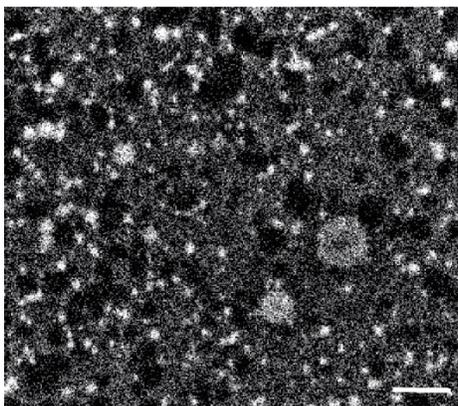


Figure 5-10: Localizations excluded from spectroscopic analysis based on photon number (Average localization uncertainty ~ 25 nm, Photon budget: 300-450 in the spatial domain)

The limited number of localizations retained by sPAINT could result in clusters with low LD, thus leading to inaccurate measurements of the size and morphology of the PS. Although sPAINT is immune to photobleaching, prolonging the acquisition time to increase the LD has the risk of altering the shape of nanocarriers (65, 155, 156). To address this issue, we developed qSPACE which improved the LD by using the spectroscopic signatures from the PS to guide the recovery of localizations without spectra.

The workflow of qSPACE is briefed in Figure 5-11a and a detailed workflow is shown in Figure 5-11b. First (Step i), all the blinking events from the spatial channel are localized after sPAINT imaging. Second (Step ii), localizations in SW with bright blinking events (>450 photons in the spatial domain) are selected and DBSCAN used to generate a high-fidelity validation map

of the spatial locations of the true PS sample. Third (Step iii), we perform DBSCAN to extract clusters using the localized events from the spatial channel. Then (Step iv), we compare the localizations identified by DBSCAN alone (Step iii) to localizations in the validation map (Step ii) to recover spatially correlated localizations. Localizations identified by DBSCAN alone, which have neighbors in the validation map, are retained and used for qSPACE image reconstruction, while localizations without neighbors are rejected. Finally (Step v), the new clusters with higher LD can be counted, sized, and visualized.

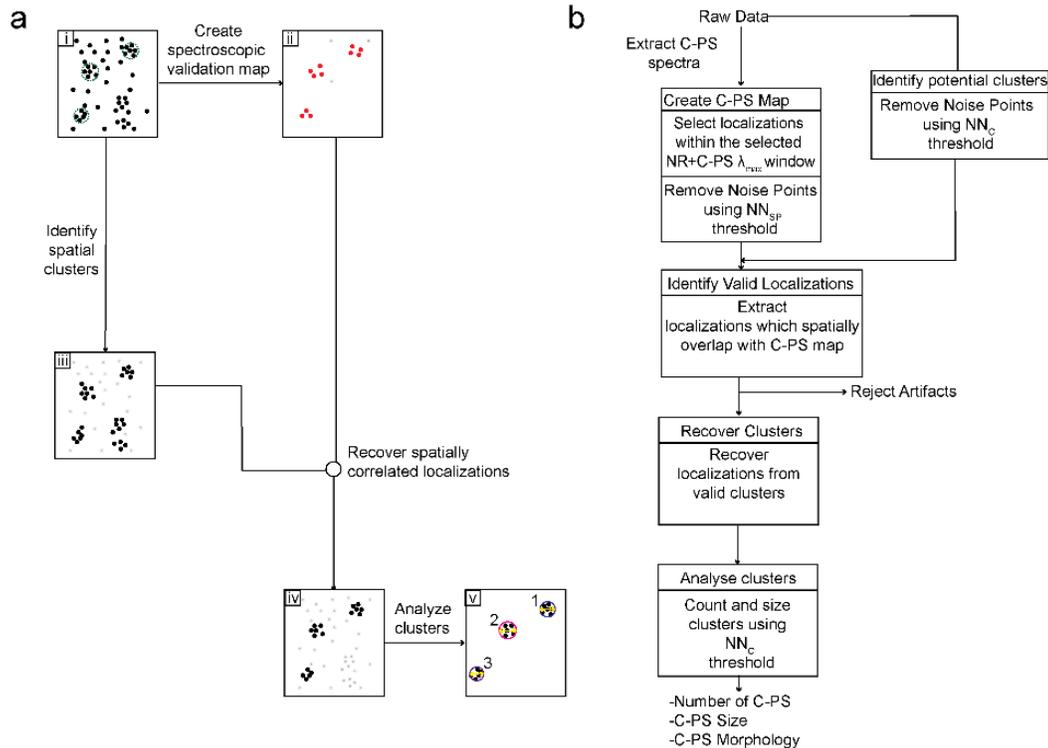


Figure 5-11: (a) The qSPACE workflow shows the (i) detected spatial localizations with the location of the sample highlighted by the green circles. A subset of localizations containing spectroscopic information is used to create a (ii) validation map that shows clusters with the selected spectra. All detected localizations are used for (iii) spatial clustering without considering spectroscopic information. (iv) Localizations from spatial clusters, which are spatially correlated with the validation map, are recovered for further analysis while artifacts are rejected. (v) The number, size, and morphology of the extracted sample can be further analyzed. (b) Detailed qSPACE flowchart outlines how spectroscopic information is used to recover localizations from clusters from the PS sample while rejecting artifacts from non-specific binding

We quantitatively compared the sizes of PS from a representative image using sPAINT and qSPACE. We measured the effect of the reduced number of localizations per cluster on size measurements by selecting 71 sPAINT clusters with at least 40 localizations. We calculated the percent variance ($P_{\Delta SIZE}$) as:

$$P_{\Delta SIZE} = \frac{|\bar{D} - D_n|}{\bar{D}}, \quad (5-1)$$

where \bar{D} is the final diameter of the cluster and the D_n is the estimated diameter for the cluster with n localizations for each cluster as the number of localizations sequentially increased. The average percent variance in cluster size reduced from 48% to <5% as the number of localizations increased (blue line Figure 5-12). This trend was further investigated by grouping the analyzed clusters based on their varying sizes. We found that the majority of clusters smaller than 100 nm could be accurately sized with ~25 localizations per cluster. Meanwhile, clusters larger than 200 nm required 38 localizations per cluster for accurate sizing.

We accounted for the variation in the size requirements for clusters of different sizes by using a threshold based on the localization density (LD)

$$LD = \frac{N}{\pi r^2}, \quad (5-2)$$

where N is the number of localizations in the cluster and r is the radius of the cluster. The red line in Figure 5-12 represents the LD for an average PS (with 113 nm diameter) as the number of localizations increased. Using the results from the size variation measurements we set a LD threshold of $3.5 \times 10^{-3} \text{ nm}^{-2}$ for accurate sizing.

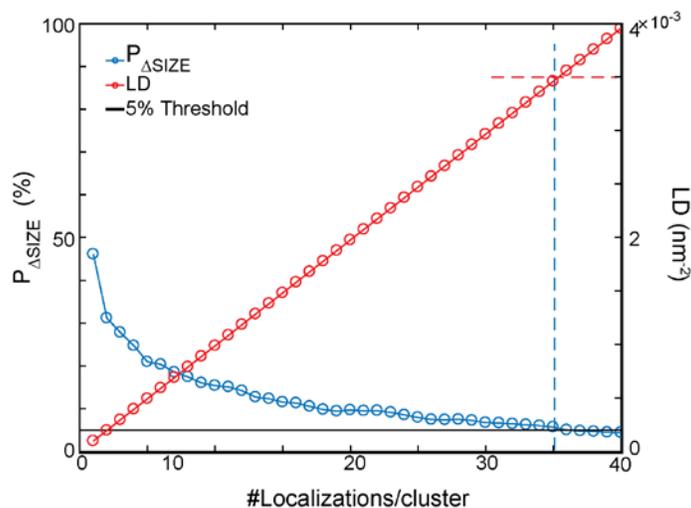


Figure 5-12: The average percent variance in size measurements ($P_{\Delta\text{SIZE}}$) and localization density for clusters as the number of localizations per cluster increases. The black line shows where $P_{\Delta\text{SIZE}} = 5\%$. The dashed blue and red lines indicate the min-NN of 35 and LD of $3.5 \times 10^{-3} \text{ nm}^{-2}$ required for accurate sizing.

We found that on average 35 localizations are required to reach a stable PS size. Since the PS ranged in size, we used the corresponding LD of $3.5 \times 10^{-3} \text{ nm}^{-2}$ as a threshold to quantify the number of PS with size variations exceeding 5%.

We tested the performance of qSPACE to reject artifacts from non-specific binding and identify PS clusters. We used the clusters identified using the sPAINT criteria previously described ($\text{SW} = [595-625]$, $\text{min-NN}_v = 10$) as a validation map. Potential clusters (PC) were identified using localizations using a min-NN_c of 35 to limit variations in size measurements to 5% (Figures 5-13 & 5-14). While traditional density-based algorithms would rely on stricter clustering threshold to remove potential artifacts, qSPACE uses a separate channel for sample validation. Using the validation channel, the specificity can be tuned by adjusting the min-NN_v independently (Figure 5-14). The spatial coordinates of localizations in the validation map are compared to the spatial coordinates of localizations which formed potential clusters. Localizations which did not have

neighbors in the validation map were rejected. All other localizations can be retained as members of true clusters. After validation, the min- NN_C (35) for selecting the potential nanocarriers threshold set was used for cluster assignment.

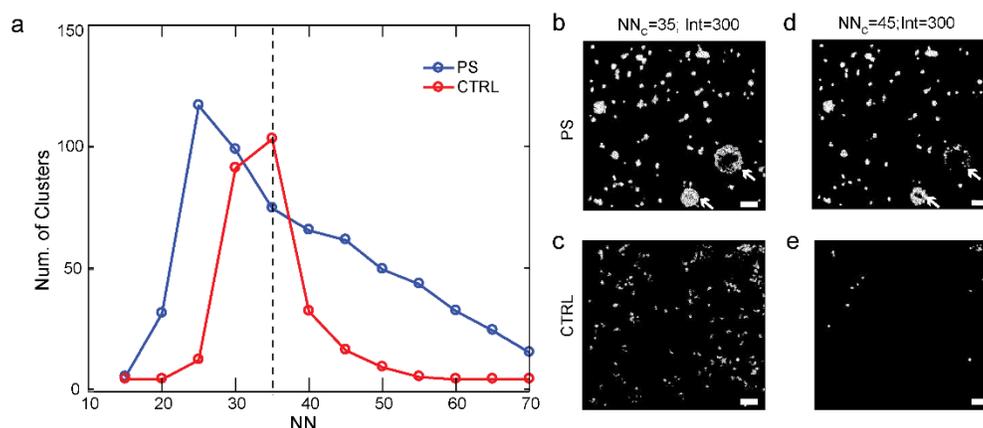


Figure 5-13: Plots showing the number of clusters at least 300 photons in the spatial domain as the minimum number of NN increases in the (a) PS and control images. The dashed black line represents the threshold used for detecting potential clusters for qSPACE. Representative super-resolution reconstructions of the (b) PS sample and (c) the control sample after application of the clustering thresholds (min- $NN_C=35$ and Int=300), (d) PS sample and (e) the control sample after application of the clustering thresholds (min- $NN_C=45$ and Int=300). The white arrows indicate morphology changes in d. (Scale bar: 500 nm)

Due to the improved LD, all 199 PS identified by qSPACE achieved the LD threshold (Figure 5-15a). In contrast, 23% of the 187 PS identified by sPAINT alone fail to meet the LD threshold for accurate sizing (Figure 5-15b).

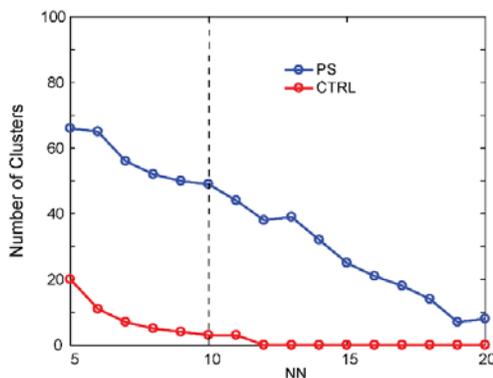


Figure 5-14: Plots showing the tuning of the min-NN_v for noise removal in the validation channel using qSPACE with a SW window of 595-625 nm and min-NN_c of 35 for potential cluster assignment. The dashed black line shows the selected min-NN_v.

Additionally, we used the Fourier ring correlation (FRC) to compare the image qualities of super-resolution reconstructions of PS identified using sPAINT alone with further qSPACE processing in a representative image. Although the localization precisions for both methods are comparable (sPAINT: ~20 nm; qSPACE: ~23 nm), the differences in LD affects image resolution (102). We calculated the global FRC resolution that takes LD into consideration as previously reported (102, 139). The FRC resolution is calculated as the inverse of the spatial frequency at a 1/7 of FRC (102). We found that qSPACE had a relatively higher FRC resolution of 69 nm while sPAINT had an FRC resolution of 96 nm, as shown in Figure 5-15c.

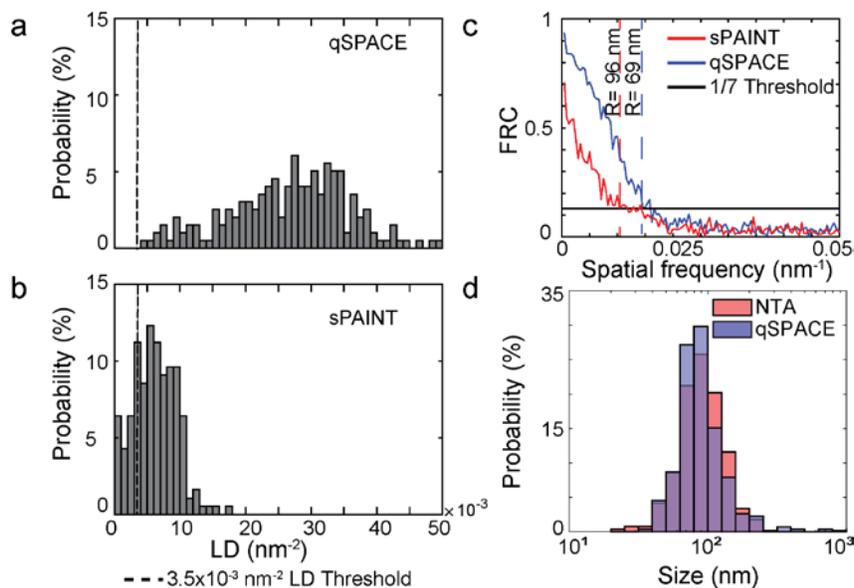


Figure 5-15: (a) Histogram of the localization densities (black dashed lines indicate the $3.5 \times 10^{-3} \text{ nm}^{-2}$ LD threshold) for qSPACE and (b) sPAINT; (c) Comparison of the global FRC curves for sPAINT (red) and qSPACE (blue). The dashed lines are the corresponding FRC resolution for each method at the 1/7 FRC threshold; (d) Size distribution of PS (0.02 mg/mL) measured by NTA and qSPACE.

Figure 5-15d compares the size distribution of 265 PS clusters (0.02 mg/mL) in 4 different images measured by qSPACE and ~5000 PS measured by NTA using 5 measurements. The average size of PS measured by qSPACE are $96.6 \pm 2.1 \text{ nm}$. These results agreed well with the average size measured by NTA ($97.6 \pm 1.4 \text{ nm}$). In addition, the cryoTEM image of PS (Figure 5-16) confirmed that the size of PS ranged from 40 nm to 200 nm. X-ray scattering analysis also showed consistent results with a core radius of $\sim 106 \text{ nm}$ and shell thickness of $\sim 14.3 \text{ nm}$ (7). Notably, PS larger than 200 nm, which resulted from the crosslinking process (7), were detected by both NTA (3%) and qSPACE (4%). These validations suggest that qSPACE provides more accurate size measurement of PS and higher resolution than sPAINT alone.

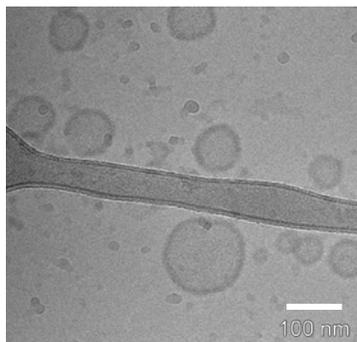


Figure 5-16: CryoTEM image of PS sample.

Quantitative characterization of PS at various concentrations

Finally, we used qSPACE to visualize, size, and count PS at different concentrations. Figures 5-17a-c show representative super-resolution reconstructions of three samples containing 0, 0.02, and 0.2 mg/mL of PS, referred to as control (CTRL), low concentration (LC) and high concentration (HC) samples, respectively. The NR concentration remained constant in these three samples. In the CTRL image (Figure 5-17a), the clusters were predominantly identified as artifacts from non-specific bindings (red color). In the reconstructed LC and HC images (Figures 5-17b & 5-17c), we identified 60 and 199 PS (cyan color), respectively. For statistical analysis, we performed four measurements using different areas of each sample. Figures 5-17d-f respectively show the size distributions of CTRL, LC and HC samples. Using qSPACE the misidentified PS in the control samples have an average size of 63 ± 7.7 nm. The average sizes of identified PS are 137.3 ± 4.8 nm and 96.6 ± 2.1 nm in the HC and LC samples, respectively. Presumably, aggregation of PS occurs more frequently in the HC samples, thus, resulting in larger PS size.

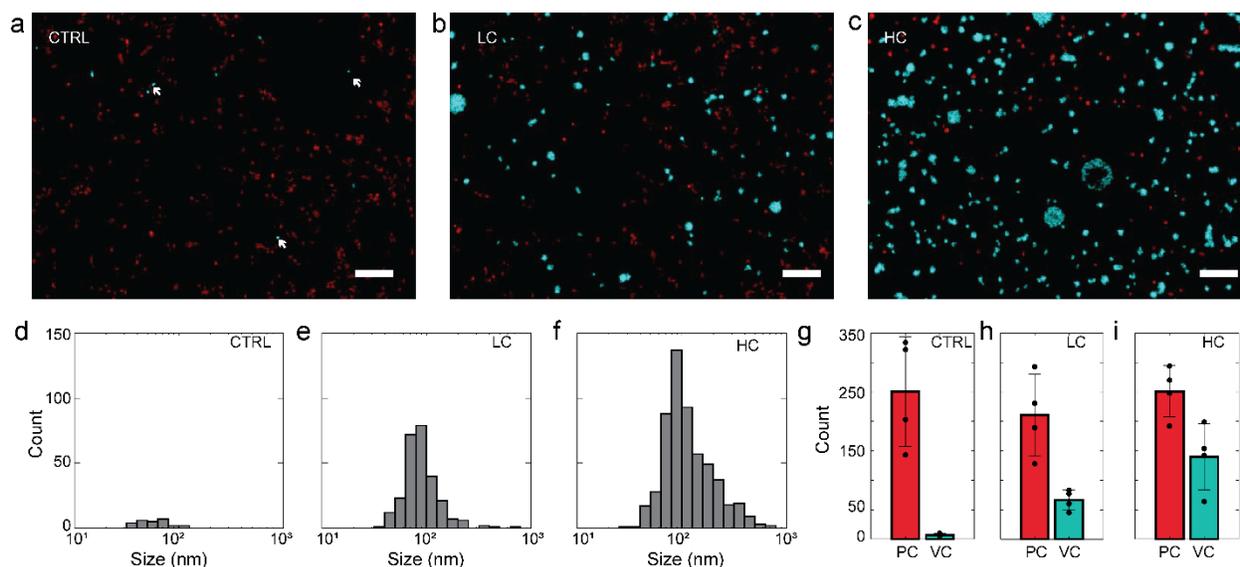


Figure 5-17: Representative super-resolution reconstructions of the (a) CTRL sample with artifacts pseudo-colored in red and misidentified PS pseudo-colored in cyan. The three white arrows highlight three examples of sample misidentification; (b) LC sample (0.02 mg/mL) with artifacts pseudo-colored in red and validated PS pseudo-colored in cyan; and (c) HC sample (0.2 mg/mL) with artifacts pseudo-colored in red and validated PS pseudo-colored in cyan (Scale bar: 1 μ m). The size distributions for the (d) misidentified PS in the CTRL samples; (e) validated PS in the LC samples; and (f) validated PS in the HC samples. Comparison between the total of number of PC found and the qSPACE VC in the (g) CTRL samples, (h) LC samples, and (i) HC samples.

The counting of PS confirmed the improved performance of qSPACE compared to the conventional DBSCAN method. Particularly, 251 ± 93 PC (red bar in Figure 5-17g) were extracted in the control sample by DBSCAN without spectral analysis while only 7 ± 3 clusters or 3% of these clusters were misidentified as PS by qSPACE (validated clusters or VC, cyan bar in Figure 5-17g). The average size of the clusters identified as artifacts (clusters correctly rejected by qSPACE) from the control sample was 120.1 ± 4.1 nm. Clusters identified as artifacts in the LC and HC samples had similar distributions (see Figure 5-18 for the size distribution of artifacts identified by qSPACE and sPAINT). Additionally, 66 ± 18 of 210 ± 70 and 138 ± 56 of 251 ± 44 PC were validated as PS in the LC and HC samples respectively using qSPACE (Figures 5-17h & 5-175i).

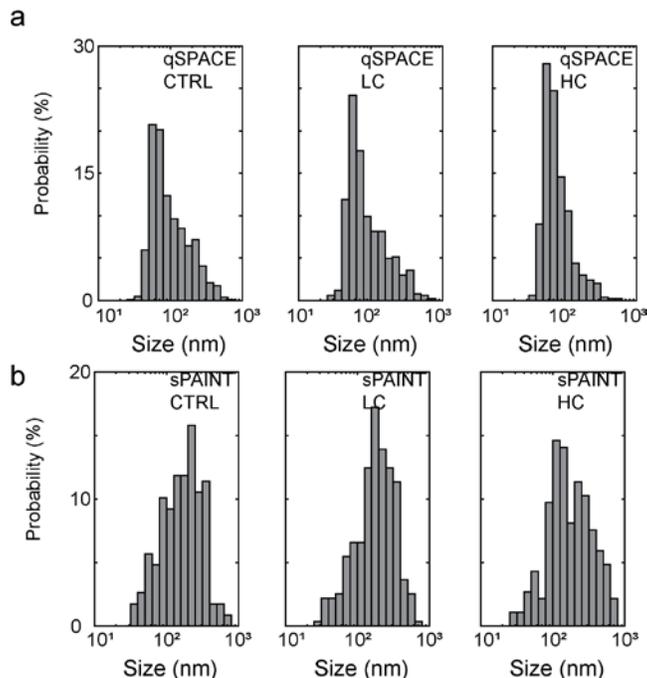


Figure 5-18: (a) Histograms showing the size distribution of artifacts identified from 4 reconstructions of the control sample (809 clusters with average size 120 ± 4 nm) and clusters identified as artifacts in the LC (504 clusters with average size 111 ± 5 nm) and HC (501 clusters with average size 90 ± 7 nm) PS samples by qSPACE. (b) Histograms showing the size distribution of artifacts identified in 4 reconstructions of the control sample (229 clusters with average size 199 ± 2 nm) and clusters identified as artifacts in the LC (273 clusters with average size 220 ± 11 nm) and HC (186 clusters with average size 211 ± 8 nm) PS samples by sPAINT (clusters formed from localizations with λ_{\max} outside the SW filter) .

We compared the relative number of PS in LC to HC using qSPACE and NTA. NTA analysis of the two samples indicates a 5-fold higher concentration in HC than LC sample. Using qSPACE, the concentration ratio between HC and LC is ~ 2.1 . The difference in concentration measurements using qSPACE and NTA may come from differences in the sample preparation steps. While NTA measured the entire solution, the samples measured by qSPACE relied on the adhesion of the PS to the PLL substrate. Rinsing steps could also change the relative concentrations

of the sample. In short, these results suggest that qSPACE can be used to accurately size and detect the relative trend in the concentration variations of nanocarrier samples.

5.3 Conclusions

We outlined a strategy to quantitatively image and characterize nanocarriers with high specificity, molecular sampling, and ~20 nm localization precision. Using sPAINT with NR dye, we successfully imaged PS nanocarriers immobilized on PLL substrates and rejected 83% of non-specific binding events by identifying the unique spectral signature of NR upon specific binding to PS. We further developed qSPACE to recover the discarded dim localizations in sPAINT to improve LD by 4-fold, reduce variations in size measurements to 5%, and improve the FRC resolution by 30%. Using qSPACE to count and size PS at various concentrations we reduced misidentification of the PS by 97% and validated our results with standard size measurements. qSPACE is not limited to spherical shape clusters and can be used to process distinct morphology. The continued development of fluorescent probes, optical systems, and imaging processing methods for sPAINT would enable the quantitative visualization of polymer assemblies and dynamics in vitro.

5.4 Methods

Polymersome preparation

Polymersomes were fabricated based on the self-assembly of amphiphilic brush block copolymers poly (oligo (ethylene glycol) methyl ether methacrylate)-b-poly (oligo (propylene sulfide) methacrylate) (POEGMA-POPSMA), which were synthesized based on reversible addition-

fragmentation chain transfer (RAFT) (7). POEGMA-POPS₇MA₈ PS were assembled using the thin-film hydration method as previously described (7). Briefly, 20 mg of POEGMA-POPS₇MA₈ copolymer was dissolved in 150 μ L dichloromethane within 1.8 mL clear glass vials (ThermoFisher). After desiccation to remove the solvent, the resulting thin films were hydrated in 1 mL of phosphate-buffered saline (PBS) under shaking at 1500 rpm overnight. The suspension was then extruded through a 0.2 μ m membrane filter. The cross-linked PS were prepared by the reaction of uncross-linked PS suspension (20mg/ml) with 1,2-ethanedithiol (36 μ L, 0.15M in ethanol) through thiol-disulfide exchange reactions under shaking (1500 rpm) at room temperature. The obtained PS were then purified by Zeba Spin Desalting Columns (7K MWCO, ThermoFisher) equilibrated with PBS solution.

Nile Red preparation

A 3 mM NR (ThermoFisher) stock solution was prepared in dimethylsulfoxide (DMSO) (Sigma Aldrich). For sPAINT imaging the solution was further diluted to 5 nM in PBS and used freshly.

Bulk fluorescence measurements

The bulk fluorescence emission spectra of 1 mL solutions containing 1 μ M of NR mixed with 100 μ L of PS (20 mg/mL in PBS), PLL solution (0.01% w/v, Sigma Aldrich), and PBS, respectively, were measured using a fluorimeter (SpectraMax M3, Molecular Devices). Measurements were recorded using 532 nm excitation.

Sample immobilization

No. 1 borosilicate four-well Lab-Tek Chambered coverglass were rinsed three times using PBS then filled with 500 μ L PLL solution (0.01% w/v, Sigma Aldrich P4707) and incubated for 30 min to coat the glass surface. The wells were then rinsed three times with PBS. PS was diluted to 0.2 mg/mL (high concentration sample) and 0.02 mg/mL (low concentration sample) using PBS. Mixtures of 500 μ L of the HC sample, LC sample along with 500 μ L of PBS (control sample) were transferred to separate 1 mL conical tubes. The three samples were each sonicated for 5 min. The samples were then added to separate wells and incubated for 5 min. The samples were rinsed using PBS. Finally, 500 μ L of freshly prepared 5 nM NR in PBS was sonicated for 5 mins and then added to each well for imaging. We also tested bovine albumin serum (BSA) and neutravidin (NEU) for sample immobilization chemicals used for the immobilization of biotinylated molecules. However, bulk measurements showed that this immobilization methods could provide a larger number of sources for sample misidentification (see Figure 5-19).

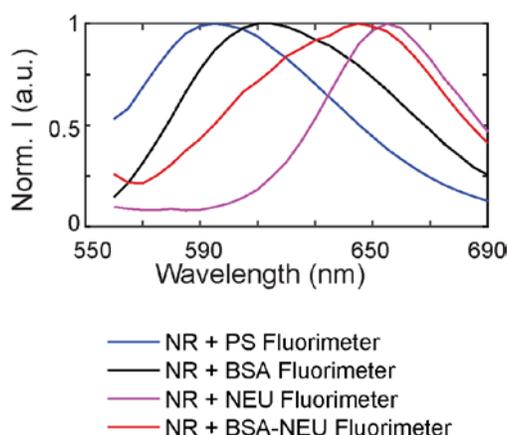


Figure 5-19: The average emission spectra of NR in solutions containing PS and sample preparation reagents (BSA and NEU) measured using a fluorimeter.

sSMLM optical setup

The optical setup contains an inverted optical microscope (Eclipse Ti-U with perfect-focus system, Nikon), equipped with a 532-nm Spectra Physics laser with 200-mW maximum output, a high numerical aperture total internal reflection (TIRF) objective lens (100 \times , NA 1.49, Nikon CFI TIRF apochromat) and home-made transmission spectrometer. The fluorescence was collected through the objective and then directed into a spectrometer consisting of a blazed transmission diffraction grating (300 grooves/mm, Edmund Optics) and two imaging lenses with focal lengths 100 mm and 80 mm respectively. 60% and 40% of the emitted photons were respectively allocated for the non-dispersed zeroth order channel and for the spectrally dispersed first order channel and recorded simultaneously on different regions of an electron multiplying charge-coupled device (EMCCD, ProEm HS 512X3, Princeton Instruments).

Imaging procedure and data analysis

The optical system was calibrated using a fluorescent lamp with the entrance slit adjusted to 10 μm . The emission peaks at 436.5 nm, 487.7 nm, 546.5 nm, and 611.6 nm were fit with a first order polynomial to calculate the system's spectral dispersion (4.3 nm/pixel). Samples were placed on the microscope stage and imaged using an angle slightly smaller than the TIRF mode. We used a 532-nm laser with a power density of 1 kW/cm² to initiate transient NR binding to PS. An EMCCD camera acquired images with an integration time of 30 ms. Four areas of the sample were randomly selected and imaged for each sample. For each FOV 50,000 frames were recorded and the spatial image was reconstructed using ThunderSTORM (121). ThunderSTORM was used to perform drift correction for each sample prior to spectral analysis. The drift correction file was saved and used

to correct all data prior to cluster analysis (109, 121). Events from overlapping binding events were removed. Additionally, localizations below the average photon count (300) in the spatial image were excluded.

sPAINT window selection

Spectroscopic analysis was performed using localizations with a minimum of 450 photons in the spatial domain using an $8 \mu\text{m} \times 7 \mu\text{m}$ section of three different control and three different PS super-resolution reconstructions. The corresponding spectroscopic signature for each localization in the spatial channel was extracted from the first-order images. The emission maximum of each spectroscopic signal was identified by fitting the normalized spectra to a gaussian (104). Localizations with emission maximum outside the calibration window (550 - 700 nm) were excluded. The histograms from 18 different PS from three different images (six PS per image) are shown in Figure 5-2c-e. Examples of emission maximum from three ROIs are shown in Figure 5-3b. Additionally, the histogram of the emission maximum from the control sample was compared to the PS histograms and a SW (595 - 625 nm) for recognizing NR interactions with PS and non-specific binding events was selected. 300 localizations with more than 500 photons in the spectral domain from the isolated PS were used to plot the average spectra in Figure 5-2a. Similarly, 300 localizations with at least 500 photons in the spectral domain and $\lambda_{\text{max}} > 645 \text{ nm}$ were used to plot the average spectra for NR interactions with PLL.

Cluster analysis

We adopted a DBSCAN algorithm optimized for clusters with heterogeneous cluster densities (93). Individual PS were identified using a density-based algorithm with a cluster radius of twice the average localization uncertainty (40 nm). NN analysis was performed to determine the minimum number of acceptable localizations within the cluster radius. Localizations that did not meet the minimum criteria were removed. The remaining localizations were sorted based on their NN, Euclidean distance between neighboring points, and spatial coordinates. To consider the influence of the cluster density on the analysis a reachability criterion is used. If localizations are members of the same cluster then the variation in the distance between those points will be less than the variation in the distance between points which are members of a different cluster. Using this sorted list, the distance to the next localization in the list is used to identify core and boundary (93) localizations of each cluster within a maximum reachable distance of 1 μm (based on the maximum PS size (7)). Localizations which were members of a separate cluster were identified when the variation in the distance to the next localization in the sorted list exceeds twice the standard deviation of reachable localizations. The diameter of each cluster was estimated as twice the average distance between the core location and all other localizations in the cluster. To select the min-NN threshold, a range of potential thresholds were tested using both the PS and control samples. The number of clusters resulting from the control relative to the PS sample was used to select the min-NN threshold. Counting accuracy for the algorithm using clusters with varying densities was estimated by comparing the reported clusters by the algorithm to clusters manually counting in 10 different $2 \mu\text{m} \times 2 \mu\text{m}$ regions of interest in a representative super-resolution reconstruction.

Quantitative sPAINT analysis

Cluster analysis using sPAINT relied both on the spectral and spatial information for localizations with at least 450 photons in the spatial domain. The localizations were then further filtered using a min-NN threshold. The performance of spectral analysis paired with spatial clustering was tested by counting the number of clusters in $5\ \mu\text{m} \times 5\ \mu\text{m}$ sections of representative PS and control super-resolution reconstructions with and without applying the SW (595 - 625 nm). A min-NN of 10 was selected for cluster analysis after application of the SW and a min-NN of 20 was selected for cluster analysis without applying the SW. The number and size of artifacts found by sPAINT (Figure 5-18b) were identified using localizations which were not in SW and not spatially correlated with any of the clusters using the SW filter.

Size variation analysis

We measured the effect of the reduced number of localizations per cluster on size measurements by selecting 71 sPAINT clusters with at least 40 localizations. We calculated the difference between the final size of the cluster and the size of the cluster as the number of localizations sequentially increased. The percent variance in the size was calculated as the ratio of the difference in size measurements as the number of localizations increased and the final size of the cluster. This trend was further investigated by grouping the analyzed clusters based on their varying sizes. To account for the different sizes of PS, we set a criterion for evaluating size measurements based on the LD (the number of localizations in a cluster divided by the PS cluster area detected using sPAINT).

qSPACE validation

The validation map for qSPACE was created using the localizations after applying the SW (595 - 625 nm) and removing isolated localizations (min-NN of 10). Potential spatial clusters were identified by applying a weak min-NN threshold of 35 to all localizations with at least 300 photons in the spatial domain. The spatial coordinates of the localizations, which formed spatial clusters were compared to those in the validation map. Localizations that did not have any neighbors in the validation map marked as artifacts. The remaining spatial localizations were used for counting and sizing as described in the cluster analysis section. For the qSPACE super-resolution reconstructions in Figure 5-16, all clusters were pseudo-colored based on their classification. The final qSPACE analysis was performed using 12 different $11 \times 8 \mu\text{m}^2$ images acquired from 3 samples. Four measurements were captured from each sample.

Nanosight measurements

The size distribution of a $1:10^4$ dilution of the stock PS solution was measured using nanoparticle tracking analysis (Nanosight NS300, Malvern). We also tested the relative concentration of the high concentration and low concentration samples by testing $1:10^3$ dilutions of the samples used in the sPAINT experiment. All Nanosight measurements were performed using an unlabeled sample and a 488-nm excitation source and five trials were performed statistical analysis using NTA.

CryoTEM measurements

Cryogenic transmission electron microscopy (cryoTEM) of the PS samples were prepared as previously described (7). 4 μL of a 5 mg/mL sample was applied to a pretreated, lacey carbon 400 mesh TEM copper grid (Electron Microscopy Sciences). The grids were plunge-frozen with a Gatan Cryoplunge freezer at room temperature. Images were collected using a JEOL 3200FSC transmission electron microscope.

CHAPTER 6: Quantitative sSMLM #3: Improving image resolution using Photon-Accumulation Enhanced Reconstruction

The spatial resolution in single-molecule localization microscopy (SMLM) is limited to around 20 nm due to the physical photon limit of individual stochastic single-molecule emissions, where all emissions are treated as independent events. Using spectroscopic SMLM (sSMLM), we observed that single-molecule emissions from the same molecular species exhibit detectable spectral heterogeneity, which establishes the foundation for molecular identification. Taking advantage of such spectral heterogeneity, we developed photon-accumulation enhanced reconstruction (PACER) for sSMLM to break the physical photon limit by accumulating photons over repeated stochastic emissions from the same dye molecule through spectral identification. Using PACER, we experimentally demonstrated a 1.7-nm localization precision and resolved quantum dots that were 6.1 nm apart by utilizing their high spectral heterogeneity (SH). We further validated the localization precision using two types of Alexa Fluor 647 (AF647) labeled DNA origami nanostructures with known feature size as small as 6 nm. This chapter is adapted from a manuscript in preparation for submission to *PNAS*.

6.1 Introduction

Pushing the limit of spatial resolution in optical nanoscopic imaging gained significant momentum since the award of the 2014 Nobel Prize in Chemistry to single molecule analysis and super-resolution imaging. Super-resolution imaging technologies, such as stimulated emission depletion (STED) microscopy and single-molecule localization microscopy (SMLM), surpassed the

diffraction limit with demonstrated resolution of 10-20 nm (17, 18, 20, 157, 158). In particular, SMLM technologies, including photoactivated localization microscopy (PALM), stochastic optical reconstruction microscopy (STORM), and point accumulation for imaging in nanoscale topography (PAINT), employ a stochastic switching or binding approach to regulate the density of fluorescent blinking within a recorded image frame. Accumulating the precise localization of individual single molecule fluorescence events from thousands of recorded image frames allows for the reconstruction of a super-resolution image. Despite being widely successful, the prerequisite stochastic process of single molecule fluorescence often produces a rather limited number of photons, which practically restricts the localization precision of SMLM techniques to around 20 nm.

To further improve the localization precision, DNA-PAINT and MINFLUX have been developed (58, 59, 159, 160). DNA-PAINT exploits transient binding between a docking strand and dye-conjugated imager strands to replace the stochastic switching process to eliminate the influence of photobleaching and allow the use of brighter non-blinking dyes. DNA-PAINT is capable of imaging discrete arrangements of fluorescent molecules with ~5 nm spacing on synthetic DNA nanostructures (159, 160) but more sophisticated labeling strategies are needed for imaging biological samples. On the other hand, MINFLUX relies on prior knowledge of the donut-shaped illumination to triangulate the spatial location of individual emitters with greatly improved precision of ~2 nm, but at a reduced field of view (FOV) (58, 59).

In fact, commonly used blinking dyes produce a large number of photons over the repeated occurrence of stochastic blinking events. However, due to the lack of molecular discrimination, the repeated stochastic blinking from the same molecule are all treated as independent events with

limited localization precision around 20 nm. Thus, a method to identify the molecular origin of the stochastic blink events and subsequently accumulate the emitted photon from the same dye molecule can significantly improve the localization precision, without the need to reinvent the labeling and imaging protocols. To illustrate this concept, let's consider a single fluorescent molecule located at the center of the $1 \times 1 \mu\text{m}^2$ FOV as a representative example (Figure 6-1a). The diffraction-limited imaging system forms a blurred image (Figure 6-1b), which can be mathematically represented by a point spread function (PSF) with a typical diffraction-limited resolution of ~ 200 nm (Figure 6-1c). The calculated centroid location of the recorded PSF is used to approximate the true location of the fluorescent molecule with a precision greatly exceeding the diffraction-limited resolution (44, 161, 162). By programmatically switching the majority of the fluorescent molecules to the "off" state, SMLM detects a small subset of the fluorescent molecules that are stochastically restored to the "on" state. The centroids of the sparsely distributed PSFs from a subset of molecules can be individually localized. These localized events are then accumulated over thousands of recorded image frames to build-up a super-resolution image. The localization precision in this process is inversely proportional to the square root of the photon number for individual blinking events (61). Figure 6-1d shows a simulated SMLM image of 100 repeated stochastic blinking events from the same molecule, with 200 photons per stochastic blinking (predefined by the NA 1.49 objective at the wavelength of 550 nm) with a pixel size of 100 nm (see Figure 6-1b). Then we superimposed Poisson noise backgrounds on these images to mimic the shot noise of a camera. The total photon count and noise level were adjusted to match the experimental conditions which can result in a localization precision of approximately 30 nm. We then performed reconstruction using a standard SMLM algorithm using the ThunderSTORM

plug-in in ImageJ (121). After reconstruction, positions of all the collected localizations were plotted as white crosses as shown in the upper row of Figure 6-1j.

Conventional SMLM fits a Gaussian kernel to the diffraction limited PSF and approximates the centroid location to improve image resolution. To give the most basic example, we first consider a single fluorescent blinking event from a single fluorophore imaged by conventional SMLM. After capturing a single blinking event on the detector, depending on the detected photons, system noises, pixel size, and other factors, the location of the single molecule can be approximated, with a certain precision of the Gaussian kernel fitting. The localization precision (σ) depends on the photon counts in this particular frame and can be approximated by the following equation (61, 62):

$$\sigma = \sqrt{\left(\frac{s_i^2 + a^2/12}{N}\right)\left(\frac{16}{9} + 4\tau\right)}, \quad (6-1)$$

where s_i is the standard deviation of the Gaussian fit in the x and y -directions; a is the pixel size of the CCD camera; N is the number of detected photons; $\tau = \frac{2\pi b^2(s_i^2 + a^2/12)}{Na^2}$; and b is the standard deviation of the camera background. As a result, the localization precision is proportional to the inverse square root of the number of detected photons.

Since many molecules can blink multiple times during the imaging process, there is a chance that the same molecule blinks at a later time and can be captured by the detector. In this scenario, the later blinking events from the same molecule will be reconsidered by the SMLM algorithm and will also be plotted to the image. However, since the parameters that were used to approximate the location of the single molecule have variations (different background, different

photon yield, etc.), the possible location calculated by the Gaussian kernel fitting can be slightly different from the original location of the first blinking event.

Despite the fact that the photons originated from the same molecule, each stochastic blinking event is individually localized at photon-number limited precision of 30 nm (Figures 6-1e-f). Since all the photons are originated from the same molecule, in principle, repeated blinking events can be aggregated to accumulate a much larger photon number to improve the localization precision as compared to localization precision being limited to each event (Figures 6-1g-i). Figure 6-1j shows a simulation results on how accumulating photons from different numbers of repeated blinking events can improve the localization precision. Their centroid, which represents the result of PACER, was further plotted as a red cross. In the middle row of Figure 6-1j, we calculated 500 randomly generated cases and plot all centroids by convolving a Gaussian kernel corresponding to the localization precision calculated from the accumulated photon number. We further plot the profile along the x-axis as shown in the lower row of Figure 6-1j. When NB=1, the result shows the precision of conventional SMLM (28.57 nm) since PACER was not used. As NB increases, the localization precision improves, enhancing the resolution to 2.85 nm (10-fold improvement) when NB=100. Figure 6-1k shows the calculated localization precision with respect to NB. In conventional SMLM, the localization precision of each blinking event scales with the inverse square root of photon counts (61). Therefore, after PACER, the fundamental resolution can be redefined by using the total photons emitted from each single molecule. The analytic curve has great agreement with the simulated results. Based on this principle, we further plot the curve calculated by

$$P_{PACER} = P_0/\sqrt{NB}, \quad (6-2)$$

where P_0 is the localization precision at $NB=1$ (Figure 6-1i). This trend suggests that PACER indeed improves the localization precision by increasing the photon counts for image reconstruction.

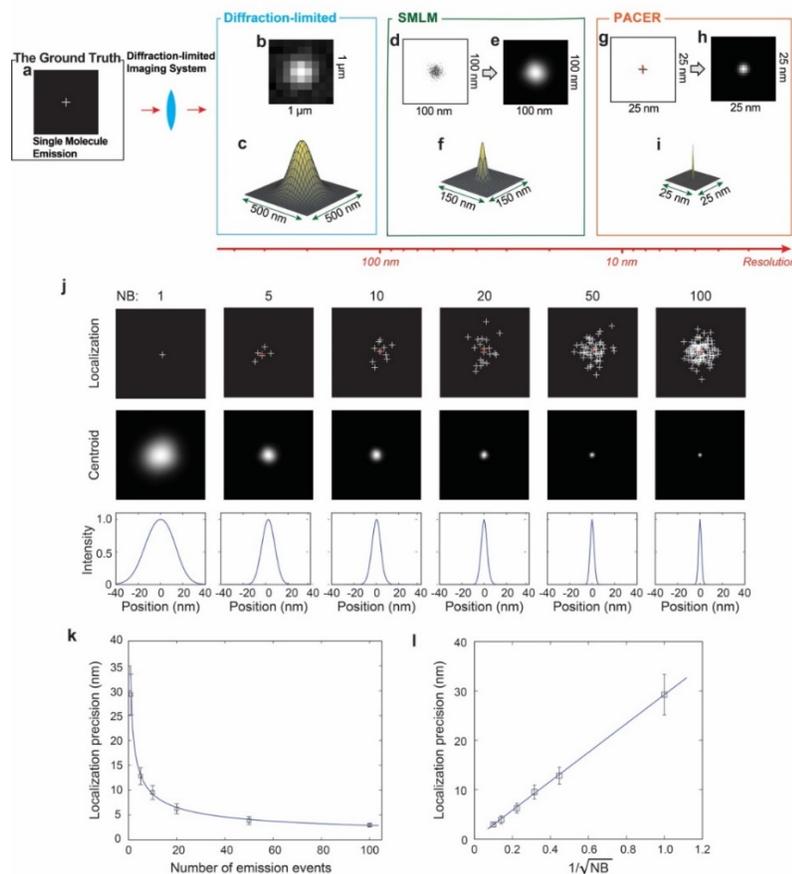


Figure 6-1: Schematic showing how image resolution is improved by molecular discrimination. (a) Due to the wave nature of light, when light comes from a point emitter focused by an optical imaging system, the interference can result in a blurred distribution of light called a PSF. (b-c) The size of the PSF sets up the fundamental resolution limit of an optical imaging system, namely Abbe diffraction limit. (d-f) In SMLM, the probable location of a single fluorescent molecule can be estimated from the centroid of the fluorescence diffraction pattern produced on a camera. The localization precision is determined by the photon count collected in each captured frame. The scatter plot and rendered image of 100 blinking events indicate a localization precision of ~ 30 nm. (g-i) Through molecular discrimination, blinking events can be combined, resulting in an improved localization precision of sub-3 nm. (j) Improved resolution using photon accumulation with respect to the number of blinking events (NB). White crosses denoted positions of all localizations in each test and red crosses denoted their centroids (upper panel). Reconstructed images only using centroids (middle panel). Localization precision is calculated along the x-axis and its intensity was then normalized for comparison, which are 28.57 nm, 13.24 nm, 9.02 nm, 6.40 nm, 3.89 nm, and 2.85 nm for NB=1, 5, 10, 20, 50, and 100, respectively (lower panel). (k) Simulated localization precision with respect to NB. Error bars are from 100 independent simulations. (l) The black line is the curve calculated by σ_0/\sqrt{NB} , where σ_0 is the localization precision at NB=1.

We simulated 100 images containing a single blinking event with 200 photons using the same method (see Figure 6-2a). The total photon count and noise level were adjusted to match the experimental conditions which can result in a localization precision of an approximately 30 nm. We then performed reconstruction using a standard SMLM algorithm. The reconstructed position of each blinking event was analyzed and plotted in the lower panel. In this case, the spatial precision is calculated to be around 30 nm if all blinking events are considered separately, which is consistent with the predefined simulation condition.

If we assume all 100 blinking events originate from the same molecule, then PACER method can be implemented to accumulate emitted photons to the improved spatial resolution. The accumulation of photons can be implemented in two alternative procedures. In option 1, we combined the photos from all 100 blinking events to create an accumulated image with total photon count of 20,000 (Figure 6-2b), which can result a much-improved localization precision (approximately 10x improvement). After performing reconstruction, the result was plotted in Figure 6-2c. However, it is worthwhile to note that this procedure is prone to the stage drifting error. To overcome this issue, we developed option 2 where all 100 blinking events (Figure 6-2d) are first localized individually. This allows for compensating stage drifting error in reference to the fiducial marker. The true location of the molecule can then be determined by calculating the centroid of all 100 localization results (Figure 6-2e). As shown in Figure 6-2f, the centroid of the combined localizations after taking option 2 matched the location of the emission events after the reconstruction based on path 1. This demonstrates that option 1 and 2 are mathematically equivalent when they share the same number of photons and background level.

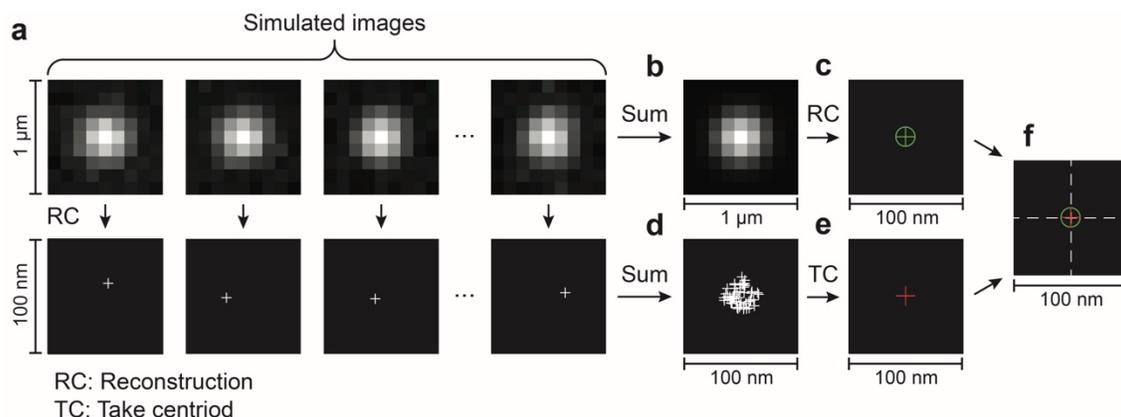


Figure 6-2: Illustration of two different paths to combine all photons from the same molecule. Simulated stochastic emission events from the same single molecule (a) Simulated diffraction-limited spatial images and their respective reconstructed locations. (b) Sum of the photons in the diffraction-limited image (c) Location of the emitter after combining all photons from different blinking events. (d) All single molecule localizations are combined after reconstruction. (e) The centroid of all localizations from the same molecule. (f) The location of the emission events after the reconstruction based on option 1 overlaps with the centroid of the combined localizations after taking option 2.

Accumulating photons to improve localization precision can be accomplished if the true origin of the stochastic blinking can be specifically identified. This can be implemented via spatial clustering when the fluorescence molecules are sparsely distributed. As previously demonstrated, aggregating photons over prolonged detector exposure time can reach a localization precision of 1.5 nm in well-separated single molecule tracking (44). However, imaging biological systems often requires much higher molecular labeling densities and thus, spatial locations alone become insufficient for identifying the sources of emitted photons to specific individual molecules among a densely packed ensemble.

In addition to their spatial locations, the fluorescence spectra of emission events may provide additional dimensionality to specifically identify individual molecules for potential photon accumulation (161, 162). Simultaneously recording the spatial and spectral information of each

stochastic fluorescent emission event was recently made possible using spectroscopic single-molecule localization microscope (sSMLM) developed by us and other groups (30, 31, 104). In our previous study, we have demonstrated the use of sSMLM to identify fluorescent molecules of different species based on their distinct emission spectra (31). Interestingly, single-molecule spectroscopy studies suggest individual fluorescent molecules of the same species also exhibits distinct dissimilarity in their emission spectra, this phenomenon is often referred to as the spectral heterogeneity (SH) (30, 109). Thus, capturing the inherent SH of individual fluorescent molecules further enables for spectroscopic discrimination of individual molecules of the same species in sSMLM (41, 163).

In this work, we report photon-accumulation enhanced reconstruction (PACER), which accounts for photons in repeated emission events from the same molecules through spectroscopic discrimination by sSMLM. Using PACER, we demonstrated a 1.7-nm localization precision using quantum dots (QDs), which exhibit high SH. We further showed that even with fluorophores with low SH (AF647), PACER resolved 6-nm spatial features in DNA origami nanostructures

6.2 Results and Discussion

Imaging Quantum dots with PACER

We first experimentally demonstrate the feasibility of PACER in achieving sub-2-nm localization precision using QDs. Figure 6-3a shows the schematic of the sSMLM system. It employs a dispersive optical component to simultaneously capture the full fluorescent emission spectrum along with the spatial location of every single-molecule emission event. We chose QDs for their high SH due to structural variation and crystalline defects (164). The high SH enabled us to easily

distinguish individual QDs using their distinct fluorescence spectra. Before imaging, we sparsely dispersed QDs (Lumidot, #694614, Sigma-Aldrich Co.) onto cover slides. Using sSMLM, the captured emission spectra are associated with the locations of individual blinking events. Figure 6-3b shows one frame of the spatial images and Figure 6-3c shows the simultaneously acquired spectral image of the QDs. Among all the recorded localization events, some “outliers” caught our attention. Specifically, the red dashed box highlights a region, within which we observed more frequent stochastic blinking events compared to the surroundings. The corresponding spectral image also appears to be much wider than the typical single QD emission spectrum. The repeated occurrence of single molecule blinking events within this highlighted region are spatially clustered in close vicinity. In conventional SMLM image reconstruction, these blinking events are treated independently, leading to a cluster of localizations without knowledge of their exact origins as shown in Figure 6-3d.

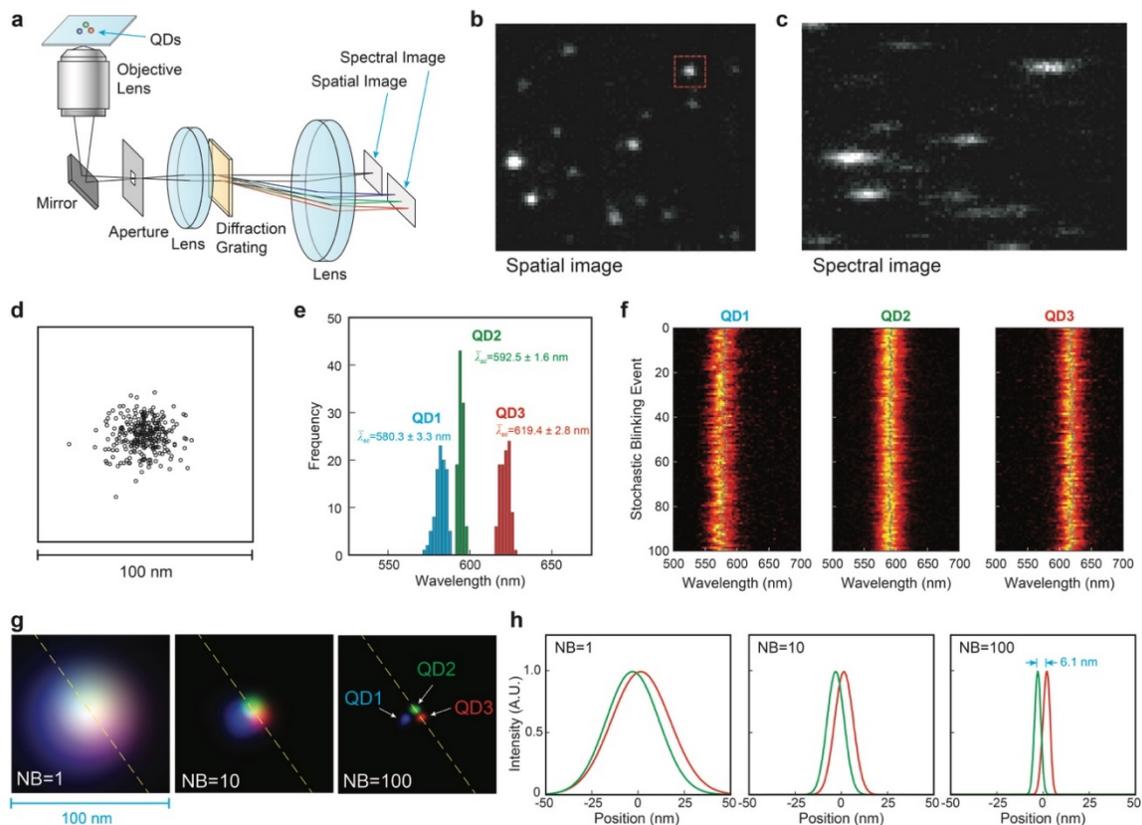


Figure 6-3: Experimental demonstration of PACER in achieving a 1.7-nm localization precision. (a) The schematic of sSMLM. (b) One frame of the spatial images and (c) one frame of the simultaneously acquired spectral images of the QD sample. (d) The scatter plot of localization events in the red dashed box in (b). (e) Histogram of the SC distribution. (f) Fluorescence spectra of three QDs after classifying by SCs using spectral intensity threshold of 300, 460, and 480 photons, respectively, and spectral windows of 575-585 nm, 585-600 nm, and 615-625 nm, respectively, as filtering criteria. The corresponding SC of each fluorescence spectrum is noted as an open circle in the plot. (g) Rendered sSMLM images after combining multiple emission events with NB of 1, 10, and 100, respectively. (h) Line profiles across two QDs in (g) with NB of 1, 10, 100.

From the spectral image, we extracted spectral centroids (SCs) (122, 165) to represent the spectroscopic signature of each QD. We calculate SC as

$$\lambda_{sc} = \frac{\sum_{\lambda} \lambda I(\lambda)}{\sum_{\lambda} I(\lambda)}, \quad (6-3)$$

where λ is the emission wavelength and $I(\lambda)$ is the spectral intensity at λ . As shown in Figure 6-3e, the SC distribution of the blinking events from the highlighted region reveals three distinct distributions. Such a SC distribution suggests that the highlighted region may contain three QDs and these QDs demonstrated distinct spectroscopic signatures, which are 580.3 ± 3.3 nm (QD1), 592.5 ± 1.6 nm (QD2), and 619.4 ± 2.8 nm (QD3), as the result of SH. For reference, individual frames showing each spatial and spectral image of QD1, QD2 and QD3 with the corresponding emission are shown in Figure 6-4.

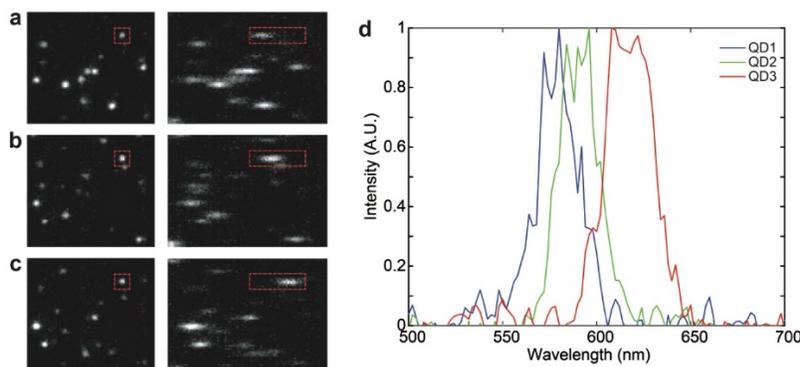


Figure 6-4: Identification of individual QDs. (a-c) Single frames with emission from QD1, QD2 and QD3 respectively occurring from the same location highlighted in the red dashed box. (d) Normalized emission spectra of QD1, QD2 and QD3 extracted from spectral images in (a-c).

Based on the SCs, we can identify and classify the origin of each detected blinking event to one of the three QDs. The first 100 repeated blinking events originating from each of the three QDs after classification were further selected to demonstrate the principle of PACER. Their corresponding spectra and SCs were shown in Figure 6-3f. The observed SH establishes the

foundation for PACER to discriminate individual molecules from the ensemble populations based on their spectroscopic signature. In PACER, the emitted photons originated from repeated blinking of the same QD can all be combined to collectively improve the localization precision. Figure 6-3g illustrates the improved localization precision with respect to increasing number of NB, where we overlay three rendered QDs with pseudo-colors. When NB=1 the localization precision is limited to 17.0 nm, corresponding to an average of 250 photons in the spatial image. Such a localization precision is insufficient to spatially distinguish the three QDs in the reconstructed image (Figure 6-3g, NB=1). After applying PACER, we can positively classify all stochastic blinking events to their respective true origins based on their unique heterogeneous fluorescence emission spectra. After classifying the origins of the blinking events, we accumulate photons from repeated blinking events that are from the same QD from multiple recorded frames to increase the total photon number and, thus, to improve the localization precision (see Figure 6-5 for rendered images with increasing NB to 5, 10, 50, and 100). Accumulating photons from 10 blinking events (Figure 6-3g, NB=10) leads to a more than 3-fold improvement in localization precision to ~4.9 nm, where three QDs can be resolved in the reconstructed sSMLM image. When NB=100, an approximated 10-fold improvement in localization precision to ~1.7 nm shows that QD2 and QD3 are 6.1 nanometers apart. Considering QDs have the mean diameter of 5.2 nm, it is possible that QD2 and QD3 are almost in contact. Figure 6-3h shows the line profiles of the reconstructed image across the QD2 and QD3 with NB of 1, 10, 100, further illustrating the improvement in localization precision as NB increases.

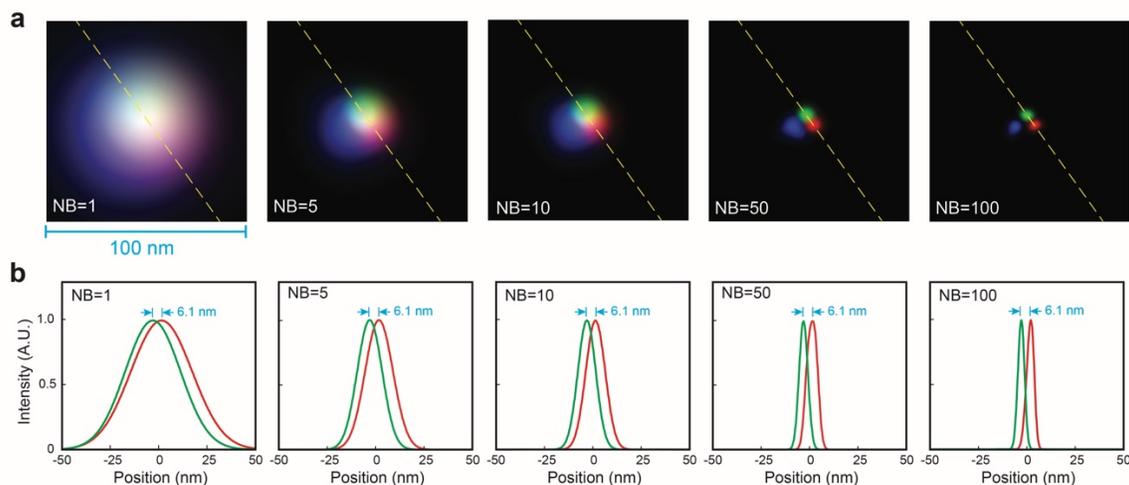


Figure 6-5: Illustration of resolution improvement using QDs through PACER. (a) Rendered sSMLM images after combining multiple emission events with NB of 1, 5, 10, 50, and 100, respectively. (b) The corresponding line profiles across two QDs.

Imaging DNA nanorulers

Upon successful demonstration of improved localization precision using QDs sample, we further experimentally validated PACER using the model system with known inter-molecular spacing. Figure 6-6 shows the results of PACER resolving up to 6-nm spacing in DNA nanorulers (GattaStorm Nanoruler, Gattaquant GMBH) labeled with AF647 fluorescent molecules. The nanoruler comprises DNA origami backbone with precise placement of two fluorescence molecules at controllable inter-molecular spacing (137, 138) (Figure 6-6a, also see details in Methods). We acquired 2,000 frames with an integration time of 10 ms per frame. The recorded spectra of all the individual stochastic blinking from the nanoruler sample labeled with two AF647 molecules are plotted in Figure 6-6b and the corresponding histogram of the SCs is shown in Figure 6-6c. Nanoruler samples with contrast in the measured SC histogram exceeding the threshold value of 20% are accepted. Two distinct peaks with mean SCs separated by approximately 10 nm, which

underlie the effect of SH of AF647, can be clearly observed in Figure 6-6c. Each individual molecule exhibits consistent emission spectra during stochastic switching, with the measured spectral precision less than 2 nm, as estimated by the Standard Deviation (S.D.) of the SC distribution. The averaged fluorescent spectra of the two AF647 molecules are shown in Figure 6-6d after classification based on Figure 6-6c.

We imaged nanorulers with inter-molecular spacing values of 23 nm (Figure 6-6e), 11 nm (Figure 6-6k), and 6 nm (Figure 6-6o). After classifying each localized event to one of the two AF647 molecules and corrected the stage drift based on the position of a fiducially marker, the scatterplot of all the localized blinking events is color-coded in red and green to represent the origins of the two AF647 molecules (Figure 6-6f). Two molecules can be readily separated with colors indicating distinct spectral signatures after classification as shown in Figure 6-6g. After PACER, the localization precision improved from 18.2 nm to 2.6 nm, as denoted by the red and green solid lines in Figure 6-6h. The measured distance between the images of the two molecules is 22.7 nm (Figure 6-6i). Figure 6-6j shows the histogram of the measured inter-molecular spacing from all the 57 nanoruler samples, which shows a mean value of 23.2 nm with a S.D. of 0.8 nm. It agrees well with the expected value of 23 ± 1 nm according to manufacturer's specification.

Figure 6-6k illustrates the nanorulers with an inter-molecular spacing of 11 nm. The color-coded scatter plot of the localization events and the reconstructed PACER image after drift correction of one representative sample are shown in Figures 6-6l and 6-6m, respectively. Figure 6-6n is the histogram of the measured inter-molecular spacing from 35 nanoruler samples, showing the mean spacing of 10.8 nm with a S.D. of 0.9 nm. Again, it agrees well with the expected value of 11 ± 1 nm according to manufacturer's specification. Finally, we performed PACER on the nanoruler

sample with inter-molecular spacing of 6 nm (Figure 6-6o). The corresponding results are shown in Figures 6-6p and 6-6q. The histogram of the measured inter-molecular spaces from 15 nanoruler samples shown in Figure 6-6r indicates the mean spacing of 6.1 nm with S.D. of 0.8 nm, which agrees well with the expected value of 6 ± 1 nm according to manufacturer's specification. Thus, even in a rather challenging case, using AF647 with rather weak SH, PACER can still successfully resolve the fluorescent molecules at the inter-molecular spacing of 6-nm distance, which represent 5-fold improvement in the localization precision under the identical experimental conditions.

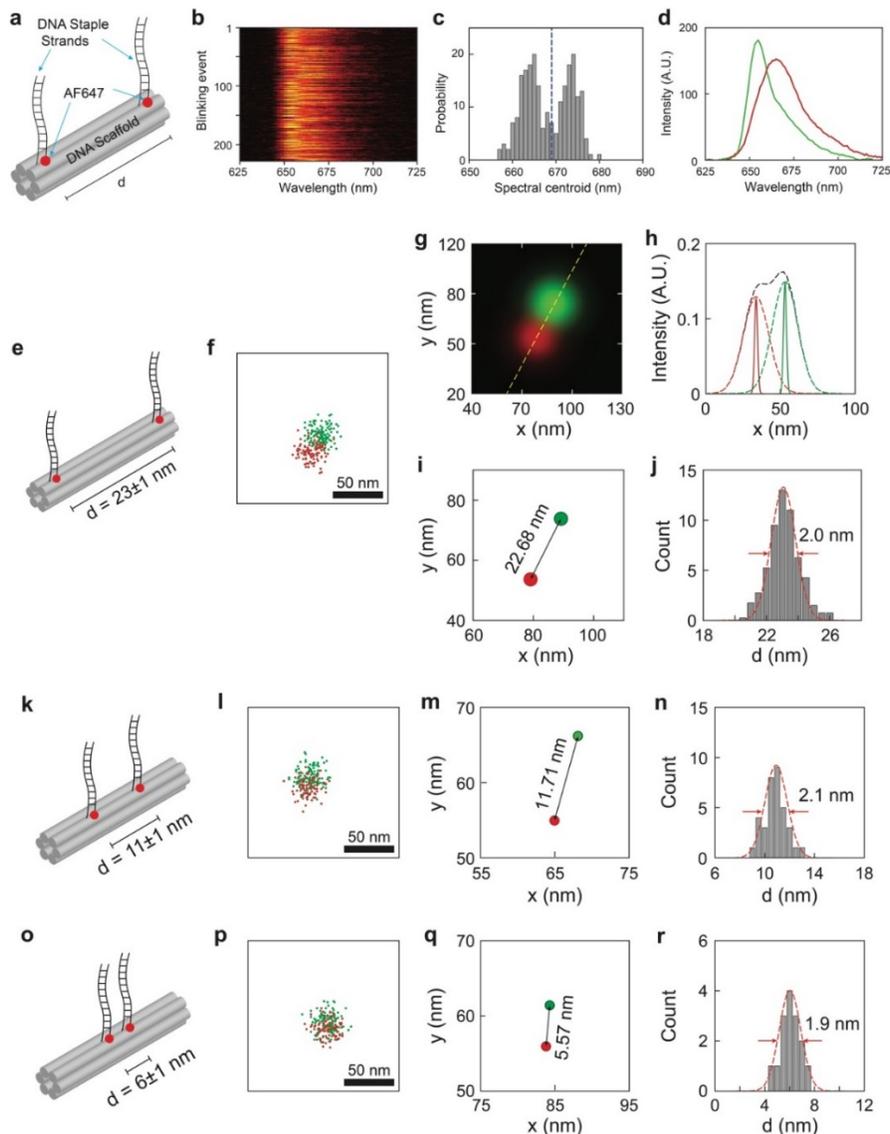


Figure 6-6: Experimental validation of PACER using DNA nanoruler samples. (a) Schematic illustration of a DNA nanoruler labeled with a pair of AF647 featuring a predefined mark-to-mark distance. (b) Representative emission spectra of two molecules on one nanoruler. (c) Histogram of SCs indicates the existence of two molecules with distinct spectral signatures on one nanoruler. (d) The average spectra of the two molecules separated by SC at the wavelength of 669 nm. (e) Schematic of a DNA nanoruler featuring mark-to-mark distance of 23 nm. (f) The scatter plot and (g) the rendered sSMLM image of localization events with colors indicating distinct spectral signatures after stage drift correction. (h) Comparison of molecule location in SMLM image (the dashed black line) and sSMLM images without (the dashed colored lines) and with (the solid colored lines) PACER. (i) Calculated location of molecules through PACER. (j) Histogram of mark-to-mark distance measured from 57 nanorulers. The representative results of a DNA nanoruler featuring mark-to-mark distance of (k-n) 11 nm and (o-r) 6 nm.

Imaging DNA origami nanogrids

Finally, we tested PACER in identifying individual fluorescent molecules in a densely packed cluster. We imaged DNA origami nanogrids (Brightness 9R, GattaQuant GMBH) consisting of a 3×3 array of AF647 fluorescent molecules with a uniform inter-molecular spacing of 11 nm (Figure 6-7a). While initially being developed as a brightness standard, the same nanogrid was used by Stefan Hell's group in the development MINFLUX method to represent densely packed fluorescent-labeled samples (58). The average number of photons recorded from AF647 molecules in the spatial image is 834 photons, leading to a localization precision of 9.21 nm and a spatial resolution of 21.69 nm. Thus, it is impossible to spatially resolve individual AF647 molecules at inter-molecular spacing of 11 nm from reconstructed SMLM image (Figure 6-7b). However, capturing the spectral signature associated with each blinking events provides additional information to better separate the otherwise overlapping localization events in the extended spatial-spectral domain. Figure 6-7c shows the scatter plot of the same localization events in the spatial-spectral domain, with SCs of each blinking events being used to represent the spectral coordinates. We applied a k-means cluster analysis, a built-in function in MATLAB to classify the localization events into seven clusters based on the spatial coordinates and SCs and then assigned pseudo-colors to seven clusters. Figure 6-6d is the corresponding sSMLM image with pseudo-colors to the seven clusters. Figure 6-6e shows the final PACER sSMLM image with an average NB of 23.7 for the seven clusters. On average, the improved localization precision is 3.7 nm and the improved spatial resolution of 8.7 nm, which resolved the majority of the AF647 molecules in the nanogrid. The gray circles show the best guesses of the locations and orientations of the nanogrid, indicates a great match of six but with one molecule mislocated (the blue dot in Figure 6-7e). This is likely

due to a manufacturing artifact or misidentification of fluorophores at neighboring labeling sites which happens to exhibit similar spectral signatures (24, 58, 159).

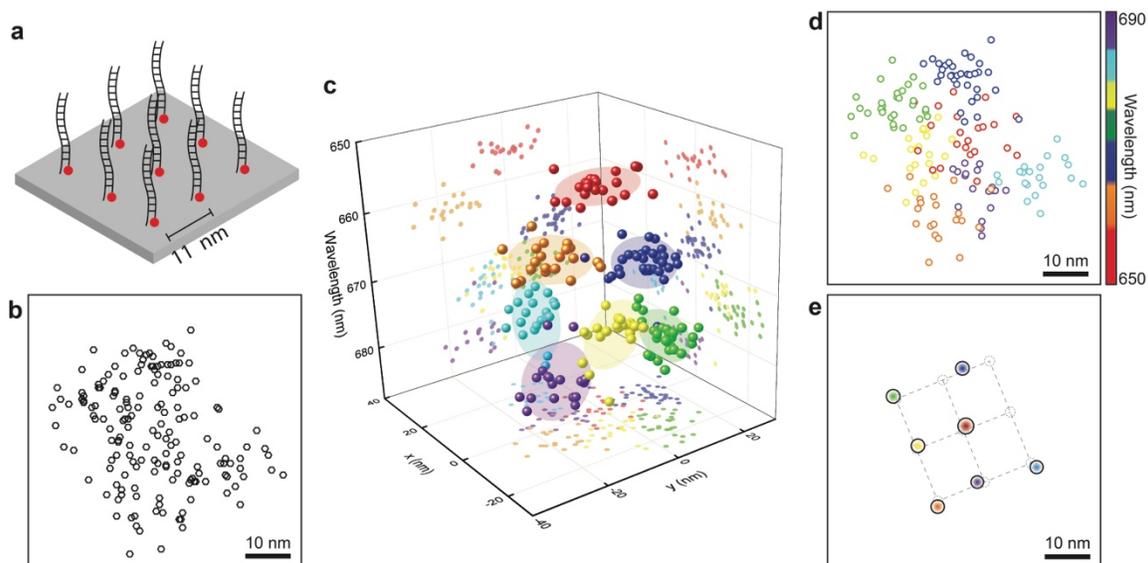


Figure 6-7: PACER imaging DNA origami nanogrids consisting of a 3×3 array of AF647 fluorescent molecules with 11-nm inter-molecular spacing. (a) Schematic illustration of the DNA origami nanogrid labeled with 3×3 array of AF647 with 11-nm inter-molecular spacing. (b) Conventional SMLM reconstruction of all stochastic blinking events. (c) All events can be separated to seven clusters based on their spatial locations and SCs in the extended spatial-spectral domain. (d) The corresponding sSMLM reconstruction with pseudo-colors assigned to seven clusters. (e) Locations of fluorophores calculated after PACER. Black circles represent the localization precision after PACER. The gray dashed circles represent the best guess of the location and orientation of the nanogrid.

6.3 Conclusions

In this study, we demonstrated a single-digit nanometer resolution that can be achieved by combining sSMLM and PACER. We first validated its feasibility of distinguishing molecules/particles of the same type by capturing the intrinsic SH. By using the spectroscopic signature as a unique identifier, photons from individual molecules can now be accumulated in

achieving greatly improved localization precision. We have experimentally validated sub-2-nm localization precision using QDs, and synthetic DNA origami nanostructures (nanorulers and nanogrids) with the smallest feature size of 6 nm. The experimental results indicate that spectroscopic signature of individual molecules would greatly benefit molecular identification and resolution improvement using PACER. This technique offers significantly improved localization precision over the conventional SMLM technique but remains compatible with the existing fluorescence labels and imaging protocols. The broad accessibility and the molecular-scale resolution can potentially provide new insights into biological phenomena and enable significant research progress to be made in the life sciences.

6.4 Materials and Methods

Optical setup

The optical setup contains an inverted optical microscope (Eclipse Ti-U with perfect-focus system, Nikon), equipped with 645-nm and 445-nm solid-state lasers with 500-mW maximum output, a high numerical aperture objective lens (100 \times , NA1.49, Nikon CFI apochromat TIRF) for TIRF illumination, and a home-built transmission spectrometer. The illumination power was controlled by a set of linear polarizers. The imaging filter set was consisted of a laser clean-up filter (FF01-642/10-25, Semrock), a dichroic mirror (FF649-DI01-25 \times 36, Semrock), and a long-pass filter (BLP01-647R-25, Semrock) at the emission port to reject the reflected laser beam. The fluorescence image was then coupled into a transmission spectrometer featuring a blazed dispersive grating (150 grooves/mm). The image further divided into a non-dispersed zeroth-order spatial image and a spectrally dispersed first-order spectral image and can be simultaneously

collected by a high-sensitivity electron multiplying charge-coupled device (EMCCD, ProEM, Princeton Instruments).

sSMLM imaging procedure

The samples were placed on the microscope stage and imaged under a TIRF objective (Nikon CFI apochromat 100 \times , 1.49 NA), with an additional magnification of 1.5 \times by a tube lens. We used a 445-nm laser to excite fluorescence from QDs and a 645-nm laser to excite fluorescence from AF647 on DNA nanorulers and nanogrids. The illumination intensity was set to be 10 kWcm⁻² and 3 kWcm⁻² to create stochastic radiations from QDs and AF647 for sSMLM imaging, respectively. The EMCCD camera acquired images from the monochromator with integrating time of 10 ms at a frame rate of 85 Hz. Unless specifically noted, 5,000 frames were recorded for imaging reconstruction using PACER.

Preparation of quantum dots samples

The QD nanoparticles used in the experiment is core-shell CdSe/ZnS LumidotsTM (product #: 694614) with mean particle size of 5.2 nm and emission peak at 610 nm. QDs were immobilized on a coverslip (#1.5, VWR) by spin-coating of their suspension (~100 ng/mL) and covered with silicon oil to protect them from photo-oxidation.

DNA nanoruler and nanogrid samples

The nanorulers was customized from Gattaquant DNA Nanotechnologies with mark-to-mark distances of 23 nm, 11 nm, and 6 nm, respectively. Each mark only contains one AF647 molecule.

All samples were delivered in solution and then immobilized on BSA-biotin-neutravidin surface in LabTek (VWR) chambers.

sSMLM image reconstruction and rendering

Experimentally acquired sSMLM images are split to spatial and spectral images for the reconstruction. We used the method described by Song et al. for background subtraction and spectral calibration (122). The spatial images are then analyzed by using the standard localization algorithm (ThunderSTORM, ImageJ plug-in (121)) to determine the locations of individual blinking events. The locations of molecules are used to establish a reference point to obtain their corresponding emission spectra from the spectral images. Image rendering is realized by superimposing positions of all the collected localizations after convolving a Gaussian kernel corresponding to the localization precision calculated from their photon number using Eqn. 6-1. After performing PACER, the localization precision of the molecule centroids was recalculated from their accumulated photon number and pseudo-color is assigned to indicate different molecules.

Immobilization of DNA nanorulers

The LabTek chamber (VWR) was washed three times with 500 μ L PBS and then incubated with 200 μ L of BSA-biotin solution (1 mg/ml in PBS) for 5 min. After removing the BSA-biotin solution, the chamber was washed 3 times with 500 μ L of PBS. Then the chamber was incubated with 200 μ L of NEU solution (1 mg/ml in PBS) for 5 min and washed three times with 500 μ L of 1x PBS supplemented with 10 mM magnesium chloride (MgCl_2). 1 μ L of the DNA origami solution were diluted with 200 μ L of 10 mM MgCl_2 . The chamber was incubated with the diluted DNA

origami solution for 5 min to achieve the desired surface density of DNA origami structures ($\sim 0.1 \mu\text{m}^{-2}$) and then washed three times with 500 μl of PBS. The chamber was filled with the imaging buffer prior to imaging.

Imaging buffer

A standard imaging buffer was freshly made and added to the DNA nanoruler samples prior to imaging. It contained TN buffer (50 mM Tris and 10 mM NaCl), an oxygen scavenging system (0.5 mg/ml glucose oxidase (Sigma-Aldrich)), 40 $\mu\text{g/ml}$ catalase (Sigma-Aldrich) and 10% (w/v) glucose (Sigma-Aldrich), and 143 mM βME (Sigma-Aldrich).

CHAPTER 7: Conclusions and Future Perspectives

In this dissertation we developed RainbowSTORM, a freely available software platform for comprehensive spectroscopic analysis and sSMLM image reconstruction. We believe that the access to this software platform will facilitate the increased adoption of sSMLM thus enabling new discoveries and applications. Additionally, we developed three quantitative methods which used the spectroscopic information to specifically remove emission events from unwanted fluorescence (142), more accurately size nanocarriers (142), and measure the distance between fluorescent molecules with sub-2nm localization precision. Overall, using the full emission spectra to better identify emission events related to the sample allows for emission from fluorescent impurities (123) and non-specific binding (142) to be removed without excluding emission from the true sample. This prevents overcounting and undercounting artifacts. Additionally, by combining the specificity of spectroscopic measurements with spatial clustering analysis using qSPACE, high utilization of localization events can be retained, thus, preventing molecular sampling issues and enabling accurate sizing of molecules (142). Further, by using the SH of dyes to enable molecular discrimination, PACER can be used to reconstruct images with sub-10 nm resolutions.

Based on these findings, the methods presented in this dissertation can be further developed for quantitative imaging of nanostructures within cells. These methods also need to be further developed to facilitate 3D sSMLM of immobilized molecules and molecules in the cell. As these techniques become more advanced, their application to live cell imaging can further improve quantitative studies of biological processes using sSMLM.

As previously shown fluorescent impurities can be detected within cells (31), therefore unlabeled cells can be used as a control during experiments to allow for the identification of sources of intrinsic fluorescence (96-98). Additionally, the development of a library of the spectra of impurities associated with common sample preparation reagents (110, 123, 125) for *in vitro* and *in situ* imaging can be helpful when analyzing sSMLM data. sSMLM would also benefit from a comprehensive evaluation of fluorophores and their associated SH (30, 31, 109). Additionally, studies designed to understand the origin of SH can help to unlock the development of methods which exploit SH to enable further analysis of multicolor images and improved application of PACER. Quantitative super-resolution imaging can be further improved by combining spectroscopic analysis with additional correlative measurements which detect temporal dynamics or polarization of fluorescent probes. We also anticipate further development of machine-learning (113) and cluster analysis (104, 142) techniques which utilize spectroscopic information to improve the performance of sSMLM. As these methods become more advanced, we hope to include them in the RainbowSTORM framework in order to encourage the continued development of quantitative analytics for sSMLM (119).

REFERENCES

1. U. Kubitscheck, "Fluorescence Microscopy From Principles to Biological Applications Preface," *Fluorescence Microscopy: From Principles to Biological Applications* Xiii-Xv (2013).
2. J. W. Lichtman, and J. A. Conchello, "Fluorescence microscopy," *Nat Methods* **2**(12), 910-919 (2005).
3. W. E. Moerner, and D. P. Fromm, "Methods of single-molecule fluorescence spectroscopy and microscopy," *Rev Sci Instrum* **74**(8), 3597-3619 (2003).
4. K. Grussmayer, K. Yserentant, and D. P. Herten, "Photons in - numbers out: perspectives in quantitative fluorescence microscopy for in situ protein counting," *Methods Appl Fluores* **7**(1), (2019).
5. E. M. Puchner et al., "Counting molecules in single organelles with superresolution microscopy allows tracking of the endosome maturation trajectory," *P Natl Acad Sci USA* **110**(40), 16015-16020 (2013).
6. S. A. Mutch et al., "Determining the number of specific proteins in cellular compartments by quantitative microscopy," *Nat Protoc* **6**(12), 1953-1968 (2011).
7. F. F. Du et al., "Sequential intracellular release of water-soluble cargos from Shell-crosslinked polymersomes," *J Control Release* **282**,90-100 (2018).
8. V. C. Coffman, and J. Q. Wu, "Counting protein molecules using quantitative fluorescence microscopy," *Trends Biochem Sci* **37**(11), 499-506 (2012).
9. S. Gite et al., "A Rapid, Accurate, Single Molecule Counting Method Detects Clostridium difficile Toxin B in Stool Samples," *Sci Rep* **8**,(2018).

10. H. Aoki et al., "Quantitative analysis of end-to-end distance of single polymer chain in ultra-thin film by super-resolution fluorescence imaging," *Chem Phys* **419**,54-58 (2013).
11. B. Y. Hua et al., "An improved surface passivation method for single-molecule studies," *Nat Methods* **11**(12), 1233-1236 (2014).
12. E. C. Heider et al., "Quantitative Fluorescence Microscopy To Determine Molecular Occupancy of Phospholipid Vesicles," *Analytical Chemistry* **83**(13), 5128-5136 (2011).
13. E. Abbe, "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung," *Archiv für Mikroskopische Anatomie* **9**(1), 413-468 (1873).
14. J. P. Patterson et al., "CryoTEM as an Advanced Analytical Tool for Materials Chemists," *Accounts Chem Res* **50**(7), 1495-1501 (2017).
15. G. Binnig, C. F. Quate, and C. Gerber, "Atomic Force Microscope," *Phys Rev Lett* **56**(9), 930-933 (1986).
16. S. Pujals et al., "Super-resolution microscopy as a powerful tool to study complex synthetic materials," *Nat Rev Chem* **3**(2), 68-84 (2019).
17. E. Betzig et al., "Imaging intracellular fluorescent proteins at nanometer resolution," *Science* **313**(5793), 1642-1645 (2006).
18. M. J. Rust, M. Bates, and X. W. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nat Methods* **3**(10), 793-795 (2006).
19. S. W. Hell, "Far-field optical nanoscopy," *Science* **316**(5828), 1153-1158 (2007).
20. M. G. L. Gustafsson, "Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution," *P Natl Acad Sci USA* **102**(37), 13081-13086 (2005).

21. F. Fricke et al., "One, two or three? Probing the stoichiometry of membrane proteins by single-molecule localization microscopy," *Sci Rep* **5**, 14072 (2015).
22. J. Gunzenhauser et al., "Quantitative Super-Resolution Imaging Reveals Protein Stoichiometry and Nanoscale Morphology of Assembling HIV-Gag Virions," *Nano Lett* **12**(9), 4705-4710 (2012).
23. X. L. Nan et al., "Single-molecule superresolution imaging allows quantitative analysis of RAF multimer formation and signaling," *P Natl Acad Sci USA* **110**(46), 18519-18524 (2013).
24. R. Jungmann et al., "Quantitative super-resolution imaging with qPAINT," *Nat Methods* **13**(5), 439-442 (2016).
25. X. Su et al., "Single-Molecule Counting of Point Mutations by Transient DNA Binding," *Sci Rep* **7**, 43824 (2017).
26. N. Ehmman et al., "Quantitative super-resolution imaging of Bruchpilot distinguishes active zone states," *Nat Commun* **5**, 4650 (2014).
27. D. M. Owen et al., "PALM imaging and cluster analysis of protein heterogeneity at the cell surface," *J Biophotonics* **3**(7), 446-454 (2010).
28. D. van der Zwaag et al., "Super Resolution Imaging of Nanoparticles Cellular Uptake and Trafficking," *ACS Appl Mater Inter* **8**(10), 6391-6399 (2016).
29. T. C. Voss, I. A. Demarco, and R. N. Day, "Quantitative imaging of protein interactions in the cell nucleus," *Biotechniques* **38**(3), 413-424 (2005).
30. Z. Y. Zhang et al., "Ultrahigh-throughput single-molecule spectroscopy and spectrally resolved super-resolution microscopy," *Nat Methods* **12**(10), 935-938 (2015).

31. B. Q. Dong et al., "Super-resolution spectroscopic microscopy via photon localization," *Nat Commun* **7**, 12290 (2016).
32. M. J. Mlodzianoski et al., "Super-Resolution Imaging of Molecular Emission Spectra and Single Molecule Spectral Fluctuations," *Plos One* **11**(3), e0147506 (2016).
33. J. F. W. Herschel, "No. II. On the Epipolic Dispersion of Light, Being a Supplement to a Paper Entitled, "On a Case of Superficial Colour Presented by a Homogeneous Liquid Internally Colourless"," *Philosophical Transactions of the Royal Society of London* **135**, 147-153 (1845).
34. J. F. W. Herschel, "No. I. On a Case of Superficial Colour Presented by a Homogeneous Liquid Internally Colourless," *Philosophical Transactions of the Royal Society of London* **135**, 143-145 (1845).
35. A. Jablonski, "General Theory of Pressure Broadening of Spectral Lines," *Phys Rev* **68**(3-4), 78-93 (1945).
36. M. Minsky, "Memoir on Inventing the Confocal Scanning Microscope," *Scanning* **10**(4), 128-138 (1988).
37. J. S. Verdaasdonk et al., "Bending the Rules: Widefield Microscopy and the Abbe Limit of Resolution," *J Cell Physiol* **229**(2), 132-138 (2014).
38. J. C. Waters, "Accuracy and precision in quantitative fluorescence microscopy," *J Cell Biol* **185**(7), 1135-1148 (2009).
39. J. C. Waters, and T. Wittmann, "Concepts in quantitative fluorescence microscopy," *Method Cell Biol* **123**, 1-18 (2014).
40. M. Fricker, J. Runions, and I. Moore, "Quantitative fluorescence microscopy: From art to science," *Annu Rev Plant Biol* **57**, 79-107 (2006).

41. W. P. Ambrose, and W. E. Moerner, "Fluorescence Spectroscopy and Spectral Diffusion of Single Impurity Molecules in a Crystal," *Nature* **349**(6306), 225-227 (1991).
42. W. E. Moerner, "A dozen years of single-molecule spectroscopy in physics, chemistry, and biophysics," *J Phys Chem B* **106**(5), 910-927 (2002).
43. W. E. Moerner et al., "Near-Field Optical Spectroscopy of Individual Molecules in Solids," *Phys Rev Lett* **73**(20), 2764-2767 (1994).
44. A. Yildiz et al., "Myosin V walks hand-over-hand: Single fluorophore imaging with 1.5-nm localization," *Science* **300**(5628), 2061-2065 (2003).
45. S. R. Jung, B. S. Fujimoto, and D. T. Chiu, "Quantitative microscopy based on single-molecule fluorescence," *Curr Opin Chem Biol* **39**, 64-73 (2017).
46. T. Forster, "*Zwischenmolekulare Energiewanderung Und Fluoreszenz," *Ann Phys-Berlin* **2**(1-2), 55-75 (1948).
47. T. Forster, "Energiewanderung Und Fluoreszenz," *Naturwissenschaften* **33**(6), 166-175 (1946).
48. L. Schermelleh, R. Heintzmann, and H. Leonhardt, "A guide to super-resolution fluorescence microscopy," *J Cell Biol* **190**(2), 165-175 (2010).
49. M. G. L. Gustafsson et al., "Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination," *Biophys J* **94**(12), 4957-4970 (2008).
50. L. Schermelleh et al., "Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy," *Science* **320**(5881), 1332-1336 (2008).
51. O. Schulz et al., "Resolution doubling in fluorescence microscopy with confocal spinning-disk image scanning microscopy," *P Natl Acad Sci USA* **110**(52), 21000-21005 (2013).

52. A. G. York et al., "Instant super-resolution imaging in live cells and embryos via analog image processing," *Nat Methods* **10**(11), 1122-1126 (2013).
53. B. Hein, K. I. Willig, and S. W. Hell, "Stimulated emission depletion (STED) nanoscopy of a fluorescent protein-labeled organelle inside a living cell," *P Natl Acad Sci USA* **105**(38), 14271-14276 (2008).
54. F. Bottanelli et al., "Two-colour live-cell nanoscale imaging of intracellular targets," *Nat Commun* **7**, 10778 (2016).
55. N. T. Urban et al., "STED Nanoscopy of Actin Dynamics in Synapses Deep Inside Living Brain Slices," *Biophys J* **101**(5), 1277-1284 (2011).
56. S. Berning et al., "Nanoscopy in a Living Mouse Brain," *Science* **335**(6068), 551-551 (2012).
57. S. W. Hell, S. Jakobs, and L. Kastrup, "Imaging and writing at the nanoscale with focused visible light through saturable optical transitions," *Appl Phys a-Mater* **77**(7), 859-860 (2003).
58. F. Balzarotti et al., "Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes," *Science* **355**(6325), 606-612 (2017).
59. Y. Eilers et al., "MINFLUX monitors rapid molecular jumps with superior spatiotemporal resolution," *P Natl Acad Sci USA* **115**(24), 6117-6122 (2018).
60. K. C. Gwosch et al., "MINFLUX nanoscopy delivers 3D multicolor nanometer resolution in cells," *Nat Methods* **17**, 217-224 (2020).
61. K. I. Mortensen et al., "Optimized localization analysis for single-molecule tracking and super-resolution microscopy," *Nat Methods* **7**(5), 377-381 (2010).
62. B. Rieger, and S. Stallinga, "The Lateral and Axial Localization Uncertainty in Super-Resolution Light Microscopy," *ChemPhysChem* **15**(4), 664-670 (2014).

63. M. Bates et al., "Multicolor super-resolution imaging with photo-switchable fluorescent probes," *Science* **317**(5845), 1749-1753 (2007).
64. B. Huang et al., "Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy," *Science* **319**(5864), 810-813 (2008).
65. A. Sharonov, and R. M. Hochstrasser, "Wide-field subdiffraction imaging by accumulated binding of diffusing probes," *P Natl Acad Sci USA* **103**(50), 18911-18916 (2006).
66. Y. Zhang et al., "Far-Red Photoactivatable BODIPYs for the Super-Resolution Imaging of Live Cells," *J Am Chem Soc* **140**(40), 12741-12745 (2018).
67. S. Y. Wang et al., "Characterization and development of photoactivatable fluorescent proteins for single-molecule-based superresolution imaging," *P Natl Acad Sci USA* **111**(23), 8452-8457 (2014).
68. M. Bates, B. Huang, and X. W. Zhuang, "Super-resolution microscopy by nanoscale localization of photo-switchable fluorescent probes," *Curr Opin Chem Biol* **12**(5), 505-514 (2008).
69. U. Endesfelder, and M. Heilemann, "Direct Stochastic Optical Reconstruction Microscopy (dSTORM)," *Advanced Fluorescence Microscopy: Methods and Protocols* **1251**(262-275 (2015).
70. G. T. Dempsey et al., "Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging," *Nat Methods* **8**(12), 1027-1036 (2011).
71. D. R. Whelan, and T. D. M. Bell, "Image artifacts in Single Molecule Localization Microscopy: why optimization of sample preparation protocols matters," *Sci Rep* **5**, 7924(2015).

72. B. Adelizzi et al., "Painting Supramolecular Polymers in Organic Solvents by Super-resolution Microscopy," *ACS Nano* **12**(5), 4431-4439 (2018).
73. A. Aloï et al., "iPAINT: a general approach tailored to image the topology of interfaces with nanometer resolution," *Nanoscale* **8**(16), 8712-8716 (2016).
74. C. K. Spahn et al., "A toolbox for multiplexed super-resolution imaging of the E. coli nucleoid and membrane using novel PAINT labels," *Sci Rep* **8**, 14768 (2018).
75. R. Jungmann et al., "Single-Molecule Kinetics and Super-Resolution Microscopy by Fluorescence Imaging of Transient Binding on DNA Origami," *Nano Lett* **10**(11), 4756-4761 (2010).
76. R. Jungmann et al., "Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT," *Nat Methods* **11**(3), 313-U292 (2014).
77. J. Schnitzbauer et al., "Super-resolution microscopy with DNA-PAINT," *Nat Protoc* **12**(6), 1198-1228 (2017).
78. A. Auer et al., "Fast, Background-Free DNA-PAINT Imaging Using FRET-Based Probes," *Nano Lett* **17**(10), 6428-6434 (2017).
79. M. F. Juetten et al., "Three-dimensional sub-100 nm resolution fluorescence microscopy of thick samples," *Nat Methods* **5**(6), 527-529 (2008).
80. C. M. Winterflood et al., "Dual-Color 3D Superresolution Microscopy by Combined Spectral-Demixing and Biplane Imaging," *Biophys J* **109**(1), 3-6 (2015).
81. G. Shtengel et al., "Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure," *P Natl Acad Sci USA* **106**(9), 3125-3130 (2009).

82. D. Aquino et al., "Two-color nanoscopy of three-dimensional volumes by 4Pi detection of stochastically switched fluorophores," *Nat Methods* **8**(4), 353-359 (2011).
83. P. R. Nicovich, D. M. Owen, and K. Gaus, "Turning single-molecule localization microscopy into a quantitative bioanalytical tool," *Nat Protoc* **12**(3), 453-460 (2017).
84. G. Hummer, F. Fricke, and M. Heilemann, "Model-independent counting of molecules in single-molecule localization microscopy," *Mol Biol Cell* **27**(22), 3637-3644 (2016).
85. U. Endesfelder et al., "Multiscale Spatial Organization of RNA Polymerase in Escherichia coli," *Biophys J* **105**(1), 172-181 (2013).
86. M. D. Lew et al., "Three-dimensional superresolution colocalization of intracellular protein superstructures and the cell surface in live *Caulobacter crescentus*," *P Natl Acad Sci USA* **108**(46), E1102-E1110 (2011).
87. E. Y. Kim et al., "Distinct pathological signatures in human cellular models of myotonic dystrophy subtypes," *JCI Insight* **4**(6), e122686 (2019).
88. S. Rajoo et al., "Stoichiometry and compositional plasticity of the yeast nuclear pore complex revealed by quantitative fluorescence microscopy," *P Natl Acad Sci USA* **115**(17), E3969-E3977 (2018).
89. P. J. Clark, and F. C. Evans, "Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations," *Ecology* **35**(4), 445-453 (1954).
90. B. D. Ripley, "Modeling Spatial Patterns," *J Roy Stat Soc B Met* **39**(2), 172-212 (1977).
91. P. Sengupta et al., "Probing protein heterogeneity in the plasma membrane using PALM and pair correlation analysis," *Nat Methods* **8**(11), 969-975 (2011).

92. A. Mazouchi, and J. N. Milstein, "Fast Optimized Cluster Algorithm for Localizations (FOCAL): a spatial cluster analysis for super-resolved microscopy," *Bioinformatics* **32**(5), 747-754 (2016).
93. M. Ankerst et al., "OPTICS: Ordering points to identify the clustering structure," *Sigmod Record, Vol 28, No 2 - June 1999* 49-60 (1999).
94. M. Ester et al., "A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, pp. 226-231, AAAI Press, Portland, Oregon (1996).
95. M. B. Stone, and S. L. Veatch, "Far-red organic fluorophores contain a fluorescent impurity," *Chemphyschem* **15**(11), 2240-2246 (2014).
96. R. L. Affleck et al., "Reduction of luminescent background in ultrasensitive fluorescence detection by photobleaching," *Anal Chem* **68**(13), 2270-2276 (1996).
97. T. K. Herman, S. A. Mackowiak, and L. J. Kaufman, "High power light emitting diode based setup for photobleaching fluorescent impurities," *Rev Sci Instrum* **80**(1), (2009).
98. B. Dong et al., "Superresolution intrinsic fluorescence imaging of chromatin utilizing native, unmodified nucleic acids for contrast," *P Natl Acad Sci USA* **113**(35), 9716-9721 (2016).
99. B. Dong et al., "Stochastic fluorescence switching of nucleic acids under visible light illumination," *Opt Express* **25**(7), 7929-7944 (2017).
100. E. Kozma, and P. Kele, "Fluorogenic probes for super-resolution microscopy," *Org Biomol Chem* **17**(2), 215-233 (2019).

101. E. R. Griffis, S. L. Xu, and M. A. Powers, "Nup98 localizes to both nuclear and cytoplasmic sides of the nuclear pore and binds to two distinct nucleoporin subcomplexes," *Mol Biol Cell* **14**(2), 600-610 (2003).
102. R. P. J. Nieuwenhuizen et al., "Measuring image resolution in optical nanoscopy," *Nat Methods* **10**(6), 557-562 (2013).
103. B. Dong et al., "Spectroscopic analysis beyond the diffraction limit," *Int J Biochem Cell B* **101**,113-117 (2018).
104. M. N. Bongiovanni et al., "Multi-dimensional super-resolution imaging enables surface hydrophobicity mapping," *Nat Commun* **7**,13544 (2016).
105. J. E. Lee et al., "Mapping Surface Hydrophobicity of alpha-Synuclein Oligomers at the Nanoscale," *Nano Lett* **18**(12), 7494-7501 (2018).
106. D. Kim, Z. Y. Zhang, and K. Xu, "Spectrally Resolved Super-Resolution Microscopy Unveils Multipath Reaction Pathways of Single Spiropyran Molecules," *J Am Chem Soc* **139**(28), 9447-9450 (2017).
107. S. Moon et al., "Spectrally Resolved, Functional Super-Resolution Microscopy Reveals Nanoscale Compositional Heterogeneity in Live-Cell Membranes," *J Am Chem Soc* **139**(32), 10944-10947 (2017).
108. L. M. Xiang et al., "Optical characterization of surface adlayers and their compositional demixing at the nanoscale," *Nat Commun* **9**,1435 (2018).
109. Y. Zhang et al., "Multicolor super-resolution imaging using spectroscopic single-molecule localization microscopy with optimal spectral dispersion," *Appl Optics* **58**(9), 2248-2255 (2019).

110. B. E. Urban et al., "Subsurface Super-resolution Imaging of Unstained Polymer Nanostructures," *Sci Rep* **6**, 28156 (2016).
111. A. Eshein et al., "Sub-10-nm imaging of nucleic acids using spectroscopic intrinsic-contrast photon-localization optical nanoscopy (SICLON)," *Optics Letters* **43**(23), 5817-5820 (2018).
112. R. Yan et al., "Spectrally Resolved and Functional Super-resolution Microscopy via Ultrahigh-Throughput Single-Molecule Spectroscopy," *Accounts Chem Res* **51**(3), 697-705 (2018).
113. Z. Zhang et al., "Machine-learning based spectral classification for spectroscopic single-molecule localization microscopy," *Opt Lett* **44**(23), 5864-5867 (2019).
114. B. Dong et al., "Parallel Three-Dimensional Tracking of Quantum Rods Using Polarization-Sensitive Spectroscopic Photon Localization Microscopy," *ACS Photonics* **4**(7), 1747-1752 (2017).
115. K. Song et al., "Three-dimensional biplane spectroscopic single-molecule localization microscopy," *Optica* **6**(6), 709-715 (2019).
116. T. Huang et al., "Simultaneous Multicolor Single-Molecule Tracking with Single-Laser Excitation via Spectral Imaging," *Biophys J* **114**(2), 301-310 (2018).
117. D. I. Danylchuk et al., "Switchable Solvatochromic Probes for Live-Cell Super-resolution Imaging of Plasma Membrane Organization," *Angew Chem Int Edit* **58**(42), 14920-14924 (2019).
118. L. Sansalone et al., "High-Throughput Single-Molecule Spectroscopy Resolves the Conformational Isomers of BODIPY Chromophores," *J Phys Chem Lett* **10**(21), 6807-6812 (2019).

119. J. L. Davis et al., "RainbowSTORM: An open-source ImageJ plugin for spectroscopic single-molecule localization microscopy (sSMLM) data analysis and image reconstruction," *bioRxiv* (2020).
120. D. Sage et al., "Super-resolution fight club: assessment of 2D and 3D single-molecule localization microscopy software," *Nat Methods* **16**(5), 387-395 (2019).
121. M. Ovesny et al., "ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging," *Bioinformatics* **30**(16), 2389-2390 (2014).
122. K. Song et al., "Theoretical analysis of spectral precision in spectroscopic single-molecule localization microscopy," *Rev Sci Instrum* **89**(12), 123703 (2018).
123. J. L. Davis et al., "Method to identify and minimize artifacts induced by fluorescent impurities in single-molecule localization microscopy," *J Biomed Opt* **23**(10), 106501 (2018).
124. X. Y. Wang et al., "Non-blinking semiconductor nanocrystals (Retraction of vol 459, pg 686, 2009)," *Nature* **527**, 544 (2015).
125. F. T. Rabouw et al., "Non-blinking single-photon emitters in silica," *Sci Rep* **6**, 21187 (2016).
126. H. H. Tuson et al., "Addressing the Requirements of High-Sensitivity Single-Molecule Imaging of Low-Copy-Number Proteins in Bacteria," *Chemphyschem* **17**(10), 1435-1440 (2016).
127. X. B. Shi et al., "Counting quantum dot aggregates for the detection of biotinylated proteins," *Chem Commun* **51**(12), 2353-2356 (2015).
128. Y. Wang, P. Penkul, and J. N. Milstein, "Quantitative Localization Microscopy Reveals a Novel Organization of a High-Copy Number Plasmid," *Biophys J* **111**(3), 467-479 (2016).

129. H. Labit et al., "A simple and optimized method of producing silanized surfaces for FISH and replication mapping on combed DNA fibers," *Biotechniques* **45**(6), 649-658 (2008).
130. J. J. Cras et al., "Comparison of chemical cleaning methods of glass in preparation for silanization," *Biosens Bioelectron* **14**(8-9), 683-688 (1999).
131. C. F. Mallinson, J. E. Castle, and J. F. Watts, "Analysis of the Be KLL Auger Transition of Beryllium Nitride and Beryllium Carbide by AES," *Surface Science Spectra* **22**(1), 71-80 (2015).
132. Q. S. Wei et al., "Imaging and Sizing of Single DNA Molecules on a Mobile Phone," *ACS Nano* **8**(12), 12725-12733 (2014).
133. K. H. Chen et al., "RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells," *Science* **348**(6233), aaa6090 (2015).
134. M. S. Zhang et al., "Rational design of true monomeric and bright photoactivatable fluorescent proteins," *Nat Methods* **9**(7), 727-U297 (2012).
135. P. Rubin-Delanchy et al., "Bayesian cluster identification in single-molecule Localization microscopy data," *Nat Methods* **12**(11), 1072-1076 (2015).
136. B. J. Beliveau et al., "Single-molecule super-resolution imaging of chromosomes and in situ haplotype visualization using Oligopaint FISH probes," *Nat Commun* **6**, 7147 (2015).
137. C. Steinhauer et al., "DNA Origami as a Nanoscopic Ruler for Super-Resolution Microscopy," *Angew Chem Int Edit* **48**(47), 8870-8873 (2009).
138. J. J. Schmied et al., "Fluorescence and super-resolution standards based on DNA origami," *Nat Methods* **9**(12), 1133-1134 (2012).
139. S. Culley et al., "Quantitative mapping and minimization of super-resolution optical imaging artifacts," *Nat Methods* **15**(4), 263-266 (2018).

140. T. Ha, and P. Tinnefeld, "Photophysics of Fluorescent Probes for Single-Molecule Biophysics and Super-Resolution Imaging," *Annu Rev Phys Chem* **63**, 595-617 (2012).
141. F. Pennacchietti, T. J. Gould, and S. T. Hess, "The Role of Probe Photophysics in Localization-Based Superresolution Microscopy," *Biophys J* **113**(9), 2037-2054 (2017).
142. J. L. Davis et al., "Super-Resolution Imaging of Self-Assembled Nanocarriers Using Quantitative Spectroscopic Analysis for Cluster Extraction," *Langmuir* **36**(9), 2291-2299 (2020).
143. S. D. Allen et al., "On the advancement of polymeric bicontinuous nanospheres toward biomedical applications," *Nanoscale Horiz* **4**(2), 258-272 (2019).
144. J. Habel et al., "Selecting analytical tools for characterization of polymersomes in aqueous solution," *Rsc Adv* **5**(97), 79924-79946 (2015).
145. C. E. Boott et al., "In Situ Visualization of Block Copolymer Self-Assembly in Organic Media by Super-Resolution Fluorescence Microscopy," *Chem-Eur J* **21**(51), 18539-18542 (2015).
146. H. B. Qiu et al., "Uniform patchy and hollow rectangular platelet micelles from crystallizable polymer blends," *Science* **352**(6286), 697-701 (2016).
147. D. Woll, and C. Flors, "Super-resolution Fluorescence Imaging for Materials Science," *Small Methods* **1**(10), (2017).
148. L. Albertazzi et al., "Probing Exchange Pathways in One-Dimensional Aggregates with Super-Resolution Microscopy," *Science* **344**(6183), 491-495 (2014).
149. A. Aloï et al., "Imaging Nanostructures by Single-Molecule Localization Microscopy in Organic Solvents," *J Am Chem Soc* **138**(9), 2953-2956 (2016).

150. E. Achtert, C. Bohm, and P. Kroger, "DeLiClu: Boosting robustness, completeness, usability, and efficiency of hierarchical clustering by a closest pair ranking," *Lect Notes Artif Int* **3918**, 119-128 (2006).
151. F. Levet et al., "SR-Tesseler: a method to segment and quantify localization-based super-resolution microscopy data," *Nat Methods* **12**(11), 1065-1071 (2015).
152. P. Greenspan, and S. D. Fowler, "Spectrofluorometric Studies of the Lipid Probe, Nile Red," *J Lipid Res* **26**(7), 781-789 (1985).
153. P. Greenspan, E. P. Mayer, and S. D. Fowler, "Nile Red - a Selective Fluorescent Stain for Intracellular Lipid Droplets," *J Cell Biol* **100**(3), 965-973 (1985).
154. A. K. Dutta, K. Kamada, and K. Ohta, "Spectroscopic studies of nile red in organic solvents and polymers," *J Photoch Photobio A* **93**(1), 57-64 (1996).
155. R. Salva et al., "Polymersome Shape Transformation at the Nanoscale," *ACS Nano* **7**(10), 9298-9311 (2013).
156. R. S. M. Rikken et al., "Shaping polymersomes into predictable morphologies via out-of-equilibrium self-assembly," *Nat Commun* **7**, 12606 (2016).
157. S. W. Hell, and J. Wichmann, "Breaking the Diffraction Resolution Limit by Stimulated-Emission - Stimulated-Emission-Depletion Fluorescence Microscopy," *Optics Letters* **19**(11), 780-782 (1994).
158. J. Fölling et al., "Fluorescence nanoscopy by ground-state depletion and single-molecule return," *Nat Methods* **5**(11), 943-945 (2008).
159. M. J. Dai, R. Jungmann, and P. Yin, "Optical imaging of individual biomolecules in densely packed clusters," *Nat Nanotechnol* **11**(9), 798-807 (2016).

160. M. Raab et al., "Fluorescence Microscopy with 6 nm Resolution on DNA Origami," *Chemphyschem* **15**(12), 2431-2435 (2014).
161. L. S. Churchman et al., "Single molecule high-resolution colocalization of Cy3 and Cy5 attached to macromolecules measures intramolecular distances through time," *P Natl Acad Sci USA* **102**(5), 1419-1423 (2005).
162. A. Pertsinidis, Y. X. Zhang, and S. Chu, "Subnanometre single-molecule localization, registration and distance measurements," *Nature* **466**(7306), 647-651 (2010).
163. J. K. Trautman et al., "Near-Field Spectroscopy of Single Molecules at Room-Temperature," *Nature* **369**(6475), 40-42 (1994).
164. F. Pinaud et al., "Advances in fluorescence imaging with quantum dot bio-probes," *Biomaterials* **27**(9), 1679-1687 (2006).
165. H. P. Lu, and X. S. Xie, "Single-molecule spectral fluctuations at room temperature," *Nature* **385**(6612), 143-146 (1997).

VITA

Personal information

Name: Janel L. Davis

Date of Birth: June 24, 1989

Place of Birth: Georgetown, Guyana

Education

PhD, Biomedical Engineering, Northwestern University (US), expected in June 2020

MS, Biomedical Engineering, Northwestern University (US), 2018

BS, Biomedical Engineering, The University of Texas at Austin (US), 2014

Patents

1. Dong, B., Davis, J.L., Sun, C., Zhang, H.F., Patel, K.J., Urban, B., Backman, V., Almassalha, L., Stypula-Cyrus, Y. & Nguyen, T. (2018). Device, Methods, and Systems relating to Super-Resolution Imaging. US 20180088048 filed May 1, 2017 and issued March 29, 2018.

Peer reviewed journal publications (* Equal contribution)

1. Davis, J.L.*, Soetikno, B.*, Song, K., Zhang, Y., Sun, C. & Zhang, H.F. (2020). RainbowSTORM: An open-source ImageJ plug-in for spectroscopic single-molecule localization microscopy (sSMLM) data analysis and image reconstruction. *BioRxiv*

2. Dong, B.*, Song, K.*, Davis, J.L.*, Sun, C. & Zhang, H.F. (2020). Sub-2-nm precision in single-molecule localization microscopy using photon-accumulation enhanced reconstruction (PACER). *Manuscript in preparation*.
3. Davis, J.L., Zhang, Y., Yi, S., Du, F., Song, K., Scott, E.A., Sun, C. & Zhang, H.F. (2020). Super-resolution imaging of self-assembled nanocarriers using quantitative spectroscopic analysis for cluster extraction. *Langmuir*. 36(9), 2291-2299.
4. Sansalone, L., Zhang, Y., Mazza, M.M.A., Davis, J.L., Song, K., Captain, B., Zhang, H.F. & Raymo, F.M. (2019). High-Throughput Single-Molecule Spectroscopy Resolves the Conformational Isomers of BODIPY Chromophores. *The Journal of Physical Chemistry Letters*. 10(21), 6807-6812.
5. Zhang, Y., Song, K., Dong, B., Davis, J.L., Wang, G., Sun, C. & Zhang, H.F. (2019). Multicolor super-resolution imaging using spectroscopic single-molecule localization microscopy with optimal spectral dispersion. *Applied Optics*. 58(9), 2248-2255.
6. Kim, E.Y., Barefield, D.Y., Vo, A.H., Gacita, A.M., Schuster, E.J., Wyatt, E.J., Davis, J.L., Dong, B., Sun, C., Page, P., Dellefave-Castillo, L., Demonbreun, A., Zhang, H.F. & McNally, E.M.(2019). Distinct pathological signatures in human cellular models of myotonic dystrophy subtypes. *JCI Insight*. 4(6), e122686.
7. Davis, J.L., Dong, B., Sun, C. & Zhang, H.F. (2018). Method to identify and minimize artifacts induced by fluorescent impurities in single-molecule localization microscopy. *Journal of Biomedical Optics*. 23(10), 106501.
8. Dong, B., Davis, J.L., Sun, C., & Zhang, H.F. (2018). Spectroscopic analysis beyond the diffraction limit. *The International Journal of Biochemistry & Cell Biology*. 101, 113-117.

9. Richards, L.M., Kazmi, S.M.S., Davis, J.L., Olin, K.E. & Dunn, A.K. (2013). Low-cost laser speckle contrast imaging of blood flow using a webcam. *Biomedical Optics Express*. 4(10), 2269-83.

Conference presentations

1. Davis, J.L., Yi, S., Zhang, Y., Song, K., Scott, E.A., Sun, C. & Zhang, H.F. (2019). Quantitative super-resolution imaging of self-assembled nanocarriers via spectroscopic single molecule localization microscopy (sSMLM). *American Chemical Society National Meeting 2019* (Orlando, Florida).

2. Davis, J.L., Dong, B., Zhang, Y., Sun, C. & Zhang, H.F. (2018). Quantitative drug resistance detection via Spectroscopic Photon Localization Microscopy. *Biomedical Engineering Society Annual Meeting 2018* (Atlanta, Georgia)

3. Davis, J.L., Dong, B., Sun, C. & Zhang, H.F. (2017). Investigation of drug resistance in bacteria using Spectroscopic Photon Localization Microscopy. *Global Grand Challenges Summit 2017* (Washington, DC).

4. Davis, J.L., Dong, B., Sun, C. & Zhang, H.F. (2017). Surface cleaning and functionalization for single molecule fluorescence imaging and spectroscopy. *SPIE Photonics West, BiOS 2017* (San Francisco, CA).