Inferring system properties from thermodynamic fluctuations: a tool development approach

by

Yoon Jung

B.S., Korea Advanced Institute of Science and Technology (2014)

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2020

© Massachusetts Institute of Technology 2020. All rights reserved.

Author . . . . . .
Department of Physics
May 12, 2020

Certified by . . . . . .
Nikta Fakhri
Assistant Professor
Thesis Supervisor

Accepted by . . . . . .
Nergis Mavalvala
Associate Department Head

Inferring system properties from thermodynamic fluctuations: a tool development approach

by

Yoon Jung

Submitted to the Department of Physics
on May 12, 2020, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Physics

Abstract

Biological systems are far from equilibrium which require novel tools for unraveling their complex behavior. This thesis focuses on developing a toolbox in order to understand properties of living systems from thermodynamic fluctuations. In the first chapter, I discuss a fluorescence imaging platform which allows 3D information combined with non-invasive and photostable probes named single-walled carbon nanotubes. The second chapter discusses an image processing algorithm for analyzing the fluorescence images acquired with the proposed custom-built microscope. I demonstrate its robust image reconstruction capability under dense scenes of fluorescence images with its inherent parallel nature which allows implementation on GPUs. Finally, I develop a framework which predicts system properties from thermodynamic fluctuations in a data-driven manner. The proposed framework uses feature extraction methods based on wavelets with recurrent neural networks for processing time series data. A combination of these tools completes a pipeline which allows studying complex behavior of biological systems.

Thesis Supervisor: Nikta Fakhri
Title: Assistant Professor

Acknowledgments

I express my deepest appreciation to my academic advisor, Nikta Fakhri, for her patience and guidance throughout my studies. Her continuous support in me finding what I love doing has made my time at MIT a truly memorable one. I also thank Professors Leonid Mirny, Nuh Gedik, and Pankaj Mehta for sparing their precious time giving guidance and kindly being on my committee.

I am extremely grateful to Tony, Kuan-Chung, and Professor Rafael Piestun for being such great collaborators and patient mentors. Many thanks to the wonderful members in the Fakhri Group for their generous support: Alex, Tzer Han, Hugh, Melis, Sebastian, Jinghui, James, Yu-Chen, Junang, Ty and the Physics of Living Systems community at MIT, including past members: Hridesh and Tommaso. My work in the group has been so enjoyable thanks to all of you. Also, I very much appreciate the kind assistance from Cathy, Sydney, Karen, and Monica, and would like to extend my sincere thanks to Ms. Inshil Kim, Professors Yong-gyu Kim, Won-Ki Cho, and Iksoo Chang for their warm support during my studies.

I gratefully acknowledge the friends I have met at MIT outside the research group, including Moo Sun, Sangyeon, Andrew H., Chanwoo, Yoon Ah, Hyodong, Sungjoon, Giyoung, Sun Jin, Kunjoong, Chaeah, Sanghoon, and the Korean physics community at MIT. I thank the people that I knew before joining MIT: Keunhwan, Hyunjin, Taeshik, Seunghee, Sunghoon, Hanjoon, Woohyun, Minju, and Alex Bae. A special thanks to Yunsu, Yonghan, Young Gul, Seongheon, Hyung Seok, Jiho, Byungsoo, Sunghyun for all of the fun we have had and their friendship. I cannot leave MIT without mentioning Choongman, Dongha, and Taeseop, who taught me the meaning of support and caring for friends. This accomplishment would not have been possible without your warm encouragement. And my dear friend Jungjune, words cannot express how grateful I am for your mental support and true friendship.

I am indebted to my family including my cousins Jaehee and Jaeyeon, who have helped me endure the hardships throughout the years. I thank my sister who has generously supported me in many ways. Finally, I thank my parents for the inspiration

and spiritual support throughout my life. I have been blessed with the greatest parents I could hope to have, who are not only my parents but also my role models.

Contents

1 Introduction **15**

2 Probing 4D Nuclear Dynamics in Live Cells with Single-Walled Carbon Nanotubes **17**

2.1 Imaging single-walled carbon nanotubes 18

2.2 HILO microscopy for optical sectioning 19

2.3 Double-Helix PSF microscopy 20

2.4 Microscope setup description 21

2.5 Introducing SWNTs into living cell nuclei 23

2.5.1 SWNT preparation 23

2.5.2 Transfection 23

2.5.3 Imaging 24

2.5.4 Pixel coregistration between EMCCD and InGaAs 25

2.5.5 3D tracking of SWNTs in HeLa cell nuclei 25

2.6 Conclusion 26

3 SOLAR-STORM: Fast and robust 3D localization of fluorophores in dense clusters **27**

3.1 Introduction 27

3.2 Theory 29

3.3 Algorithm 32

3.4 Simulations 36

3.5 Experimental data 38

3.6 Benchmarking 40
3.7 Conclusion 42

4 Inferring system properties from thermodynamic fluctuations: a data-driven approach 43

4.1 Introduction 43
4.2 Problem formulation 45
4.3 Invariant scattering for feature extraction 46
4.4 Training recurrent neural networks 48
4.5 Mechanically coupled beads 49
4.5.1 Simulations 49
4.5.2 Training with simulated data 50
4.6 Experimental data 52
4.6.1 Fluorescent cortical granules in a starfish oocyte 52
4.6.2 Silica microspheres in bacterial bath 53
4.7 Conclusion 57

5 Conclusion and outlook 59

List of Figures

2-1 Fluorescence microscopy involves a trade off between different factors that determine image quality and sample integrity 18

2-2 SWNT at different axial positions imaged using a DH-PSF microscope 21

2-3 Schematic of the custom-built microscope. Quarter wave plates are placed downstream to the lasers to generate circularly polarized light. Mirrors, dichroic mirrors, and lenses are denoted as "M", "DM", and "L", respectively. The yellow line downstream to the sample plane denotes the fluorescence collected through the objective. 22

2-4 HaloTag combined with SWNTs for labeling nuclear proteins 24

2-5 Electroporation process for introducing SWNTs into cells 24

2-6 SWNT localized in the HeLa cell nucleus imaged with a DH-PSF microscope 26

3-1 A schematic of the localization process. (a) An example of the acquired raw data showing fluorophores in dense clusters. (b) SOLAR-STORM can process multiple frames in parallel and return the coordinates of the fluorophores which are marked in red. (c) An example of the localization results rendered in 3D, which corresponds to the region indicated by the white box in (b). (d) Mean intensity projection of the acquired images. (e) Localizations from individual frames are combined to obtain a single super-resolution image, where the depth is color coded. 31

3-2 Performance comparison between Easy-DHPSF (blue) and SOLAR-STORM (red) for difference densities of fluorophores. Error bars denote one standard deviation. 37

3-3 Comparison of 3D super-resolution image quality between Easy-DHPSF (a) and SOLAR-STORM (b). SOLAR-STORM is capable of reconstructing overlapping images of multiple emitters, returning more localizations than Easy-DHPSF in densely labeled regions. (c), (f) Enlarged view of the boxed region with solid lines in (a), (b), respectively. (d), (g) Enlarged view of the boxed region with dashed lines in (a), (b), respectively. (e), (h) Enlarged view of the boxed region with solid lines in (d), (g), respectively, showing a section of a microtubule which is further analyzed in Figure 3-4 39

3-4 A detailed analysis of a microtuble section shown in Figure 3-3 (e), (h). (a), (b) A depth-color coded 2D scatter plot of the localizations obtained from Easy-DHPSF and SOLAR-STORM. The localizations are fit with a straight line from which the shortest vector is computed and decomposed into axial and lateral directions for each localization. The norm of these vectors are separately shown in (c) and (d) as the localization precision differs between lateral and axial directions. . . . 41

4-1 Schematic of the framework where the scattering transform is first applied on a time series dataset using a cascade of wavelet transforms and modulus rectifiers along frequency-decreasing paths. The resulting multivariate time series is paired with quantities representing the system properties to train the recurrent neural network. 49

4-2 Prediction results on simulations of two beads. (a) Schematic of the simulation consisting of two mechanically coupled beads in contact with reservoirs at different temperatures T_h and T_c. (b) Training and validation loss plotted for training LSTM networks. (c) RMSE plotted as a function of trajectory length for the given simulated data. (d) RMSE plotted as a function of the number of trajectories in the training data. 50

4-3 Description of experimental setting. (a) Fluorescence imaging of silica microspheres in a bacterial bath. The inset indicates the beads coming in different shapes and sizes. (b) Infrared fluorescence imaging of cortical granules within a starfish ooctye. (c) Enlarged view of the confined microsphere using brightfield imaging. (d) A microsphere within a bacterial bath confined in a optical trap. 52

List of Tables

4.1 Classification results of models trained on trajectories of microspheres in bacterial bath 55

Chapter 1

Introduction

The capability to observe minute particles and their behavior when suspended in liquids using microscopes combined with stochastic thermodynamics has expanded our understanding to such microscopic systems. The invention of optical microscopes has led to the discovery of Brownian motion while an attempt for a theoretical description has brought the emergence of stochastic thermodynamics, demonstrating the importance of both experimental and theoretical tools for studying systems at small scales.

Stochastic thermodynamics is crucial as it allows quantifying fluctuations while defining thermodynamic quantities such as work, heat, and dissipation [73]. It also serves as the basis of useful tools such as the fluctuation dissipation theorem which describes the relaxation of a system perturbed from equilibrium in terms of its fluctuations in the equilibrium state [13, 38]. Einstein's relation can be thought of as an example of the fluctuation dissipation theorem which relates the particle diffusion constant and the fluid viscosity [17]. Despite its successful applications in systems taken out from equilibrium, biological systems are inherently out of equilibrium and thus limit usage of the fluctuation dissipation theorem. Thus, understanding living systems requires a novel set of theoretical approaches combined with tools that can analyze measurements for unraveling their complexity.

In this thesis, I propose a toolbox for studying systems out of equilibrium in order to better understand their complex behavior. First, I discuss an fluorescence

imaging platform with novel features with an application in living cell nuclei. The complex dynamics of nuclear proteins span a large range of length and time scales which can be studied with fluorescence microscopy. However, challenges remain as data acquisition is limited to short time scales due to fluorophore instability and photodamage by excitation light sources. I introduce an experimental technique using single-walled carbon nanotubes combined with near-infrared excitation to circumvent photodamage.

The following chapter discusses a novel image processing algorithm which is designed for a wide range of fluorescence imaging applications. This software is benchmarked against other existing methods with an application in super-resolution imaging which is an essential tool for acquiring high-resolution profiles of biological samples. The reconstructed images not only breaks the diffraction limit but also extends the visualization from 2D to 3D which is essential for unraveling the high complexity within living systems.

Finally, I discuss a data-driven framework for inferring system properties based on thermodynamic fluctuations. In contrast to model-based approaches that have focused on quantifying the system's dissipation rate, the proposed framework is based on a specialized type of neural network named recurrent neural networks and a novel feature extraction method using wavelet transforms to statistically infer a variety of system properties. It utilizes the symmetries within the data for robust inference while operating in a probe-independent manner. Combined with the microscopy and image processing tools developed in this thesis, this toolbox will serve as an effective pipeline to infer system properties from thermodynamic fluctuations in living systems.

Chapter 2

Probing 4D Nuclear Dynamics in Live Cells with Single-Walled Carbon Nanotubes

Fluorescence microscopy is a vital technique for studying biological systems across a wide range of length and time scales. Equipped with a high degree of biomolecular specificity amidst non-fluorescing structures and capability of imaging biological processes in real time, fluorescence microscopy remains an essential tool for observing rapidly occurring events in biological systems.

Observing biomolecular processes with fluorescence microscopy brings a trade-off between several factors that determine the image quality and sample integrity [47]. For example, as shown in Figure 2-1, increasing the power of the excitation light source leads to a higher amount of fluorescence which can be used to increase signal-to-noise ratio (SNR) and temporal resolution. However, the power of the light source and the settings of the detector must be optimized to fully utilize the dynamic range of the detector while not wasting the limited photon budget of the fluorophores or generate phototoxic chemicals. These settings in the microscope therefore must be carefully tuned prior to conducting experiments. In order to mitigate the experimental constraints imposed by the imaging modality, I demonstrate usage of a custom-developed imaging platform with several features.

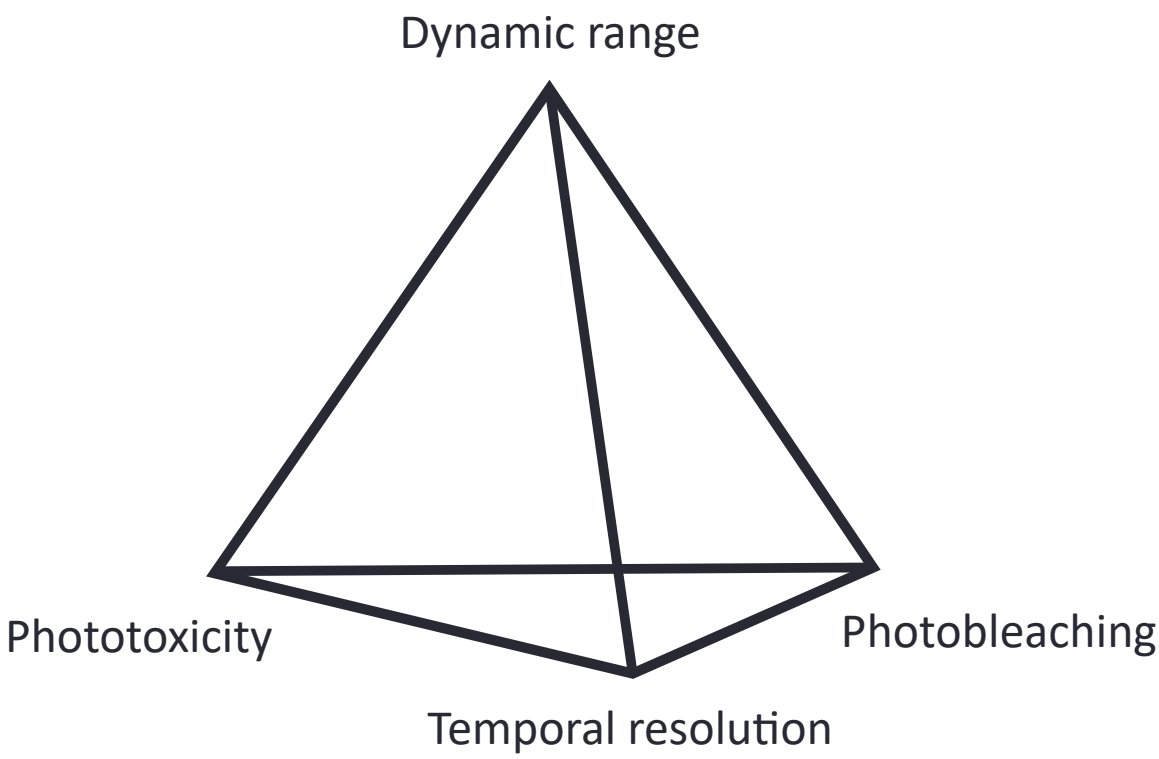

Figure 2-1: Fluorescence microscopy involves a trade off between different factors that determine image quality and sample integrity

2.1 Imaging single-walled carbon nanotubes

The microscope is optimized for imaging single-walled carbon nanotubes (SWNTs). They are cylindrical structures with high aspect ratio that can be visualized as rolled-up sheets of graphene. SWNTs can be characterized with a chiral vector $C_h = na_1 + ma_2$ where a_1, a_2 are the unit lattice vectors of the graphene sheet and n, m are integers that define the chirality [5]. It is shown that a fraction of these SWNT structures are semiconductors which under surfactant-stabilized conditions show photoluminescence in the near-infrared regime (900-1400 nm) when excited with light sources in the visible range [59]. The chirality dependent-emission wavelength allows using different types of SWNTs for multicolor imaging. The large Stokes shift in the near infrared (NIR) regime allows imaging SWNTs at high SNR as this window is free from the autofluorescence of biological tissues. The wavelength of the excitation light sources used for imaging SWNTs spans a wide range which overlaps with the emission wavelength window of a Ti:Sapphire laser. I observed that a certain chirality

of SWNTs can be imaged through excitation at a wavelength around 705 nm which is longer than the excitation wavelength of commonly used fluorophores in biological applications. This is advantageous as it induces less photodamage to biological systems and exhibits a longer scattering depth. Furthermore, SWNTs exhibit stable emission without blinking or photobleaching being suitable for long-term imaging in biological systems which has been demonstrated by tracking kinesin-1 motor proteins in COS-7 cells [19]. Finally, due to their high aspect ratio, SWNTs can diffuse into tight spaces and report on the architecture and dynamics of stiff environments in a non-perturbing manner [18].

2.2 HILO microscopy for optical sectioning

Commonly used fluorescence microscopes use a single objective to both excite fluorophores and collect the resulting fluorescence. However, the excited fluorophores which are out of focus increase the background level and thus reduce the SNR. Total internal reflection fluorescence (TIRF) microscopy is a commonly used technique to reduce the background level by utilizing the evanescent wave immediately adjacent to the interface between the sample and the coverslip [52]. However, this is limited to imaging specimens that are approximately 200 nm away from the surface. Instead, the proposed microscope configuration uses HILO, an optical sectioning technique more suitable for live cell imaging [77]. In HILO microscopy, a thin optical sheet is created by propagating the light at an offset from the center of the lens located at the back focal plane of the objective lens. This optical sheet enables exciting only fluorophores within a small axial range, leading to higher signal but lower background noise, increasing the SNR. This comes at a cost of the field of view: the larger the field of view, the thicker the optical sheet. Therefore, the optomechanics along the beam path was assembled to easily insert and remove lenses to allow different beam expansion factors in order to easily switch between different excitation areas.

2.3 Double-Helix PSF microscopy

The fluorescence of a point source collected by the objective lens passes downstream to the tube lens to be displayed through a detector. This response function of the microscope is termed the point spread function (PSF) whose shape is known as the Airy disk for conventional fluorescence microscope setups. For an Airy disk PSF, the image on the detector is radially symmetric where a Gaussian-shaped function which approximates the PSF can be used to determine the lateral position of the point source up to nanometer resolution. Finding the axial position of the point source, however, can be challenging due to several reasons. Firstly, the Airy disk PSF has a simple radially symmetric shape. This prevents accurate localization of emitters as it is nearly infeasible to determine whether the emitter is above or below the focal plane. Secondly, the Airy disk PSF shape diverges rapidly as the emitter moves along the axial direction. This not only limits the axial range where one can localize the fluorophore with reasonable accuracy, but also results in accuracy depending on the axial position of the fluorophore as the signal can drop below the background noise as the PSF shape expands beyond a certain size.

Several 3D single-molecule localization microscopy modalities have been developed in order to achieve higher localization accuracy in the axial direction. The most widely used methods are biplane imaging [34], astigmatic imaging [32], and double-helix PSF (DH-PSF) microscopy [64, 65]. Biplane imaging uses a 50:50 beam splitter to form fluorescence images of two object planes that are separated by a few hundred nanometers. Astigmatic imaging inserts a cylindrical lens in the emission arm to create a PSF that elongates along different directions depending on which side the point source is located with respect to the focal plane. For a DH-PSF microscope, the response of a point source is a PSF that encodes the axial position by having two lobes that rotate around each other when varying the axial position of the point source. The fluorescence images of a single point source using a DH-PSF microscope are shown in Figure 2-2. Each modality has its own advantages and disadvantages and must be chosen depending on the experimental requirements. In the proposed

setup, the DH-PSF microscope was implemented according to several reasons. Firstly, the DH-PSF has a high 3D localization accuracy due to its high Fisher information. Secondly, the implementation is based on a 4F system which consists of a pair of identical lenses that are spaced at a distance twice their focal length and a transparent phase mask in between. This phase mask is manufactured through lithography whose resulting pattern generates the DH-PSF. This configuration allows easy switching to the Airy disk PSF for 2D imaging by simply removing the phase mask in the Fourier plane of the 4F system. Finally, the DH-PSF has a long axial range that enables 3D localization, suitable for collecting long trajectories as particles can easily diffuse along the axial direction.

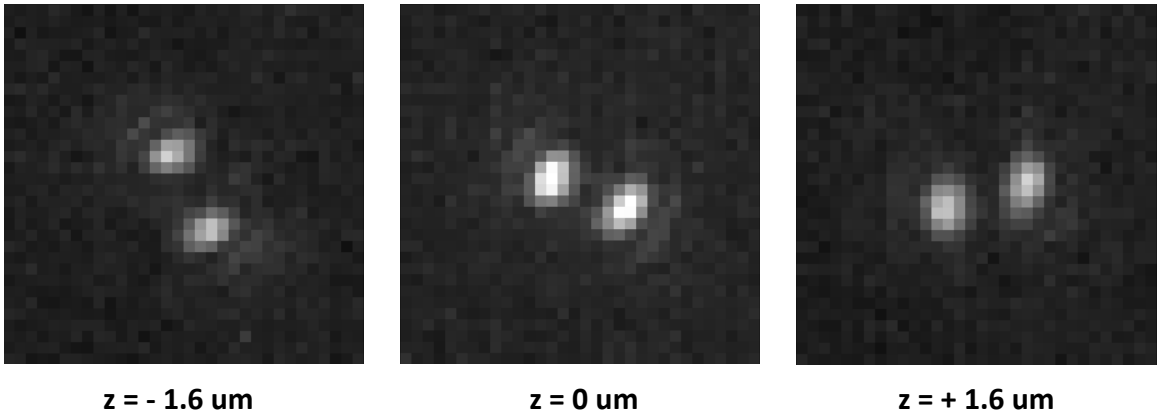

Figure 2-2: SWNT at different axial positions imaged using a DH-PSF microscope

2.4 Microscope setup description

The schematic of the custom-built fluorescence microscope is shown in Figure 2-3. Three lasers are combined into a single collinear beam path with the usage of long-pass dichroic mirrors. 488 nm and 561 nm diode OBIS lasers (Coherent) and an additional Ti:Sapphire laser (Coherent Mira-HP) are used for excitation. A fieldstop (S) is located in a plane conjugated to the sample plane which enables optical sectioning at the cost of a smaller field of view when closed to a certain degree. The fluorescence collected by the objective enters the two-channel 4F system which contains two DH-PSF phase masks for 3D imaging. This allows 3D imaging by measuring the tilt of a

line connecting the two lobes in the fluorescence image. Overall, the optical sectioning and 3D tracking feature in both the visible and NIR spectrum allows imaging of a wide range of applications when studying biological systems.

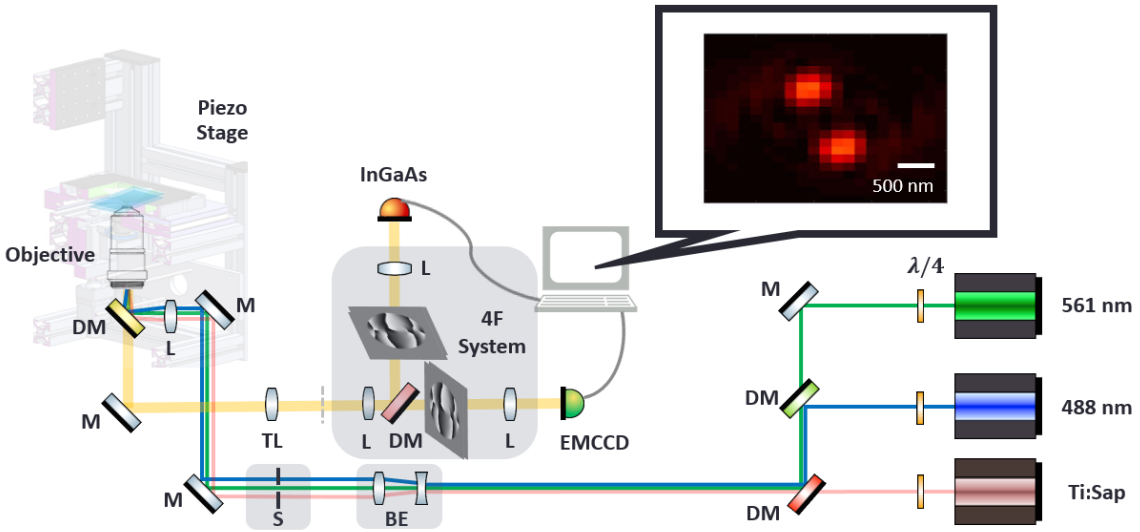

Figure 2-3: Schematic of the custom-built microscope. Quarter wave plates are placed downstream to the lasers to generate circularly polarized light. Mirrors, dichroic mirrors, and lenses are denoted as "M", "DM", and "L", respectively. The yellow line downstream to the sample plane denotes the fluorescence collected through the objective.

A number of components that make up the microscope are customized in order to match the unique features of the SWNTs. First, as most fluorophores widely used for biological applications emit light in the visible spectrum, a number of dichroic mirrors and filters are custom-designed in order to match the excitation/emission spectrum of the SWNTs. As the fluorophores used for experiments either emit in the visible range or NIR, achromatic lenses must be placed in the 4F system to minimize aberration. Next, the phase mask located upstream of the detectors is custom-designed with a pattern optimized for the NIR spectrum since the response on the detector of a fluorophore for the DH-PSF modality depends on the emission wavelength in terms of rotation rate of the two lobes and their sizes. Furthermore, an InGaAs detector is used for imaging the SWNTs due to its superior quantum efficiency in the NIR regime compared to EMCCDs [80]. The InGaAs detector requires additional cooling as the

image quality is limited by dark-current noise which can be reduced only by lowering the detector temperature. Finally, in order to reduce unnecessary phototoxicity and photobleaching of the fluorophores, mechanical shutters are placed downstream to the lasers in order to synchronize excitation with image acquisition. These mechanical shutters are controlled through a data acquisition (DAQ) card (National Instruments) via a custom LabVIEW script.

2.5 Introducing SWNTs into living cell nuclei

In order to observe intracellular dynamics of SWNTs, I develop a biochemical platform for introducing SWNTs into living HeLa cells for fluorescence imaging. HaloTag labeling is based on a covalent bond between a reporter protein and a specific ligand which forms under general physiological conditions [48]. d(T)30 oligonucleotides with 5'-Amine modification (Invitrogen) and HaloTag® Succinimidyl Ester Ligands (Promega) are used to functionalize SWNTs.

2.5.1 SWNT preparation

SWNTs are provided by Rice University's Richard E. Smalley Institute for Nanoscale Science and Technology. 1 mg of SWNTs (Rice University, batch number 189.2) produced through the high-pressure carbon monoxide (HiPco) process is mixed with 2 mg of Halo-ligand d(T)30 oligonucleotides (Invitrogen) in 2 ml of deionized water which is tip-sonicated at 20 kHz at a power of 20 W for 90 minutes. The sonicated sample is ultracentrifuged at 16000 g for 90 minutes where 400 µL from the surface is collected and stored at 4 °C.

2.5.2 Transfection

Genetically modified stable HeLa cell lines expressing Halo and NeonGreen conjugate to H2B are transfected with SWNTs. Cells at 50 % confluency are washed with PBS and trypsinized for 3 minutes at 37 °C followed by adding 5 ml of GlutaMAX

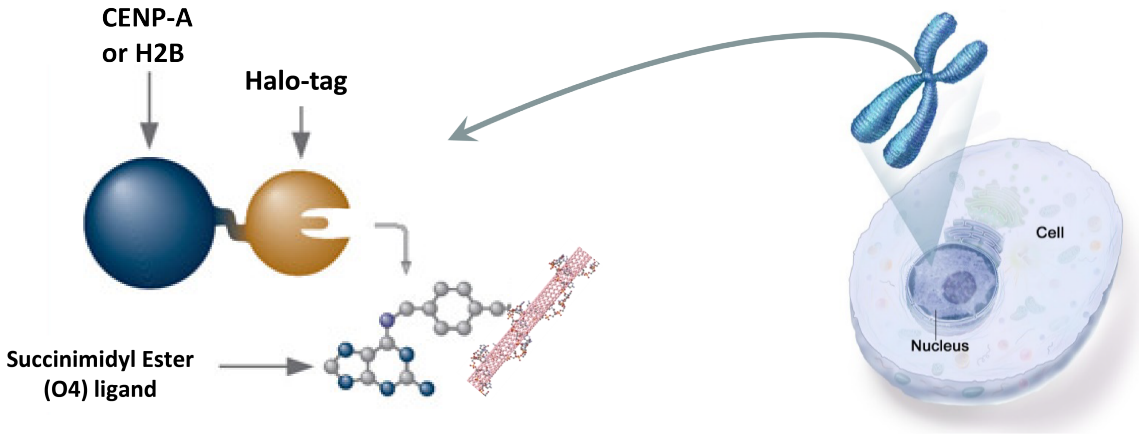

Figure 2-4: HaloTag combined with SWNTs for labeling nuclear proteins

DMEM (Thermofisher). A volume corresponding to 1 million cells determined by a hemocytometer is centrifuged at 1000 rpm for 5 minutes for resuspension in a mixture of 82 µL of the SE solution, 18 µL of the supplement (Lonza) and 2 µL of the prepared SWNT stock. 100 µL of the mixture is transferred to an electroporation cuvette and electroporated using the X unit nucleofector (Lonza). Cells are incubated in glass-bottom petri dishes (Matec) at 37 °C for 12 to 18 hours prior to imaging.

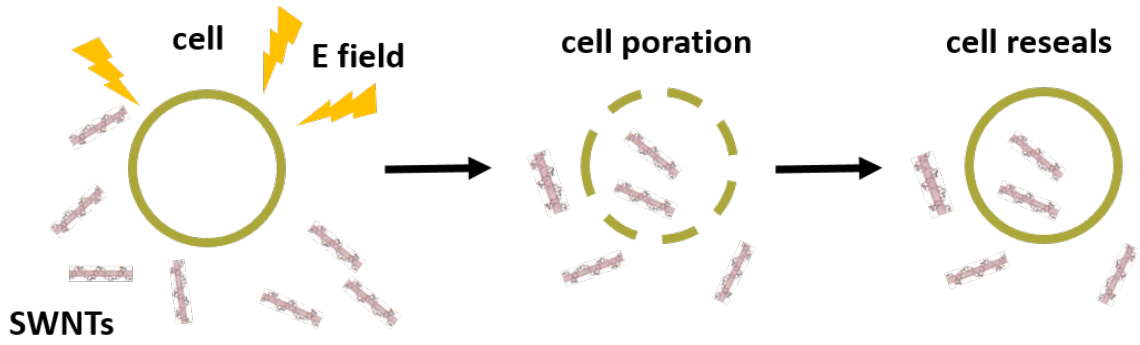

Figure 2-5: Electroporation process for introducing SWNTs into cells

2.5.3 Imaging

Incubated cells are washed with fresh PBS and GlutaMAX DMEM and mounted on the microscope enclosed with a heating system (ibidi) to maintain a temperature

close to 37 °C. Cells are imaged on the custom-built dual-channel DH-PSF microscope while the phase mask is removed from only the visible channel in the emission arm for conventional fluorescence microscopy. The sample is illuminated with either the 561 nm or Ti:Sapphire laser at 705 nm for imaging SWNTs. NeonGreen-labeled H2B is imaged with 488 nm excitation in order to identify the nucleus. The EMCCD (Andor iXon 897) installed in the visible channel of the emission arm is set with an exposure time of 500 ms with 2 MHz readout rate. The InGaAs (Princeton Instruments) is set with an exposure time of 1 s with readout rate 2 MHz as the SWNT dynamics only displayed distinctive motion at long time scales.

2.5.4 Pixel coregistration between EMCCD and InGaAs

In order to overlay images acquired from the two channels, NeonGreen conjugated to SWNTs through HaloTag labeling is used as markers to coregister pixels between the EMCCD in the visible channel and the InGaAs in the NIR channel. (x, y) coordinates of 4 to 5 identified complexes are used to calculate the linear mapping between the coordinates in the EMCCD frame and the InGaAs frame using least squares minimization. The linear mapping obtained from these markers are then used to rotate, scale, and translate the InGaAs images with respect to the EMCCD images to overlay in ImageJ software.

2.5.5 3D tracking of SWNTs in HeLa cell nuclei

An example snapshot of the SWNTs localized within the nucleus of a HeLa cell is shown in Figure . Images are recorded for time periods up to several hours, during which the fluorescence level of SWNTs does not change, demonstrating the high photostability of SWNTs in biological tissues. Furthermore, some SWNTs exhibit abrupt jumps in position likely due to active processes within the nucleus. The red channel displays high SNR, which is due to the high specificity in the NIR window suggesting SWNTs are effective probes for imaging biological tissues. The phase mask in the NIR channel is inserted to allow 3D localization of SWNTs whose trajectory is

processed with a custom-written software described in the following chapter.

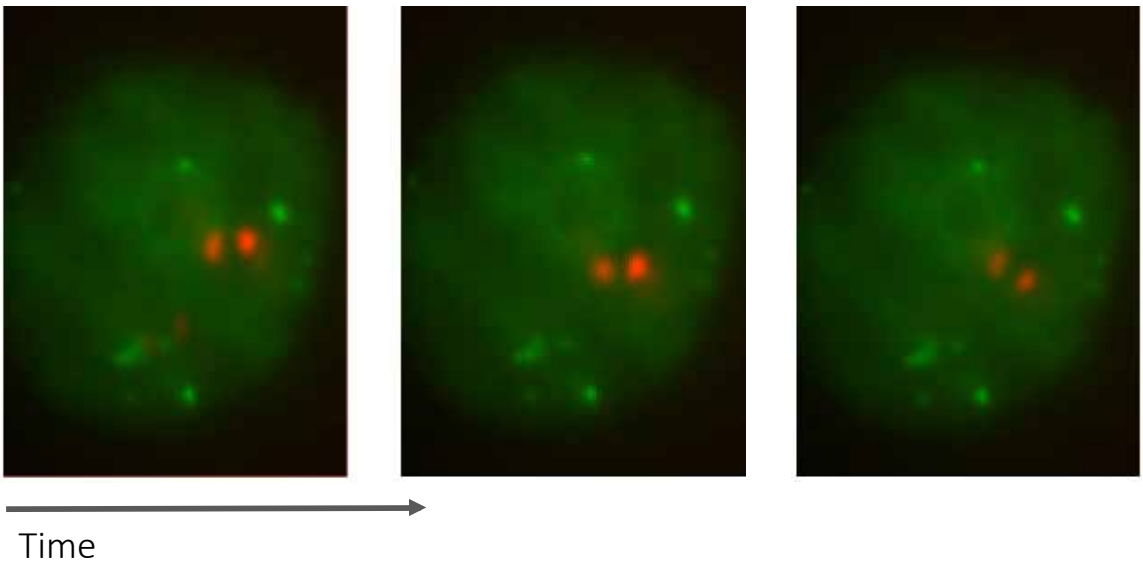

Figure 2-6: SWNT localized in the HeLa cell nucleus imaged with a DH-PSF microscope

2.6 Conclusion

Fluorescence microscopy for living biological samples can be challenging due to several constraints such as the limited photon budget of commonly used fluorophores and phototoxicity. In this chapter, I demonstrated implementation of an imaging modality that extends a conventional fluorescence microscope to mitigate these constraints. The custom-built setup allows imaging SWNTs in 3D with less photodamage from excitation light sources at reduced background noise for improved image quality. This non-invasive imaging modality may allow tracking SWNT-labeled biomolecules for long periods of time at high temporal resolution which will be a great benefit to studies involving optical imaging of live cells.

Chapter 3

SOLAR-STORM: Fast and robust 3D localization of fluorophores in dense clusters

This chapter is adopted with minor modifications from a manuscript prepared for submission to Optics Express with coauthors Anthony Barsic, Rafael Piestun, and Nikta Fakhri.

3.1 Introduction

Super-resolution localization microscopy is based on sparse activation of fluorophores whose spatial coordinates can be calculated from their non-overlapping fluorescence images [69, 27, 9]. By combining thousands to tens of thousands of frames, a single super-resolution image can be reconstructed. However, limiting the number of activated fluorophores per frame to only a few requires longer total acquisition time, imposing constraints on the experimental design. These experimental constraints can be relaxed by using advanced reconstruction algorithms that can localize fluorophores in dense clusters, allowing fewer frames to be collected to reconstruct a single super-resolution image. The underlying approaches of these localization methods include maximum-likelihood estimation, sparsity, and deep learning

[14, 30, 33, 83, 57, 3, 56, 55, 58, 60]. By returning the positions of fluorophores in dense clusters, fewer frames are required to reconstruct a single super-resolution image, resulting in higher temporal resolution and enabling observation of nanoscale dynamics in living systems. These methods can also be applied to 3D localization using engineered point spread functions (PSFs) with higher axial Fisher information [64, 32, 34, 66, 26, 2, 24, 43, 54, 62, 25, 6, 61, 4, 10, 46, 51, 82, 45]. However, extending localization in dense clusters to 3D can be challenging. Unlike the commonly used 2D Gaussian approximation of the airy-disk PSF, many engineered PSFs for 3D imaging cannot be described with an analytic expression, especially in the presence of always-present aberrations [4]. Although the full 3D PSF can be experimentally acquired for calibration, its larger axial range and complex shape for encoding the axial position requires larger memory and longer localization runtime. This also leads to a higher chance of mislocalization in dense scenes since images of adjacent fluorophores are more likely to overlap with each other and signal-to-noise ratio (SNR) is degraded due to the spread of photons throughout the axial and lateral directions.

Here, I develop algorithms and software based on compressive sensing theory termed Spline fitting and Orthogonal matching pursuit-based Localization Algorithm for Reconstructing STochastic Optical Reconstruction Microscopy images (SOLAR-STORM), which allows robust 3D localization of fluorophores in dense clusters. I evaluate the localization performance of SOLAR-STORM using simulated fluorescent emitters with varying density, and synthetic STORM data sets modeling microtubule structures provided by the 2016 single-molecule localization microscopy (SMLM) challenge to compare with other sofware packages. I also run SOLAR-STORM on STORM images of fluorescent protein-labeled microtubules in PtK1 cells. Evaluation results indicate that, compared to other algorithms and software packages, SOLAR-STORM provides robust image reconstruction at superior speed across different types of data sets. The full package can be found at `https://github.com/yoonjung0705/solar-storm`.

3.2 Theory

Motivated by compressive sensing theory, fluorescent image reconstruction for localizing fluorophores can be formulated into the following mathematical problem called constrained basis pursuit denoising (CBPDN) [83].

$$\text{Minimize: } \|\mathbf{x}\|_1, \text{ Subject to: } \mathbf{x} \geq 0, \ \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon\sqrt{\|\mathbf{y}\|_1} \tag{3.1}$$

$\mathbf{A}$ is the dictionary matrix where different columns correspond to fluorescence images of a single-molecule located at different positions. If the PSF is acquired experimentally, cubic splines are used to interpolate along the axial direction for a given z-stack of a fluorescent emitter. The dictionary matrix can then be constructed for any given 3D position within the z-stack range. For shifts along the axial direction, The elements of $\mathbf{x}$ are the photon counts contributed by the corresponding columns $\mathbf{A}$ in the given fluorescence image $\mathbf{y}$, and a positive value of ϵ is to allow signal recovery in noisy conditions. In 3D localization, it is not practical to directly solve the above global optimization problem using basis pursuit such as interior point methods, or l1-homotopy due to long runtime [3]. This computational complexity remains challenging even for a small patch size or using a coarse grid-based PSF matrix. Instead, I use a greedy algorithm termed orthogonal matching pursuit (OMP) which has shown efficient signal reconstruction in linear inverse problems [15, 79]. From a small number of noisy linear measurements, OMP can reconstruct the exact support of a high-dimensional sparse signal with high probability with a shorter runtime compared to basis pursuit. Still, OMP must be highly parallelized in order to efficiently reconstruct tens of thousands of images with a reasonable time and space complexity.

In order to reconstruct multiple images in parallel, each image is first divided into patches that slightly overlap with each other. Based on a previous study, I develop a parallel version of OMP executed on a graphics processing unit which allows simultaneous localization in hundreds or tens of hundreds of patches to overcome challenges in 3D localization [71]. SOLAR-STORM also uses various stopping criteria such as the Bayesian information criterion (BIC) to perform localization with no thread di-

vergence on patches containing different numbers of fluorescent emitters [67]. This systematic model selection allows SOLAR-STORM to set its maximum number of expected fluorophores to a fixed value which does not require tuning hyperparameters for different data sets. False-positives can be further removed based on local density filtering or eliminating localizations with photon counts below a given threshold.

In the above optimization problem, the dictionary matrix $\mathbf{A}$ is formulated on a discrete grid and therefore the 3D coordinates obtained from OMP can only have discretized values. To output a solution on a continuous space, SOLAR-STORM performs debiasing from the discrete grid using the Taylor expansion of the PSF which is a modification of an approach used in previous studies [56, 54]. Using the Taylor expansion of the PSF allows a coarser dictionary and thus dramatically reduces the space and time complexity of the algorithm. However, this debiasing step in the proposed method differs from previous studies as it is performed concurrently for all patches with no thread divergence, maintaining the overall parallel structure of the algorithm.

A schematic of SOLAR-STORM's localization process is displayed in the simulated images in Figure 3-1. A point source is convolved with the microscope's PSF which in this case is the double-helix PSF (DH-PSF). For a DH-PSF microscopy setup, the response of a point source in the sample is a fluorescence image consisting of two lobes which rotate around each other as the point source moves along the axial direction. Thus, the axial position of the point source is encoded in the rotation amount of the two lobes. Remarkably, SOLAR-STORM allows on-the-fly processing of fluorescence images, thereby returning the 3D coordinates of the fluorescent emitters during data acquisition ((a)-(c)). These locations are shown in 3D in (c), before and after the Taylor expansion-based debiasing as orange and red marks, respectively. Since this Taylor expansion is only valid for a small region around a point on the 3D grid, the debiasing is not performed if the calculated correction amount is larger than the step size of the dictionary. The orange boxes show the threshold on the amount of correction allowed. A fraction of the localizations are removed based on criteria including BIC and photon counts in order to reduce the number of false

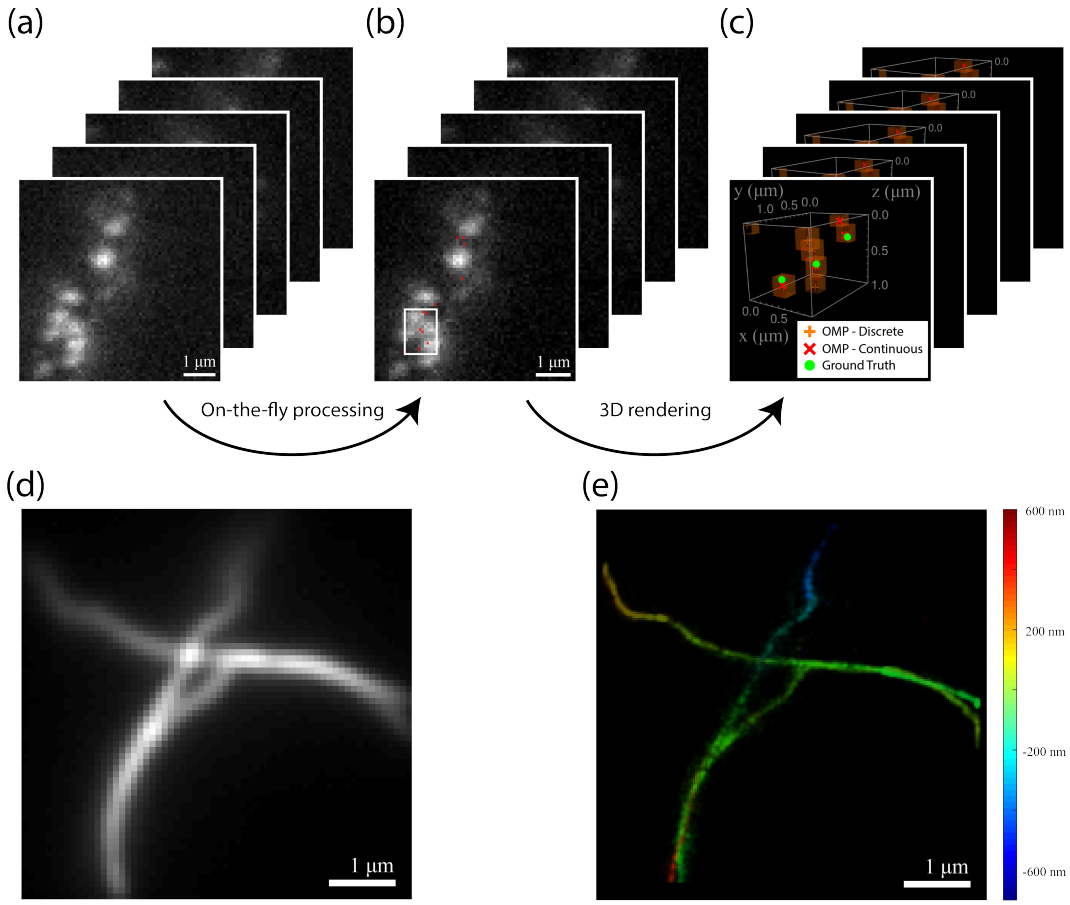

Figure 3-1: A schematic of the localization process. (a) An example of the acquired raw data showing fluorophores in dense clusters. (b) SOLAR-STORM can process multiple frames in parallel and return the coordinates of the fluorophores which are marked in red. (c) An example of the localization results rendered in 3D, which corresponds to the region indicated by the white box in (b). (d) Mean intensity projection of the acquired images. (e) Localizations from individual frames are combined to obtain a single super-resolution image, where the depth is color coded.

positives. Finally, these localizations from individual frames are displayed altogether to produce a single depth-color coded super-resolution image shown in (e). (d) shows the mean intensity projection of the raw data images. I use the publicly available training data sets obtained from the 2016 SMLM challenge website which was labeled as MT0.N1.HD (high density, low noise data set) which corresponds to simulated 3D data of fluorescently labeled microtubules [70].

3.3 Algorithm

In the CBPDN problem, the set of possible linear combinations of atoms that are equivalent to the measured signal within the noise level is considered. Within this set, the linear combination of atoms that contains the minimum number of nonzero coefficients is the solution to this CBPDN problem. In this section, the pseudo-code of SOLAR-STORM is discussed which provides an approximate solution to this CBPDN problem. I first discuss the serial version of OMP and show how this extends to the parallel implementation for multiple data sets with the debiasing step. In the serial version of OMP shown in Algorithm 1, the correlation is computed between the residual and the columns (atoms) in the dictionary for each iteration. The atom that shows the largest correlation is selected and added to the support, followed by a least squares minimization to find the coefficients that give the optimal approximation of the data set using the atoms in the support. The residual is computed by subtracting the approximation from the data set, while the atom added to the support is removed from the dictionary so that an atom already in the support does not get selected again. Updating the dictionary by removing atoms in the support ensures the support matrix being full rank which is required for using the closed form solution when solving the least squares problem in Algorithm 1 step 5. The process terminates when 1) negative coefficients are found, or 2) when the maximum number of iterations is reached. The serial version of the overall algorithm is shown below, where m and n denote the number of pixels in the patch and number of atoms in the dictionary, respectively.

Algorithm 1 Orthogonal Matching Pursuit (serial version)

Inputs: L_2 normalized Dictionary $\mathbf{A} \in \mathbb{R}^{m \times n}$, measurement $\mathbf{y} \in \mathbb{R}^{m \times 1}$, maximum number of iterations K

Ouputs: Sparse coefficient vector $\mathbf{x} \in \mathbb{R}^{n \times 1}$

1: **Initialization:** Support $\mathcal{I}_0 = \emptyset$, iteration number $k = 0$, residual $\mathbf{r}_0 = \mathbf{y}$

2: **while** $k < K$ & $\max\limits_{i \in \mathcal{I}_k^{\complement}}(\mathbf{a}_i^T \mathbf{r}) > 0$ **do**

3: $\quad i_{k+1} = \arg\max\limits_{i \in \mathcal{I}_k^{\complement}}(\mathbf{a}_i^T \mathbf{r})$

4: $\quad \mathcal{I}_{k+1} = \mathcal{I}_k \bigcup \{i_{k+1}\}, \mathbf{A}_{k+1} = [\mathbf{A}_k \ \mathbf{a}_{i_k}]$

5: $\quad \mathbf{x}_{k+1} = \arg\min\limits_{\mathbf{x} \geq \mathbf{0}} \|\mathbf{y} - \mathbf{A}_{k+1}\mathbf{x}\|_2$

6: $\quad \mathbf{r}_{k+1} = \mathbf{y} - \mathbf{A}_{k+1}\mathbf{x}_{k+1}$

7: $\quad k \leftarrow k + 1$

8: **end while**

The above steps can be modified to process multiple data sets simultaneously as shown in Algorithm 2 where q is an index for the patches. The data sets are represented as a matrix instead of a single column vector and the dictionary is $L2$ normalized which allows computing the correlation with a single matrix-matrix multiplication that can be efficiently performed on a GPU. The atoms that give maximum correlation values are stored in the support for each patch simultaneously. Unlike the case of a single patch, atoms in the support cannot be removed from the dictionary since the dictionary is shared among multiple patches. Furthermore, all data sets must be processed simultaneously even when the stopping criteria has been reached for a fraction of the patches. Therefore, I maintain the Gram-Schmidt-based QR-factorization of the matrix containing the atoms in the support for all patches at every iteration. In this factorization, the corresponding diagonal entry of the diagonal matrix R is 0 for an atom that appears more than once in the support. Therefore, the least squares solution can be computed without attempting to calculate the inverse of the matrix $\mathbf{A}_{k+1}^T \mathbf{A}_{k+1}$ while still being able to process multiple patches simultaneously. This property of the QR-factorization can also be utilized to effectively terminate OMP for a fraction of the patches. For a patch that meets one of

the stopping criteria, SOLAR-STORM adds an existing atom to the support, which automatically sets the coefficients of newly added atoms to zero, bringing no effective change to the solution. This not only solves the issue of having two identical atoms in the support but also allows efficient calculation of the least squares solution for all patches without thread divergence on GPUs.

Algorithm 2 Orthogonal Matching Pursuit (parallel version)

Inputs: L_2 normalized Dictionary $\mathbf{A} \in \mathbb{R}^{m \times n}$, measurement $\mathbf{y}_q \in \mathbb{R}^{m \times 1}$, maximum number of iterations K

Ouputs: Sparse coefficient vector $\mathbf{x}_q \in \mathbb{R}^{n \times 1}$

Initialization: Support $\mathcal{I}_{q,0} = \emptyset$, iteration number $k = 0$, residual $\mathbf{r}_{q,0} = \mathbf{y}_q$
while $k < K$ **do**
  $i_{q,k+1} = \arg\max_{i \in \mathbb{U}} \left| \mathbf{a}_i^T \mathbf{r}_{q,k} \right|$
  $\mathcal{I}_{q,k+1} = \mathcal{I}_{q,k} \bigcup \{i_{q,k+1}\}, \mathbf{A}_{q,k+1} = [\mathbf{A}_{q,k} \ \mathbf{a}_{i_{q,k}}]$
  if $k > 0$ **then**
    $\mathbf{w}_q = \mathbf{Q}_{q,k}^T \mathbf{a}_{i_{q,k}}$
    $\mathbf{Q}_{q,k+1} = \begin{bmatrix} \mathbf{Q}_{q,k} & \frac{\mathbf{a}_{i_{q,k}} - \mathbf{Q}_{q,k}\mathbf{w}_q}{\sqrt{\|\mathbf{a}_{i_{q,k}}\|^2 - \|\mathbf{w}_q\|^2}} \end{bmatrix}$
    $\mathbf{R}_{q,k+1} = \begin{bmatrix} \mathbf{R}_{q,k} & \mathbf{w}_q \\ \mathbf{0}^T & \sqrt{\|\mathbf{a}_{i_{q,k}}\|^2 - \|\mathbf{w}_q\|^2} \end{bmatrix}$
  else
    $\mathbf{Q}_{q,k+1} = \mathbf{a}_{i_{q,k}}, \mathbf{R}_{q,k+1} = 1$
  end if
  $\mathbf{x}_{q,k+1} = \mathbf{R}_{q,k+1}^{-1} \mathbf{Q}_{q,k+1}^T \mathbf{y}_q$
  $\mathbf{r}_{q,k+1} = \mathbf{y}_q - \mathbf{A}_{q,k+1} \mathbf{x}_{q,k+1}$
  $k \leftarrow k + 1$
end while

After obtaining an approximate solution through OMP, SOLAR-STORM performs debiasing using the Taylor expansion of the PSF. For each iteration of SOLAR-STORM, this Taylor expansion is computed with respect to the spatial coordinates $\widetilde{x}, \widetilde{y}, \widetilde{z}$, and concatenated to form matrix $\widetilde{\mathbf{A}}_{q,k}$. The newly added atom in the current

iteration is also included in $\widetilde{\mathbf{A}}_{q,k}$ which serves as a partial derivative of the PSF with respect to the number of photon counts. In order to maintain the parallel structure of the algorithm, the QR-factorization of $\widetilde{\mathbf{A}}_{q,k}$ is computed to solve the least squares problem through back substitution. The final 3D coordinates on a continuous space are obtained by applying the correction to the coordinates given by OMP on the predetermined grid.

Algorithm 3 Orthogonal Matching Pursuit (parallel version, continuous space)

Inputs: L_2 normalized Dictionary $\mathbf{A} \in \mathbb{R}^{m\times n}$, measurement $\mathbf{b}_q \in \mathbb{R}^{m\times 1}$, lookup table $\mathcal{L}$, maximum number of iterations K

Ouputs: Sparse coefficient vector $\mathbf{x}_q \in \mathbb{R}^{n\times 1}$

Initialization: Support $\mathcal{I}_{q,0} = \emptyset$, iteration number $k = 0$, residual $\mathbf{r}_{q,0} = \mathbf{y}_q$
while $k < K$ **do**
  $i_{q,k+1} = \arg\max_{i\in\mathbb{U}} \left|\mathbf{a}_i^T \mathbf{r}_{q,k}\right|$
  $\mathcal{I}_{q,k+1} = \mathcal{I}_{q,k} \bigcup \{i_{q,k+1}\}, \mathbf{A}_{q,k+1} = [\mathbf{A}_{q,k}\ \mathbf{a}_{i_{q,k}}]$
  if $k > 0$ **then**
    $\mathbf{w}_q = \mathbf{Q}_{q,k}^T \mathbf{a}_{i_{q,k}}$
    $\mathbf{Q}_{q,k+1} = \begin{bmatrix} \mathbf{Q}_{q,k} & \frac{\mathbf{a}_{i_{q,k}} - \mathbf{Q}_{q,k}\mathbf{w}_q}{\sqrt{\|\mathbf{a}_{i_{q,k}}\|^2 - \|\mathbf{w}_q\|^2}} \end{bmatrix}$
    $\mathbf{R}_{q,k+1} = \begin{bmatrix} \mathbf{R}_{q,k} & \mathbf{w}_q \\ \mathbf{0}^T & \sqrt{\|\mathbf{a}_{i_{q,k}}\|^2 - \|\mathbf{w}_q\|^2} \end{bmatrix}$
  else
    $\mathbf{Q}_{q,k+1} = \mathbf{a}_{i_{q,k}}, \mathbf{R}_{q,k+1} = 1$
  end if
  $\mathbf{x}_{q,k+1} = \mathbf{R}_{q,k+1}^{-1} \mathbf{Q}_{q,k+1}^T \mathbf{y}_q$
  $\mathbf{r}_{q,k+1} = \mathbf{y}_q - \mathbf{A}_{q,k+1}\mathbf{x}_{q,k+1}$
  $\widetilde{\mathbf{d}}_{q,k+1} = [\ \widetilde{x}_{q,k+1}\ \ \widetilde{y}_{q,k+1}\ \ \widetilde{z}_{q,k+1}\ ]^T = \mathcal{L}(i_{q,k+1})$
  $\widetilde{\mathbf{A}}_{q,k+1} = \left[\left.\frac{\partial \mathbf{a}_{i_{q,k}}}{\partial \widetilde{x}}\right|_{\widetilde{\mathbf{d}}_{q,k+1}} \ \left.\frac{\partial \mathbf{a}_{i_{q,k}}}{\partial \widetilde{y}}\right|_{\widetilde{\mathbf{d}}_{q,k+1}} \ \left.\frac{\partial \mathbf{a}_{i_{q,k}}}{\partial \widetilde{z}}\right|_{\widetilde{\mathbf{d}}_{q,k+1}} \ \mathbf{a}_{i_{q,k}} \right]$
  $\widetilde{\mathbf{A}}_{q,k+1} = \widetilde{\mathbf{Q}}_{q,k+1}\widetilde{\mathbf{R}}_{q,k+1}$
  $\boldsymbol{\delta}_{q,k+1} = [\ \Delta\widetilde{x}_{q,k+1}\ \ \Delta\widetilde{y}_{q,k+1}\ \ \Delta\widetilde{z}_{q,k+1}\ \ \Delta\widetilde{h}_{q,k+1}\ ]^T$
    $= \widetilde{\mathbf{R}}_{q,k+1}^{-1} \widetilde{\mathbf{Q}}_{q,k+1}^T \mathbf{r}_{q,k+1}$
  $\mathbf{r}_{q,k+1} \leftarrow \mathbf{r}_{q,k+1} - \widetilde{\mathbf{A}}_{q,k+1}\boldsymbol{\delta}_{q,k+1}$
  $\widetilde{x}_{q,k+1} \leftarrow \widetilde{x}_{q,k+1} + \Delta\widetilde{x}_{q,k+1}$
  $\widetilde{y}_{q,k+1} \leftarrow \widetilde{y}_{q,k+1} + \Delta\widetilde{y}_{q,k+1}$
  $\widetilde{z}_{q,k+1} \leftarrow \widetilde{z}_{q,k+1} + \Delta\widetilde{z}_{q,k+1}$
  $k \leftarrow k + 1$
end while

Finally, BIC can be evaluated for each patch in order to perform systematic model selection [67]. Several stopping criteria have been implemented in the parallel version of OMP to reduce the number of false positives. Still, due to its greedy nature, OMP generates many false positives compared to the solution obtained from basis pursuit. I derive a BIC formula specifically for OMP and select the iteration number that has the lowest value for each patch. This model selection process can be combined with other stopping criteria such as setting photon counts threshold. BIC calculation involves computing the likelihood, which was derived below assuming Poisson noise for the photon counts. Here, $\mathbf{y}$ and $\hat{\mathbf{y}}$ denote the measured photon counts in the patch, and the photon counts from the least-squares approximation using the support subspace in OMP, respectively. Note that OMP runs after subtracting the estimated background from the raw data, therefore the estimated background was added prior to the likelihood evaluation.

$$\mathcal{L}(\hat{\mathbf{y}}|\mathbf{y}) = \prod_{j=1}^{m} \frac{e^{-\hat{y}_j}\hat{y}_j^{y_j}}{y_j!} \tag{3.2}$$

$$\begin{aligned} BIC &= k \ln m - 2 \ln \mathcal{L}(\hat{\mathbf{y}}|\mathbf{y}) \\ &\approx k \ln m - 2 \sum_{j=1}^{m} \left[(y_j - \hat{y}_j) + y_j \ln \left(\hat{y}_j / y_{q,j} \right) \right] \end{aligned} \tag{3.3}$$

3.4 Simulations

I assess the localization performance of SOLAR-STORM by simulating images of fluorescent emitters in a 25.6 µm × 25.6 µm region. For the PSF, I use the z-stack images of the DH-PSF provided by the 2016 SMLM challenge. For each frame, a fixed number of emitters were randomly distributed within a 19 µm × 19 µm region to mitigate boundary effects prior to being convolved with the DH-PSF which approximately spanned a 6.6 µm × 6.6 µm region. All simulated emitters were set to be activated for each frame. The final images had 256 × 256 pixels where each pixel had a width of 100 nm. I simulate 50 frames for each fluorophore density

ranging from 0.1 emitters / µm^2 to 2 emitters / µm^2. As emitters were placed in a 19 µm × 19 µm region, 0.1 emitters / µm^2 roughly corresponded to 38 emitters per frame. Figure 3-2 shows the measured performance metrics used for evaluating the localization results of SOLAR-STORM and a readily available algorithm Easy-DHPSF which is based on a single-molecule fitting approach [43]. Parameter settings for the Easy-DHPSF software were tuned until the best results were achieved. I further discuss the comparative performance of SOLAR-STORM and other multi-emitter fitting algorithms in the benchmarking section.

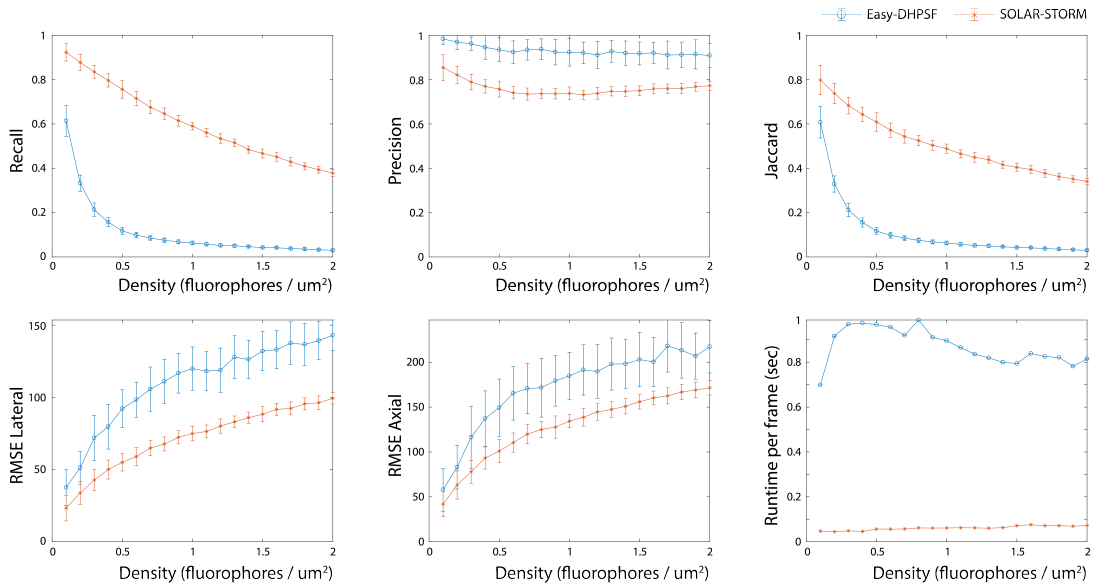

Figure 3-2: Performance comparison between Easy-DHPSF (blue) and SOLAR-STORM (red) for difference densities of fluorophores. Error bars denote one standard deviation.

The metrics in Figure 3-2 are defined as follows where TP, TN, FP, FN denote the number of true positives, true negatives, false positives, and false negatives, respectively.

$$
\begin{aligned}
(Recall) &= \frac{TP}{TP + FN} \\
(Precision) &= \frac{TP}{TP + FP} \\
(Jaccard) &= \frac{TP}{TP + FP + FN}
\end{aligned}
\tag{3.4}
$$

Figure 3-2 shows that SOLAR-STORM outperforms Easy-DHPSF for the given

range of fluorophore densities. In comparison to Easy-DHPSF, SOLAR-STORM shows higher recall, due to its multi-emitter fitting nature. This comes at a cost of having more false positives indicated by the precision metric. However, as this depends on the parameter settings, false positives can be further removed by lowering the number of iterations and the changing the photon counts threshold in SOLAR-STORM. The BIC-based systematic model selection also removes false positives and therefore does not require the user to configure parameters specifically to different data sets. Due to the trade-off between recall and precision I provide the Jaccard index to better assess the performance of the two algorithms. As SOLAR-STORM effectively reduces the number of false positives through model selection, the Jaccard index plot is largely determined by the number of true positives that are found and thus follows the trend of the recall plot. The root-mean-square error (RMSE) computed along the lateral and axial direction is lower for SOLAR-STORM compared to Easy-DHPSF for the entire range of fluorophore densities. This is majorly due to SOLAR-STORM's PSF Taylor expansion-based correction when the emitter density is low, and its capability of localizing multiple emitters with high precision when the emitter density is high.

The runtime for processing a single frame is also shown for both algorithms. The software runtime was measured using an NVIDIA TITAN X GPU and Intel(R) Core(TM) i7-4790 CPU. The figure shows that SOLAR-STORM completes image reconstruction in a shorter amount of time compared to Easy-DHPSF not only because of its execution on a GPU but mainly due to the algorithm's highly parallel structure.

3.5 Experimental data

To demonstrate the robust performance of SOLAR-STORM for experimental data, I compare with Easy-DHPSF on a super-resolution imaging data set of fluorescently labeled microtubules shown in Figure 3-3. The SPINDLE microscope system was used [24] with a 1.45 NA 100x objective (Nikon) and an electron multiplying charged coupled device (Andor iXon DU897E CS0 BV). For activation, the 488 nm line of

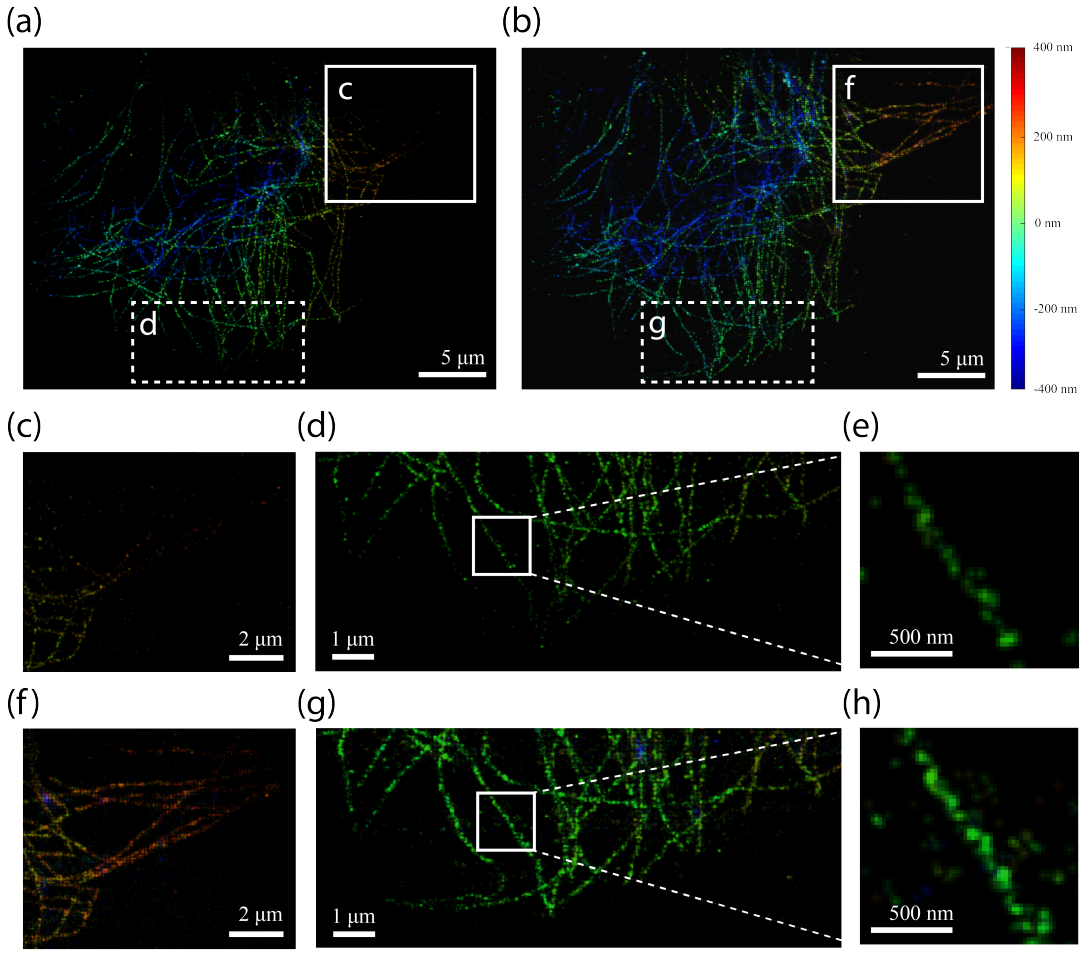

Figure 3-3: Comparison of 3D super-resolution image quality between Easy-DHPSF (a) and SOLAR-STORM (b). SOLAR-STORM is capable of reconstructing overlapping images of multiple emitters, returning more localizations than Easy-DHPSF in densely labeled regions. (c), (f) Enlarged view of the boxed region with solid lines in (a), (b), respectively. (d), (g) Enlarged view of the boxed region with dashed lines in (a), (b), respectively. (e), (h) Enlarged view of the boxed region with solid lines in (d), (g), respectively, showing a section of a microtubule which is further analyzed in Figure 3-4

Argon ion laser (Coherent) was used, while the 641 nm diode laser (Coherent Cube) was used for deactivation. The data set had 28582 frames of a 35.2 µm × 25.6 µm-sized region and was used in previous studies for performance evaluation [24, 6]. A tubulin antibody labeled with Alexa-647 and Alexa-488 dyes serving as a photo-switch was used to image microtubules of fixed PtK1 cells (Rat-kangaroo kidney epithelial cells). The regions indicated with white boxes in (a), (b) are enlarged for better visualization.

Both subregions are more visible for the image reconstructed using SOLAR-STORM not only due its capability of localizing fluorescent emitters in dense clusters, but also its PSF-independent localization scheme. Using the experimentally acquired PSF is favorable as it fully describes the optical imaging system without setting up a model. Furthermore, as SOLAR-STORM uses template matching for initial localization, there is no need to correct depth-dependent lateral shifts in model-based approaches [39]. Finally, the runtime per frame was 25.4 ms for SOLAR-STORM while Easy-DHPSF being 325.1 ms, demonstrating the capability of SOLAR-STORM to process images in real time.

In Figure 3-4, the localizations of a small section of a microtubule in Figure 3-3 (e), (h) are fit to a straight line using principal component analysis. For each localization, the lateral and axial direction components of the vector from the localization to its perpendicular foot on the line were computed. The histogram of these values are plotted for both softwares (c)-(e), indicating that the superior image quality is likely due to the higher recall rate of SOLAR-STORM. This was further confirmed by estimating the diameter of the microtubule by simulating emitters on a surface of a cylinder with different radii. The Kullback-Leibler divergence was measured between the histogram of the experimental data and the simulated data for each value of the radius. The estimate of the diameter was 74 nm for SOLAR-STORM whereas Easy-DHPSF gave 84 nm. The deviation from the true value which is roughly 25 nm might be attributed to the microtubule deviating from a straight line and the finite precision of the localizations. However, as SOLAR-STORM gives a better estimate, this suggests that it provides an improved image reconstruction quality at various conditions of optical imaging systems.

3.6 Benchmarking

Data analysis packages for localization microscopy tend to require a number of user inputs whose optimal values may depend on data sets. For an objective comparison with other 3D localization software, I run the software on publicly available data set

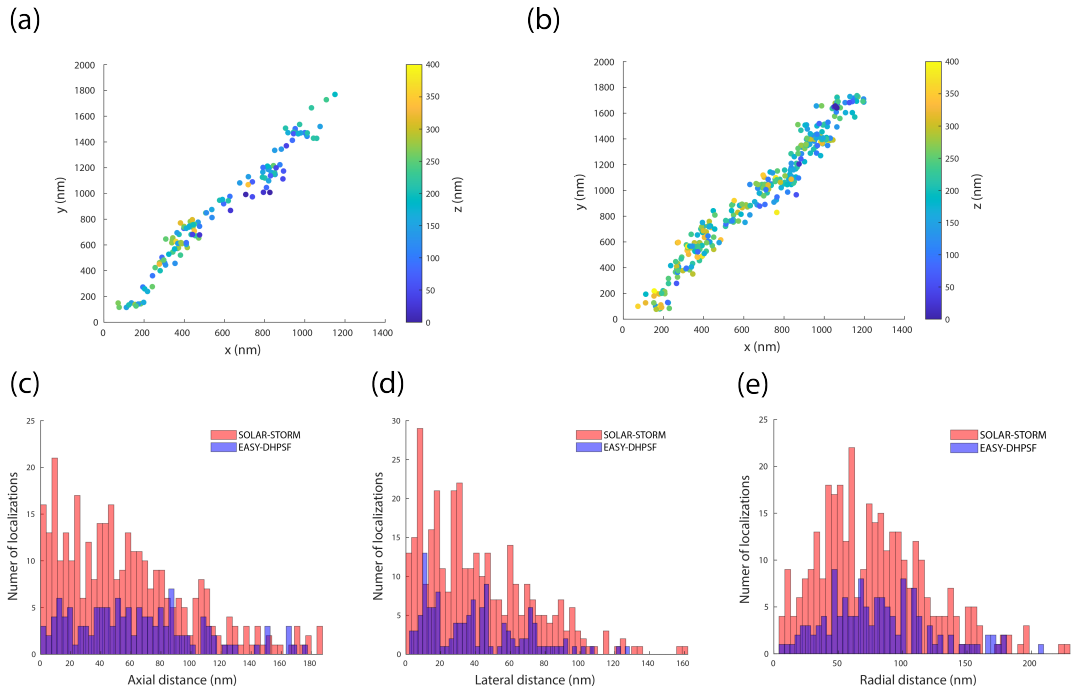

Figure 3-4: A detailed analysis of a microtuble section shown in Figure 3-3 (e), (h). (a), (b) A depth-color coded 2D scatter plot of the localizations obtained from Easy-DHPSF and SOLAR-STORM. The localizations are fit with a straight line from which the shortest vector is computed and decomposed into axial and lateral directions for each localization. The norm of these vectors are separately shown in (c) and (d) as the localization precision differs between lateral and axial directions.

provided by the 2016 SMLM challenge [70]. Participants were asked to run their software on simulated synthetic super-resolution imaging data sets modeling microtubule structures for 2 different imaging modes: long sequence and high density. For each imaging mode, a low SNR and high SNR data set was provided, resulting in 4 different data sets in total. The list of localized particles submitted by participants were evaluated using metrics including the Jaccard index, RMSE, and runtime. SOLAR-STORM ranked 4th position among 9 algorithms for the dataset with low density and 3rd among 8 algorithms for the dataset with high density. The algorithms were ranked using algorithm efficiency which is a metric that combines the RMSE and Jaccard index. Note that the results of SOLAR-STORM were based on an earlier version of the software which did not utilize the Taylor-expansion based

debiasing. Since the competition is now open permanently whose website is continuously updated with contributions from new packages, researchers may refer to the link `http://bigwww.epfl.ch/smlm/challenge2016/` where the results of the updated version of the software are available.

3.7 Conclusion

3D localization of fluorescent emitters in dense clusters is a challenging task due to various issues such as the complex shape of 3D PSFs, large amount of calculations and depth-dependent corrections required. SOLAR-STORM provides precise localization results from dense and noisy data sets in real time which was validated on simulated and experimental data sets. I emphasize that its high speed does not come from simply implementing it on a GPU. Rather, the algorithm runs efficiently due to its parallel structure, based on a parallel version of OMP combined with the PSF's Taylor expansion-based debiasing and model selection using BIC. Using SOLAR-STORM, multiple patches containing different numbers of fluorescent emitters can be processed concurrently without any thread divergence on GPUs. This is a major advantage over other software where the localization process needs to be further accelerated by scaling on multiple GPUs. Being a fast and precise algorithm with robust localization results, SOLAR-STORM can benefit research in the super-resolution microscopy community as well as general fluorescence microscopy studies.

Chapter 4

Inferring system properties from thermodynamic fluctuations: a data-driven approach

This chapter is adopted with minor modifications from a manuscript prepared for submission to Physical Review Letters with coauthors Junang Li and Nikta Fakhri.

4.1 Introduction

Biological systems inherently operate out of equilibrium orchestrated by intricate dynamic processes which can be described by stochastic thermodynamics [73]. At the mesoscopic scale, stochastic thermodynamics unravels the thermodynamic content of fluctuations without tracking every degree of freedom. For example, a recent development of a class of inequalities termed thermodynamic uncertainty relations have shown the potential for quantifying how far the system is from equilibrium through fluctuations without tracking all degrees of freedom [21, 44, 31]. Despite the advent of these tools, a characterization of biological systems from fluctuating quantities at the mesoscopic scale remains challenging as an understanding of the dynamics and a clear definition of system internal states is required.

Alternatively, a data-driven framework can be developed instead of attempting

to fully characterize the system from first principles when examples of thermodynamic fluctuations are accessible whose underlying system conditions are known. In these settings, a mapping from nonequilibrium fluctuations to system properties can be learned via supervised learning [68, 40, 41], a branch of machine learning which has shown success in a wide range of applications including computer vision [42, 37, 20, 78], speech recognition [81, 53, 28, 22], and natural language processing [8, 16, 75]. Thus, either a regressive function or a classifier can be learned by pairing experimentally accessible fluctuating quantities with system properties.

However, a variety of biological processes such as metabolism, cell division, and embryonic development occur concurrently on a wide range of time scales from nanoseconds to days which are reflected collectively in fluctuations at the mesoscopic scale. Therefore, inferring properties of the system with respect to a specific biological process requires identifying corresponding features buried in the complex and stochastic thermodynamic fluctuations. Furthermore, experimentally acquired mesoscopic fluctuations entail hallmarks of the experimental protocols, obscuring features emerging from processes whose properties are of interest. Finally, collecting vast amounts of data can be oftentimes infeasible. The time span of the measurements are limited by the short time scales of dynamic and transient biological processes. The amount of data that can be acquired is also subject to experimental constraints such as the limited photon budget of fluorophores when used as probes for sensing the system [47].

Here, I develop a data-driven framework that addresses these concerns using a wavelet transform-based operator named the scattering transform, combined with a type of recurrent neural networks termed long short-term memory (LSTM) networks [49, 11, 1, 50, 12]. Applying the scattering transform on fluctuations captures the features of a certain underlying process which linearizes deformations while the variability in time translation is removed for processes occurring at shorter time scales. I demonstrate its usage in simulations of two mechanically coupled beads at a temperature gradient and an experimental system consisting of an optically trapped silica microsphere embedded in a bacterial bath. Finally, trajectories of endogenous cor-

tical granules in starfish oocytes are analyzed which establishes the technique as a non-invasive tool for understanding biological processes within living systems in their native state.

4.2 Problem formulation

I describe a data-driven framework that extracts essential features from the mesoscopic fluctuations which capture hallmarks of the underlying biological process. The extracted features are then paired with labels to train a mapping via supervised learning which returns properties of the system. I develop this framework while recalling the challenges mentioned in the previous section. First, as biological processes occur concurrently at multiple time scales, the resulting mesoscopic fluctuations are non-stationary. However, although the system slowly evolves over time, the thermodynamic fluctuations can be assumed to be approximately stationary within a short time window due to the separation of time scales. Therefore, the extracted features must remain in a format of time series in order to reflect its non-stationary nature over long time scales while only the variability in time shift being removed at shorter time scales. Second, the function that maps measured time series to system properties must be robust upon different experimental protocols. For example, if the fluctuations are measured using probe particles that sense the environment of a biological system, the outcomes of the mapping should not depend on the shape or sizes of the particles. To illustrate how this constrains the feature extraction operator, I resort to a description of the mesoscopic fluctuations based on Langevin dynamics formulation. In the overdamped limit, the thermodynamic fluctuations of the system can be described with a Langevin equation:

$$\gamma \frac{dx}{dt} = F + \xi(t) \tag{4.1}$$

x is the dynamical variable of the system, e.g., the position of the probe particle, and γ is the friction constant. $\xi(t)$ is the thermal noise at time t which satisfies the

following relations assuming a Gaussian and white correlated stochastic process.

$$< \xi(t) >= 0, < \xi(t)\xi(t') >= 2\gamma k_B T \delta(t - t') \tag{4.2}$$

Finally, F is the sum of all remaining forces acting on the probe particle. Assume F only being a function of x which does not depend on the friction coefficient γ or time t and define $x_1(t)$ as a stochastic process which satisfies Eq. (4.1) with $\gamma = \gamma_1$. A time-dilated stochastic process $x_2(t) = x_1((1 - \epsilon)t)$ is then a solution to Eq. (4.1) with $\gamma = 1/(1 - \epsilon)\gamma_1$. This implies that in the case where F is only a function of x, a time dilation of measured mesoscopic fluctuations is statistically identical to that acquired using a probe with a different size. Based on this observation, I require the feature extraction operator to be Lipschitz continuous upon time dilation so that the proposed framework remains robust upon probes with different sizes. For biological systems where probes also experience forces that do not solely depend on the position x, these additional terms are treated as small perturbations. By learning a nonlinear mapping that serves either as a regressive function or a classifier, I demonstrate that the system properties can be inferred from stochastic fluctuations.

4.3 Invariant scattering for feature extraction

Effective feature extraction from fluctuations is at the core of the proposed framework to infer system properties in a robust manner. In order to remove the variability in time shift at short time scales while being Lipschitz continuous upon deformations, I introduce the scattering transform which is a cascade of wavelet filter banks and modulus nonlinearities. Combined with linear classifiers, the scattering transform is an effective method for audio and texture classification particularly when training with limited amounts of data. Define x as the time series of the mesoscopic fluctuations, and ψ_j as the filters which are constructed by dilating a single filter ψ with octave bandwidth being $1/Q$. When setting $a = 2^{1/Q}$ and a^J as the maximum wavelet scale, the set of filters are then expressed as $\psi_j(t) = a^{-j}\psi(a^{-j}t)$ for $j \leq J$. Using these set of

filters, the goal is to build translation invariant coefficients from a wavelet transform $x * \psi_j$. This can be done by taking the integral $\int x * \psi_j(u)du$, which, however, results in a trivial invariant 0. The usage of a linear operator M that commutes with translations for computing $\int M(x * \psi_j)(u)du$ also returns 0, suggesting that nonlinearity must be introduced to build translation invariant coefficients. The modulus operator is used as it commutes with diffeomorphism and is stable upon additive noise. This results in an invariant descriptor $\int |x * \psi_j(u)|du$. This removes all the nonzero frequency content of the wavelet transform, which is undesirable as the high frequency content of the mesoscopic fluctuations will be lost. In order to partially recovered this information, additional coefficients are computed by applying another wavelet transform prior to time averaging: $\int |x * \psi_{j_1}(u)| * \psi_{j_2}(u)du$. This iterative process allows one to build invariant coefficients where the variability in time translation is removed while preserving information at different time scales.

Finally, in order to keep the variability at time scales larger than 2^J, the coefficients invariant to local translations smaller than 2^J are computed. The integral is replaced with a convolution with the scaling function $\phi(t)$, which defines the scattering transform for a given sequence $p = (j_1, j_2, \ldots, j_m)$:

$$S[p]x(t) = |||x * \psi_{j_1}| * \psi_{j_2}| \cdots | * \psi_{j_m}| * \phi_J \tag{4.3}$$

The result is a scattering network where the above transform is computed for each node. Due to averaging with ϕ whose window size is 2^J, the number of coefficients can be reduced by subsampling the time series with a period proportional to this scale. Furthermore, the scattering transform is computed only for frequency-decreasing paths as $|x * \psi_j| * \psi_{j'}$ is nonnegligible only when $j < j'$. This significantly reduces computations when later training models using scattering transform coefficients.

4.4 Training recurrent neural networks

The scattering transform outputs a time series with multiple components which can be paired with quantities that represent the system properties in order to train a recurrent neural network (RNN). An RNN maintains an internal state vector to process sequential inputs with variable length and has shown success in fields such as natural language processing and speech recognition. A basic RNN architecture consists of a cell which connects to itself and processes sequential data one at a time. However, this structure suffers from the vanishing gradient problem and thus precludes learning long-term dependencies which is problematic as signatures of a biological process may appear across a range of time scales [7]. This can be alleviated by usage of long short-term memory (LSTM) networks [29]. In LSTMs, the flow of information into and out from the LSTM unit cell is controlled by a set of gates which effectively prevents the vanishing gradient problem. The software was written in Python and LSTMs were implemented with the Pytorch library [63]. The scattering transform for a typical size of the acquired datasets (2048 timepoints, 2000 trajectories with 2 components) took on average 5.8 seconds when processed on a single core of an Intel(R) Core(TM) i7-8565U CPU @ 1.80GHz which was negligible compared to the time training LSTMs. LSTM training is accelerated by running on 4 NVIDIA Tesla V100 GPUs using half-precision through implementation with the Python Horovod library functions [74]. The Adam optimizer [35] is used with batch size of 64 or 128. For training a regressive function for simulated data, the output of the network is a single number that represents the system properties. For classification with experimental data, the result is a layer of multiple nodes with a softmax function corresponding to the probability of being assigned to the corresponding class. A wide range of network sizes are first considered with the number of layers being 2 or 3, hidden state sizes ranging from 20 to 300. For the given simulation and experimental data discussed in later sections, this is narrowed down to the number of layers being 3 and hidden state sizes ranging from 50 to 250 with a step size 50. The trained models are then cross-validation on the validation dataset for model selection.

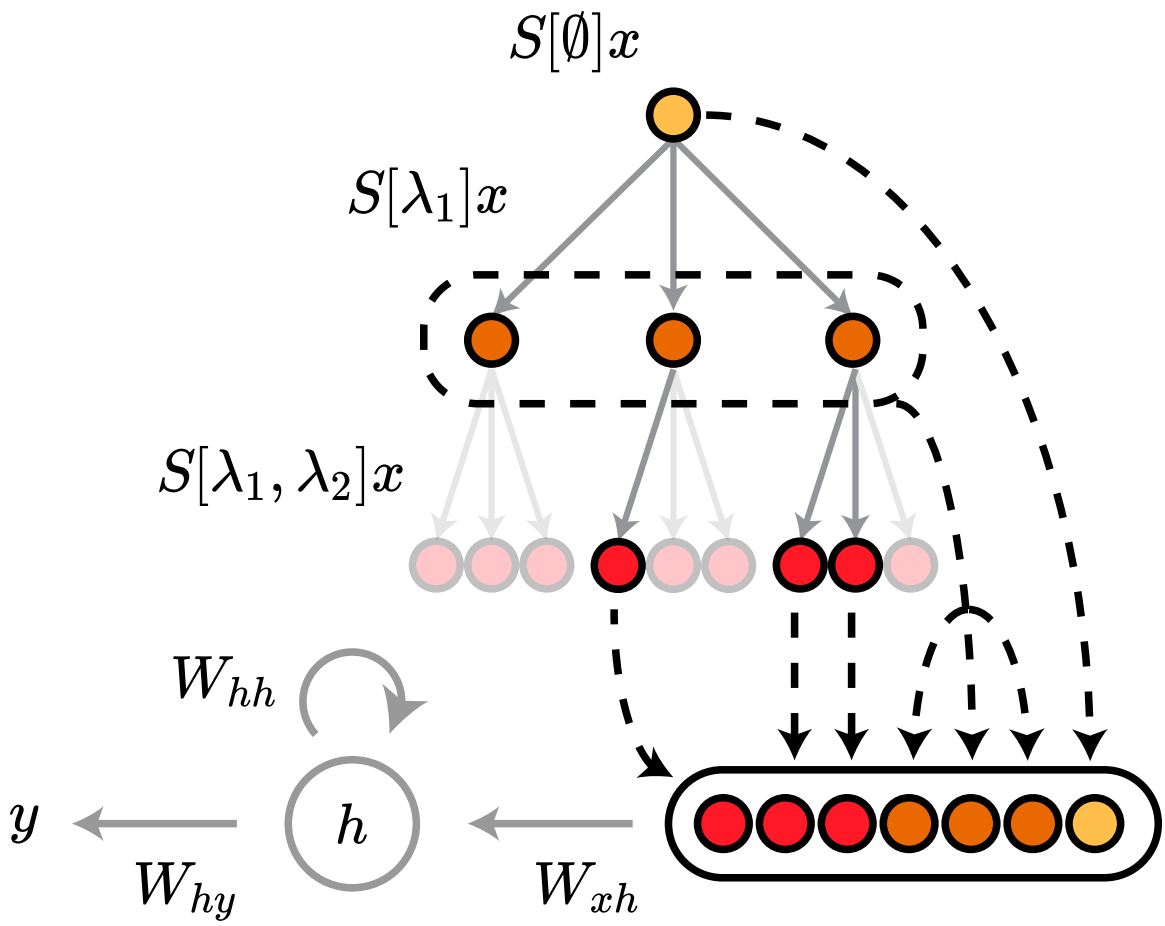

Figure 4-1: Schematic of the framework where the scattering transform is first applied on a time series dataset using a cascade of wavelet transforms and modulus rectifiers along frequency-decreasing paths. The resulting multivariate time series is paired with quantities representing the system properties to train the recurrent neural network.

4.5 Mechanically coupled beads

4.5.1 Simulations

I test the proposed method on trajectories of simulated overdamped, tethered beads. As illustrated in Figure 4-2 (a), two beads are mechanically coupled through a spring with stiffness k_2 while the springs attached to the walls have stiffness k_1. This system is driven out of equilibrium yet exhibits a steady state by having the beads in contact with two reservoirs with different temperatures T_h and T_c.

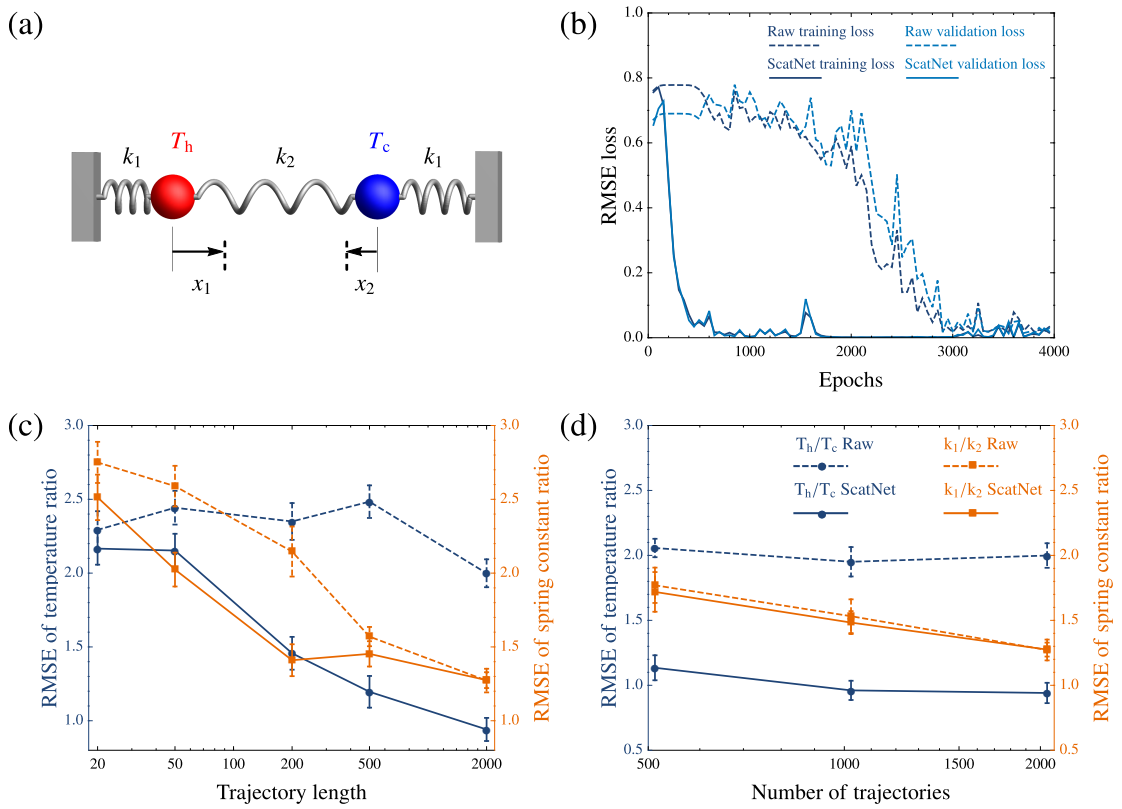

Figure 4-2: Prediction results on simulations of two beads. (a) Schematic of the simulation consisting of two mechanically coupled beads in contact with reservoirs at different temperatures T_h and T_c. (b) Training and validation loss plotted for training LSTM networks. (c) RMSE plotted as a function of trajectory length for the given simulated data. (d) RMSE plotted as a function of the number of trajectories in the training data.

4.5.2 Training with simulated data

Trajectories of the two beads following Langevin dynamics in Eq. (4.1) are simulated by randomly sampling a spring constant ratio k_1/k_2 from a uniform distribution with interval $[1, 7)$ and a temperature ratio T_h/T_c from a uniform distribution with interval $[4, 10)$. Finally, in order to confirm the predictions being robust upon usage of different probe particles, a drag coefficient γ is randomly sampled from a uniform distribution on an interval $[1, 3)$ for each trajectory. Trajectories are generated for 512, 1024, and 2048 timepoints which are processed with the scattering transform resulting in 16 timepoints and 22, 29, and 37 components for each bead, respectively. Therefore the output is a time series with length 16 and 44, 58, and 74 components

for the 3 cases. These multivariate time series are fed into a 3-layer bidirectional LSTM [72, 23] to train a regressive function for predicting either T_h/T_c or k_1/k_2. For a given trajectory length for the simulation, I experiment with a different number of trajectories for training LSTMs in order to measure the performance in data-scarce regimes. The proposed method is benchmarked against an LSTM with the same architecture whose inputs are the raw time series of the simulated two beads. I apportion the data into a training dataset and a validation dataset with a 80-20 split. Although the time required for the loss to reach a plateau varies among different network sizes, training completes within 12 minutes for all models that use the scattering transform. For LSTM models trained with the raw time series, the training takes up to 7 hours. On average, training LSTMs for the proposed method is faster with 1 to 2 orders of magnitude. The performance of the selected model is computed on a new simulated test dataset consisting of 1000 trajectories. Figure 4-2 displays the root-mean-square error (RMSE) results of the temperature ratios and the spring constant ratios. For the given range of trajectory lengths and number of trajectories, the LSTM trained with scattering transform coefficients outperforms the LSTM trained with the raw time series in terms of both the spring constant ratios and temperature ratios. As trajectories are generated with different drag coefficients, these results can be attributed to the scattering transform's stability upon deformations and local invariance under time translation. The LSTM trained with scattering coefficients being the inputs shows a decrease in the RMSE as the simulated trajectory increases. In contrast, the LSTM trained with raw inputs does not show a decrease which suggests that it fails to find a proper mapping from the fluctuations to the system properties. In Figure 4-2(d), the RMSE is plotted as a function of the number of trajectories in the training dataset. The LSTM trained with scattering coefficient inputs shows a decrease in the RMSE when more data is available. This is not the case for predicting the temperature ratios with LSTM trained with raw inputs which again suggests that the LSTM directly trained with raw inputs does not properly learn to extract signatures of system properties from the stochastic fluctuations.

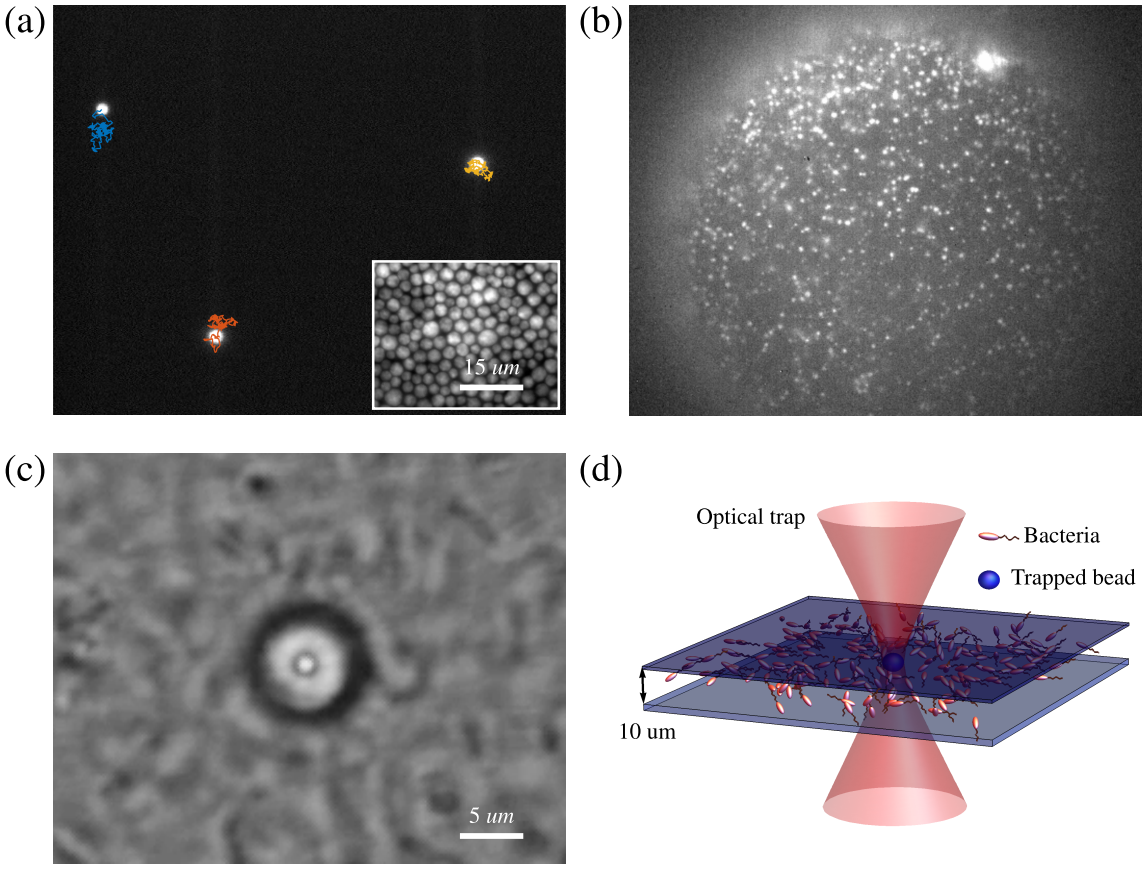

Figure 4-3: Description of experimental setting. (a) Fluorescence imaging of silica microspheres in a bacterial bath. The inset indicates the beads coming in different shapes and sizes. (b) Infrared fluorescence imaging of cortical granules within a starfish ooctye. (c) Enlarged view of the confined microsphere using brightfield imaging. (d) A microsphere within a bacterial bath confined in a optical trap.

4.6 Experimental data

4.6.1 Fluorescent cortical granules in a starfish oocyte

In order to validate the proposed method on biological systems, I collect trajectories of endogeneous fluorescent granules embedded in the actin cortex of a starfish oocyte at different levels of activity as shown in Figure 4-3(b). The actin cortex consists of actin filaments and myosin motors which generates forces powered by ATP. I therefore track an additional group of cortical granules from an oocyte with a lower level of activity through ATP depletion treatments. In the same manner, another set of

granule trajectories are collected from an oocyte with a higher level of activity by over-expressing myosin motors. The cortical granules are imaged on a custom-built microscope whose centroid positions are determined using a single molecule tracking algorithm. Trajectories with variable length are acquired for low, native, and high levels of activity, resulting in 741, 606, and 615 trajectories, respectively. 20 % of this dataset is used as a test dataset, while the remaining amount is further divided into a training set and a validation set with a 80-20 split.

Cortical granule trajectories are scattering-transformed whose resulting multivariate time series are fed into a 3-layer bidirectional LSTM to train a classifier to distinguish different oocyte conditions from the position fluctuations. A baseline is set by training an LSTM with an identical architecture whose inputs are the trajectories without any preprocessing. The training completes within 5 minutes for models based on the scattering transform while training with the raw time series finishes in an hour, making the proposed method roughly an order of magnitude faster.

The ratio of the correct classifications over the total number of predictions is 54.96 % for the baseline and 78.88 % for the proposed method. The higher accuracy of the proposed method is attributed to two reasons. First, the distribution of the cortical granule sizes has a broad range which typically range between 1-5 um in diameter where the resulting scattering coefficients remain stable. Second, the force fluctuations arising from the ATP-consuming molecular motors are exerted on probe particles through the highly cross-linked actomyosin network in a manner irrespective of the probe particle properties. This effectively allows learning a decision surface when the fluctuations are preprocessed with the scattering transform.

4.6.2 Silica microspheres in bacterial bath

The proposed method was initially inspired by a model where nonequilibrium forces acting on the probe can be approximated as functions that solely depend on the particle position x, while additional terms being treated as small perturbations. In order to understand to what extent the proposed method can be generalized, a system consisting of silica microspheres in a bacterial bath is studied where the initial

assumption no longer holds.

A strain of light-activated pseudomonas aurantiaca is grown overnight in 5 ml of LB broth having ampicillin resistance at 30 °C. After 12 hours, a 1000-fold dilution is performed in 5 ml Tryptone broth and incubated for 4 hours at 30 °C. 5 ul of 1 M arabidose and 5 ul of all-trans-retinal is added and incubated at 30 °C for another 7 hours.

The bacterial suspension is aliquoted into 2 microcentrifuge tubes each having 800 ul which are centrifuged for 3 minutes at 3000 g at room temperature. For an experiment that involves usage of the optical tweezer, 600, 750 ul of the supernatant is removed and vortexed in order to create 4x, 16x concentrated bacteria suspension samples, respectively. These samples are later diluted to 2x, 8x, respectively, when mixed with microspheres suspension at a 1:1 volume ratio. For an additional experiment which does not use the optical tweezer, 750 ul is removed for only one aliquot to create a final concentration of 0.5x, 8x.

For the experiment which does not involve usage of the optical trap, a suspension of fluorescent microspheres is used where the particles come in a range of shapes and sizes. This is shown in the inset of 4-3(a) while the average diameter of the microspheres is 3 um. For the experiment with the optical trap, microspheres with diameters 5, 6, and 7 um (Microparticles GmbH) are mixed with equal amounts to mimic a polydisperse suspension where this size range is chosen considering the stiffness of the optical trap. For each aliquot of the bacterial suspension, 20 ul of the bacterial suspension sample is mixed with 20 ul of the suspension of microspheres. The concentration of the microspheres suspension is determined by requiring roughly 5 microspheres appearing in the field of view in the final mixed sample. 1.5 ul of 10 um-diameter microspheres are then added to the mixed sample containing microspheres in bacterial suspension. The concentration of the 10 um-diameter microspheres suspension is determined by requiring roughly 1 or 2 of these silica beads appearing in the field of view when imaging the final sample. 5 ul of this suspension is sandwiched between a passivated slide and coverslip. A 500 g weight is applied to press down the coverslip for 1 minute to bring the sample chamber thickness close to

a constant 10 um in order to make the initial bacterial concentration represent the effective bacterial density in the sample during imaging.

For samples prepared using fluorescent silica beads of size 3 um, beads are imaged excited with a 561 nm laser (Coherent) using a custom-built microscope whose fluorescence is collected using an EMCCD (Andor iXon 897). Images are recorded for the 4 different system conditions corresponding to the 2 different concentrations of the bacterial suspension and 2 different levels of the bacterial activity controlled by tuning the excitation laser intensity. Using a fluorescence microscopy setup is beneficial as only microspheres appear in the field of view which enables tracking multiple particles simultaneously in a reliable manner.

An optical tweezer kit (Thorlabs) is used to image and trap silica beads. For each concentration, two different intensities are used to modulate the speed of the light-activated bacteria, resulting in 4 different classes of conditions. Multiple microspheres are optically trapped and imaged one at a time at a frame rate of 50 Hz for each of these conditions.

Figure 4-3(a) displays the fluorescent beads in the bacterial suspension while (c) and (d) show an example of a bead in the optical trap and the setup schematic. In both experiments, the trajectories of the microspheres are determined using a particle tracking algorithm designed for radially symmetric objects. To generate enough samples for training and testing, the acquired trajectories are divided into multiple shorter trajectories with 256 timepoints whose overall duration is roughly 5 seconds. For the experiment without the optical trap, this resulted in on average

Table 4.1: Classification results of models trained on trajectories of microspheres in bacterial bath

	Accuracy (%)	Training time (seconds)
without optical trap		
raw	67.80	339.4
scat	**79.66**	**8.6**
with optical trap		
raw	56.90	3312.1
scat	**78.74**	**483.0**

60 trajectories for each class, which is divided into a training/validation dataset and a test dataset with a 75-25 split. For the experiment involving the optical trap, the entire dataset consists of 174 trajectories per class, which is equally divided into a training/validation dataset and a test dataset. The training/validation dataset is further divided into a training set and a validation set with a 80-20 split. The training time ranges from 10 seconds to 10 minutes for models based on the scattering transform while training with the raw time series ranges from 5 minutes to an hour. On average, training with the proposed method is an order of magnitude faster.

The classification results for the 4 classes (2 bacterial activity levels and 2 concentrations) are displayed in Table 4.1. For both experiments, the proposed method combining the scattering transform with LSTMs outperforms the LSTMs trained with the positional fluctuations directly even though the assumptions on the active forces introduced earlier do not hold. The capability of removing variability in time shift and stability upon deformations is crucial when the system is probed with particles that come with a range of shapes or sizes. A noticeable bacterial speed decay was observed over a few hours and thus the total sum of the duration of all trajectories is limited to a few tens of minutes. Therefore, in order to understand to what extent the accuracy of both the proposed and baseline methods further improve, a different experimental setting would be required to acquire more data. For example, microspheres are trapped and imaged one at a time with the optical trap and therefore a setup with multiple traps could help understand this question. In the experiment with the microspheres without optical trapping, multiple beads in the field of view are tracked simultaneously, although the duration is limited to a few minutes since longer trajectories have higher chances of the microspheres interacting with each other which would correspond to anomalous events in the time series. However, in the starfish oocyte dataset, fluorescent granules are embedded in the actin cortex which allows us to safely disregard the possibility of interactions between neighboring granules. Therefore, a sufficient amount of trajectories can be collected easily in parallel for a wide range of biological processes where probe particles sense their environment while position fluctuations are spatially limited to a finite region.

4.7 Conclusion

A data-driven framework is proposed for inferring system properties from stochastic fluctuations which combines a scattering transform with LSTM networks. The scattering transform removes local variability in time shift and is stable upon deformations while preserving any nonstationary behavior in the biological system, allowing robust inference upon usage of probe particles with different properties. The proposed method is benchmarked against LSTMs trained directly with the raw format of the time series to show preprocessing with the scattering transform predicts system properties in much more reliable manner for both simulations and experimental data. With its additional advantage of quick convergence within tens of minutes for models trained on GPUs, the proposed method can be effectively used in a wide scope of biological processes occurring across multiple time scales for predicting system properties in a data-driven manner.

Chapter 5

Conclusion and outlook

The framework developed in this thesis has focused on learning properties of living systems in their native state using experimental, data analysis tools in a non-invasive manner. Observing and analyzing fluctuations in living systems while minimizing external perturbations is crucial for understanding the principles governing biological processes. However, not only noninvasive studies but also manipulation of systems out of equilibrium can be useful for certain applications from a practical point of view such as designing microscopic heat engines.

Heat engines convert thermal energy to mechanical work where maximizing its efficiency has been a long standing topic for engineers. This can be enabled through a combination of stochastic thermodynamics, experimental tools for manipulating microscopic objects, and certain data-driven methods. For example, stochastic thermodynamics allows defining thermodynamic quantities such as work or dissipation at a fluctuating level for describing the efficiency of a microscopic heat engine. Experimental tools for microscopic object manipulation such as optical tweezers combined with living systems such as a bacterial bath can be used to set up a microscopic heat engine [36]. Finally, certain data-driven methods such as reinforcement learning can provide instructions on the external control knobs in order to let the system evolve in a specific manner. As a Langevin dynamics formulation and an optical tweezer setup have been briefly introduced in the previous chapter, I provide a discussion on reinforcement learning and its potential application for maximizing the efficiency of

a microscopic heat engine.

Designing a heat engine with maximal efficiency can be thought of as finding the optimal trajectory in the space of the system's control parameters. The problem setting resembles that of a field of machine learning called reinforcement learning [76]. Reinforcement learning is an area of machine learning that aims to understand how to map situations to actions. It is formulated such that a learning agent interacts with the environment and learns what actions to take in order to maximize a numerical reward signal. The agent learns through experience whose one aspect is the exploration-exploitation dilemma; in order to find the best state-action mapping which is termed as a policy, the agent must explore to make better action selections in the future, but at the same time exploit the experience the agent has gained so far. This trade-off between exploration and exploitation is an important feature of reinforcement learning which allows the agent to learn actions that maximize the long-term return rather than exploiting immediate rewards.

By defining the state as a set of numbers that represent the external control parameters of the system, a heat engine can be characterized as a trajectory in state space. For the colloidal particle immersed in a bacterial bath and trapped with an optical tweezer, the state can be described with two quantities: the light intensity which controls the bacterial activity, and the stiffness of the optical trap while the actions correspond to changes in these quantities. By defining the efficiency as the numerical reward, the given experimental setting will learn a trajectory in the state space of the control knobs to maximize the efficiency of the microscopic heat engine.

In this thesis, I have demonstrated a set of tools ranging from data collection to data analysis and inference of system properties based on thermodynamic fluctuations. The toolbox consists of a microscopy platform with 3D imaging in infrared for noninvasive imaging of living systems, an image reconstruction algorithm for 3D localization, and a supervised learning based inference framework for predicting system properties from mesoscopic fluctuations. In addition, a reinforcement learning-based framework for designing a microscopic heat engine was discussed as an extension of this toolbox to not only observe but also manipulate living systems in a desired

manner.

Biological systems are inherently out of equilibrium and therefore pose many challenges from data acquisition to preprocessing and analysis. This set of novel tools will help tackle challenging problems in non-equilibrium statistical mechanics to better understand living systems from a physics point of view.

Bibliography

[1] Joakim Andén and Stéphane Mallat. Multiscale scattering for audio classification. In *ISMIR*, page 657, 2011.

[2] Hazen Babcock, Yaron M. Sigal, and Xiaowei Zhuang. A high-density 3d localization algorithm for stochastic optical reconstruction microscopy. *Optical Nanoscopy*, 1(1), Jul 2012.

[3] Hazen P. Babcock, Jeffrey R. Moffitt, Yunlong Cao, and Xiaowei Zhuang. Fast compressed sensing analysis for super-resolution imaging using l1-homotopy. *Opt. Express*, 21(23):28583–28596, Nov 2013.

[4] Hazen P. Babcock and Xiaowei Zhuang. Analyzing single molecule localization microscopy data using cubic splines. *Scientific Reports*, 7(1), Apr 2017.

[5] Sergei M. Bachilo, Michael S. Strano, Carter Kittrell, Robert H. Hauge, Richard E. Smalley, and R. Bruce Weisman. Structure-assigned optical spectra of single-walled carbon nanotubes. *Science*, 298(5602):2361–2366, 2002.

[6] Anthony Barsic, Ginni Grover, and Rafael Piestun. Three-dimensional super-resolution and localization of dense clusters of single molecules. *Scientific Reports*, 4(1), Jun 2014.

[7] Y. Bengio, P. Simard, and P. Frasconi. Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5:157, 1994.

[8] Yoshua Bengio, Réjean Ducharme, and Pascal Vincent. A neural probabilistic language model. In *Advances in Neural Information Processing Systems 13*, page 932. 2001.

[9] Eric Betzig, George H. Patterson, Rachid Sougrat, O. Wolf Lindwasser, Scott Olenych, Juan S. Bonifacino, Michael W. Davidson, Jennifer Lippincott-Schwartz, and Harald F. Hess. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642–1645, Sep 2006.

[10] Nicholas Boyd, Eric Jonas, Hazen Babcock, and Benjamin Recht. Deeploco: Fast 3d localization microscopy using neural networks. *bioRxiv*, Feb 2018.

[11] Joan Bruna and Stéphane Mallat. Classification with scattering operators. In *CVPR 2011*, page 1561, 2011.

[12] Joan Bruna and Stéphane Mallat. Invariant scattering convolution networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35:1872, 2013.

[13] Herbert B. Callen and Theodore A. Welton. Irreversibility and generalized noise. *Phys. Rev.*, 83:34–40, 1951.

[14] Jerry Chao, E. Sally Ward, and Raimund J. Ober. A software framework for the analysis of complex microscopy image data. *IEEE Trans. Inf. Technol. Biomed.*, 14(4):1075–1087, Jul 2010.

[15] Scott Shaobing Chen, David L. Donoho, and Michael A. Saunders. Atomic decomposition by basis pursuit. *SIAM Rev.*, 43(1):129–159, Jan 2001.

[16] Ronan Collobert, Jason Weston, Léon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. Natural language processing (almost) from scratch. *J. Mach. Learn. Res.*, 12:2493, 2011.

[17] A. Einstein. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der Physik*, 322(8):549–560, 1905.

[18] Nikta Fakhri, Frederick C. MacKintosh, Brahim Lounis, Laurent Cognet, and Matteo Pasquali. Brownian motion of stiff filaments in a crowded environment. *Science*, 330(6012):1804–1807, 2010.

[19] Nikta Fakhri, Alok D. Wessel, Charlotte Willms, Matteo Pasquali, Dieter R. Klopfenstein, Frederick C. MacKintosh, and Christoph F. Schmidt. High-resolution mapping of intracellular fluctuations using carbon nanotubes. *Science*, 344(6187):1031–1035, 2014.

[20] C. Farabet, C. Couprie, L. Najman, and Y. LeCun. Learning hierarchical features for scene labeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35:1915.

[21] Todd R. Gingrich, Jordan M. Horowitz, Nikolay Perunov, and Jeremy L. England. Dissipation bounds all steady-state current fluctuations. *Phys. Rev. Lett.*, 116:120601, 2016.

[22] A. Graves, A. Mohamed, and G. Hinton. Speech recognition with deep recurrent neural networks. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, page 6645, 2013.

[23] Alex Graves and Jürgen Schmidhuber. Framewise phoneme classification with bidirectional lstm and other neural network architectures. *Neural Networks*, 18:602, 2005.

[24] Ginni Grover, Keith DeLuca, Sean Quirin, Jennifer DeLuca, and Rafael Piestun. Super-resolution photon-efficient imaging by nanometric double-helix point spread function localization of emitters (spindle). *Opt. Express*, 20(24):26681–26695, Nov 2012.

[25] Lusheng Gu, Yi Sheng, Yan Chen, Hao Chang, Yongdeng Zhang, Pingping Lv, Wei Ji, and Tao Xu. High-density 3d single molecular analysis based on compressed sensing. *Biophysical Journal*, 106(11):2443–2449, Jun 2014.

[26] Ricardo Henriques, Mickael Lelek, Eugenio F Fornasiero, Flavia Valtorta, Christophe Zimmer, and Musa M Mhlanga. Quickpalm: 3d real-time photoactivation nanoscopy image processing in imagej. *Nature Methods*, 7(5):339–340, May 2010.

[27] Samuel T. Hess, Thanu P.K. Girirajan, and Michael D. Mason. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical Journal*, 91(11):4258–4272, Dec 2006.

[28] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, and B. Kingsbury. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29:82, 2012.

[29] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9:1735, 1997.

[30] Seamus J Holden, Stephan Uphoff, and Achillefs N Kapanidis. Daostorm: an algorithm for high-density super-resolution microscopy. *Nature Methods*, 8(4):279–280, Apr 2011.

[31] Jordan M. Horowitz and Todd R. Gingrich. Thermodynamic uncertainty relations constrain non-equilibrium fluctuations. *Nature Physics*, 16:15, 2020.

[32] Bo Huang, Wenqin Wang, Mark Bates, and Xiaowei Zhuang. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science*, 319(5864):810–813, 2008.

[33] Fang Huang, Samantha L. Schwartz, Jason M. Byars, and Keith A. Lidke. Simultaneous multiple-emitter fitting for single molecule super-resolution imaging. *Biomed. Opt. Express*, 2(5):1377–1393, May 2011.

[34] Manuel F Juette, Travis J Gould, Mark D Lessard, Michael J Mlodzianoski, Bhupendra S Nagpure, Brian T Bennett, Samuel T Hess, and Joerg Bewersdorf. Three-dimensional sub–100 nm resolution fluorescence microscopy of thick samples. *Nature Methods*, 5(6):527–529, 2008.

[35] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2014.

[36] Sudeesh Krishnamurthy, Subho Ghosh, Dipankar Chatterji, Rajesh Ganapathy, and A. K. Sood. A micrometre-sized heat engine operating between bacterial reservoirs. *Nature Physics*, 12:1134–1138, 2016.

[37] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems 25*, page 1097. 2012.

[38] Ryogo Kubo. Statistical-mechanical theory of irreversible processes. i. general theory and simple applications to magnetic and conduction problems. *Journal of the Physical Society of Japan*, 12(6):570–586, 1957.

[39] Carlini L, Holden SJ, Douglass KM, and Manley S. Correction of a depth-dependent lateral distortion in 3d super-resolution imaging. *PLoS ONE*, 10(11), Nov 2015.

[40] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86:2278, 1998.

[41] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521:436, 2015.

[42] Yann LeCun, Bernhard E. Boser, John S. Denker, Donnie Henderson, R. E. Howard, Wayne E. Hubbard, and Lawrence D. Jackel. Handwritten digit recognition with a back-propagation network. In D. S. Touretzky, editor, *Advances in Neural Information Processing Systems 2*, page 396. Morgan-Kaufmann, 1990.

[43] M. D. Lew, von Diezmann, A. R. S., and W. E. Moerner. Easy-dhpsf open-source software for three-dimensional localization of single molecules with precision beyond the optical diffraction limit. *Protoc. Exch.*, Feb 2013.

[44] Junang Li, Jordan M. Horowitz, Todd R. Gingrich, and Nikta Fakhri. Quantifying dissipation using fluctuating currents. *Nature Communications*, 10:1666, 2019.

[45] Luchang Li, Bo Xin, Weibing Kuang, Zhiwei Zhou, and Zhen-Li Huang. Divide and conquer: real-time maximum likelihood fitting of multiple emitters for super-resolution localization microscopy. *Opt. Express*, 27(15):21029–21049, Jul 2019.

[46] Yiming Li, Markus Mund, Philipp Hoess, Joran Deschamps, Ulf Matti, Bianca Nijmeijer, Vilma Jimenez Sabinina, Jan Ellenberg, Ingmar Schoen, and Jonas Ries. Real-time 3d single-molecule localization using experimental point spread functions. *Nature Methods*, 15(5):367–369, May 2018.

[47] Zhe Liu, Luke D. Lavis, and Eric Betzig. Imaging live-cell dynamics and structure at the single-molecule level. *Molecular Cell*, 58:644, 2015.

[48] Georgyi V. Los, Lance P. Encell, Mark G. McDougall, Danette D. Hartzell, Natasha Karassina, Chad Zimprich, Monika G. Wood, Randy Learish, Rachel Friedman Ohana, Marjeta Urh, Dan Simpson, Jacqui Mendez, Kris Zimmerman, Paul Otto, Gediminas Vidugiris, Ji Zhu, Aldis Darzins, Dieter H. Klaubert, Robert F. Bulleit, and Keith V. Wood. Halotag: A novel protein

labeling technology for cell imaging and protein analysis. *ACS Chemical Biology*, 3(6):373–382, 2008.

[49] Stéphane Mallat. *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. Academic Press, Inc., USA, 3rd edition, 2008.

[50] Stéphane Mallat. Group invariant scattering. *Communications on Pure and Applied Mathematics*, 65:1331, 2012.

[51] K. J. A. Martens, A. N. Bader, S. Baas, B. Rieger, and J. Hohlbein. Phasor based single-molecule localization microscopy in 3d (psmlm-3d): An algorithm for mhz localization rates using standard cpus. *J. Chem. Phys.*, 148(12):123311, 2018.

[52] Alexa L. Mattheyses, Sanford M. Simon, and Joshua Z. Rappoport. Imaging with total internal reflection fluorescence microscopy for the cell biologist. *Journal of Cell Science*, 123(21):3621–3628, 2010.

[53] T. Mikolov, A. Deoras, D. Povey, L. Burget, and J. Černocký. Strategies for training large scale neural network language models. In *2011 IEEE Workshop on Automatic Speech Recognition Understanding*, page 196, 2011.

[54] Junhong Min, Seamus J. Holden, Lina Carlini, Michael Unser, Suliana Manley, and Jong Chul Ye. 3d high-density localization microscopy using hybrid astigmatic/ biplane imaging and sparse image reconstruction. *Biomed. Opt. Express*, 5(11):3935–3948, Nov 2014.

[55] Junhong Min, Kyong Hwan Jin, Michael Unser, and Jong Chul Ye. Grid-free localization algorithm using low-rank hankel matrix for super-resolution microscopy. *IEEE Transactions on Image Processing*, 27(10):4771–4786, Oct 2018.

[56] Junhong Min, Cédric Vonesch, Hagai Kirshner, Lina Carlini, Nicolas Olivier, Seamus Holden, Suliana Manley, Jong Chul Ye, and Michael Unser. Falcon: fast and unbiased reconstruction of high-density super-resolution microscopy data. *Scientific Reports*, 4(1), Apr 2014.

[57] Eran A. Mukamel, Hazen Babcock, and Xiaowei Zhuang. Statistical deconvolution for superresolution fluorescence microscopy. *Biophysical Journal*, x102x(10):2391–2400, May 2012.

[58] Elias Nehme, Lucien E. Weiss, Tomer Michaeli, and Yoav Shechtman. Deepstorm: super-resolution single-molecule microscopy by deep learning. *Optica*, 5(4):458–464, Apr 2018.

[59] Michael J. O'Connell, Sergei M. Bachilo, Chad B. Huffman, Valerie C. Moore, Michael S. Strano, Erik H. Haroz, Kristy L. Rialon, Peter J. Boul, William H. Noon, Carter Kittrell, Jianpeng Ma, Robert H. Hauge, R. Bruce Weisman, and Richard E. Smalley. Band gap fluorescence from individual single-walled carbon nanotubes. *Science*, 297(5581):593–596, 2002.

[60] Wei Ouyang, Andrey Aristov, Mickaël Lelek, Xian Hao, and Christophe Zimmer. Deep learning massively accelerates super-resolution localization microscopy. *Nature Biotechnology*, 36(5):460–468, May 2018.

[61] Martin Ovesný, Pavel Křížek, Zdeněk Švindrych, and Guy M. Hagen. High density 3d localization microscopy using sparse support recovery. *Opt. Express*, 22(25):31263–31276, Dec 2014.

[62] Martin Ovesný, Pavel Křížek, Josef Borkovec, Zdeněk Švindrych, and Guy M. Hagen. ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. *Bioinformatics*, 30(16):2389–2390, Aug 2014.

[63] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. dAlché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, page 8024. Curran Associates, Inc., 2019.

[64] Sri Rama Prasanna Pavani and Rafael Piestun. High-efficiency rotating point spread functions. *Opt. Express*, 16(5):3484–3489, Mar 2008.

[65] Sri Rama Prasanna Pavani and Rafael Piestun. Three dimensional tracking of fluorescent microparticles using a photon-limited double-helix response system. *Opt. Express*, 16(26):22048–22057, Dec 2008.

[66] Sri Rama Prasanna Pavani, Michael A. Thompson, Julie S. Biteen, Samuel J. Lord, Na Liu, Robert J. Twieg, Rafael Piestun, and W. E. Moerner. Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proceedings of the National Academy of Sciences*, 106(9):2995–2999, Jan 2009.

[67] Tingwei Quan, Hongyu Zhu, Xiaomao Liu, Yongfeng Liu, Jiuping Ding, Shaoqun Zeng, and Zhen-Li Huang. High-density localization of active molecules using structured sparse model and bayesian information criterion. *Opt. Express*, 19(18):16963–16974, Aug 2011.

[68] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, 323:533, 1986.

[69] Michael J Rust, Mark Bates, and Xiaowei Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm). *Nature Methods*, 3(10):793–796, Oct 2006.

[70] Daniel Sage, Thanh-An Pham, Hazen Babcock, Tomas Lukes, Thomas Pengo, Jerry Chao, Ramraj Velmurugan, Alex Herbert, Anurag Agrawal, Silvia Colabrese, Ann Wheeler, Anna Archetti, Bernd Rieger, Raimund Ober, Guy M. Hagen, Jean-Baptiste Sibarita, Jonas Ries, Ricardo Henriques, Michael Unser, and Seamus Holden. Super-resolution fight club: assessment of 2d and 3d single-molecule localization microscopy software. *Nature Methods*, 16(5):387–395, May 2019.

[71] Prasanna Sattigeri, Jayaraman J. Thiagarajan, Karthikeyan N. Ramamurthy, and Andreas Spanias. Implementation of a fast image coding and retrieval system using a gpu. *2012 IEEE ESPA*, pages 5–8, Feb 2012.

[72] M. Schuster and K. K. Paliwal. Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, 45:2673, 1997.

[73] Ken Sekimoto. Langevin Equation and Thermodynamics. *Progress of Theoretical Physics Supplement*, 130:17–27, 1998.

[74] Alexander Sergeev and Mike Del Balso. Horovod: fast and easy distributed deep learning in tensorflow. *CoRR*, abs/1802.05799, 2018.

[75] Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. Sequence to sequence learning with neural networks. In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, page 3104, 2014.

[76] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. A Bradford Book, Cambridge, MA, USA, 2018.

[77] Makio Tokunaga, Naoko Imamoto, and Kumiko Sakata-Sogawa. Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nature Methods*, 5(2):159–161, 2008.

[78] Jonathan Tompson, Arjun Jain, Yann LeCun, and Christoph Bregler. Joint training of a convolutional network and a graphical model for human pose estimation. In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 1*, page 1799, 2014.

[79] Joel A. Tropp and Anna C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Trans. Inf. Theor.*, 53(12):4655–4666, Dec 2007.

[80] Dmitri A. Tsyboulski, Sergei M. Bachilo, and R. Bruce Weisman. Versatile visualization of individual single-walled carbon nanotubes with near-infrared fluorescence microscopy. *Nano Letters*, 5(5):975–979, 2005.

[81] A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, and K. J. Lang. Phoneme recognition using time-delay neural networks. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37:328, 1989.

[82] P. Zelger, K. Kaser, B. Rossboth, L. Velas, G. J. Schütz, and A. Jesacher. Three-dimensional localization microscopy using deep learning. *Opt. Express*, 26(25):33166–33179, Dec 2018.

[83] Lei Zhu, Wei Zhang, Daniel Elnatan, and Bo Huang. Faster storm using compressed sensing. *Nature Methods*, 9(7):721–723, Jul 2012.