Low SNR Computational Pattern Detection Applied to Multi-Spectral 3D Molecular Dynamics

A dissertation presented

by

Frederick Chang

to

The Department of Molecular and Cellular Biology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biochemistry

Harvard University

Cambridge, Massachusetts

January 2018

Dissertation Advisor: Professor Nancy Kleckner Frederick Chang

Low SNR Computational Pattern Detection Applied to Multi-Spectral 3D Molecular Dynamics

Abstract

We describe an algorithm for computational pattern detection with expanded performance into low Signal-to-Noise Ratio (SNR) regimes. Our approach consists of a Two-Stage Likelihood Pipeline (TSLP) and is applicable to large N-dimensional datasets. We apply this approach to the detection and localization of 3D fluorescent point sources (spots) in datasets generated by fluorescent microscopy. We demonstrate by in silico benchmarking that our approach is both maximal in sensitivity and selectivity in detection and minimal in error in photometry and localization of spots. Most importantly, the probability distribution of the Likelihood ratio is empirically derived, therefore the detection of spots in varying background conditions consists of one parameter. Since imaging in low SNR regimes corresponds to imaging with low excitation energies, our approach can enable long timescale imaging of 3D fluorescent spots at high temporal resolution. We illustrate this capability by analyzing 3D in vivo dynamics of fluorescently tagged single molecules and oligomeric complexes. Additionally, imaging in 3D and/or multiple colors (or ND) is achieved with a minimal increase in excitation energy as compared to 2D single color imaging.

Acknowledgments

Without the support from my parents, my wife and my PI, nothing would have been accomplished. With the birth of my son, this PhD probably took longer to accomplish (but then again he is the best thing I accomplished).

Contents

1 Introduction **1**

1.1 Biology through the lens of imaging 1

1.2 Optimal signal processing 4

1.3 The method of Likelihood 6

2 Description of the Two Stage Likelihood Pipeline **11**

2.1 Overview 11

2.2 Description of Stage I 12

2.3 Candidate Region Selection 15

2.4 Description of Stage II 16

3 Description of Multi-Spectral and Multi-Emitter Generalization **20**

3.1 The problem of multi-spectral imaging 20

3.2 Description of multi-spectral multi-emitter Stage I 22

3.3 Description of multi-spectral multi-emitter Candidate Region Selection 22

3.4 Description of recursive multi-spectral multi-emitter Stage II 24

4 Benchmarking **26**

4.1 Synthetic benchmarking of detection and localization 26

5 Distribution of the Log Likelihood Ratio **33**

5.1 Deriving the theoretical distribution for the Log Likelihood Ratio 33

6 Mathematical Derivation of the Two Stage Likelihood Pipeline 38

6.1 Introduction 38

6.2 Background 38

6.3 The Noise Model, Intensity Hypothesis and the Likelihood . . 39

6.4 Derivation of Stage I 41

6.5 Derivation of the False Positive Rate 47

6.6 Derivation of the Multi Spectral Multi Emitter Stage I 51

6.7 Calculating the Multi Spectral Multi Emitter FPR 59

6.8 Candidate Selection 59

6.9 Derivation of Stage II 59

6.10 Derivation for the recursive Multi Spectral Multi Emitter Stage II 62

6.11 Deriving the Standard Error 64

7 3D Single Molecule Imaging in a Living System 68

7.1 MreB dynamics in B. *subtilus* 68

8 Dynamical Localization of a Thylakoid Membrane Binding Protein is Required for Acquisition of Photosynthetic Competency 72

8.1 Abstract 72

8.2 Introduction 73

8.3 Results 77

8.3.1 Vipp1 forms transient puncta at regions of high thylakoid curvature at the cell periphery 77

8.3.2 Spatio-temporal distribution of Vipp1 enables photosynthetic competency 87
8.4 Discussion 101
8.5 Experimental Procedures 105
8.5.1 Strains and Growth Conditions 105
8.5.2 Genetic Manipulations 107
8.5.3 Light microscopy and image analysis 108

9 Conclusion **112**

A Appendix **113**
A.1 Publications 113

Statement of Contributions

Chapter 1,2,3,4	Frederick Chang (FC) & Nancy Kleckner (NK) wrote the text. FC & NK designed the research. FC wrote the code, analyzed the data, performed the research.
Chapter 5,6,7	FC wrote the text, analyzed the data, and derived the mathematics.
Chapter 8	Andrian Gutu (AG)[†], FC[†] & Erin O'Shea (EO) conceived and designed the research. AG performed the experiments. AG and FC analyzed the data, prepared the figures and wrote the initial manuscript draft. EKO supervised research and revised manuscript draft.

List of Figures

1 Noise model for 3D fluorescent spot detection 8

2 Pattern detection framework 14

3 Two Stage Likelihood pipeline framework 18

4 Pattern detection logic and Likelihood implementation for fluorescence spot detection 19

5 Multi-spectral description 23

6 Recursive Two-Stage Likelihood Pipeline 26

7 Visualizing the performance of the single color version of the Two Stage Likelihood Pipeline 31

8 Two Stage Likelihood Pipeline framework 32

9 Benchmarking the performance of Multi Color Stage I filter at different levels of Backgrounds (B) and Amplitudes (A) . . 34

10 One threshold to rule them all 37

11 Approximate LLRatio values at different (B)ackground values 48

12 Gamma PDF fit to the histograms of background values at various (A)mplitude and (B)ackground Values from the simulations defined above 48

13 Scale parameter of the Gamma fits 49

14 Shape parameter of the Gamma fits 49

15 Linear fit of the Scale vs (B)ackground values 49

16 -Log(FPR) as calculated by the gamma pdf fit at different (B)ackground Values 50

17 Gamma PDF fit to the histograms of signal values at various (A)mplitude and (B)ackground Values 50
18 Scale parameter of the Gamma fits to the signal histograms . 51
19 Shape parameter of the Gamma fits to the signal histograms 51
20 Two color intensity model 52
21 Three color intensity model 58
22 Scale parameter of the Gamma fits for 2 spectra LLRatio . . 59
23 Shape parameter of the Gamma fits for 2 spectra LLRatio . . 59
24 Linear fit of the Scale vs (B)ackground values 59
25 Schematic of the Two Stage Likelihood Pipeline 66
26 mreb 71
27 Characterization of Vipp1 expression and cellular localization in strains with epitope-tagged versions of Vipp1. 78
28 Vipp1 forms peripheral puncta that are located at regions of high thylakoid curvature 81
29 Vipp1 puncta colocalize with CurT enrichments at thylakoid edges 84
30 Detection and quantification of Vipp1-GFP by immuno-electron microscopy 86
31 Vipp1 forms transient puncta of limited mobility at the cell periphery 88
32 Perturbation of Vipp1 localization and characterization . . . 89
33 Time-course of rapamycin-induced Vipp1 relocalization . . . 92

34 Relocalization of Vipp1 to the nucleoid has no significant effect on Dark or *Light Growth* but elicits a severe growth defect during *Dark-to-Light Transition* 95
35 Vipp1 relocalization produces no growth or thylakoid morphology defect in Dark or *Light Growth* cultures but elicits a severe growth defect in the *Dark-to-Light Transition*. 97
36 Representative transmission EM image of dark- and light-grown with and without rapamycin 99
37 The early stages of *Dark-to-Light Transition* are more prone to the effect of Vipp1 perturbation. 100

1 Introduction

1.1 Biology through the lens of imaging

Time lapse imaging of living systems provides a glimpse into the non-equilibrium universe that life depends on. At the cellular level, macroscopic components ranging from the spindle to chromatin are made up of a multitude of single molecule components that are dynamic, interacting, and often catalyzing chemical reactions. It is an open question in molecular biology how to explain macroscopic biological phenomenology based on the emergent effects of these molecules and their interactions. If these molecules were in equilibrium, time lapse information is not necessary since static snapshots are sufficient to completely characterize their dynamics. But draw any boundary between the living system and its environment and there is always a flux of energy and mass. This violates the equilibrium assumption of conservation.

For single molecule time lapse imaging in living systems, 3D information is necessary as molecular agents are organized differently along x, y, or z dimensions. Long time lapse information is important so as to reliably extract information from noisy dynamic trajectories that is not just diffusion. High temporal resolution is essential as this timescale relates to its biophysical context such as binding and transport. Labeling multiple molecular agents is important since this allows the interaction between different molecular actors to be measured over space and time. However, imaging in these desirable time and length scales has been prohibitive since the excitation energy used for fluorescent imaging is toxic for both fluor and specimen.

Recently, a renaissance of new techniques emerged that seek to reduce this excitation energy cost. A diversity of new optical approaches have been published, and they are all fundamentally about controlling the excitation energy to limit unnecessary damage (Chen et al., 2014; Voie et al., 1993; Fadero et al., 2017). Light sheet microscopy is one such example, in which the excitation plane is optimized to only align with the focal plane. This is in contrast to traditional imaging, like in epi-fluorescence and confocal, in which the entire volume of the specimen is illuminated. However, other recent innovations reduce the excitation energy required by increasing the sensitivity of the microscope. Camera sensor improvements are one such example, recently maturing to nearly their theoretical limits in sensitivity.

The final component in imaging to be optimized is signal processing, the set of algorithms used to extract the relevant information from the measured dataset. For single molecule imaging, signal processing is used to detect and localize the digitized pattern manifested by the imaged molecule. If the algorithms used can allow imaging to be done at lower Signal to Noise Ratios (SNRs), this corresponds to lower excitation energies. Yet, the bulk of signal processing used in imaging has remained the same for decades. For example, for the task of detecting a particular pattern in a dataset, an algorithm called matched filtering is typically used. This algorithm is theoretically proven to be optimal in enhancing the signal of interest over that of noise (Rothstein, 1954). But the caveat is that the assumptions underpinning this optimality does not include interference from other patterns and noise sources, and these are important features to address when imaging living systems.

Since signal processing is a mature field, my initial assessment was that

these commonly used algorithms were sufficient. However, in 2010, the Sedat Lab published an observation that would challenge my understanding (Carlton et al., 2010). They showed that they can dramatically lower the excitation energy used by using a recently described patch-based de-noising algorithm. This algorithm lowered the excitation energy used by allowing very low SNR datasets to be measured while still maintaining the relevant signal of interest[1].

However, this paper was more a proof-in-principle that signal processing is an important, yet oft-neglected component to the imaging system. The formidable barrier to the practical application of their algorithm was in the computational overhead of applying a patch-based de-noising algorithm to datasets commonly encountered in imaging, e.g. large 3D timelapse datasets. With respect to single molecule imaging, in which sub-pixel localization is desired, the patch-based de-noising algorithm destroys all relevant information used for this process.

Nonetheless, this Sedat et.al observation encouraged me to ask what are the limits to signal processing, especially with regards to single molecule detection and localization. The rest of my thesis describes how I reached these limits, and how I applied this new algorithm to the domain of 3D single molecule imaging in-vivo.

[1]The assumption of the algorithm is based on the fact that in most examples in imaging, there exists a plethora of redundant information, be it in within a timepoint or neighboring timepoints. The algorithm collates all of this redundant information as similar patches and assumes they are the same. Then the algorithm extracts the denoised version by averaging across the whole set. Finally, the algorithm puts this denoised version back to where it was collated from.

1.2 Optimal signal processing

That border between the ability to detect or not detect a faint signal typically defines the edge between scientific discovery and the unknown. From detecting delicate gravitational waves in astronomy to the detection of faint fluorescent single molecules in biology, the important task is to detect and localize a weak signal of interest that is often embedded in noise from multiple sources, including other interfering signals. An essential component that contributes to the performance of this task are the computational algorithms used for detection and localization of the pattern manifested by the signal in the measured dataset.

In biology, imaging fluorescently labeled single molecules or objects smaller than the diffraction limit (hereafter spots) is becoming increasingly important. These signals represent the spatial positions of labeled molecules in a living system, and understanding how Life structurally and dynamically organizes these molecules is informing the next generation of questions (Chen et al., 2014; Liu et al., 2014). One of the holy grails for molecular biology is the 3D molecule dynamics of different molecular agents in living systems, over long timescales. It is this key measurement that will open up the invisible non-equilibrium world that is the foundation for Life.

However, there are several challenges to this endeavor: (i) The excitation energy used to image the fluorescent signal is both toxic to the living system and to the fluorophore (Laissue et al., 2017). (ii) The fluorescent signal of interest is typically embedded in complex background signals (Ettinger and Wittmann, 2014). (iii) Generalizing spot detection from 2D to 3D, let

alone multi spectral 3D, in which different molecular actors can be imaged, severely exacerbates this excitation energy cost.

It is known that lowering the amount of excitation energy used is critical for productive imaging, especially with living systems (Chen et al., 2014; Balzarotti et al., 2017; Carlton et al., 2010). But, given a particular optical setup, lowering the excitation energy corresponds to a lower Signal-To-Noise Ratio (SNR) in the measured dataset. Consequently, this degrades the performance of the algorithm used for spot detection and localization, and the SNR regime in which the algorithm begins to fail defines the fundamental lower bound on the minimum excitation energy that can be productively used for that particular optical setup.

We solve these challenges by deriving and implementing an optimal and general pattern detection and localization algorithm that outperforms all other numerical methods. It is optimal in the sense that this algorithm is derived using the well-known Likelihood Principle, which is a privileged statistical framework notable for being able to extract all information with respect to the pattern of interest that is available in the dataset (Edwards, 1993)[2]. For the domain of fluorescent spots, we show that our algorithm is an optimal 3D multi-spectral and multi-emitter algorithm that is maximal in sensitivity and selectivity in detection and minimum in error for photometry and localization. Since this is a computational improvement, all imaging modalities that produce a dataset, from epi fluorescent to lattice-light

[2]Edwards defined the Likelihood Principle as "Within the framework of a statistical model, all of the information which the data provide concerning the relative merits of two hypotheses is contained in the likelihood ratio of those hypotheses." (Edwards 1972, 1992 p. 30)

sheet, can be improved by this low SNR extension. Additionally, we show that our approach outperform previous theoretically proven computational filters such as Convolution, Normalized Cross Correlation and Laplacian of Gaussian in detection by statistical measures of sensitivity, and selectivity. Consequently, for the case of epi fluorescent imaging, we can image 3D or multi spectral 3D diffraction limited spots for long timescale and at high temporal resolution.

1.3 The method of Likelihood

In general, detecting and localizing a pattern in a given dataset amounts to fitting a statistical model of the pattern to the given dataset, with localization given by the position parameters of the model and detection given by the goodness of fit between the model and the dataset. In terms of the Likelihood Approach, the goodness of fit between the model and the dataset is called the Likelihood, and the operation of finding the parameter set that returns the best Likelihood is the Maximum Likelihood Estimate (MLE) (Edwards, 1993). The Likelihood is proportional to the probability of observing the data given the model. Querying whether or not different hypothetical patterns best fit a given dataset amounts to comparing their different Likelihood values found at their MLEs by comparing their Likelihood Ratios (LR). This LR at their corresponding MLEs is called the Maximum Likelihood Ratio (MLR). The MLR is the most stringent criterion to be used when comparing for the presence or absence of different types of patterns.

The Likelihood Approach derives its power from its ability to proba-

bilistically model the noise encountered in the dataset when detecting and localizing the pattern of interest. For example, in fluorescent microscopy, the noise typically encountered comes from various sources, from the quantum emission process of fluorescence to the measurement read out noise. By modeling the noise, the fitting of a given model can take into consideration that some datum are noisier than others and thus should be given less weight when determining the outcome of a fit. Mathematically and empirically, this operation is proven to be theoretically optimal in providing the least error in the estimates of parameter values (Smith et al., 2010a; Cramer, 1946). Given that a defining characteristic of a low SNR regime is the prominence of noise, it follows that the Likelihood Approach is uniquely powerful for pattern localization at low SNR regimes.

The power of any pattern detection and localization approach is defined on two fronts: Its detection performance and its localization performance. In terms of detection performance, the Neyman-Pearson Lemma states that using the Likelihood Ratio is the most stringent criterion when discriminating between two statistical hypotheses (Neyman and Pearson, 1933). In terms of localization performance, the Likelihood Approach is mathematically proven to give the least error in the parameter fit of a given model as given by the Cramer-Rao Lower Bound (Rao, 1945; Cramer, 1946).

Yet, directly applying the Likelihood Approach as described above is computationally intractable for two reasons. (i) Many noise models, such as those encountered in imaging, are non-trivial in numerical implementation, meaning that the algorithms used are iterative in computation (Huang et al., 2013a). In general, an iterative algorithm must begin with an initial guess,

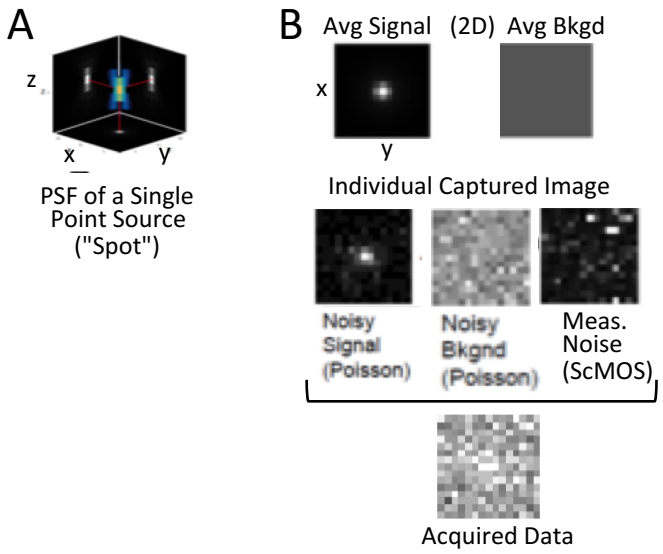

Figure 1: A. 3D intensity distribution of a Point Spread Function (PSF, aka spot) typically encountered in fluorescent microscopy, rendered with floating 3D voxels color coded by intensity. Maximum Intensity Projections for each dimension along X, Y or Z are displayed within each principal axis plane, with the dimension of each projection highlighted by a red line.

B. A schematic of the noise sources typically encountered in fluorescent imaging. Top: the signal to be detected is a fluorescent 3D spot and it is embedded in a fluorescent background (shown is average over many image acquisitions). Middle: Signal and Background levels at each pixel fluctuate from image to image (= "noise"). Spot signal and photon background are intrinsically noisy due to the quantum nature of light. The noisy signal and background are modeled as a Poisson process, parameterized by the average signal and background. Measurement noise is non-photon signal that appears in the data during conversion to digitized for and) is modeled using statistics typical of an CMOS sensor, with pixel dependent Gaussian read noise. The acquired data of the 3D spot is the sum of all three noisy components. For data captured in a single image, a weak spot signal will be obscured by noise/background.

e.g. the initial parameter values for the hypothetical pattern, then climb the local Likelihood peak by iteratively choosing at each step the next parameter set that increases the Likelihood value. The iterative nature of this process is already a challenge since this can be a computational bottleneck with respect to analyzing large datasets. (ii) However, the most damning problem is that in low SNR regimes, the Likelihood landscape is characterized to be rocky, in which there are many local maximums. As a result, many initial parameter guesses will never converge to the global maximum Likelihood, thus efficiently finding the optimal MLE at low SNR is an open question (Jin et al., 2016).

Because of these challenges, all pattern detection approaches including state of the art spot detection software described thus far either (a) avoid utilizing Likelihood methods at all or (b) apply it only after earlier "ad-hoc" computational steps (Sage et al., 2015a). Ad-hoc meant that non-optimal algorithms were used to find potential candidate pattern(s) and then roughly estimate their intensities and/or locations. Sage, et al. highlighted that this ad-hoc computation expanded the number of parameters the user needed to tune, and without expert understanding of such parameters, this ignorance formed a barrier between the user and effective utilization of the spot detection software.

Furthermore, many current spot detection software are limited to analyzing 2D datasets. However, there are many applications for 3D spot detection, in which the detection of the pattern consists of a 3D spot within a 3D dataset. A review of all state-of-the-art spot detection approaches in (Sage et al., 2015a) were mostly limited to 2D spot detection in 2D datasets.

Because of this limitation, many labs used sub-optimal 3D spot detection in which they adapt 2D spot detection to 3D data. For example, as described in a recent publication (Finn et al., 2017), *The x and y coordinates of the brightest pixel in each spot were calculated [from the maximum intensity z projection]. The z coordinate of the spot center was then calculated by identifying the slice in the z-stack with the highest value in fluorescence intensity for each of the spot centers.*

Additionally, a generalization of the pattern detection and localization framework for the application of multi spectral and multi emitter applications is difficult since it is not clear how to generalize from these ad-hoc approaches. Even if these sub-optimal approaches were forcibly generalized, the concomitant number of parameters to be tuned would explode, thus severely limiting its use.

Thus, there is in no case in which the full power of the Likelihood Approach is used for the critical step of primary pattern detection, and for the case of multi spectral and multi emitter detection and localization.

2 Description of the Two Stage Likelihood Pipeline

2.1 Overview

Our solution to these problems is by using the Likelihood Principle from the start rather than the end of the computational pipeline, thus ensuring that all computational steps are baptized by Likelihood Principles. To make this computationally tractable for large datasets, the algorithm is structured as a Two Stage Likelihood Pipeline (TSLP), in which the first stage (Stage I) is derived from a minimally invasive approximation to the fully detailed Likelihood function (see Mathematics of TSLP in Chapter 6). Given this principled framework, we generalize the Likelihood methods to handle multi spectral and multi pattern detection and localization.

The TSLP takes advantage of the fact that in most cases the pattern of interest to be detected and localized in a given dataset exists in a small subset of the dataset, e.g. in fluorescent imaging, spots take up a relatively sparse subset of the dataset. Therefore, Stage I first detects and localizes candidate spots in a computationally efficient way, but at the cost of some stringency. These candidate regions of the dataset are selected by thresholding the approximated LR calculated by Stage I - that is retaining the set of pixels in which their LR values are above a user defined threshold. In Stage II, a fully detailed Likelihood Analysis is applied to these candidates regions, with the initial parameter estimates given by the outputs of Stage I. The final detection of spots may be done by thresholding the fully detailed LR of Stage II. Thus, the computationally expensive Stage II is applied only to a small subset of the dataset, and it is armed with the most stringent

initial parameter estimates given by Stage I. This guarantees convergence even at low SNR.

For clarity, the TSLP will be first explained in terms of one spectral measurement and one fluorescent spot detection and localization. Afterwards, the generalization to multi spectral and multi emitter will be discussed. For full generality, spot patterns and background patterns can be replaced with any numeric pattern of interest. Atomic units of the dataset will be referred as pixels, even though it can be of any dimension, e.g. voxels. Illustrations will also be in terms 1D patterns and datasets even though it can be of any dimension. The following is a colloquial description of the TSLP, with the mathematical description located in the next chapter. The LR can be used interchangeably with its log transformed version the Log Likelihood Ratio (LLR). LLR is preferable because this is both numerically stable on the computation side and informative on the information theoretic side.

2.2 Description of Stage I

For purposes of Stage I, the data set is divided into a series of small "patches", sized so as to cover the pattern of interest, with one patch centered at every position in the data set. For every patch, two intensity hypotheses are tested – the signal and the null hypothesis. The signal hypothesis tests how well the dataset in a given patch match the intensities predicted by assuming there exists a *pattern + background* distribution of intensities. The null hypothesis tests *background* only. Each hypothesis carries with it a set of parameters that need to be optimized by MLE to best match the given data in a patch, e.g position(s) of pattern(s) and their amplitude(s). The key dif-

ference between the two hypotheses is whether or not assuming the presence of the pattern better fits the dataset or not.

In Stage I, solving the MLEs for each hypothesis is made computationally tractable for large datasets, but at the cost of precision:

(i) The position of the pattern is fixed within each patch, e.g. position is removed as a variable in the model for the signal hypothesis, and instead position is defined by the position of the patch. This limits the precision with which the pattern is localized within the dataset to the resolution of a single pixel.

(ii) For each patch, the MLEs for each hypothesis is determined by a single algebraic step, meaning there is no iterative computation. This is accomplished by combining assumption (i) and an approximation to the fully detailed noise model, with this approximation exact in very low SNR regimes. This approximation linearizes the equations in a way that allows for the MLE of all the amplitude parameters to be solved for in one algebraic step (See Mathematical Derivation of Stage I in Chapter 6.4). This avoids the problem of iterative algorithms in which they can get trapped at false global maximas in the rocky Likelihood landscape. Additionally, this approximation structures the MLE algorithm into a set of mathematical operations that scale computationally well with large data sets, i.e. {+, -, multiplication, division, convolution}.

These simplifications of (i) and (ii) define the "minimally invasive approximation" to the fully detailed MLR. Given these simplifications, Stage I LLR, or any of the MLEs are termed *approximate* LLR, *approximate* MLE of A, etc.

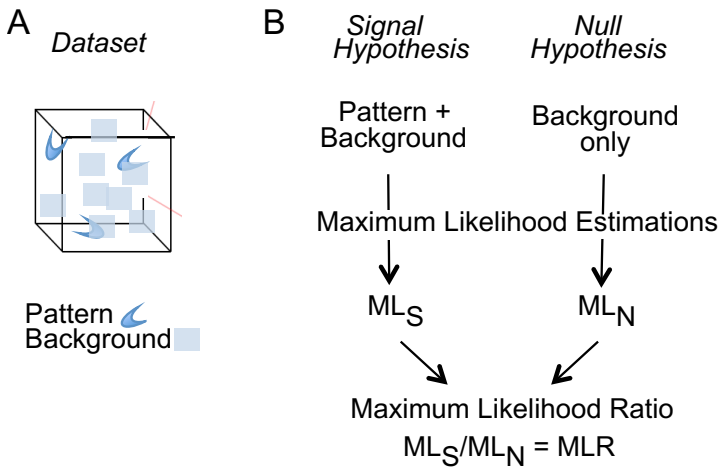

Figure 2:
A. Numerical data set characterized by pattern(s) plus background.

B. Pattern detection in a Likelihood framework. Mathematical expressions are formulated to describe two hypothetical conditions: the pattern of interest plus the background (the "Signal Hypothesis) and the background alone (the "Null Hypothesis".

Each hypothesis is evaluated by the Likelihood approach. Each such Maximum Likelihood Estimation defines a "Maximum Likelihood" (ML). The ML defies the set of parameter values at which the hypothesis could best explain the data and gives the "Likelihood" that, at those parameter values, the data are explained by the model for that hypothesis. The ratio of MLs for the two hypotheses, defined here as the Maximum Likelihood Ratio (MLR), is proportional to the probability that a pattern (+ background) is present in the data, rather than only background.

Given an input dataset, Stage I outputs several datasets that are the same size:

- Approximate LLR reports the fold probability difference between signal versus null intensity hypothesis for every position in the dataset.
- Approximate MLE of A and Bs, which correspond to the best fit parameters for the signal hypothesis, solved for every position in the dataset.
- Approximate MLE of Bn, which correspond to the best fit parameters in the null hypothesis, solved for every position in the dataset.

The approximate MLE of A for the signal hypothesis or the approximate LLR are essentially filtered version of the dataset given, in which filter means to enhance the signals coming from the pattern of interest only. These provide a quantitative visualization of the pattern information alone, separate from any background contributions.

2.3 Candidate Region Selection

Outputs of Stage I are used to select for candidate regions in the dataset to be further examined using a fully detailed Likelihood method. For example, the approximate LLR can be thresholded for selection as relatively high values of LLR indicate that the pattern is likely present in that location. Since the LLR is interrogating the *pattern + background* hypothesis over that of just *background* only, the LLR metric is a background subtracted measure. This is crucial because in real-world scenarios complex background signals

interfere with the detection and localization of the fluorescent spot, and this problem has been a challenge to previous spot detection approaches (Sage et al., 2015a).

2.4 Description of Stage II

In Stage II, each candidate region is subject to a fully detailed Likelihood analysis which relieves the approximations made for Stage I analysis.

(i) The position of the pattern(s) is now an adjustable parameter, thus allowing for its sub-pixel localization.

(ii) A fully-detailed noise model is used, therefore solving the MLE is an iterative hill climbing algorithm; given an initial parameter guess, the algorithm selects another parameter set that increases the Likelihood value, then repeats. The algorithm terminates when the Likelihood value no longer increases, therefore defining the peak ML and its corresponding MLE of the parameters.

The initial parameter guess for the iterative hill-climbing MLE algorithm are given by the outputs of the Stage I analysis. Since Stage I is a minimally invasive approximation to the fully detailed Likelihood function, this initial parameter guess is very close to the optimal MLE in the fully detailed Likelihood landscape. It is this feature that guarantees robust MLE convergence, especially in low SNR regimes. This initial parameter guess is extracted from Stage I outputs that correspond to the candidate region examined. That is, for the signal hypothesis, there exists 5 parameters for a 3D dataset and 3D pattern: Amplitude of the pattern, its position in x, y, and z, and the background. The initial xyz position is taken from the position of the

maximum LLR value in the candidate region. The initial amplitude and background estimate is taken from the corresponding approximate MLE of A and Bs values at the xyz position respectively.

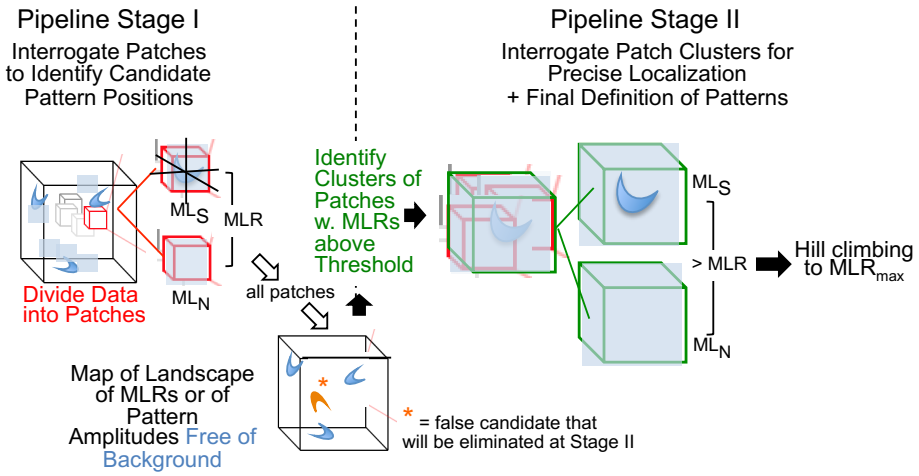

Figure 3: The Two-Stage Likelihood Pipeline allows to scan a data set by the pattern detection logic described in Figure 1 in a computationally tractable form.

Stage I. The data set is divided into patches that are slightly larger than the pattern of interest, with each such patch centered on every position in the data set. In the model for the pattern, the pattern is present at a constant specified position within the patch. For each patch, MLs and their MLR is determined as in Figure 1B, using approximated versions of the mathematical expressions for the noise model. Below: the values of the MLRs for each patch (or the values obtained from the MLR or by some other filter) provide a quantitative landscape of the data that can be presented as a visual output if desired. Individual patches or clusters of patches with high MLRs (e.g. above a defined threshold) indicate the possible presence of a pattern, thus defined as a "candidate pattern".

Stage II. For each such high MLR position, a corresponding region that encompasses all patches with high MLRs, is defined and re-interrogated by Stage II in a more rigorous way. At this stage, for each candidate pattern region, a hill-climbing exercise is performed separately for the Signal and Null Hypotheses (Figure 1), guided by Likelihood functions that now utilize fully-detailed noise models and, for the Signal Hypothesis, with the position(s) of the pattern(s) no longer fixed. For the hill-climbing exercise and for each hypothesis, each parameter (e.g. position and amplitude) is defined relative to its preceding value by the displacement that gives the biggest increase in Likelihood. This process is continued until an ML is reached. An MLR is determined from the MLs of the two hypothesis (Figure 1). If the value of this MLR is above a desired threshold, a pattern is deemed to be present, with the position of that pattern in the dataset concomitantly determined. Extensions of this approach can allow to identify multiple partially-overlapping patterns within a dataset and/or multiple instances of pattern(s) within a single patch, as described in Chapter 3.

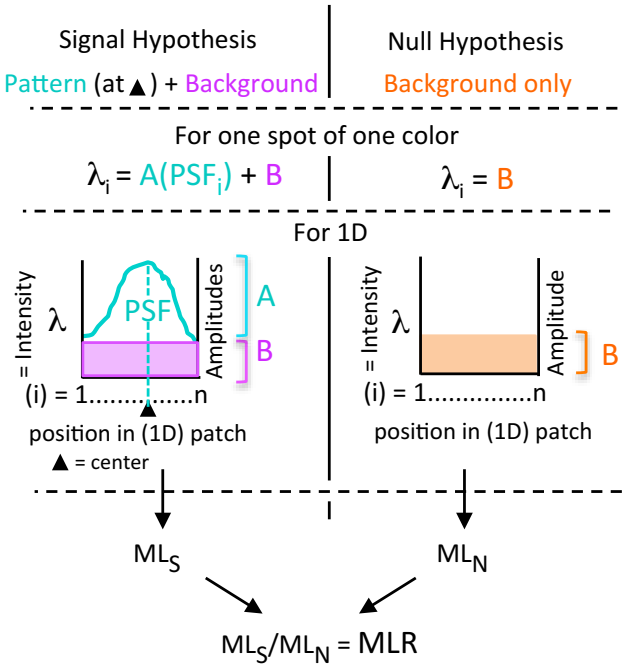

Figure 4: Pattern detection logic and Likelihood implementation for 3D fluorescence spot detection. General approach (Figure 1B) as applied to a diffraction-limited signal where the pattern is the Point Spread Runction (PSF) and the background is uniform. Illustrated for a 1D data set for simplicity.

3 Description of Multi-Spectral and Multi-Emitter Generalization

3.1 The problem of multi-spectral imaging

It is often desirable to be able to simultaneously detect and follow more than one type of molecule or object in a single sample so as to examine their spatio-temporal relationships and interactions. This can be achieved by labeling different entities of interest with different fluorophores, aka fluorophores of different "colors". Multi-color imaging is achieved by collecting data in a number of "channels", each involves excitation of the sample in one wavelength window and collection of resulting photon emissions in another window, where each of the two windows is defined by an appropriate spectral filter. Often (but not always) data is collected in a set of c channels whose excitation/emission windows are each appropriate to one of c fluorophore colors in the sample. Such approaches are referred to as multi-spectral imaging with a corresponding multi-spectral dataset (Zimmermann et al., 2014). Each datum in this multi spectral dataset is indexed by its position and also its spectral channel, e.g. [x,y,z,j], with j specifying the index of the channel. Therefore 3D multi-spectral data sets are 4D datasets.

The challenge of multi-spectral imaging is that the excitation and/or emission windows of different channels may partially overlap. For example, the signal detected in a particular channel designed to excite and detect emissions of one particular color may also include signals from other colors. This is spectral crosstalk. For multi-color fluorescent spot detection and

localization, the information related to a spot is now spread across multiple channels, as well as interference coming from multiple channels. Thus, the optimal algorithm used for detection and localization must take advantage of this property (Cutrale et al., 2017). However, current state of the art spot detection and localization treat each channel independently, in which they first detect and localize the spot in each individual channel, then find the correspondence of the detected spots between the channels (Bossi et al., 2008). Consequently, if there is a probability of a certain error for the detection of a single spot, this error is roughly multiplied by the number of channels in this scheme. This limitation is among the reasons that multi-spectral imaging is currently largely limited to relatively high SNR regimes and/or low number of colors that are well separated.

The basic premise of our multi-spectral generalization of the TSLP is based on calculating how a given intensity hypothesis manifests in a multi spectral dataset with spectral crosstalk. Note that the starting intensity hypothesis has no consideration of spectral crosstalk. For the example of two colors, a multi-spectral intensity hypothesis can be of the form, a green spot with its green amplitude, position and background, and a red spot with its red amplitude, position and background. The task is to infer this original entity, the intensity hypothesis, that is devoid of spectral crosstalk. Therefore, only when the likelihood of this intensity hypothesis is calculated, e.g. how this hypothesis manifests in a mulit-spectral dataset, is spectral crosstalk considered. These spectral crosstalk coefficients are typically constants that are calibrated beforehand and are specific to the optical setup and colored fluors used (Zimmermann et al., 2014).

3.2 Description of multi-spectral multi-emitter Stage I

The same patch based schema described above are used, in which there is a patch centered at every position in the dataset. However, instead of assuming one spot, there are c spots, with one spot of each color present at some specified position within the patch. The model for the signal hypothesis is thus a multi-spectral multi-emitter model in which all spots are overlapping and the background is assumed to be uniform for all colors. The null hypothesis is assumed to be composed of background only for all the colors. The spectral crosstalk coefficients are given as constants. The same Stage I approximations are applied here, and the multi-spectral multi-emitter signal and null hypothesis are then evaluated for their approximate MLE and their corresponding approximate ML.

The outputs of a multi-spectral multi-emitter Stage I therefore are generalized to be,

- Approx LLR
- Approx MLE of A_1,...,A_c *multi spectral multi emitter signal hypothesis*
- Approx MLE of Bs_1,...,Bs_c *multi spectral multi emitter signal hypothesis*
- Approx MLE of Bn_1,....,Bn_c *multi spectral multi emitter null hypothesis*

in which, 1 through c indexes the channel

3.3 Description of multi-spectral multi-emitter Candidate Region Selection

If the approximate multi-spectral LLR is high, it means there is likely some combination of colors and spots at this position in the data set; e.g. if there is no spot

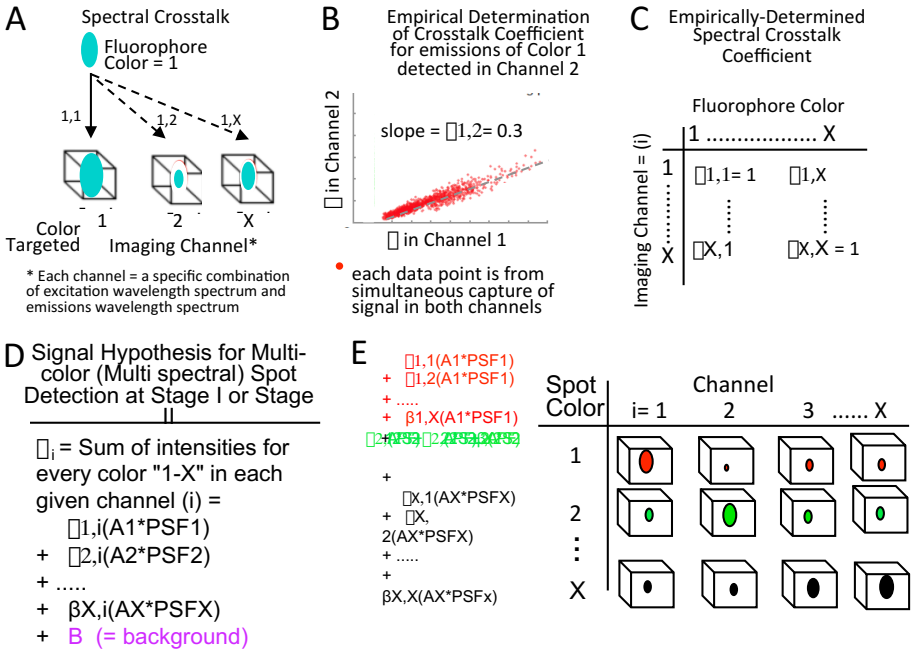

Figure 5: (A) To image a fluorophore of a particular color, the sample is excited in one wavelength window and resultant emissions are detected in a different wavelength window. Each such combination comprises a "channel" whose wavelength windows are appropriately defined so as to target that particular color. Samples containing fluorophores of more than one color are imaged in a corresponding number of specifically-targeted channels. This regime is known as "multi-spectral imaging". In this regime, emissions from a given fluorophore may also be detected in (an)other channel(s) whose excitation and emissions detection windows partially overlap those of the targeted channel. This phenomenon is known as spectral crosstalk. (B) For any given fluorophore, the level of signal for a particular color as detected in a channel other than the "target channel", as a fraction of the level seen in the target channel, can be determined empirically (details in text). Fluor is tdTomato, analyzed using a Pinkel configured multi- band pass filter cube. (C) In a given multi-color sample that is imaged in multiple channels, emissions for each color may occur in any/all channels. Empirical determinations (panel B) define the matrix of spectral crosstalk coefficients (b) for all combinations of colors and channels. The crosstalk coefficient for a color in its cognate targeted channel is defined as 1.

(Continued) (D) The Stage I Signal Hypothesis for Multi-color (Multi-spectral) fluorescence imaging assumes that a spot of each color is present at the same, specified position. The model sums: [for each channel, the signal from each fluorophore color], for every channel. Uniform background is assumed for all channels (= B). (E) Model written out explicitly (left) in relation to visual representation of the data (right).

of a certain color, the amplitude corresponding to that color is close to zero. But if there is a spot of a certain color, the corresponding amplitude is greater than zero. Since multi-spectral multi-emitter LLR is the same LLR as for the single emitter case above, candidate region selection follows the same protocol.

3.4 Description of recursive multi-spectral multi-emitter Stage II

The same relaxation of the assumptions to Stage I are applied here, in which the noise model is fully detailed and the spot(s) positions are now adjustable parameters. However, instead of fitting just one spot, Stage II can fit a multiple number of spots to a given candidate region, with these spots coming from any number of channels.

However, to ensure proper convergence for the task of multi-emitter fitting, the algorithm is recursive, building up a multi-emitter fit by sequentially subtracting and then fitting the single most likely spot to the residues.

Step 1. The initial parameter guess for the iterative hill-climbing MLE algorithm is given by the outputs of the multi-spectral multi-emitter Stage I analysis that corresponds to the candidate region. The first single most likely emitter's xyz position is taken from the position of the maximum approximate LLR value in the candidate region. The color index of the emitter is chosen by selecting the channel containing the maximum MLE amplitude from the set of different colored MLE amplitudes, at that xyz position. Given this color, the amplitude and background

corresponding to that color and xyz position are extracted. A fully detailed MLE and its ML of this initial emitter are found given these initial parameter estimates.

Step 2. The above fully detailed MLE corresponds to a hypothetical intensity distribution for that first emitter, and this intensity distribution is then then subtracted from the multi-spectral data in the candidate region. What is remaining is the residue, with its composition mostly removed of that emitter.

Step 3. This residual data is also subject to noise and may contain additional emitters. Hence the residue is then processed by Stage I again to determine the approximate LLR (and thus possible existence) of another emitter of the same or another color. The next potential emitter position is identified by the highest approximate LLR, with its spectral component chosen from the one that contains the maximum from the list of spectral MLE amplitudes solved for that position. This emitter candidate with a sufficiently high approximate LLR in this analysis is then jointly further analyzed by Stage II along with the previously analyzed spots.

Step 4. Steps *2* and *3* are repeated until the highest MLR for any remaining signal is too low to indicate the presence of a spot or until the analysis reaches a preprogrammed stop which defines a limit to the number of spots to be solved for.

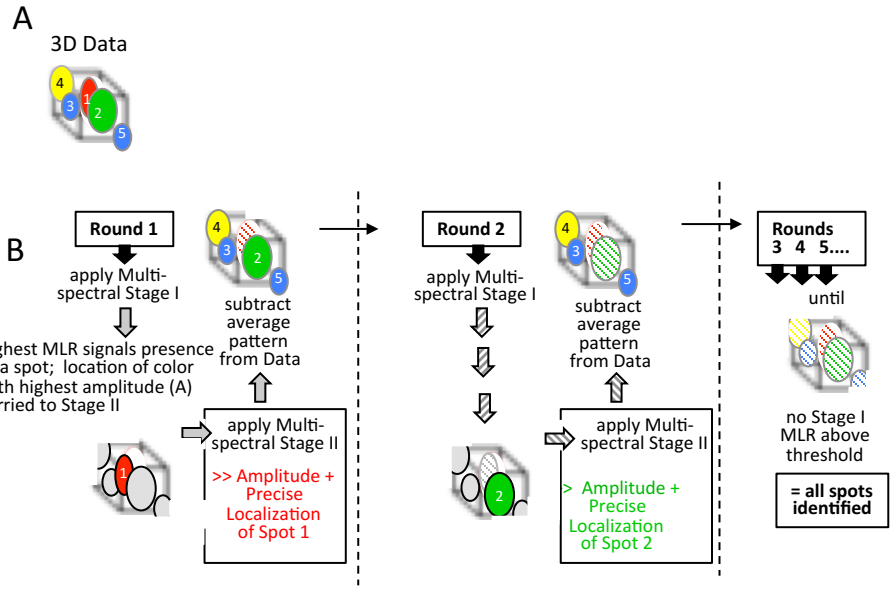

Figure 6: Implementation of Recursive Two-stage Likelihood Pipeline for sensitive spot detection and precise localization of one or more spots of more than one color in a multi-spectral dataset (or for one or more spots of one color in a single color dataset).

4 Benchmarking

4.1 Synthetic benchmarking of detection and localization

In fluorescent microscopy, the SNR regime in which the performance of spot detection and localization is deemed unacceptable defines the corresponding minimum excitation energy that can be used for imaging. What is deemed acceptable or unacceptable is user-specific, based on a complex mix of constraints on localization and photometry performance alongside detection performance. For live cell imaging, it is commonly desirable to image at the lowest possible SNR regime in which detection is the limiting metric (Carlton et al., 2010). Additionally, it is desirable to distinguish between overlapping spots, be it in one channel or two or more channels. Distinguishing overlapping spots is dependent on SNR and their distances.

Here, we summarize the performance of detection and localization of spots at low SNR regimes, in which the spot intensity is low relative to the background. This is accomplished by synthetically generating datasets with realistic noise models that are typical in fluorescent imaging. Since the benchmark is synthetically generated, ground truth information regarding the spots intensity and position are known.

The synthetic benchmarking models the noise and image formation typical in fluorescent microscopy of diffraction limited spots. A range of different spot intensity and background levels at low levels are presented. The spots are modeled as a 3D Gaussian intensity distributions that is centered in a given position in the dataset. The intensity model for the dataset is composed of this 3D Gaussian with amplitude A added to a uniform background B in units of average photon emissions. Therefore, different A and B levels set the mean photon emission rate per pixel, which is then sampled as a Poisson Distribution (Hecht et al., 1942). The camera measurement process is parameterized by its quantum efficiency coefficient (QE) and its per-pixel read noise variance. The QE describes the efficiency of converting a photon to an electron, and the read noise reflects the random process of converting that captured electron to digital units. The read noise process models CMOS sensor pixel statistics, with pixel specific read noise modeled as a Gaussian parameterized by a variance and zero mean (Fossum et al., 1995). The read noise is modeled per-pixel to account for the diversity of read noise variances that typical CMOS sensors display; the distribution of read noise variance resembles a skewed distribution with a very long tail, which means there exists a minority of pixels that are very noisy relative to their neighbors (Huang et al., 2013a). Thus, to simulate a CMOS measurement process, given a Poisson-sampled photon distribution over a set of pixels, these photons are multiplied by the QE coefficient and the per pixel read noise is added. The gain, which converts electron units to digital units, is multiplied and a constant offset added. The final units are in analog to digital units (ADU).

Localization and photometry performance are measured as the error in those corresponding parameter estimates with respect to the ground truth (Smith et al., 2010a). Even if the parameters that describe a synthetic spot are constant, e.g. the spot is placed in a given x, y and z coordinate with a certain intensity A and background level B, each random sampling of this process produces different intensity distributions that are consistent with this random process. The consequence is that when fitting an intensity model to this noisy dataset, the parameters of the fit will also vary with the differing intensity distributions.

Detection performance depends on how well the criteria that are used for selecting spots from non-spots can distinguish true spots from spurious contributions due to background and noise. In general, computational filters for pattern detection are designed with these criteria in mind, in which a given input dataset is filtered and the output intensities relate to the strength that a pattern of interest exists in that position (Turin, 1960). When the signal for the pattern of interest is weak relative to the noise and/or background contributions, the reported intensity criterion can overlap between positions where there is no spot and where there is a spot. Therefore, detection performance amounts to measuring the distribution of intensities in non-spot versus spot positions of the dataset. In synthetic datasets the spot positions are known in advanced, and the filtered intensities corresponding to non-spot and spot positions are defined as background and signal measurements, respectively. The detection performance is essentially defined as the amount of overlap between background and signal intensities. This can be measured for any computational filter, including the outputs of Stage I. Detection performance, through the lens of overlap in background and signal intensities, can be quantitatively described by two measurements, sensitivity and selectivity (Peterson et al., 1954). Sensitivity is equivalent to the True Positive Rate (TPR), which is the probability the algorithm reports a spot when there is truly a spot. Selectivity is equivalent to the True Negative Rate (TNR), which is the probability the algorithm successfully reports

no spot, when there is truly no spot.

There are also additional statistical measures related to the TPR and TNR. The False Positive Rate (FPR) is the probability the algorithm incorrectly reports a spot, when there is truly no spot. Lastly, the False Negative Rate (FNR) is the probability the algorithm incorrectly reports no spot, when there is truly a spot. The FPR and FNR are not independent measures but sister statistics to the TPR and TNR, in which FPR = 1 - TNR and the FNR = 1- TPR.

The operation of determining these detection metrics is as follows: first define hypothetical dataset parameterized by the (A)mplitude of the spot and the (B)ackground levels. A noisy sample is taken using the realistic noise model described above. This is repeated many times and the background and signal intensities of the filtered output are collected corresponding to regions that contain no-spot and centered at the spot, respectively. The background and signal intensities will be a probability distribution of intensities. The detection metrics are calculated from the cumulative distribution functions of the background and signal intensities, with the FPR = 1 - CDF(background intensities) and the FNR = CDF(signal intensities).

There is another complication to consider. Selection of a pattern in a filtered datasets amounts to selecting a threshold, and positions in the dataset that exceed that threshold are deemed candidates. Different thresholds select for different stringencies on the sensitivity and selectivity of the computational filter used. For example, if the threshold is set too low, all available patterns can be selected but at the cost of selecting many spurious contributions that do not come from the pattern. This means that given a particular filter, there is a tradeoff to sensitivity and selectivity, in which increasing one will decrease the other.

To fairly benchmark a given filters detection performance, a consistent criterion for selecting the filter specific threshold is needed. The Equal Error Rate (EER) is one such criterion, and defines the threshold at the level in which sensitivity

equals the selectivity (when the FPR equals the FNR). The EER is then defined as the value of the FPR or FNR at this particular threshold. Graphically, the EER is a measure of the overlap between the distribution of background and signal intensity measurements, with complete overlap mapping to EER = 0.5 and complete separation to EER = 0.

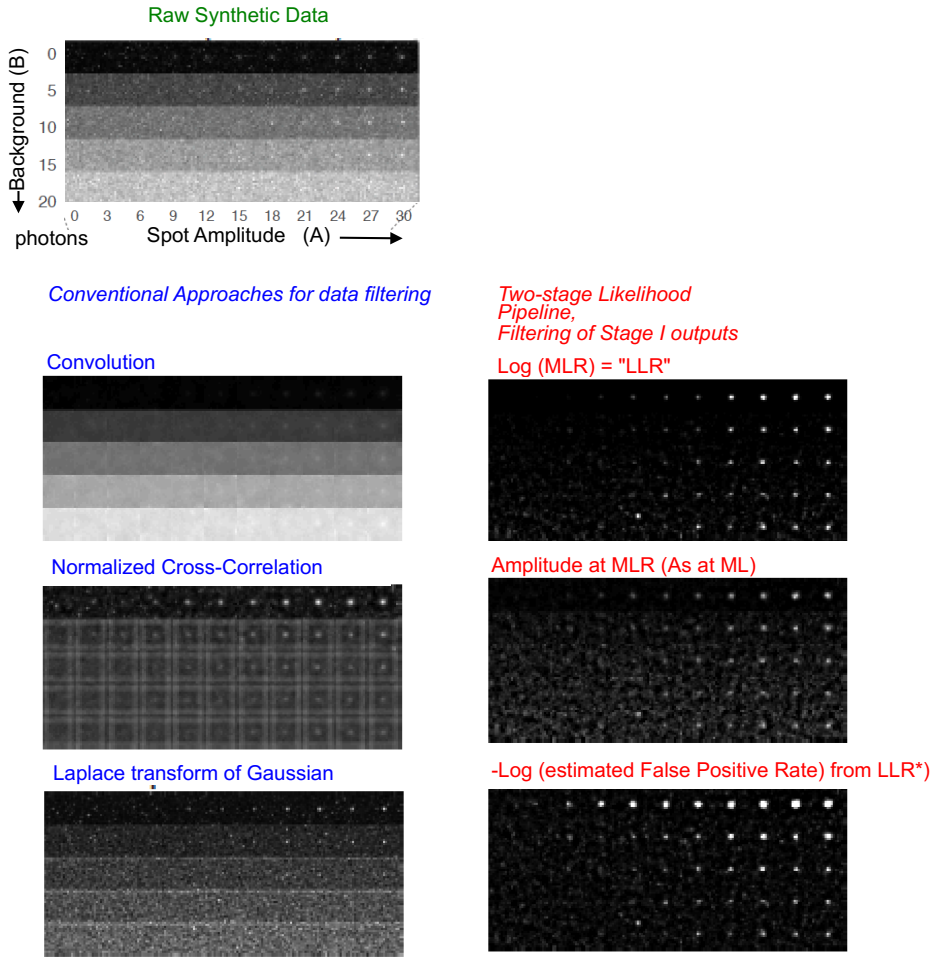

Figure 7: Visualizing the performance of the single color version of the Two Stage Likelihood Pipeline for low SNR detection and localization of fluorescent spots: alternative filters and comparison with conventional detection Part I.

(Top) Regions of raw synthetic data, each containing a centrally-located spot, were generated at each of the indicated levels of signal and background amplitudes (A and B, respectively). All combinations tested represent low SNR conditions.

(Left) Visual performance of other commonly used computational filters as applied to the raw data.

(Right) Visual performance of various Stage I outputs applied to the raw data. The design of -Log(estimated FPR) filter is described in the next chapter.

Greyscale presentations of the quantitative landscape for each region in each SNR condition are shown for raw data (top left) and output of each filter (below)

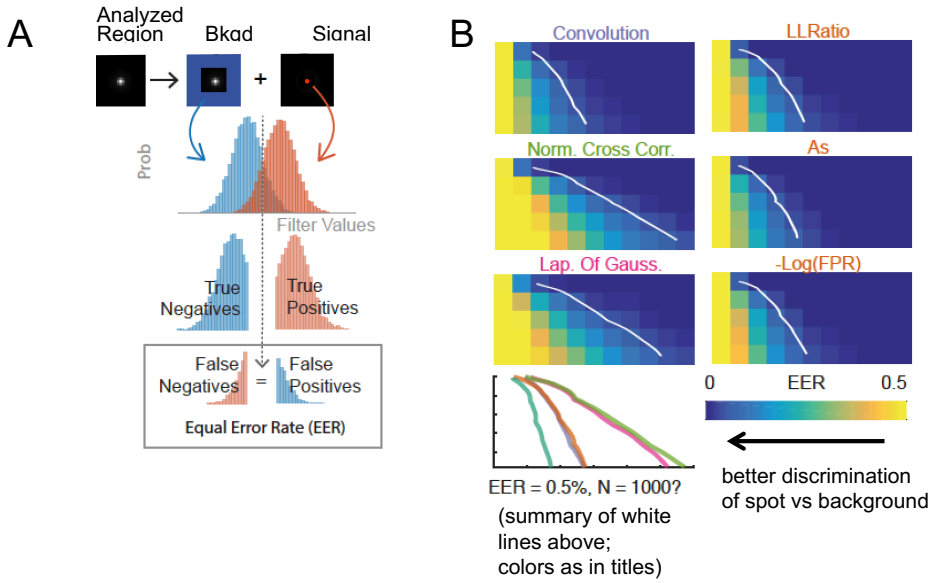

Figure 8: Benchmarking the performance of the single color version of the Two Stage Likelihood Pipeline for low SNR detection and localization of fluorescent spots: alternative filters and comparison with conventional detection Part II.

A. A method for determining the discrimination between Signal and Background in a particular set of raw data or of the output of an applied filter for synthetic data sets. Each region shown in each of the greyscale presentations of Figure 5 was analyzed quantitatively to asses the distributions of intensities for Signal and Background. For this purpose, specific definitions of "Signal" and "Background" were used. The "Signal" is the voxel selected as the center of the spot in the synthetic raw data set (Red spot). The "Background" is a region located outside of a square region that contains the Signal (blue area).

For each region analyzed, the distribution of pixel intensities in the Signal and Background sub-regions are determined.

Greater overlap between the two histograms implies less success in discriminating signal from background.

A convenient threshold for defining overlap, and thus success in discrimination, is the Equal Error Rate (EER). The EER is defined as the threshold where the False Negative Rate (FNR) equals the False Positive Rate (FPR). EER = 0 means no overlap, EER = 0.5 means 100% overlap. (FNR = proportion of cases where signal is wrongly considered as background; FPR = proportion of cases where background is wrongly considered as signal.)

(Continued) B. A heatmap of the EER measured for each A and B (and thus SNR) combination is shown for raw data and for each filter output as displayed in Figure 5. A contour line of EER=0.5% is presented as a white line.
LLRatio, As, -Log(est. FPR) and Convolution display the largest permissive area for detection (ie better performances at lower SNRs). However, only -Log(est. FPR) filter is robust to changing background conditions and this is shown in Figure 10.

5 Distribution of the Log Likelihood Ratio

5.1 Deriving the theoretical distribution for the Log Likelihood Ratio

At low SNR, the distinction between background fluctuations and the intensity of the signal of interest is small. The purpose of computational filtering is to enhance the signal of interest over that of background so as to further expand low SNR distinction between background and signal. The essential next task after filtering is to determine a threshold in which all background signals are eliminated, only leaving the signal of interest. In the schema of the detection metrics discussed above, this is equivalent to finding the threshold where the False Positive Rate is low. To do so, we need a theoretical link between the Likelihood Ratio and the detection metrics discussed above.

In 1938, Wilks derived a theorem that relates the Likelihood Ratio to these detection metrics, known as the Likelihood Ratio test (Wilks, 1938). The theorem states that the probability distribution of the possible Likelihood Ratios observed is asymptotically Chi-squared distributed. Given this distribution, a p-value can be calculated for the observed Likelihood Ratio, and the magnitude of this p-value relates to the False Positive Rate.

However, applying Wilks theorem to our situation is misleading since we violate several assumptions demanded by Wilks. The most egregious violations are: 1) Our Likelihood Ratio is calculated for a small sample size that comes from a patch

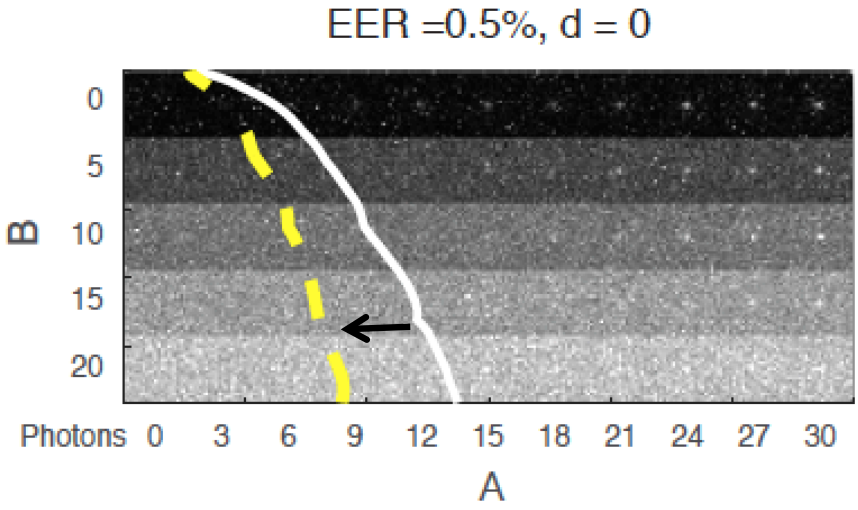

Figure 9: Figure 11 Benchmarking the performance of Multi Color Stage I filter at different levels of Backgrounds (B) and Amplitudes (A) for the case of two spectral measurements.

Synthetic data were generated in which signals comprising the PSFs for GFP and tdTomato were placed 0 distance apart. Each of the respective spot intensities and background are set by the corresponding A and B, assumed to be the same for both spots.

Data were filtered with Stage I of the multi-color spot detection algorithm. LLR landscape is shown for one of the two colors (which give the same result). Significance of the signal, compared to background (as in Figure 7) , is indicated by EER = 0.5% (yellow). Significance obtained by analysis of the same color in a single channel, From Figure 7, is shown in (white). Difference shows that multi-color analysis outperforms single color analysis (arrow).

or candidate region. Wilks demands that the sample size is large or infinite. 2) Our null hypothesis is essentially a simpler version of the signal hypothesis, in which the pattern of interest is set to zero amplitude. Wilks demands that the parameters of the null hypothesis is not on the border of the parameters of the signal hypothesis. Given our assumption that there is no negative photon emission rate, a zero amplitude pattern is on the lower bound border of amplitude possibilities.

The limitations described above are general to many different applications, therefore applying Wilks theorem results in erroneous estimates of the FPR, and other detection statistics. This problem is general to many fields, ranging from astronomy to biology (Protassov et al., 2002a; Huang et al., 2013a).

A clue to the theoretical Likelihood Ratio distribution is apparent when examining the derivation to Stage I (see Chapter 6.4). The approximate Log Likelihood Ratio (LLR) in Stage I is composed of a difference between two mean squared error components, with each mean squared error component corresponding to the signal and null hypothesis, e.g. one component is the squared error between the data and the best fit intensity model of the signal hypothesis. It is well known that the distribution of a random variable that is composed of a squared error is given by the Chi-squared distribution. Furthermore, a Chi-square distribution is one specific parameterization within the Gamma distribution family. Because the LLR is composed of a difference of two squared error components, how the LLR is distributed is determined by a difference of two Gamma random variables. The closed form solution for the difference of two Gamma random is described here (Mathai, 1993) as Theorem 2.1.

For computational simplicity and the fact that a difference of two Gamma random variables may be another Gamma distribution (depends on the parameters), I have fit a Gamma distribution to the signal and background measurements of the Stage I benchmarking. In other words, for every (A)mplitude and (B)ackground condition in the Stage I benchmark, there corresponds a best-fit Gamma distribu-

tion describing either the background or signal measurements of the approximate LLR distribution. Therefore the parameters of the Gamma distribution, in which one describes the signal and the other describes the background distributions, depend on the (A) and (B) values of the benchmark. (see Chapter 6.5 for the Gamma fits to the benchmarking)

However, when processing a given dataset, there is only one instance of the data. Stage I filters this dataset and estimates the (A)mplitude and (B)ackground MLEs and the LLR at every position in the dataset. Since we are armed with how (A) and (B) parameters correspond to the Gamma parameters, we can estimate the theoretical distributions of the LLR of the background and signal measurements. The FPR and FNR can then be calculated with high fidelity.

Future work will complete the closed form proof of the distributions of the Log Likelihood Ratio of two hypotheses.

The most important aspect of this statistically aware transformation of the Log Likelihood Ratio is that it becomes a background robust way of thresholding the data. This is important because in many biological applications, the pattern to be detected and localized, e.g. the spot, is typically embedded in a wide variety of different background conditions. Given that datasets are scaling faster in size than the ability of an experimentalist to manually curate them, it is essential to define a robust threshold for all of the encountered conditions. *One threshold to rule them all.* The one threshold can be defined by the False Positive Rate (or one of the other statistical metrics).

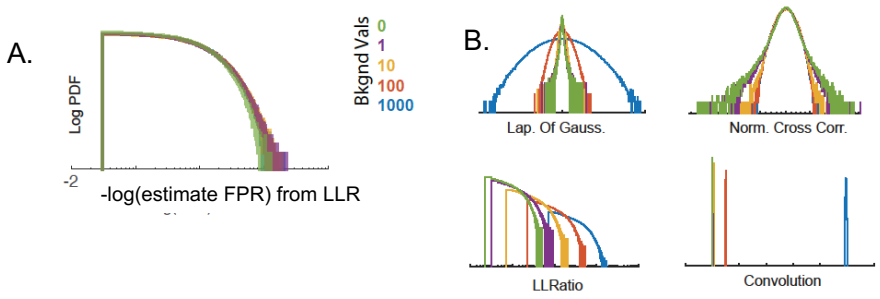

Figure 10: Identification of -log(estimate FPR) of LLR as the only filter output that is insensitive to the level of background.

A-B Synthetic data sets like those presented in Figure 7 top were prepared at a constant level of signal (0 photons) and varying levels of background from 0-1000 photons. Data were filtered in various ways as in Figure 7 bottom. The (log of) the Probability Density Function (PDF) for the value obtained for the Background Component at different levels of background levels is plotted for each filter.

Comparison of (A) vs (B) shows that only the filter output based on the FPR of the LLR is robust to different background levels. Thus this single parameter can be used to threshold spots in varying background conditions.

6 Mathematical Derivation of the Two Stage Likelihood Pipeline

6.1 Introduction

To find a fluorescent spot with the most sensitivity and selectivity, and with the least possible localization error, especially in low SNR regimes, consists of using the Likelihood Approach with all gory details of measurement noise modeled in[3]. However, this approach as-is is computationally intractable when you apply it to every position in a large data set, especially those commonly generated by CMOS sensors. The reason is because the fully detailed Likelihood Approach is iterative in nature, thus not only computationally expensive, it requires initial conditions that are already close to the correct solution, with this constraint exaggerated at low SNR regimes. To address this challenge, we describe a Two Stage Likelihood Pipeline in which the First Stage finds candidates spots and their initial conditions (using an Approximate Likelihood), and then the Second Stage applies the fully detailed Likelihood. This end-2-end Likelihood pipeline retains the stringency of the Likelihood Approach from beginning to end (which captures all information with respect to the spot), thus enables extremely low SNR spot detection and localization. The calculation of the False Positive Rate (FPR) and Multi Emitter and Multi Spectral Generalization are also discussed.

6.2 Background

Localization is defined as the fitting procedure to acquire sub-pixel localization. Detection is defined as spot discrimination between when a spot exists and when it doesn't.

[3] *Within the framework of a statistical model, all of the information which the data provide concerning the relative merits of two hypotheses is contained in the likelihood ratio of those hypotheses.* (Edwards 1972, 1992 p. 30)

Benefits of the Likelihood Approach for 2D spot localization is discussed in (Mortensen et al., 2010) and (Smith et al., 2010b). 2D spot localization with sCMOS cameras, in which there is pixel dependent noise heterogeneity discussed in (Huang et al., 2013b). Limitations of current 2D spot detection discussed in a survey of most 2D spot detection algorithms, (Sage et al., 2015b). In summary, limitations stem from the fact that there exists many ad-hoc computations in the state of-the-art spot detection softwares, with a large variety of user tuneable parameters that are difficult to understand and tune. The cause of this problem stems from application specific problems of handling interfering background signals obscuring the spot. Currently, all known Likelihood based spot detection software deploy Likelihood methods at the end of the computational pipeline.

Discussion below applies to ND pattern detection, with fluorescent spot detection and localization a 3D instance of an ND pattern, and multi spectral fluorescent spot detection and localization a 4D instance of an ND pattern.

6.3 The Noise Model, Intensity Hypothesis and the Likelihood

The fully detailed noise model typical in imaging is composed of:

Gaussian read noise per pixel i.

$$P(X_i = x_i|\mu_i, \sigma_i^2) = \frac{1}{\sigma_i\sqrt{2\pi}} e^{\frac{-(x_i-\mu_i)^2}{2\sigma_i^2}}$$

Poisson shot Noise per pixel i.

$$P(Y_i = y_i|\lambda_i) = \frac{\lambda_i^{y_i}}{y_i!} e^{\lambda_i}$$

These distributions convolve to create the measurement noise, $Poisson{*}Gaussian$

per pixel i[4].

$$P_{full}(D_i = d_i|\mu_i, \sigma_i^2, \lambda_i) = \sum_{x_i=0}^{\infty} P(X_i = x_i|\mu_i, \sigma_i) \cdot P(Y_i = d_i - x_i|\lambda_i)$$
$$= (Gaussian(\mu_i, \sigma_i^2) * Poisson(\lambda_i))[d_i]$$

For a patch of pixels, you define an intensity *hypothesis* that defines the average intensity λ_i per pixel i. For example, the *signal hypothesis* contains the *signal* of interest f_i with amplitude A that is embedded in an interfering background g_i with amplitude B. The *null hypothesis* is when the amplitude of the *signal* is zero. pos_A or pos_B are the vectors that specify that position of the patterns f_i and g_i respectively.

$$signal \text{ hypothesis} = \lambda_i(A, B, \vec{pos}_A, \vec{pos}_B) = Af_i(\vec{pos}_A) + Bg_i(\vec{pos}_B)$$

$$null \text{ hypothesis} = \lambda_i(B, \vec{pos}_B) = Bg_i(\vec{pos}_B)$$

In the case of fluorescent spot detection, in which the spot is a diffraction limited signal, f_i is the point spread function and g_i is uniform.

$$signal \text{ hypothesis} = \lambda_i(A, B, \vec{pos}) = Af_i(\vec{pos}) + B$$

$$null \text{ hypothesis} = \lambda_i(B) = B$$

The Likelihood Function is proportional to the probability of observing the data given the *noise model* and intensity *hypothesis*[5]. So discrimination of whether or not a spot exists is akin to asking what is the Likelihood of different intensity

[4] described in Section 1.3

[5] k_j in the Likelihood function is a proportionality constant that is function of the given dataset. Here, it is a reminder that Likelihood inference is not Bayesian – a system of inference that is strictly probability based.

hypotheses as manifested in a Likelihood Ratio between the *signal* vs *null*.

$$\ell(hypothesis|\vec{d_j})_{noiseModel} = k_j P_{noiseModel}(\vec{d_j}|hypothesis)$$

These Likelihood Functions contain parameters specific to each *hypothesis* and the optimal combination needs to be found before a Likelihood comparison. This called a Maximum Likelihood Estimate (MLE). For example, for the case of the diffraction limited *signal hypothesis*, the Likelihood function requires the MLE of $\{A, B, p\vec{o}s\}$ to be found.

$$\begin{aligned}\ell(A, B, p\vec{o}s|\vec{d_j})_{noiseModel} &= k_j P_{noiseModel}(\vec{d_j}|\vec{\lambda}(A, B, p\vec{o}s)) \\ &= k_j P_{noiseModel}([d_1, d_2, d_3, ..., d_i, ...d_n]_j|\vec{\lambda}(A, B, p\vec{o}s)) \\ &= k_j \prod_{i=1}^{n} P_{noiseModel}(d_i|\lambda_i(A, B, p\vec{o}s))\end{aligned}$$

6.4 Derivation of Stage I

Problems: 1) Fluorescent spots typically sparsely populate a data set, so we would like to limit the computationally expensive iterative approach to be limited to that sparse subset. 2) We need estimates of the initial MLE conditions for iterative solving as close to the final solution as possible. This ensures robust convergence at low SNR.

Answer: Stage I is an approximated version of the Likelihood Approach that accounts for sCMOS pixel dependent noise. It is the most selective and specific candidate selection filter, while also providing the required initial conditions for robust iterative processing at low SNR regimes.

As noted by a previous publication describing sCMOS 2D spot localization, (Huang et al., 2013b), the $Poisson * Gaussian$ Distribution can be approximated by a $Poisson * Poisson$ Distribution with a change of variables. This $Poisson * Poisson$ approximation of the original $Poisson * Gaussian$ Distribution is very good, therefore both versions are called the fully detailed noise model.

$$\begin{aligned}\ell(hypothesis|\vec{d_j})_{full} &= k_j \prod_{i=1}^{n} P_{full}(d_i|\sigma_i, \lambda_i(A, B, p\vec{o}s)) \\ &= k_j \prod_{i=1}^{n} (Gaussian(0, \sigma_i^2) * Poisson(\lambda_i(A, B, p\vec{o}s)))[d_i]\end{aligned}$$

$$\begin{aligned}P_{full}(d_i|\mu_i = 0, \sigma_i^2, \lambda_i) &= (Gaussian(0, \sigma_i^2) * Poisson(\lambda_i))[d_i] \\ P_{full}(d_i + \sigma_i^2|\mu_i = 0, \sigma_i^2, \lambda_i) &= (Gaussian(\sigma_i^2, \sigma_i^2) * Poisson(\lambda_i))[d_i] \\ &\approx (Poisson(\sigma_i^2) * Poisson(\lambda_i))[d_i] \\ P_{PoissPoiss}(d_i|\mu_i = 0, \sigma_i^2, \lambda_i) &= Poisson(\sigma_i^2 + \lambda_i)[d_i + \sigma_i^2]\end{aligned}$$

Let us start with the definition of the fully detailed Likelihood function based on a $Poisson * Poisson$ noise model:

$$\begin{aligned}\ell(hypothesis|\vec{d_j})_{PoissPoiss} &= k_j \prod_{i=1}^{n} (Poisson(\sigma_i^2 + \lambda_i)[d_i + \sigma_i^2]) \\ &= k_j \prod_{i=1}^{n} \frac{(\lambda_i + \sigma_i^2)^{d_i + \sigma_i^2}}{(d_i + \sigma_i^2)!} e^{-(\lambda_i + \sigma_i^2)}\end{aligned}$$

And then its corresponding Log Likelihood function:

$$\begin{aligned}\mathcal{L}(hypothesis|\vec{d}_j)_{PoissPoiss} =& Log(\ell_{PoissPoiss}) \\ =& \sum_{i=1}^{n} \log\left(\frac{1}{(d_i+\sigma_i^2)!}\right) \\ &+ \sum_{i=1}^{n}\left(d_i+\sigma_i^2\right)\log\left(\lambda_i+\sigma_i^2\right) - \sum_{i=1}^{n}\lambda_i - \sum_{i=1}^{n}\sigma_i^2 + log(k_j)\end{aligned}$$

To motivate the derivation of Stage I, we will start with the *signal hypothesis*, and the goal is to find the MLE of $\{A, B, p\vec{o}s\}$ for every position of the data set. For now, we will solve a subset of the parameters $\{A, B\}$ for a given patch $\vec{d}_j$ in the data set and a given pattern position. Later, we will generalize to solve for the MLEs of all parameters for the entire dataset.

$$\max_{\{A,B\}} \quad \mathcal{L}(signal|\vec{d}_j)_{PoissPoiss}$$

The MLE of $\{A, B\}$ is found by the following system of equations:

$$\begin{aligned}0 &= \frac{\partial \mathcal{L}_{PoissPoiss}}{\partial A} = \frac{\partial \mathcal{L}_{PoissPoiss}}{\partial \lambda_i}\frac{\partial \lambda_i}{\partial A} \\ 0 &= \frac{\partial \mathcal{L}_{PoissPoiss}}{\partial B} = \frac{\partial \mathcal{L}_{PoissPoiss}}{\partial \lambda_i}\frac{\partial \lambda_i}{\partial B}\end{aligned}$$

with

$$\frac{\partial \mathcal{L}_{PoissPoiss}}{\partial \lambda_i} = \sum_{i=1}^{n}\frac{d_i-\lambda_i}{\lambda_i+\sigma_i^2}$$

At low SNR, camera noise σ_i^2 dominates.

$$\left.\frac{\partial \mathcal{L}_{PoissPoiss}}{\partial \lambda_i}\right|_{approx} = \sum_{i=1}^{n}\frac{d_i-\lambda_i}{\lambda_i+\sigma_i^2} \approx \sum_{i=1}^{n}\frac{d_i-\lambda_i}{\cancel{\lambda_i}+\sigma_i^2}$$

Thus the above sequence of approximations: $Poisson * Poisson$ approximation,

fixing the pattern position, solving the MLEs of $\{A, B\}$ for a given patch $\vec{d_j}$ define the minimally invasive approximations to the fully detailed Likelihood method. Now the MLEs of $\{A, B\}$ for a given patch $\vec{d_j}$ by linear algebra.

$$0 = \frac{\partial \mathcal{L}_{PoissPoiss}}{\partial A} \approx \frac{\partial \mathcal{L}_{PoissPoiss}}{\partial \lambda_i}|_{approx} \frac{\partial \lambda_i}{\partial A} = \sum_{i=1}^{n} \frac{d_i - (A f_i(p\vec{o}s) + B)}{\sigma_i^2} f_i(p\vec{o}s)$$

$$0 = \frac{\partial \mathcal{L}_{PoissPoiss}}{\partial B} \approx \frac{\partial \mathcal{L}_{PoissPoiss}}{\partial \lambda_i}|_{approx} \frac{\partial \lambda_i}{\partial B} = \sum_{i=1}^{n} \frac{d_i - (A f_i(p\vec{o}s) + B)}{\sigma_i^2}$$

$$A \sum_{i=1}^{n} \frac{f_i(p\vec{o}s)^2}{\sigma_i^2} + B \sum_{i=1}^{n} \frac{f_i(p\vec{o}s)}{\sigma_i^2} - \sum_{i=1}^{n} \frac{f_i(p\vec{o}s) d_i}{\sigma_i^2} = 0$$

$$A \sum_{i=1}^{n} \frac{f_i(p\vec{o}s)}{\sigma_i^2} + B \sum_{i=1}^{n} \frac{1}{\sigma_i^2} - \sum_{i=1}^{n} \frac{d_i}{\sigma_i^2} = 0$$

The system of equations in matrix form.

$$\begin{bmatrix} \sum_{i=1}^{n} \frac{f_i(p\vec{o}s)^2}{\sigma_i^2} & \sum_{i=1}^{n} \frac{f_i(p\vec{o}s)}{\sigma_i^2} \\ \sum_{i=1}^{n} \frac{f_i(p\vec{o}s)}{\sigma_i^2} & \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} \frac{f_i(p\vec{o}s) d_i}{\sigma_i^2} \\ \sum_{i=1}^{n} \frac{d_i}{\sigma_i^2} \end{bmatrix}$$

To apply this MLE $\{A, B\}$ operation to every position in the data set, the $p\vec{o}s$ is left fixed, and the dot product between f_i and the patch of data is applied to every position in the data set via convolution. The result is that the approximate MLE of $\{A, B, x, y, z\}$ is known for every position in the data set. S now indexes over the patch of data at position i.

$$\begin{bmatrix} \sum_{k \in S} \frac{f_k^2}{\sigma_{i-k}^2} & \sum_{k \in S} \frac{f_k}{\sigma_{i-k}^2} \\ \sum_{k \in S} \frac{f_k}{\sigma_{i-k}^2} & \sum_{k \in S} \frac{1}{\sigma_{i-k}^2} \end{bmatrix} \begin{bmatrix} A_i \\ B_i \end{bmatrix} = \begin{bmatrix} \sum_{k \in S} f_k \frac{d_{i-k}}{\sigma_{i-k}^2} \\ \sum_{k \in S} \frac{d_{i-k}}{\sigma_{i-k}^2} \end{bmatrix}$$

For the general case in which the interfering signal pattern is g_k the matrix form is:

$$\begin{bmatrix} \sum_{k\in S} \frac{f_k^2}{\sigma_{i-k}^2} & \sum_{k\in S} \frac{f_k g_k}{\sigma_{i-k}^2} \\ \sum_{k\in S} \frac{f_k g_k}{\sigma_{i-k}^2} & \sum_{k\in S} \frac{g_k^2}{\sigma_{i-k}^2} \end{bmatrix} \begin{bmatrix} A_i \\ B_i \end{bmatrix} = \begin{bmatrix} \sum_{k\in S} f_k \frac{d_{i-k}}{\sigma_{i-k}^2} \\ \sum_{k\in S} g_k \frac{d_{i-k}}{\sigma_{i-k}^2} \end{bmatrix}$$

And for the generalization to $n-1$ interfering signals, or n overlapping spots, it is easy to see the general solution from with the three signal example below, in which A_i, B_i, C_i are the amplitudes of the three pattern of interest.

$$\begin{bmatrix} \sum_{k\in S} \frac{f_k^2}{\sigma_{i-k}^2} & \sum_{k\in S} \frac{f_k g_k}{\sigma_{i-k}^2} & \sum_{k\in S} \frac{f_k h_k}{\sigma_{i-k}^2} \\ \sum_{k\in S} \frac{f_k g_k}{\sigma_{i-k}^2} & \sum_{k\in S} \frac{g_k^2}{\sigma_{i-k}^2} & \sum_{k\in S} \frac{g_k h_k}{\sigma_{i-k}^2} \\ \sum_{k\in S} \frac{f_k h_k}{\sigma_{i-k}^2} & \sum_{k\in S} \frac{g_k h_k}{\sigma_{i-k}^2} & \sum_{k\in S} \frac{h_k^2}{\sigma_{i-k}^2} \end{bmatrix} \begin{bmatrix} A_i \\ B_i \\ C_i \end{bmatrix} = \begin{bmatrix} \sum_{k\in S} f_k \frac{d_{i-k}}{\sigma_{i-k}^2} \\ \sum_{k\in S} g_k \frac{d_{i-k}}{\sigma_{i-k}^2} \\ \sum_{k\in S} h_k \frac{d_{i-k}}{\sigma_{i-k}^2} \end{bmatrix}$$

Solving the system of equations for our original *signal hypothesis* results in a convolution based filter for the approximate MLE of $\{A, B\}$ for every position in the data set (hence optimal fits for the signal hypothesis is known for every position in the data set). For fluorescent imaging of diffraction limited structures (fluorescent spots), the pattern of interest is defined by the point spread function of the optics, $f_k = PSF(k)$, and the background pattern is uniform, $g_k = 1$.

$$A_i = \frac{\left(\sum_{k\in S} \frac{d_{i-k} g_k}{\sigma_{i-k}^2}\right) \sum_{k\in S} \frac{f_k g_k}{\sigma_{i-k}^2} - \left(\sum_{k\in S} \frac{d_{i-k} f_k}{\sigma_{i-k}^2}\right) \sum_{k\in S} \frac{g_k^2}{\sigma_{i-k}^2}}{\left(\sum_{k\in S} \frac{f_k g_k}{\sigma_{i-k}^2}\right)^2 - \left(\sum_{k\in S} \frac{f_k^2}{\sigma_{i-k}^2}\right) \sum_{k\in S} \frac{g_k^2}{\sigma_{i-k}^2}}$$

$$B_i = \frac{\left(\sum_{k\in S} \frac{f_k^2}{\sigma_{i-k}^2}\right) \sum_{k\in S} \frac{d_{i-k} g_k}{\sigma_{i-k}^2} - \left(\sum_{k\in S} \frac{d_{i-k} f_k}{\sigma_{i-k}^2}\right) \sum_{k\in S} \frac{f_k g_k}{\sigma_{i-k}^2}}{-\left(\sum_{k\in S} \frac{f_k g_k}{\sigma_{i-k}^2}\right)^2 + \left(\sum_{k\in S} \frac{f_k^2}{\sigma_{i-k}^2}\right) \sum_{k\in S} \frac{g_k^2}{\sigma_{i-k}^2}}$$

To recover the Likelihood Function from the approximation we integrate.

$$\mathcal{L}_{approx} = \int \frac{\partial \mathcal{L}_{PoissPoiss}}{\partial \lambda_i}\bigg|_{approx} \partial\lambda_i + C \approx -\sum_{i=1}^{n} \frac{(d_i - \lambda_i)^2}{\sigma_i^2} \equiv \text{W.L.S.}$$

The Likelihood function is Weighted Least Squares (W.L.S) weighed by pixel dependent read noise.

Since As and Bs are already found, $\lambda_i = Af_i + B$ can be plugged in to $\mathcal{L}_{approx}$ to derive the convolution based Log Likelihood Function of the *signal hypothesis*.

$$\begin{aligned}\mathcal{L}^{signal}_{approx,i} = &- A_i \left(A_i \sum_{k\in S} \frac{f_k^2}{\sigma^2_{i-k}} + 2B_i \sum_{k\in S} \frac{f_k}{\sigma^2_{i-k}} - 2\sum_{k\in S} \frac{f_k d_{i-k}}{\sigma^2_{i-k}} \right) \\ &+ B_i^2 \left(-\sum_{k\in S} \frac{1}{\sigma^2_{i-k}} \right) + 2B_i \left(\sum_{k\in S} \frac{d_{i-k}}{\sigma^2_{i-k}} \right) - \sum_{k\in S} \frac{d^2_{i-k}}{\sigma(u-x)^2}\end{aligned}$$

Similarly, the *null hypothesis* can be tackled the same way. This is the convolution filter for the MLE $\{B\}$ of the *null hypothesis*

$$B_i = \frac{\sum_{k\in S} \frac{d_{i-k}}{\sigma^2_{i-k}}}{\sum_{k\in S} \frac{1}{\sigma^2_{i-k}}}$$

This is the convolution based Log Likelihood Function of the *null hypothesis*.

$$\mathcal{L}^{null}_{approx,i} = B_i^2 \left(-\sum_{k\in S} \frac{1}{\sigma^2_{i-k}} \right) + 2B_i \left(\sum_{k\in S} \frac{d_{i-k}}{\sigma^2_{i-k}} \right) - \sum_{k\in S} \frac{d^2_{i-k}}{\sigma^2_{i-k}}$$

The LLRatio is defined as

$$\mathcal{L}Ratio_{approx,i} = \mathcal{L}^{signal}_{approx,i} - \mathcal{L}^{null}_{approx,i}$$

Thus Stage I is a Likelihood derived and convolution based computational filter

that outputs the MLEs of both signal and null hypotheses and their LLRatio.

6.5 Derivation of the False Positive Rate

So far the user can threshold the Approximate LLRatio derived in Stage I to define candidates for stage II iterative processing. However, the threshold needed to define candidates shifts depending on the background conditions, see figure 11 on the following page. This means the LLRatio is not a universal threshold parameter. Additionally, the LLRatio is a cryptic parameter to set for the typical end user.

Converting LLRatio to a False Positive Rate solves both challenges, but to get there, currently people invoke Wilks Theorem, in which under a list of conditions, the LLRatio is asymptotically Chi-squared distributed. However, this distribution is misleading since we violate the assumptions of Wilks Theorem, namely 1) small sample size 2) parameters tested is not interior to the parameter set, in other words having the null hypothesis a zero amplitude version of the signal hypothesis violates assumption 2). This problem has been noted in astronomy by (Protassov et al., 2002b).

We take advantage of numerical simulation and measure the distribution of LLRatios where there is no signal by measuring the background LLRatio fluctuations. We measure this at different Amplitude and Background levels for a typical diffraction limited spot and fit a gamma PDF, figure 12 on the next page.

We plot the two parameters of the gamma PDF, the scale and the shape versus amplitude and background and see that there is a linear relationship between the scale and the intensity of the background levels.

Now we can predict the distribution of LLRatios given the background level, thus this sets the False Positive Rate (FPR).

$$FPR(\mathcal{L}Ratio_{approx,i}) = 1 - CDF(Gamma(0.5, 2.6B_i + 2))[\mathcal{L}Ratio_{approx,i}]$$

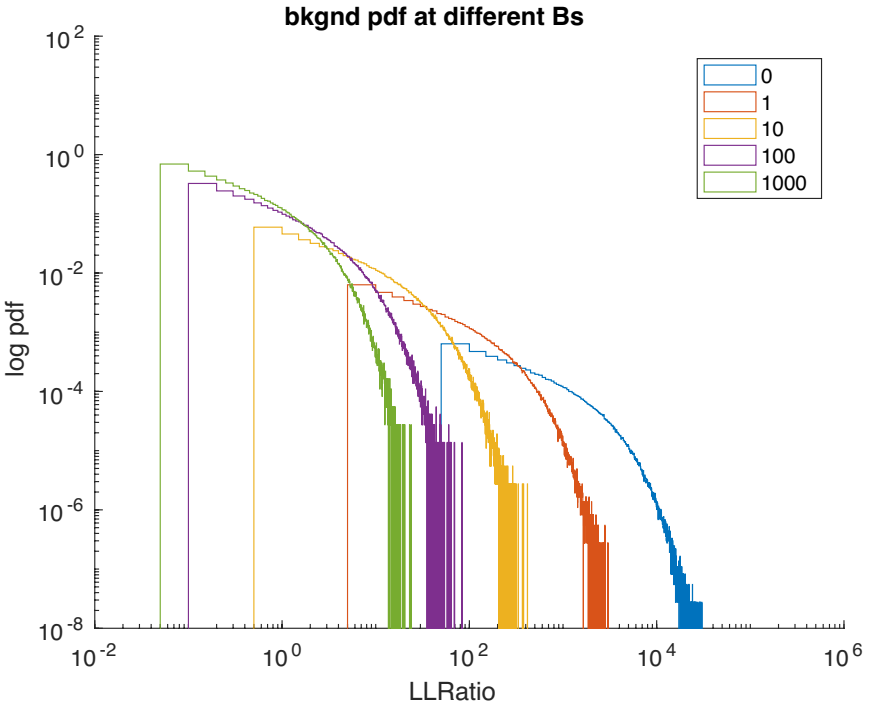

Figure 11: Approximate LLRatio values at different (B)ackground values

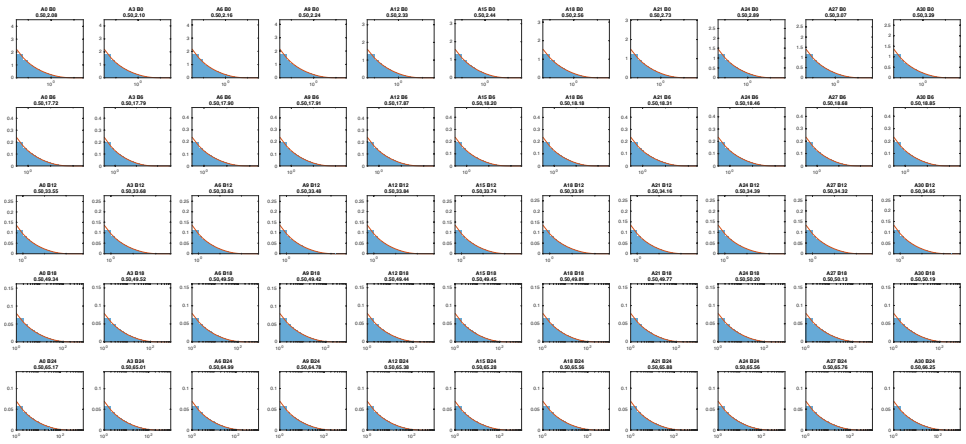

Figure 12: Gamma PDF fit to the histograms of background values at various (A)mplitude and (B)ackground Values from the simulations defined above

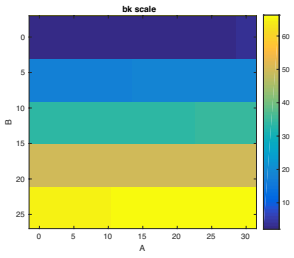

Figure 13: Scale parameter of the Gamma fits

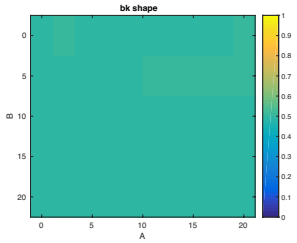

Figure 14: Shape parameter of the Gamma fits

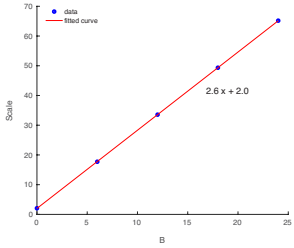

Figure 15: Linear fit of the Scale vs (B)ackground values

The FPR is now a universal threshold parameter that can be used in many different conditions. With such a nice fit, this hints at a closed form derivation for the theoretical distribution for the approximate LLRatio.

This gamma fitting can be done to the signal portion of the numerical simulation, figure 17 on the following page, with excellent fits of the scale and shape parameters versus the amplitude and background levels. We currently do not use this information, but this predicted LLRatio distribution of the signal component can predict the True Positive Rate, and possibly set the threshold for multi spot fitting.

The FPR values are then $-Log$ transformed in order for high intensities values to map to likely positions of the pattern of interest. In other words, low FPR indicates that that the LLRatio value observed at this position is likely free from spruious background contributions. This output dataset is labeled as $-Log(FPR)$. Note that other statistical metrics, such as the FNR, etc. could also have been used.

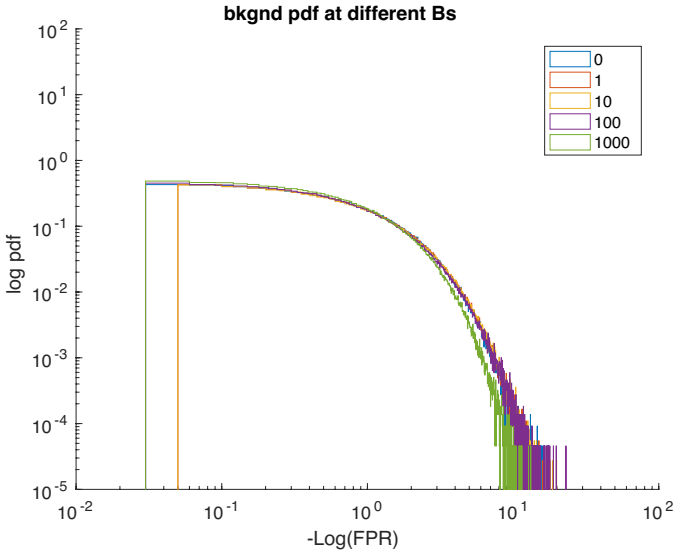

Figure 16: -Log(FPR) as calculated by the gamma pdf fit at different (B)ackground Values

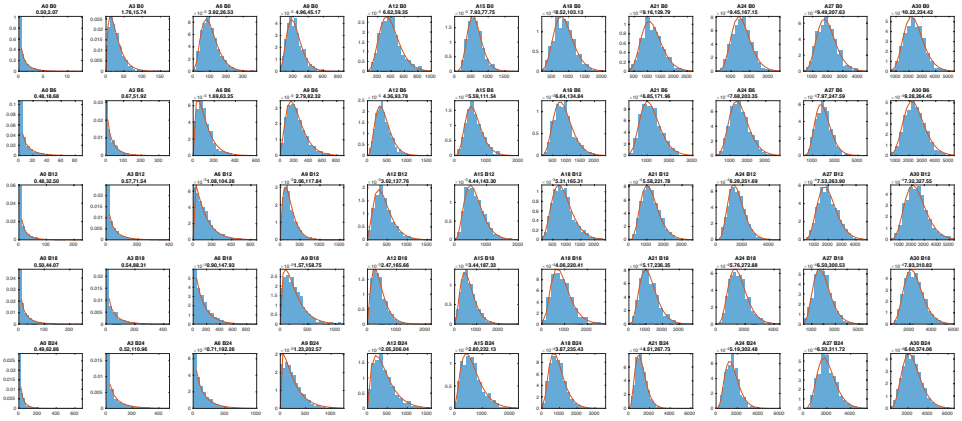

Figure 17: Gamma PDF fit to the histograms of signal values at various (A)mplitude and (B)ackground Values

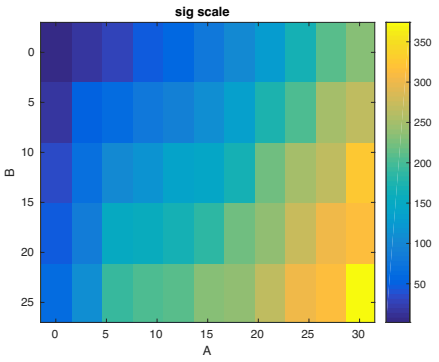

Figure 18: Scale parameter of the Gamma fits to the signal histograms

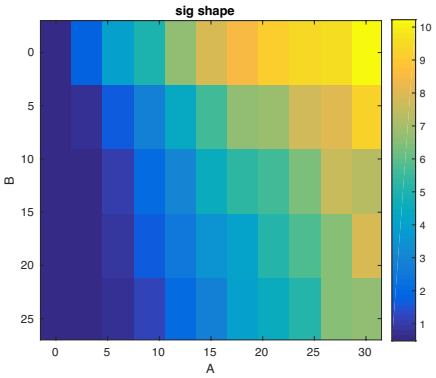

Figure 19: Shape parameter of the Gamma fits to the signal histograms

6.6 Derivation of the Multi Spectral Multi Emitter Stage I

For multi spectral spot detection, the signal to be detected is spread over the multi spectral data set, appropriately weighed by the spectral bleed thru coefficients[6]. To motivate the derivation, we will first do a two color case, green and red, which results in having to handle two data sets with two corresponding diffraction limited spots.

$$\mathbf{d} = \left\{ \vec{d}_G, \vec{d}_R \right\}$$

$$\mathbf{f} = \left\{ \vec{f}_G, \vec{f}_R \right\}$$

[6] All fluorescent spot detection software surveyed so far detect spots independently for each spectral data set, then take the intersection of each detected set. This results in a loss of sensitivity and selectivity

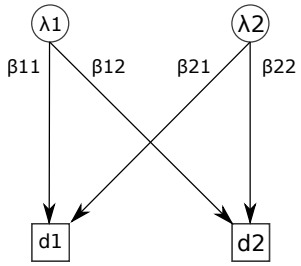

Figure 20: Two color intensity model, λ's represent the acquisition protocol for that color, β's represent the spectral bleed through, d's represent the dataset captured for that acquisition protocol

The two spectra signal hypothesis for two spectral data sets $\{\vec{d}_G, \vec{d}_R\}$:

$$
\begin{aligned}
\textit{two signal hypothesis} &= \left\{\Lambda_G, \Lambda_R\right\} \\
&= \left\{\beta_{11}\lambda_1 + \beta_{21}\lambda_2, \beta_{12}\lambda_1 + \beta_{22}\lambda_2\right\} \\
&= \Big\{\beta_{11}(A_G f_{G,i}(p\vec{o}s) + B_G) + \beta_{21}(A_R f_{R,i}(p\vec{o}s) + B_R), \\
&\qquad \beta_{12}(A_G f_{G,i}(p\vec{o}s) + B_G) + \beta_{22}(A_R f_{R,i}(p\vec{o}s) + B_R)\Big\}
\end{aligned}
$$

The Log Likelihood Function sums over the two spectral data sets, appropriately weighted:

$$
\begin{aligned}
&\mathcal{L}(\textit{two signal hypothesis}|\mathbf{d})_{PoissPoiss} = \\
&\quad \mathcal{L}(\beta_{11}(A_G f_{G,i}(p\vec{o}s) + B_G) + \beta_{21}(A_R f_{R,i}(p\vec{o}s) + B_R)|\vec{d}_G)_{PoissPoiss} \\
&+ \mathcal{L}(\beta_{12}(A_G f_{G,i}(p\vec{o}s) + B_G) + \beta_{22}(A_R f_{R,i}(p\vec{o}s) + B_R)|\vec{d}_R)_{PoissPoiss}
\end{aligned}
$$

Similar to stage I derivation above, we solve for the linear parameters $\{A, B\}$.

$$
\begin{aligned}
0 &= \frac{\partial \mathcal{L}_{PoissPoiss}}{\partial A_G} \\
0 &= \frac{\partial \mathcal{L}_{PoissPoiss}}{\partial A_R} \\
0 &= \frac{\partial \mathcal{L}_{PoissPoiss}}{\partial B_G} \\
0 &= \frac{\partial \mathcal{L}_{PoissPoiss}}{\partial B_R}
\end{aligned}
$$

With a a large amount of algebraic manipulation, this system of equations is equivalent to a matrix formulation with a general form of

$$\beta_{i,j}^T \mathbf{M} \beta_{i,j} = \beta_{i,j}^T \vec{b}$$

with $\mathbf{M}$ in block diagonal form and $\beta_{i,j}$ in block diagonal form (after some permutations). This matrix form of these equations provides the structure for the general n-color solution.

The $\beta_{i,j}^T \mathbf{M} \beta_{i,j} = \beta_{i,j}^T \vec{b}$ form of the two color case is shown below, with the $\beta_{i,j}^T$ terms canceled out[7].

[7]This conversion from algebra to matrix form is validated for a 2 color case using Mathematica in the document $2ColorMatrixFactor.nb$.

$$\begin{bmatrix} \beta_{11} & 0 & \beta_{21} & 0 \\ 0 & \beta_{11} & 0 & \beta_{21} \\ \beta_{12} & 0 & \beta_{22} & 0 \\ 0 & \beta_{12} & 0 & \beta_{22} \end{bmatrix}^T \begin{bmatrix} \sum_{k\in S} \frac{f_{G,k}^2}{\sigma_{i-k}^2} & \sum_{k\in S} \frac{f_{G,k}}{\sigma_{i-k}^2} & 0 & 0 \\ \sum_{k\in S} \frac{f_{G,k}}{\sigma_{i-k}^2} & \sum_{k\in S} \frac{1}{\sigma_{i-k}^2} & 0 & 0 \\ 0 & 0 & \sum_{k\in S} \frac{f_{R,k}^2}{\sigma_{i-k}^2} & \sum_{k\in S} \frac{f_{R,k}}{\sigma_{i-k}^2} \\ 0 & 0 & \sum_{k\in S} \frac{f_{R,k}}{\sigma_{i-k}^2} & \sum_{k\in S} \frac{1}{\sigma_{i-k}^2} \end{bmatrix}$$

$$\begin{bmatrix} \beta_{11} & 0 & \beta_{21} & 0 \\ 0 & \beta_{11} & 0 & \beta_{21} \\ \beta_{12} & 0 & \beta_{22} & 0 \\ 0 & \beta_{12} & 0 & \beta_{22} \end{bmatrix} \begin{bmatrix} A_{G,i} \\ B_{G,i} \\ A_{R,i} \\ B_{R,i} \end{bmatrix} =$$

$$\begin{bmatrix} \beta_{11} & 0 & \beta_{21} & 0 \\ 0 & \beta_{11} & 0 & \beta_{21} \\ \beta_{12} & 0 & \beta_{22} & 0 \\ 0 & \beta_{12} & 0 & \beta_{22} \end{bmatrix}^T \begin{bmatrix} \sum_{k\in S} f_{G,k} \frac{d_{G,i-k}}{\sigma_{i-k}^2} \\ \sum_{k\in S} \frac{d_{G,i-k}}{\sigma_{i-k}^2} \\ \sum_{k\in S} f_{R,k} \frac{d_{R,i-k}}{\sigma_{i-k}^2} \\ \sum_{k\in S} \frac{d_{R,i-k}}{\sigma_{i-k}^2} \end{bmatrix}$$

Canceling out the spectral bleed thru components we see the problem of multi spectral spot detection is composed of two steps.

$$\begin{bmatrix} \sum_{k\in S}\frac{f_{G,k}^2}{\sigma_{i-k}^2} & \sum_{k\in S}\frac{f_{G,k}}{\sigma_{i-k}^2} & 0 & 0 \\ \sum_{k\in S}\frac{f_{G,k}}{\sigma_{i-k}^2} & \sum_{k\in S}\frac{1}{\sigma_{i-k}^2} & 0 & 0 \\ 0 & 0 & \sum_{k\in S}\frac{f_{R,k}^2}{\sigma_{i-k}^2} & \sum_{k\in S}\frac{f_{R,k}}{\sigma_{i-k}^2} \\ 0 & 0 & \sum_{k\in S}\frac{f_{R,k}}{\sigma_{i-k}^2} & \sum_{k\in S}\frac{1}{\sigma_{i-k}^2} \end{bmatrix} \begin{bmatrix} \beta_{11} & 0 & \beta_{21} & 0 \\ 0 & \beta_{11} & 0 & \beta_{21} \\ \beta_{12} & 0 & \beta_{22} & 0 \\ 0 & \beta_{12} & 0 & \beta_{22} \end{bmatrix} \begin{bmatrix} A_{G,i} \\ B_{G,i} \\ A_{R,i} \\ B_{R,i} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{k\in S} f_{G,k}\frac{d_{G,i-k}}{\sigma_{i-k}^2} \\ \sum_{k\in S}\frac{d_{G,i-k}}{\sigma_{i-k}^2} \\ \sum_{k\in S} f_{R,k}\frac{d_{R,i-k}}{\sigma_{i-k}^2} \\ \sum_{k\in S}\frac{d_{R,i-k}}{\sigma_{i-k}^2} \end{bmatrix}$$

The First Step is to solve for the MLE of each channels $\{A, B\}$ without regard to the bleed thru coefficients since $\mathbf{M}$ is in block diagonal form, with each block equivalent to the stage I solution above.

$$\begin{bmatrix} \beta_{11} & 0 & \beta_{21} & 0 \\ 0 & \beta_{11} & 0 & \beta_{21} \\ \beta_{12} & 0 & \beta_{22} & 0 \\ 0 & \beta_{12} & 0 & \beta_{22} \end{bmatrix} \begin{bmatrix} A_{G,i} \\ B_{G,i} \\ A_{R,i} \\ B_{R,i} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{k\in S} \frac{f_{G,k}^2}{\sigma_{i-k}^2} & \sum_{k\in S} \frac{f_{G,k}}{\sigma_{i-k}^2} & 0 & 0 \\ \sum_{k\in S} \frac{f_{G,k}}{\sigma_{i-k}^2} & \sum_{k\in S} \frac{1}{\sigma_{i-k}^2} & 0 & 0 \\ 0 & 0 & \sum_{k\in S} \frac{f_{R,k}^2}{\sigma_{i-k}^2} & \sum_{k\in S} \frac{f_{R,k}}{\sigma_{i-k}^2} \\ 0 & 0 & \sum_{k\in S} \frac{f_{R,k}}{\sigma_{i-k}^2} & \sum_{k\in S} \frac{1}{\sigma_{i-k}^2} \end{bmatrix}^{-1} \begin{bmatrix} \sum_{k\in S} f_{G,k} \frac{d_{G,i-k}}{\sigma_{i-k}^2} \\ \sum_{k\in S} \frac{d_{G,i-k}}{\sigma_{i-k}^2} \\ \sum_{k\in S} f_{R,k} \frac{d_{R,i-k}}{\sigma_{i-k}^2} \\ \sum_{k\in S} \frac{d_{R,i-k}}{\sigma_{i-k}^2} \end{bmatrix}$$

$$\begin{bmatrix} \beta_{11} & 0 & \beta_{21} & 0 \\ 0 & \beta_{11} & 0 & \beta_{21} \\ \beta_{12} & 0 & \beta_{22} & 0 \\ 0 & \beta_{12} & 0 & \beta_{22} \end{bmatrix} \begin{bmatrix} A_{G,i} \\ B_{G,i} \\ A_{R,i} \\ B_{R,i} \end{bmatrix} = RHT \text{ (Right Hand Terms)}$$

Then the second step is to apply the inverse of the bleed thru coefficient to each respective $\{A, B\}$, which is equivalent to linear unmixing the As and then the Bs.

$$\begin{bmatrix} \beta_{11} & \beta_{21} & 0 & 0 \\ \beta_{12} & \beta_{22} & 0 & 0 \\ 0 & 0 & \beta_{11} & \beta_{21} \\ 0 & 0 & \beta_{12} & \beta_{22} \end{bmatrix} \begin{bmatrix} A_{G,i} \\ A_{R,i} \\ B_{G,i} \\ B_{R,i} \end{bmatrix} = \widehat{RHT}$$

$$\begin{bmatrix} A_{G,i} \\ A_{R,i} \\ B_{G,i} \\ B_{R,i} \end{bmatrix} = \begin{bmatrix} \beta_{11} & \beta_{21} & 0 & 0 \\ \beta_{12} & \beta_{22} & 0 & 0 \\ 0 & 0 & \beta_{11} & \beta_{21} \\ 0 & 0 & \beta_{12} & \beta_{22} \end{bmatrix}^{-1} \widehat{RHT}$$

In summary, to solve for the approximate MLE of multi spectral $\{A, B\}$, 1) treat eat spectral dataset independently, and apply the convolution filters of stage I with its corresponding diffraction limited PSF. 2) Linear Unmix the A's and B's with the spectral bleed thru coefficients.

The approximate LLRatio for multi spectral data sets is found by noticing that the multi spectral signal hypothesis can be effectively a single signal

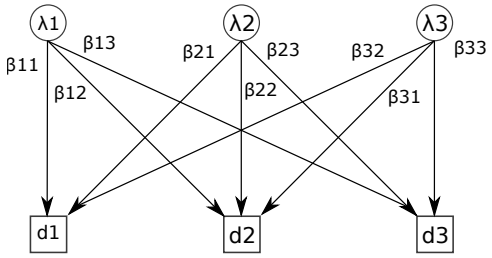

Figure 21: Three color intensity model, λ's represent the acquisition protocol for that color, β's represent the spectral bleed through, d's represent the dataset captured for that acquisition protocol

hypothesis for each channel.

$$
\begin{aligned}
&\mathcal{L}(\textit{two signal hypothesis}|\mathbf{d})_{PoissPoiss} = \\
&\quad \mathcal{L}(\beta_{11}(A_G f_{G,i}(p\vec{o}s) + B_G) + \beta_{21}(A_R f_{R,i}(p\vec{o}s) + B_R)|\vec{d}_G)_{PoissPoiss} \\
&\quad + \mathcal{L}(\beta_{12}(A_G f_{G,i}(p\vec{o}s) + B_G) + \beta_{22}(A_R f_{R,i}(p\vec{o}s) + B_R)|\vec{d}_R)_{PoissPoiss} \\
&= \quad \mathcal{L}(\widehat{A}_G f_{G,i}(p\vec{o}s) + \widehat{B}_G|\vec{d}_G)_{PoissPoiss} \\
&\quad + \mathcal{L}(\widehat{A}_R f_{R,i}(p\vec{o}s) + \widehat{R}_G|\vec{d}_R)_{PoissPoiss}
\end{aligned}
$$

The $\{\widehat{A}, \widehat{B}\}$ are the As and Bs before linear unmixing was applied. Each term, e.g. $\mathcal{L}(\widehat{A}_G f_{G,i}(p\vec{o}s) + \widehat{B}_G|\vec{d}_G)_{PoissPoiss}$ is the Log Likelihood Ratio of that spectral dataset without the unmixing applied. So the total LLRatio for a multi spectral dataset is equal to applying the single spectra LLRatio filters defined above to each spectra then summing across the spectral datasets.

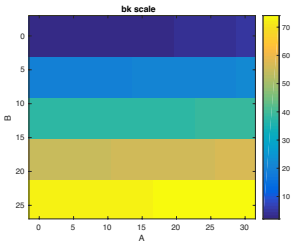

Figure 22: Scale parameter of the Gamma fits for 2 spectra LLRatio

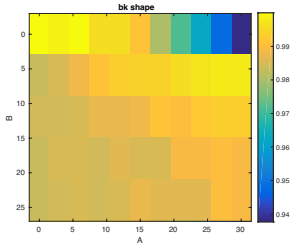

Figure 23: Shape parameter of the Gamma fits for 2 spectra LLRatio

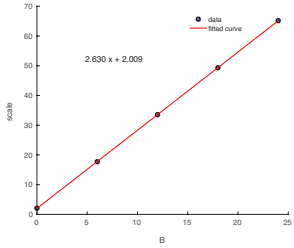

Figure 24: Linear fit of the Scale vs (B)ackground values

6.7 Calculating the Multi Spectral Multi Emitter FPR

Similar to the single channel FPR fits above, The two channel LLRatio histograms were fit to a gamma PDF to calculate the two channel FPR. The only thing that changes is the scale parameter changes from 0.5 to 1, which may reflect the summation of two LLRatios.

6.8 Candidate Selection

After Stage I, the input dataset is processed to produce an LLRatio (LLR) or $-Log(FPR)$ dataset, in which the intensities in either output relate to the fold probability increase the pattern of interest exists at a given position. The user may threshold at a given value, thus selecting a subset of the dataset as candidation regions to be further examined.

6.9 Derivation of Stage II

Stage II is a fully detailed MLE solver, thus it is iterative in nature and requires initial parameters θ that depend on the given intensity hypothesis. Since the intensity hypothesis can include multi spot models, stage II

can jointly solve for multi spot models[8]. For robust convergence at low SNR regimes, stage II first begins with Gradient Ascent then finishes with Newton Raphson, which is spiritually similar to the Levenberg-Marquardt algorithm. This requires that the Gradient and Hessian (curvature) of the Log Likelihood with respect the parameters of the hypothesis be defined.

To motivate the Stage II algorithm, let us examine a two spectra case.

$$\mathcal{L}(\textit{two signal hypothesis}|\mathbf{d}) =$$
$$\mathcal{L}(\beta_{11}\lambda_G(A_G f_{G,i}(p\vec{os}_G) + B_G) + \beta_{21}\lambda_R(A_R f_{R,i}(p\vec{os}_R) + B_R)|\vec{d}_G)$$
$$+ \mathcal{L}(\beta_{12}\lambda_G(A_G f_{G,i}(p\vec{os}_G) + B_G) + \beta_{22}\lambda_R(A_R f_{R,i}(p\vec{os}_R) + B_R)|\vec{d}_R)$$

$$\frac{\partial \mathcal{L}}{\partial \theta_l} = \beta_i \sum \frac{\partial \mathcal{L}(\vec{d}_G)}{\partial \Lambda_G} \frac{\partial \Lambda_G}{\partial \theta_l} + \beta_j \sum \frac{\partial \mathcal{L}(\vec{d}_R)}{\partial \Lambda_R} \frac{\partial \Lambda_R}{\partial \theta_l}$$

The gradient of the Log Likelihood w.r.t. θ_l is composed of a weighted sum between the gradients found in both channels, and each gradient component is a dot product between $\frac{\partial \mathcal{L}}{\partial \lambda}$ and $\frac{\partial \lambda}{\partial \theta}$. $\frac{\partial \mathcal{L}}{\partial \lambda}$ is essentially the weighted error (weighed by the details of the noise model) between the data and the intensity hypothesis, and $\frac{\partial \lambda}{\partial \theta}$ is how the intensity hypothesis changes w.r.t θ_l.

[8]So far all programs surveyed do multi spot fitting by first fitting a spot, then subtracting the fit. This introduces biases as the fit is distorted by the presence of another spot. For multi spectral spot fitting, all programs surveyed so far fit a spot independently for each spectral dataset.

$$\frac{\partial^2 \mathcal{L}}{\partial \theta_l \partial \theta_m} = \beta_i \sum \Big(\frac{\partial^2 \mathcal{L}(\vec{d}_G)}{\partial \Lambda_G^2} \frac{\partial \Lambda_G}{\partial \theta_m} \frac{\partial \Lambda_G}{\partial \theta_l} + \frac{\partial \mathcal{L}(\vec{d}_G)}{\partial \Lambda_G} \frac{\partial^2 \Lambda_G}{\partial \theta_l \partial \theta_m} \Big)$$
$$+ \beta_j \sum \Big(\frac{\partial^2 \mathcal{L}(\vec{d}_R)}{\partial \Lambda_R^2} \frac{\partial \Lambda_R}{\partial \theta_m} \frac{\partial \Lambda_R}{\partial \theta_l} + \frac{\partial \mathcal{L}(\vec{d}_R)}{\partial \Lambda_R} \frac{\partial^2 \Lambda_R}{\partial \theta_l \partial \theta_m} \Big)$$

Similarly, the Hessian of the Log Likelihood is composed of a weighted sum between the Hessians found in both channels.

Therefore information from both channels are being used to define the Gradient and Hessian. This makes sense as information pertaining to the pattern/spot is also spread across both channels when there is spectral bleed thru.

For Stage II, this is the Gradient update to θ

$$\vec{\theta}_{next} = \vec{\theta}_{prev} + k \frac{\partial \vec{\mathcal{L}}}{\partial \theta} \Big|_{prev}$$

This is the Newton Raphson update to θ.

$$\vec{\theta}_{next} = \vec{\theta}_{prev} - \Big[\frac{\partial^2 \mathcal{L}}{\partial \theta_l \partial \theta_m} \Big]^{-1} \vec{\theta}_{prev}$$

6.10 Derivation for the recursive Multi Spectral Multi Emitter Stage II

Given a candidate region, multi spectral multi emitter fitting is accomplished in Stage II by recursively fitting intensity models that contain zero or more patterns of interest. To motivate the general structure of this portion of the algorithm, let us examine a two spectra case for fluorescent spot detection.

$$\begin{aligned} \textit{two signal hypothesis} &= \Big\{\Lambda_G, \Lambda_R\Big\} \\ &= \Big\{\beta_{11}\lambda_1 + \beta_{21}\lambda_2, \beta_{12}\lambda_1 + \beta_{22}\lambda_2\Big\} \end{aligned}$$

Given a *two signal hypothesis*, the task is to assume different intensity hypotheses for each spectral channel $\{\lambda_1, \lambda_2\}$. e.g. λ_1 describes the intensity hypothesis in spectral channel 1, and can consists of background only to one or more spots in that channel. λ_2 is similar. The criterion of selecting a given intensity hypothesis, no matter how complex, is via one metric, the LLR of the given signal hypothesis fit over that of the null only.

A candidate region is processed as thus:

1. Fit a background only model: in which $\{\lambda_1, \lambda_2\}$ equals a channel specific constant. This is the null intensity hypothesis of uniform background for each channel. Within the given candidate region that is located in the corresponding Stage I MLE outputs, there exists a particular maximum value of the LLR output and its corresponding po-

sition. The initial background parameter estimates are taken from this position of the Stage I MLE, in the B component of each spectral channel. Iterative MLE with a fully detailed noise model is then applied to this initial condition, with the goodness of fit given by the Log Likelihood of this intensity hypothesis (LL_0) with 0 signifying a 0 spot assumption.

2. Fitting a spot: i) with the chosen initial position dictated by the position of the maximum LLR (from stage I) in that candidate region. ii) with the chosen spectral component dictated by the channel containing the largest MLE stage I amplitude iii) with the initial amplitude of the spot set to this amplitude. Iterative MLE with a fully detailed noise model is then applied to this initial condition, with the goodness of fit given by the LL_1, with 1 signifying the 1 spot assumption. the MLR of this intensity hypothesis over that of background only is calculated as the MLR_1, with 1 signifying a 1 spot hypothesis.

3. Fitting the next spot: The hypothetical intensities that correspond to the MLE fit from the previous step is subtracted from the original data. This leaves a residual error that may or may not contain additional spots. Since this residual is also subject to similar noise as before the subtraction, Stage I is applied to this residual so as to optimally determine the position and the amplitude of the next spot to be fit. The steps outline in 2) are then applied to initialize the next spot to be fit. the MLR of this intensity hypothesis over that of background only is calculated as the MLR_2, with 2 signifying a 2 spot hypothesis.

4. Repeat 3) until a user defined stop condition, e.g. number of spots to be fit, or a MLR_n threshold reached

6.11 Deriving the Standard Error

The standard error in the parameter estimates given by MLE is related to the Expected Fisher Information Matrix, which in turn defines the Cramer-Rao lower bound on this error (Rao, 1945; Cramer, 1946).

The Fisher Information Matrix is defined as:

$$I(\boldsymbol{\theta}) = -E\left[\frac{\partial^2 \mathcal{L}}{\partial\theta_l \partial\theta_m}\right] = E\left[\frac{\partial \mathcal{L}}{\partial\theta_l}\frac{\partial \mathcal{L}}{\partial\theta_m}\right]$$

To motivate a general derivation of the Fisher Information for the 3D multi spectral case, let us examine a two spectra case. The Fisher Information is additive across the two spectral measurements since the *LogLikelihoods* are additive across these two measurements.

$$I(\boldsymbol{\theta}) = -E\left[\frac{\partial^2 \mathcal{L}(\vec{d_G})}{\partial\theta_l \partial\theta_m}\right] - E\left[\frac{\partial^2 \mathcal{L}(\vec{d_R})}{\partial\theta_l \partial\theta_m}\right]$$

Given the Poisson approximated form of the $\mathcal{L}$, this Expectation simplifies to

$$I(\boldsymbol{\theta}) = \sum_i \frac{1}{\Lambda_G}\frac{\partial \Lambda_G}{\partial\theta_l}\frac{\partial \Lambda_G}{\partial\theta_m} + \sum_i \frac{1}{\Lambda_R}\frac{\partial \Lambda_R}{\partial\theta_l}\frac{\partial \Lambda_R}{\partial\theta_m}$$

with the Λ's defining the intensity hypothesis for each channel, thus they

account for the spectral bleedthru components.

$$\begin{aligned}\Big\{\Lambda_G, \Lambda_R\Big\} &= \Big\{\beta_{11}\lambda_G + \beta_{21}\lambda_R, \beta_{12}\lambda_G + \beta_{22}\lambda_R\Big\} \\ &= \Big\{\beta_{11}(A_G f_{G,i}(\vec{pos}) + B_G) + \beta_{21}(A_R f_{R,i}(\vec{pos}) + B_R), \\ &\qquad \beta_{12}(A_G f_{G,i}(\vec{pos}) + B_G) + \beta_{22}(A_R f_{R,i}(\vec{pos}) + B_R)\Big\}\end{aligned}$$

Geometrically, the Fisher Information for our situation is defined as the dot product between the gradients of Λ and $\frac{1}{\Lambda}$ and can be evaluated at the MLE.

$$Cov(\boldsymbol{\theta}) = I(\boldsymbol{\theta})^{-1}$$

With the standard error of θ the square root of the diagonal of the covariance matrix.

In general, multi spectral measurements are the sum of the Fisher Information for each respective channel. Multi emitter estimates are accounted for in this calculation by modifying the λs for multi emitter versions.

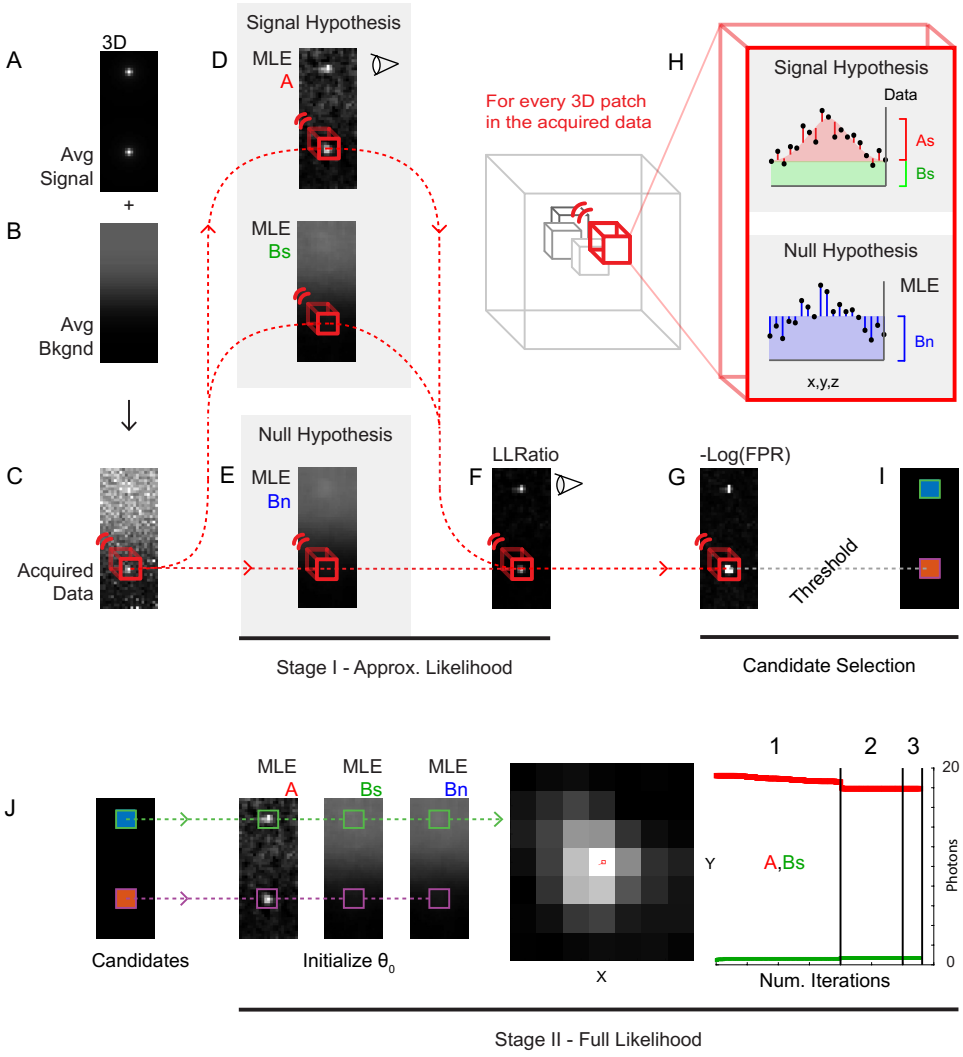

Figure 25: Schematic of the Two Stage Likelihood Pipeline (TSLP) for detecting and localizing a 3D Point Spread Function (PSF) in noisy datasets generated by fluorescent microscopy.

(Continued) 3D datasets are presented as 2D Max Z-Intensity Projections (MZP) and image contrast is quantitatively presented. Stage I is based on an approximate Likelihood model, whereas Stage II is fully detailed. For this example, the ground truth consists of A) Two 3D PSFs with equal intensities embedded in B) Background composed of an intensity gradient in 3D. C) An example noisy dataset sampled from the summed ground truths using an sCMOS camera model. H) Stage I is a computationally efficient filtering step, in which every patch of voxels in the noisy dataset, represented as a sliding red cube, have two operations carried out: D) First operation is to optimally fit the shape of the PSF to the intensities found in a patch by finding the best combination of the PSFs intensity (MLE As) and background level (MLE Bs), E) Second operation is to optimally fit the background component without the PSF (MLE Bi) to the intensities found in the same patch; F) Then, a comparison between the goodness of fit between the two operations is defined by the Log Likelihood Ratio (LLRatio). G) The False Positive Rate (FPR) is calculated from its corresponding LLRatio and Bs. I) The LLRatio or the FPR is thresholded to define the candidates. J) For each candidate, the initial conditions for Stage II are extracted from its corresponding MLE As, Bs,and Bi values. Joint multi spot fitting occurs in Stage II. Filtered datasets such as MLE As, LLRatio, and FPR can be used for denoised quantitative visualization.

7 3D Single Molecule Imaging in a Living System

7.1 MreB dynamics in B. *subtilus*

The difficulty in imaging 3D single molecule dynamics in living systems is compounded by several factors. One limitation is that the single molecule is attached to a single fluor, and this one fluor can only emit very little photons. The performance of any fluor is dominated by its photostability and its quantum yield (QY) with QY defined as the conversion efficiency between the number of photons emitted versus absorbed. Even with modern fluors that have been optimized to maximize QY and photostability, e.g. Janelia Fluors (Grimm et al., 2015), the rate limiting constraint is that there is only a single fluor (Dickson et al., 1997).

This problem of limited photon output is exacerbated by the fact that in living systems there is non-specific fluorescent background that is typically not uniformly distributed. This non-specific background could be from the specimen itself, e.g auto-fluorescence, or from non-specific labeling or spectral crosstalk from other fluorescent labels.

Since the TSLP algorithm is fundamentally a sensitive 3D-fluorescent spot detection and localization approach, optimized for low signal and high background conditions, I imagined that it could also facilitate in-vivo 3D single molecule measurements. In collaboration with the Garner Lab, we used this approach to image 3D single molecule dynamics of a prokaryotic cytoskeletal protein (MreB) that is essential for organizing non-spherical cell morphology. Up to now, taking these measurements has not been possible, even when deploying state-of-the-art 3D spot detection optics and algo-

rithms, i.e. Astigmatic or Double Helix PSF (Huang et al., 2008; Pavani et al., 2009).

Currently, these state-of-the-art 3D single molecule measurements are accomplished using engineered Point Spread Functions (PSFs). The PSF is how the optical apparatus manifest the image of small objects around the length scale of the diffraction limit, and PSF engineering modifies the native PSF to encode 3D information given a 2D measurement. Thus, 3D information is mapped to a distorted 2D PSF, with a systematic relationship between the type of distortion and the z position of that small object. Measurements of that small object are then taken with 2D datasets, in which the 3D position of that object can be inferred from the distortion. There are two dominant approaches currently in use, the Astigmatic and the Double Helix PSF approach. The reason for using engineering PSFs is that if measuring 2D datasets is already costly in terms of excitation energy, then taking 3D datasets, such as from the acquisition of multiple focal planes, will be prohibitive. This is a true sentiment if the TSLP algorithm did not exist.

The problem with PSF engineering is the loss in sensitivity of the microscope via two modes. The first mode is the loss in photon transmission efficiency of the imaging system since engineering the PSF requires the insertion of additional optical components in the emission pathway. The second mode is the loss in peak signal intensity of the distorted PSF. The native PSF resembles a 3D Gaussian in intensity, with a concentrated peak at the center, closely corresponding to the position of the imaged fluor. In contrast, all engineered PSFs distort this 3D Gaussian shape to encode z information.

Therefore, the concentrated peak is significantly lowered and distributed to the lateral areas. At low SNRs, the spreading of the peak signal results in loss of information to the read noise floor.

To image 3D single molecule trajectories of MreB in B. subtilis, we incubated MreB-Halo with J549 Halo fluor at pM concentrations[9]. Growth conditions was set to log phase. Cells were placed on a glass bottom dish with an agarose pad on top. Temperature was held at 30°C throughout the entire procedure.

An epifluorescent microscope with a 30°C chamber coupled to LED illumination was used. Mechanical z-stacks were taken, in which a 3D dataset is acquired by sequentially imaging different focal planes at different z positions.

Given just one initial measurement, unanticipated MreB dynamics were observed, including complex 3D trajectories that were missed by 2D methods and fixed observations. The next steps are to link this microscopic complexity to macroscopic morphology.

[9]These MreB experiments were done in collaboration with Yingjie Sun from the Garner Lab

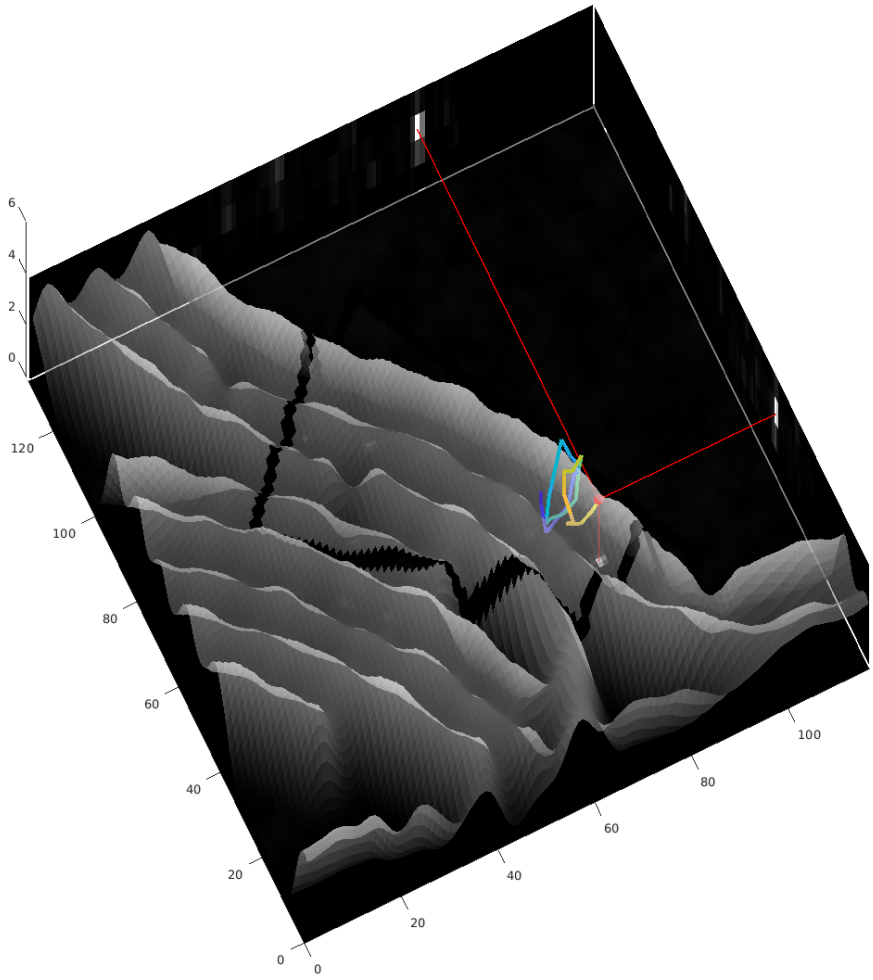

Figure 26: Imaging of a single molecule of MreB as it moves around the circumference of a cell of the bacterium B. *subtilis* (defined by phase image). Images collected every 5 sec for 2 min. Positions over time, as defined by full two-stage Likelihood Pipeline, coded from blue to yellow.

8 Dynamical Localization of a Thylakoid Membrane Binding Protein is Required for Acquisition of Photosynthetic Competency

8.1 Abstract

Vipp1 is highly conserved and essential for photosynthesis, but its function is unclear as it does not participate directly in light-dependent reactions. We analyzed Vipp1 localization in live cyanobacterial cells and show that Vipp1 is highly dynamic, continuously exchanging between a diffuse fraction that is uniformly distributed throughout the cell and a punctate fraction that is concentrated at high curvature regions of the thylakoid located at the cell periphery. Experimentally perturbing the spatial distribution of Vipp1 by relocalizing it to the nucleoid causes a severe growth defect during the transition from non-photosynthetic (*Dark*) to photosynthetic (light) growth. However, the same perturbation of Vipp1 in dark alone or light alone growth conditions causes no growth or thylakoid morphology defects. We propose that the punctuated dynamics of Vipp1 at the cell periphery in regions of high thylakoid curvature enable acquisition of photosynthetic competency, perhaps by facilitating biogenesis of photosynthetic complexes involved in light-dependent reactions of photosynthesis.

8.2 Introduction

Oxygenic photosynthesis is a metabolic process that must be dynamically modulated to accommodate varying light conditions and cellular needs (Eberhard et al., 2008). If extraneous light energy is not productively engaged in a timely manner, phototoxic damage can arise and lead to cellular death (Apel and Hirt, 2004). These light-dependent reactions are catalyzed by photosynthetic protein complexes that are assembled in a specialized membrane system which forms the thylakoid compartment in both chloroplasts and cyanobacterial cells (Rast et al., 2015). Although much is known about the structure and function of individual photosynthetic complexes, it is still unclear how the thylakoid membrane system is formed or maintained, and how the photosynthetic complexes are assembled in the thylakoid membrane (Pribil et al., 2014; Rast et al., 2015). These two issues are interrelated thylakoid structure depends on the presence of photosynthetic complexes (Pribil et al., 2014; Zhang et al., 2014) and assembly of photosynthetic complexes depends on the presence of thylakoid membrane (Yang et al., 2015; Nickelsen and Rengstl, 2013). Thus, experimentally uncoupling the factors required for thylakoid membrane formation from those required for photosynthesis is difficult.

Vipp1 (Vesicle-inducing protein in plastids 1), a conserved protein in plants, algae and cyanobacteria, has been proposed to play a role in thylakoid formation (Vothknecht et al., 2012; Heidrich et al., 2017; Kroll et al., 2001). Vipp1 is not known to participate in photosynthetic reactions and its deletion in plants causes loss of photosynthesis and thylakoid organiza-

tion, as well as loss of viability (Kroll et al., 2001; Westphal et al., 2001). The biophysical properties of Vipp1 are consistent with a membrane-related function as in vitro studies of Vipp1 revealed that it is a soluble protein with high affinity for lipid components (McDonald et al., 2015; Otters et al., 2013). On membrane surfaces prepared in vitro, recombinant Vipp1 forms higher order oligomers and, under certain conditions, mediates vesicle fusion (Hennig et al., 2015). Thus, Vipp1 may facilitate thylakoid membrane formation, with one recent model proposing that oligomeric Vipp1 is an intermembrane lipid transporter, possibly transporting lipids from the plasma (or chloroplast inner) membrane to the thylakoid membrane (Heidrich et al., 2017).

However, not all experiments are consistent with the above hypothesis. When expression of Vipp1 is reduced to a low level in cyanobacteria (Gao and Xu, 2009), photosynthetic output was severely affected but thylakoid morphology remained intact, suggesting that Vipp1 plays a role in photosynthesis in addition to, or instead of, being required for thylakoid formation. This observation, combined with the fact that deletion of a core photosynthetic complex induces a thylakoid morphology defect (Zhang et al., 2014), raises the possibility that Vipp1 may not directly facilitate thylakoid formation but instead its function may be related to photosynthesis. If this is true, knockdown or deletions of Vipp1 could generate a thylakoid morphology defect as a secondary consequence. Given this pleiotropy of vipp1 phenotypes and the contrasting conclusions reached by other studies (Aseeva et al., 2007; Fuhrmann et al., 2009b), separating the effects on thylakoid membrane formation from the effects on photosynthetic complexes formation is thought

to be hard, and potentially impossible (Heidrich et al., 2017).

Previous localization studies found that fluorescently labeled Vipp1 manifest as both diffuse and/or concentrated signals of various shapes that are largely immobile (Nordhues et al., 2012; Zhang et al., 2012; Bryan et al., 2014). In isolated chloroplasts exposed to low osmolarity stress conditions, the concentrated form of Vipp1 becomes more mobile which was suggested to be important for membrane protection (Zhang et al., 2012). In cyanobacteria exposed to damaging high light stress conditions, the diffuse form of Vipp1 was found to convert into long-lived peripheral foci at the plasma membrane region which was thought to be important for light-induced stress protection (Bryan et al., 2014). However, native dynamics of Vipp1 in normal growing conditions has not been observed or fully examined. Furthermore, without a perturbation that only targets Vipp1 localization, it is still unclear the importance of this dimension in the function of Vipp1.

In the present study we show that in normal living cells (non-stressful conditions), fluorescently labeled Vipp1 is highly dynamic, continuously exchanging between two fractions – a punctate fraction at the cell periphery that is concentrated at high curvature regions of the thylakoid, and a diffuse fraction that is uniformly distributed in the cytoplasm. Further, by rapidly perturbing the spatial distribution of Vipp1 in living cells, we show that native Vipp1 localization is not required for thylakoid membrane formation, but is essential during the transition from non-photosynthetic to photosynthetic metabolism. We propose that the punctate fraction of Vipp1 is the cytological manifestation of the oligomeric and membrane-bound form of Vipp1 whose role is to enable acquisition of photosynthetic competence. We

hypothesize that Vipp1 facilitates the functional assembly of photosynthetic complexes.

8.3 Results

8.3.1 Vipp1 forms transient puncta at regions of high thylakoid curvature at the cell periphery

To gain insight into the cellular role of Vipp1 we sought to investigate its localization and and dynamics in live cells. In low light growing conditions (8 μE light intensity) fluorescently tagged Vipp1 was previously found to localize predominantly as a diffuse signal in the cytoplasm, which upon exposure to stress via high light (600 μE light intensity) induces Vipp1 mobilization to the cell periphery where they form large puncta (Bryan et al., 2014). Since both of these light conditions are suboptimal for growth (Kopecna et al., 2012), we sought to investigate Vipp1 localization and dynamics in optimal log-growth conditions at moderate light intensity of 100 μE (see Methods).

To fluorescently label endogenous Vipp1 in *Synechocystis* sp. PCC6803 cells we integrated a Vipp1-mGFPmut3 fusion construct (hereafter Vipp1-GFP) at the native vipp1 locus via homologous recombination. Both the expression level of Vipp1-GFP protein and the bulk growth rate of the strain harboring the construct were similar to that of the wild type parental strain (Fig. 27A-B), as was reported previously (Bryan et al., 2014).

In actively growing cells Vipp1-GFP is present in two fractions: as fluorescent diffraction limited spots (hereafter puncta), and as a diffuse signal in the cytoplasm (Fig. 28A and Fig. 27C). Since Vipp1 puncta form when Vipp1 is fused to a known monomeric GFP variant that is least prone to induce multimerization of various bacterial targets (Landgraf et al., 2012) and also when fused to the SNAP-tag (Keppler et al., 2003) (Fig. 27D), which

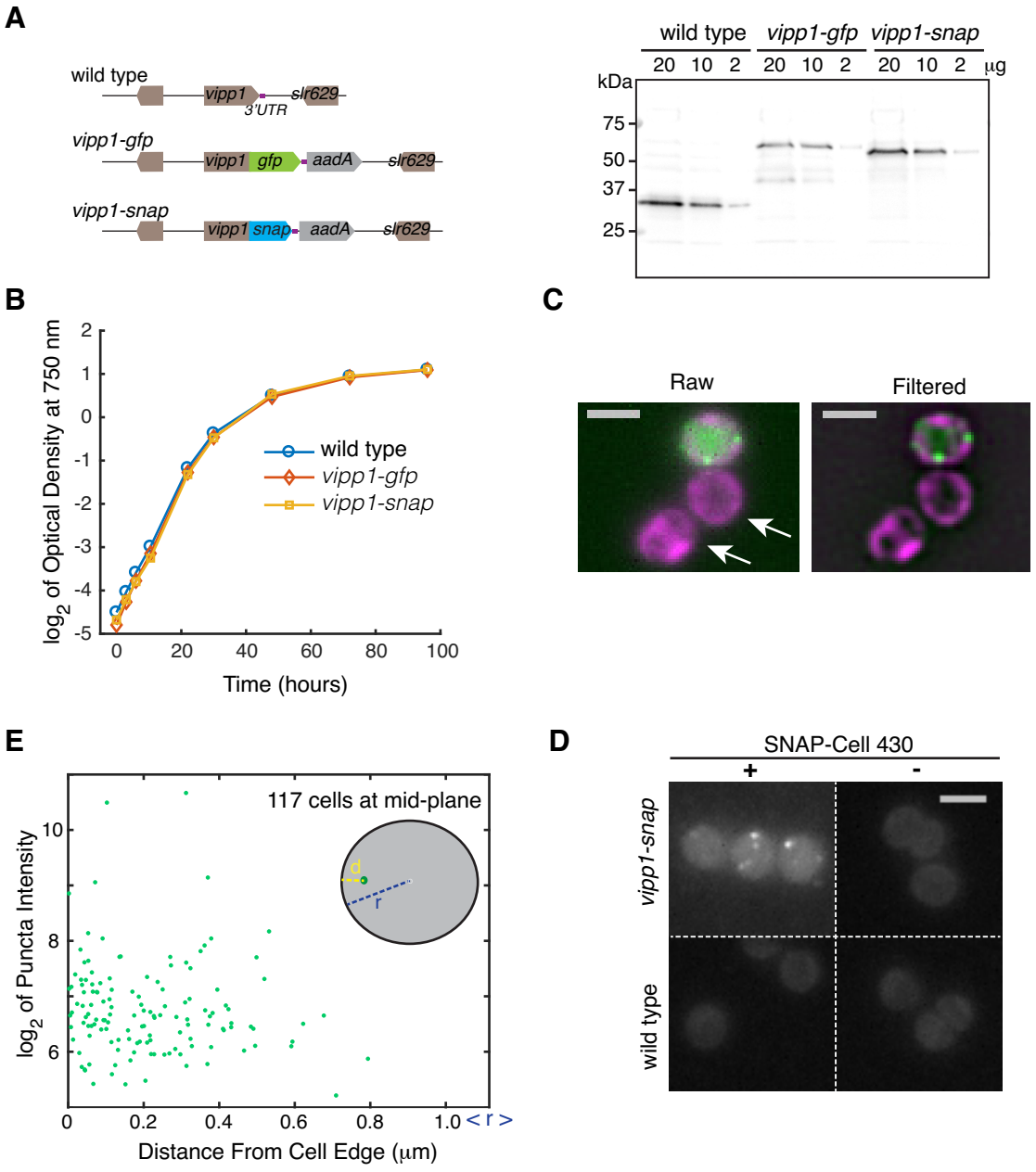

Figure 27: Characterization of Vipp1 expression and cellular localization in strains with epitope-tagged versions of Vipp1.

(Continued) A) Immunodetection of Vipp1 in cell lysates obtained from wild type, vipp1-gfp and vipp1-snap cells. Western blotting was performed as previously described (Gutu and OShea, 2013). Anti-Vipp1 rabbit polyclonal antibodies raised against Vipp1 from Chlamydomonas reinhardtii (Agrisera AS06145) were used. Left are the diagrams of the Vipp1-tagged DNA constructs used in this study (see Methods). B) Growth of wild type, vipp1-gfp and vipp1-snap strains in liquid culture bubbled with a mix of air and 1% CO2 at 100 μE intensity light. C) Merged Vipp1 (green) and thylakoid (magenta) images of Maximum Z-Intensity Projections of a mix of vipp1-gfp and wild type cells (arrows) showing the difference in cytosolic signal intensity (i.e. diffuse fraction of Vipp1) between the two cell types. On the right, an illustration of the filtering procedure (see Methods) used to enhance visualization of puncta and thylakoid features is shown. Scale Bar = 2 μm. D) Images of Maximum Z-Intensity Projections (raw epifluorescence) of wild type and vipp1-snap cells labeled with 5 μM SNAP-Cell 430 substrate (NEB S9109S) (after 20 min incubation in growth conditions) showing that Vipp1 puncta form independently of the GFP-tag. To facilitate comparison in signal intensity, the same grayscale contrast was applied to all four images. Scale bar = 2 μm. E) Intensities and distances of Vipp1 puncta relative to the cell edge. For any Vipp1 puncta found at or near the mid-Z plane of a cell, the intensity and the Euclidian distance (d) between its xy coordinates and the nearest point on the cell boundary edge obtained from segmentation in the brightfield channel was extracted and plotted. The $\langle$r$\rangle$on the x-axis is the average cell radii of the 117 analyzed cells.

is not known to induce artificial clustering, we conclude that the observed Vipp1 puncta formation and diffuse signal reflect native spatial localization.

Next we investigated the distribution and localization of Vipp1 puncta in cells growing on the microscope stage (see Methods). We find that the number of Vipp1 puncta per cell is well described by a Poisson distribution with a mean and variance of 1.36 (Fig. 28B), suggesting that the formation of each punctum is an independent event. By measuring the positioning of each punctum relative to the cell boundaries we find that most Vipp1 puncta localizes near the cell periphery (Fig. 28C and Fig. 27E). At this region, the thylakoids are highly abundant (Fig. 28C), as estimated by the fluorescence emitted by the endogenous photosynthetic proteins in the far-red portion of the visible spectrum (Vermaas et al., 2008).

Within the thylakoids, we observed that Vipp1 puncta tended to localize to regions where the thylakoid signal is low. These low thylakoid signal regions correspond to the edges of the thylakoid stacks and are known as zones of high thylakoid membrane curvature (Heinz et al., 2016). To confirm localization at sites of high thylakoid curvature, we asked whether Vipp1 puncta co-localize with CurT, a membrane protein enriched at these regions (Heinz et al., 2016). For this, we imaged at high resolution the simultaneous localization of Vipp1, CurT, and the thylakoid signal in living cells (see Methods). We find that in any given cell, Vipp1 puncta localize at the edge of thylakoid stacks where CurT is concentrated (Fig. 28 C-E). To quantify the relationship between Vipp1 puncta and CurT, we extracted and compared the intensity profiles of 76 arc lines that are each centered on a Vipp1 punctum found at the mid cell plane (Fig. 29A). We found

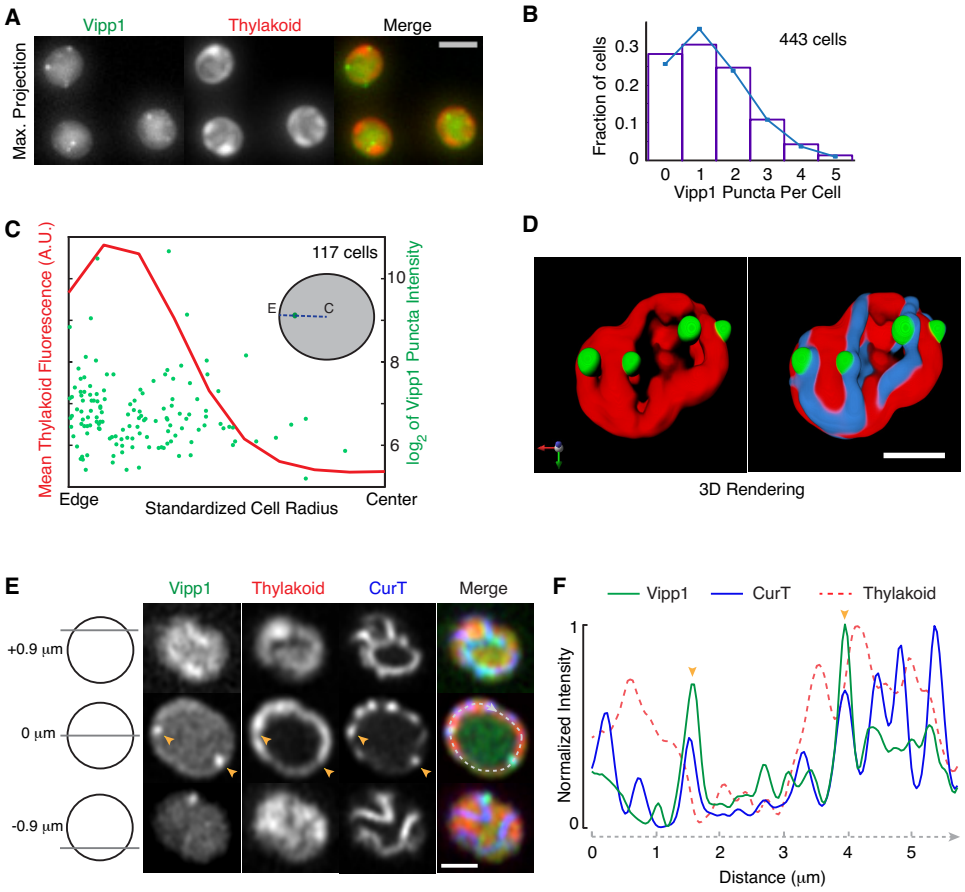

Figure 28: Vipp1 forms peripheral puncta that are located at regions of high thylakoid curvature. A) Representative live-cell epifluorescence images in the Vipp1 (GFP) and the thylakoid (far-red) channels. Images represent Maximum Z-Intensity Projections of 3D Z-stacks. Scale bar = 2 μm. B) Distribution of the number Vipp1 puncta per cell in a population of cells obtained from an exponentially growing culture. The mean and the variance of the distribution is 1.36. Puncta were identified by thresholding and segmenting filtered 3D Z-stacks.

(Continued) C) Analysis of the positioning of Vipp1 puncta relative to the cell radius. For each Vipp1 puncta identified at a given mid-cell plane (see inset diagram for an illustration), the corresponding distances to the cell center (labeled as C) and to the closest cell edge (labeled as E), as defined by brightfield segmentation masks, were obtained. The sum of the two distances defined the radius (blue dashed line) that was scaled from 0 to 1 on the x-axis. On the left y-axis, the mean of the thylakoid signal intensity along the line profiles (i.e. radii) connecting the cell centers, the Vipp1 puncta and the closest cell edge points is shown. The right y-axis shows the range of intensities (sizes) of all Vipp1 puncta analyzed. To limit the variability in size coming from dividing cells only cells whose area shape met an eccentricity value of 0.6 or less were analyzed.D) 3D rendering of a super-resolution image of a live cell of Synechocystis PCC6803 expressing Vipp1-GFP (green) and CurT-mTurquoise2 (blue). Thylakoids (red) are distributed at the cell periphery which by fluorescence microscopy show up as peripheral sheets. Vipp1 puncta are localized at the edge of thylakoid enrichments, at the same regions where the thylakoid membrane protein CurT is concentrated. Images were obtained with a laser scanning confocal system equipped with an Airyscan detector (Zeiss LSM880) which affords increased spatial resolution. Bar = 1 μm. For a full rendering of the same cell) Representative live-cell confocal fluorescence image of a cell expressing both Vipp1-GFP and CurT-CFP obtained by Airyscan imaging. As diagramed on the left, the rows show Vipp1, thylakoid and CurT channels at the top, middle and the bottom of a cell. Two Vipp1 puncta are shown at mid-cell slice (orange arrowheads) which co-localize with CurT enrichments and with gaps in the thylakoid signal. A profile line (dashed in light grey) running circumferentially through the peripheral thylakoids was used to extract the intensities of Vipp1, CurT and thylakoid fluorescence and plotted in panel F. Bar = 1 μm. F) Intensities of the Vipp1, CurT and thylakoid signals along the mid-cell circumferential curved line profile traced in panel E. Raw signal intensities were normalized from 0 (minimum) to 1 (maximum) on the y-axis. Orange arrowheads indicate colocalization of Vipp1 puncta with CurT, which is enriched at regions of high thylakoid signal changes (edges).

that in general, Vipp1 signal positively correlated with the CurT signal and negatively with the thylakoid signal (Fig. 29A). We also used automatic 3D object-based colocalization analysis to evaluate the spatial relationship between Vipp1 puncta with CurT enrichments and find that the majority of puncta overlap with the volume of CurT objects (Fig. 29B).

To better localize Vipp1 at the cell periphery and determine its relation to the thylakoid membranes we used immunoelectron microscopy. For immunodetection we used anti-GFP and gold-conjugated secondary antibodies to stain ultrasections (60 nm thickness) of freeze-substituted cells obtained from vipp1-gfp and wild type strains. We find that immunogold signals specific for Vipp1-GFP were enriched near the edges of thylakoids (high curvature regions which appear as tips in 2D representations) which typically converge near the plasma membrane (Fig. 30). Even though the antibodies can only bind the antigens from the surface of ultrasections, thus only probing a sparse subset of Vipp1-GFP, occasional clusters of 2-3 nanogold signals were observed near the thylakoid edges (Fig. 30C. This distribution is consistent with the idea that a fluorescent Vipp1 punctum consists of multiple Vipp1-GFP molecules closely juxtaposed in space. Based on these fluorescent and electron microscopy observations we conclude that the regions of high thylakoid membrane curvature are the sites of Vipp1 puncta localization.

At a given snapshot in time, many cells appear to contain no Vipp1 puncta, raising the question of whether only a fraction of cells form Vipp1 puncta or if Vipp1 puncta are relatively short lived. We used time-lapse fluorescent microscopy to investigate Vipp1 dynamics in cells growing pho-

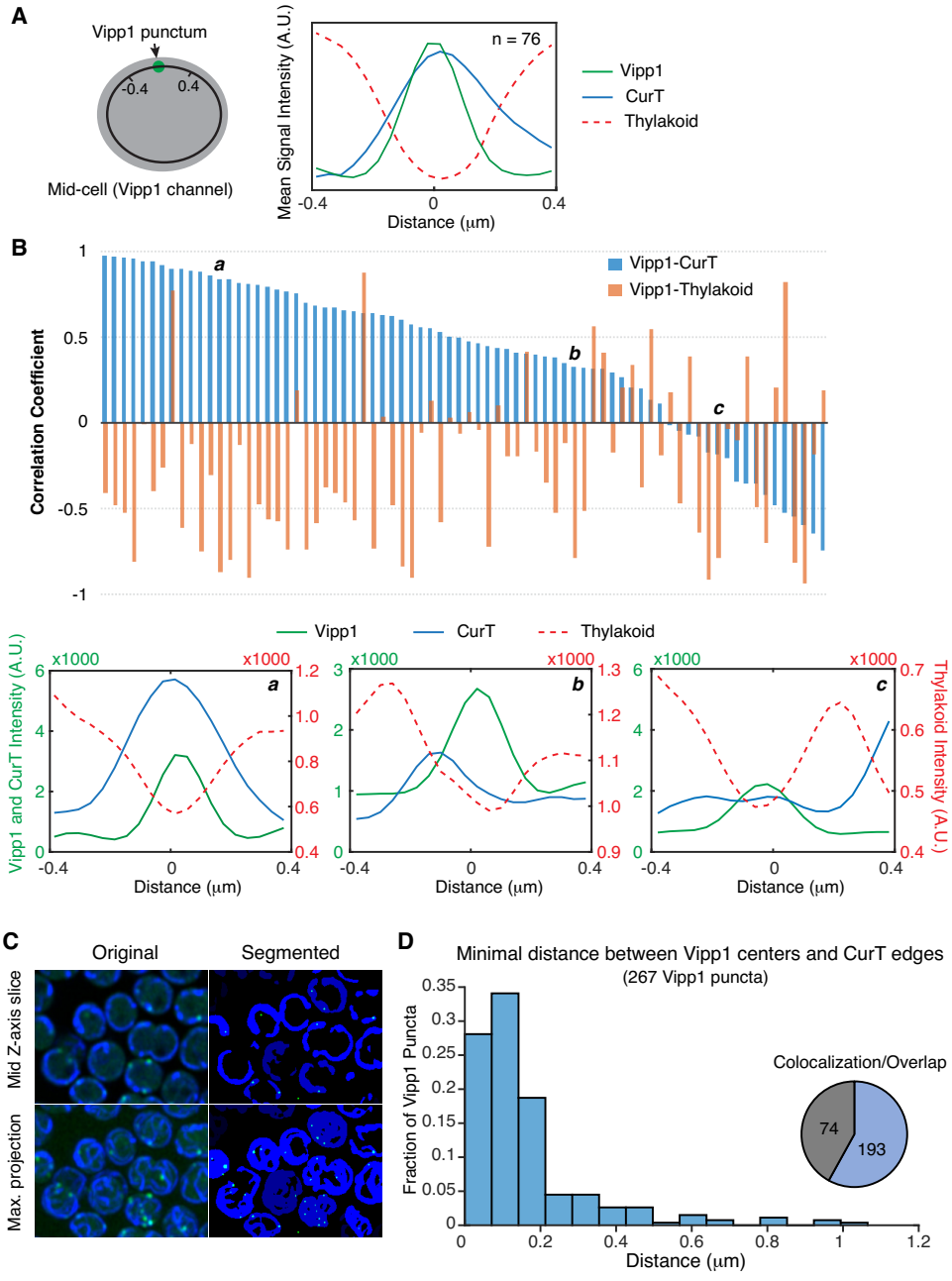

Figure 29: Vipp1 puncta colocalize with CurT enrichments at thylakoid edges.

(Continued) A) Top - examples of circumferential mid-cell intensity profiles in the Vipp1, CurT and thylakoid channels. The arc lines were manually traced parallel to the cell boundaries, through any given Vipp1 puncta situated at mid-cell (similarly to Figure 1E-F). 3D stacks of live-cells were imaged in the Vipp1, CurT and thylakoid channels on a confocal system equipped with the Airyscan detector (Zeiss LSM 880) which enhances the spatial resolution by a factor of 1.7x (Korobchevskaya et al., 2017). For any given Vipp1 puncta identified as a peak in the profile, the intensity of the CurT and thylakoid signals were also collected for a 0.4 μm arc segment on either side of the Vipp1 peak. 76 Vipp1 puncta (identified as peaks in 2D profiles - see Methods) obtained from 38 independent cell profiles were analysed. Bottom - Pearsons linear correlation between Vipp1 and CurT, and Vipp1 and the thylakoid signals were calculated for each of the 76 datasets (plotted as bar graph below) revealing mostly a positive relationship between Vipp1 and CurT and a negative one between Vipp1 and thylakoid signals (aggregated p-values: 1.22*10-6 and 5.57x10-62 respectively, null hypothesis tested being there is no relationship between the observed measurements). The correlation values of the examples shown on top are labeled with lowercase a, b and c above their corresponding bars. The weaker correlation values are obtained when the long axes of the thylakoid edges are quasi-parallel to the intensity profiles. B) Object-based colocalization analysis of Vipp1 puncta with the CurT enrichments. Left - example of Airyscan merged images of Vipp1 (green) and CurT (blue) channels and the corresponding segmentation masks obtained by iterative thresholding of the 3D stacks as implemented in the 3D ImageJ Suite plugin (Ollion et al., 2013). Both mid Z-slice and Maximum Intensity Projection along the Z axis are shown. Right - histogram of the distance distribution of each Vipp1 puncta to its most adjacent CurT object (Vipp1 puncta centers to CurT object edges). 267 Vipp1 puncta obtained from an image stack containing 70 cells were analysed. Note that this analysis includes all distances regardless if the Vipp1 centers are inside or outside of the CurT objects . Majority of the distances between Vipp1 puncta and CurT edges fall under the optical resolution resolution limit afforded by the Airyscan confocal imaging system (140 nm lateral and 400 nm axial). Inset pie chart shows how many of the Vipp1 puncta volumes overlap with the CurT objects.

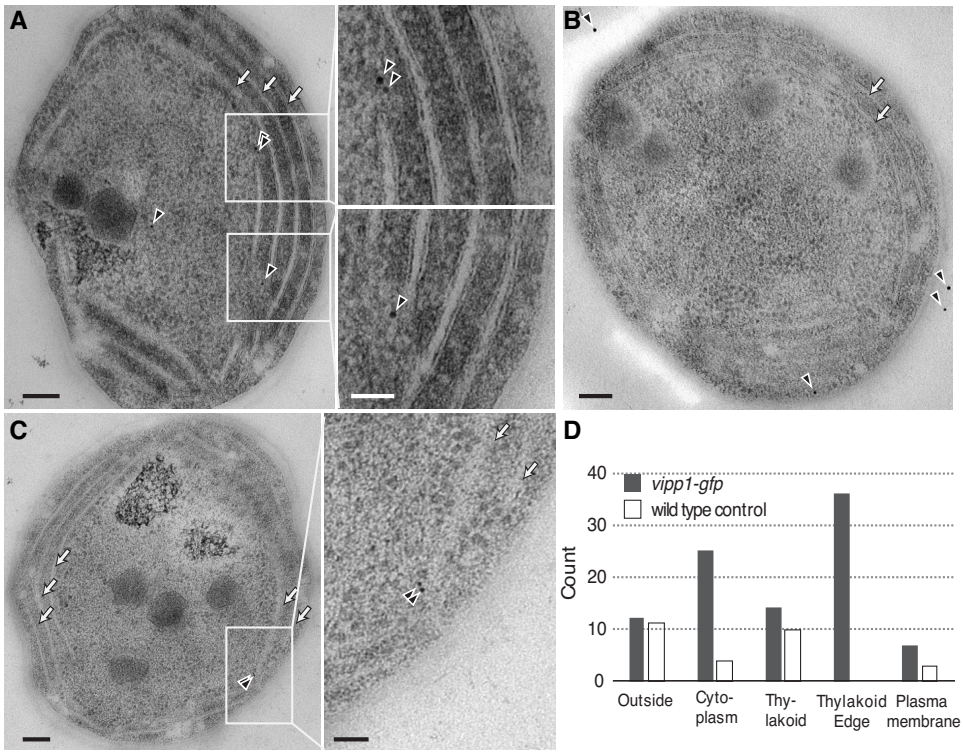

Figure 30: Detection and quantification of Vipp1-GFP by immuno-electron microscopy. Representative immunoelectron microscopy image of vipp1-gfp (A and C) or wild type cells (B) stained with anti-GFP primary antibodies detected by gold-conjugated secondary antibodies (black arrowheads). Arrows show the stacked thylakoids which run parallel to the plasma membrane. Scale bar = 100 nm (inset scale bar = 50 nm). C) Count of gold particles categorized by the proximity to the nearest cellular feature (plasma membrane, thylakoid edges, thylakoid (excluding edges), cell outside and the cytoplasm) in 35 independent whole cell sections obtained from the same immunodetection experiment on vipp1-gfp and wild type cells.

tosynthetically on the microscope stage and find that sparse Vipp1 puncta dynamically appear and disappear in all cells over time. Thus, this punctuated Vipp1 dynamics exists in all growing cells over time, explaining the zero puncta bin seen in the distribution of Vipp1 counts per cell (Fig. 28B).

To determine if Vipp1 puncta form on the membrane or if they form in the cytoplasm and diffuse to the membrane, we imaged photosynthetically growing cells at high temporal resolution to capture the appearance and disappearance as well as the mobility of Vipp1 puncta. We find that the majority of Vipp1 puncta rise and fall in intensity on a time scale of 1-2 min (Fig. 31A-B) and their mobility at the cell periphery is greatly limited (Fig. 31C). This constrained mobility and the rise and fall in fluorescence within a diffraction limited volume is consistent with a Vipp1 assembly process occurring on the membrane. Given that in vitro Vipp1 displays membrane binding property as a homo-oligomeric complex, and in cell lysates Vipp1 exists as a distribution of oligomeric states (Fuhrmann et al., 2009a; McDonald et al., 2015; Hennig et al., 2015; Heidrich et al., 2016), we infer that the dynamically forming Vipp1 puncta, occurring at or near the highly curved regions of the thylakoid compartment, represent events of punctuated oligomerization and de-oligomerization between its cytosolic and its membrane-bound forms.

8.3.2 Spatio-temporal distribution of Vipp1 enables photosynthetic competency

To test the functional importance of the spatial distribution and dynamics of Vipp1 in the cell, we designed a perturbation in which Vipp1 was rapidly and

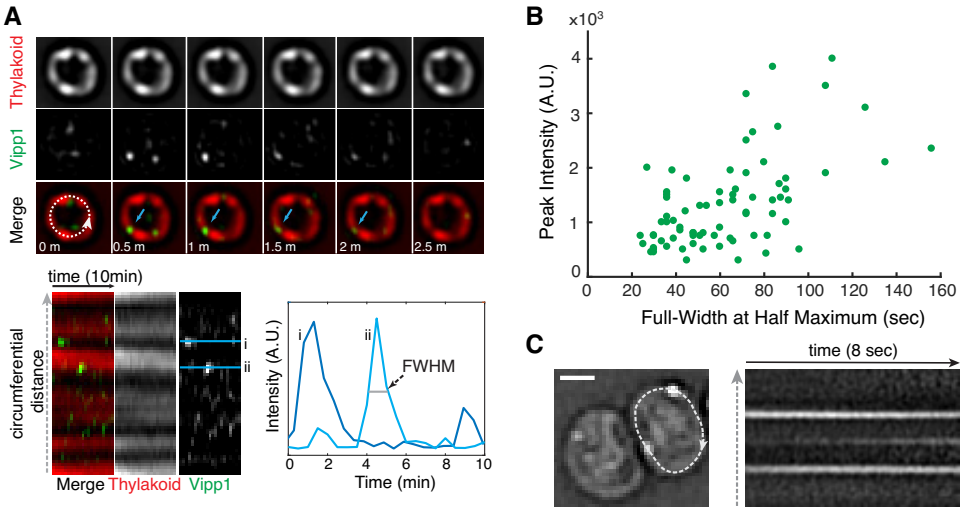

Figure 31: Vipp1 forms transient puncta of limited mobility at the cell periphery. A) Time-series montage of a live cell growing photosynthetically on the microscope stage. The images were acquired every 30 seconds in the Vipp1 and thylakoid channels for a duration of 10 minutes. The montage shows only the mid-cell Z plane for the first 2.5 minutes. The blue arrows highlight the appearance and disappearance of a Vipp1 puncta in consecutive time points. The dashed circumferential line was used to extract the intensity profiles of the two channels and build the kymogram displayed below the montage. The Vipp1 puncta appear and disappear at the edge the thylakoids which manifest as horizontal streaks. Two examples of Vipp1 intensity profiles in time (i and ii) are graphed on the right to illustrate how the full-width at half-maximum (FWHM) durations and peak intensities were obtained for a given punctum. B) Distribution of Full-Width at Half Maximum (FWHM) durations relative to their peak intensities for a representative subset of Vipp1 puncta. The values were extracted from the individual kymograms of 15 growing cells that were imaged every 30 seconds for a duration of 20 min. C) Example of Vipp1 puncta mobility recorded by continuous imaging (frame rate 81 milliseconds). On the left, the Laplacian-of-Gaussian filtered image of a cell containing two peripheral Vipp1 puncta is shown. A circumferential line profile traversing the cell periphery (dashed line) was used to extract the kymogram on the right to illustrate the constrained mobility of Vipp1 puncta over time. The kymogram also captures the birth of a new Vipp1 punctum as evidenced by the appearance of a middle horizontal streak over time. Scale bar =1 μm.

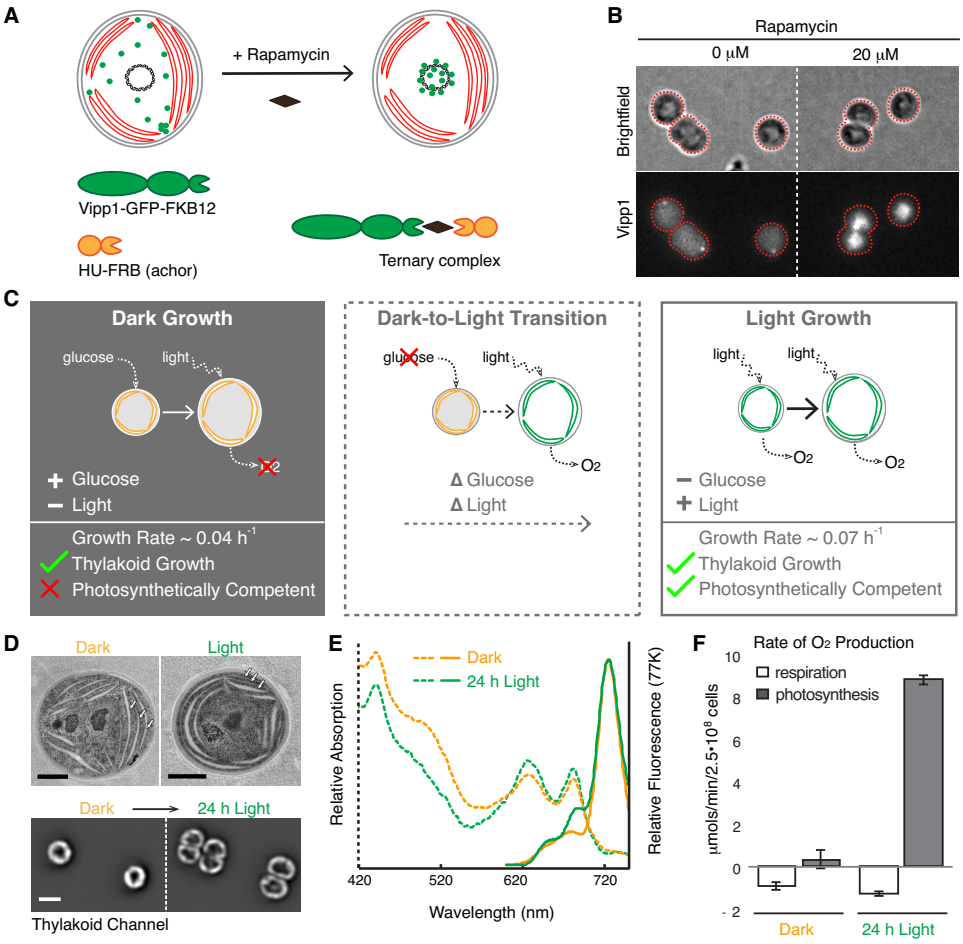

Figure 32: Perturbation of Vipp1 localization and characterization of the growth conditions in which the perturbation was applied.

(Continued) A) Diagram illustrating how Vipp1 can be relocalized to the nucleoid, away from the cell periphery using the anchor-away approach. A schematic cell is shown, with red disks representing thylakoids, the black circle in the center representing the nucleoid, and green dots representing the diffuse and punctate fractions of Vipp1. Human FKBP12 (12 kDa FK506-binding protein) was fused to Vipp1-GFP and the FRB (FKBP-rapamycin binding) domain was fused to the histone-like HU protein. In the presence of the cell permeable drug rapamycin, the Vipp1-GFP-FKBP12 chimera tethers to HU-FRB and forms a ternary complex enriched at the nucleoid due to the high binding affinity of HU to DNA. B) Rapamycin-dependent relocalization of Vipp1 to the nucleoid region at the cell center is effective and rapid. Raw images of Maximum Z-Intensity Projections in the brightfield and Vipp1 channels of exponentially growing cells expressing the Vipp1-GFP anchor-away system in the presence or absence (solvent only) of rapamycin after 20 min incubation. Nucleoid-localized Vipp1-GFP manifests as bright concentrations Vipp1-GFP signal at the cell centers. The grayscale contrast is the same for both conditions. Cell outlines based on the brightfield channel are overlaid in red. C) The three growth conditions used for testing the functional importance of native Vipp1 localization via anchor-away: *Dark Growth*, based on glucose, during which thylakoids are multiplied while photosynthetic competency is highly reduced; *Dark-to-Light Transition*, during which dark-grown cells are shifted to light in glucose-free media so that cells can regain their photosynthetic competency over time; and Light Growth, during which the cells grow in constant photosynthetic conditions. In each panel, a cell is shown growing from a smaller size to a larger size, and thylakoids are shown in orange if not competent for photosynthesis and in green if they do support photosynthesis. Photosynthetic competency is illustrated as the ability to produce O2 when light is applied. Approximate growth rates shown were estimated by monitoring the optical density of liquid cultures and are consistent with previously published measurements (Anderson and McIntosh, 1991). D) Thylakoids are present in both dark- and light-grown cells. Top - representative electron microscopy images of dark- and light grown cells showing a similar thylakoids arrangement consisting of several cell peripheral stacks of 2-4 thylakoid sheets (arrows). Scale bar = 0.5 μm. Bottom - thylakoid fluorescence images of two dark-grown cells before (left) and after 24 hours of growth in light (right) directly on the microscope stage in glucose-free media. In both conditions, the thylakoid signal is enriched at the cell periphery with occasional gaps in fluorescence intensity that correspond to the thylakoid stacks edges. The same grayscale contrast applied for both conditions. Only the mid-Z planes of filtered image stacks are shown. Scale bar = 2 μm. The dark-grown cells were obtained by growing a culture in darkness on glucose for 144 h (8 generations), as described in Methods.

(Continued) E) Both dark- and light-grown cells contain photosynthetic complexes and pigments. Whole-cell absorption (dashed lines) and low-temperature (77K) fluorescence emission (solid lines) spectra of dark-grown (144 h in darkness) and light-grown (24 h post-dark) cultures. The chlorophyll (445 and 680 nm) and the phycocyanin (625 nm) peaks in the absorption spectra are present in both growth conditions, albeit at slightly lower level in dark-grown culture. The fluorescence emission spectra (excitation at 445 nm) reveal the signature peak of the photosystem I at 725 nm present in both cultures, and the peaks associated with active photosystem II (685 nm and 690 nm) which appear conspicuous in the light-grown culture only, as was reported previously (Barthel et al., 2013). F) Dark-grown cells are photosynthetically incompetent. Rates of whole-cell oxygen production of dark- and light-grown cultures. Light-grown cells were obtained by shifting the dark-grown cells to light for 24 h in glucose-free medium. Linear rates of oxygen production were obtained from cells maintained in darkness (respiration, i.e. oxygen consumption) or in saturating light (photosynthesis, i.e. oxygen evolution) for the duration of the measurement. Respiratory activity which originates from the protein complexes that also localize in thylakoid membranes is similar between the dark- and light-grown samples. Error bars are standard error of the mean obtained from three biological replicates.

reversibly localized to the cell center, away from the cell periphery, without changing its overall concentration. To achieve this perturbation we used the anchor-away technique (Liberles et al., 1997; Haruki et al., 2008) in which a protein of interest can be rapidly moved away from its native location by inducing a drug-dependent dimerization to another protein, known as the anchor, that is in a different sub-cellular location (Fig. 32A). For the anchor we used the histone-like HU protein that binds to DNA in the nucleoid and is localized at the center of the cell (Wang et al., 2011). In the presence of the drug rapamycin, cells expressing the components of the anchor-away system rapidly alter native localization of Vipp1 by coalescing the entire pool of Vipp1 to the nucleoid region of the cell, away from the cell periphery (Fig. 32B, Fig. 33). We note that this perturbation of Vipp1 localization inherently alters the dynamics of puncta formation occurring at the cell

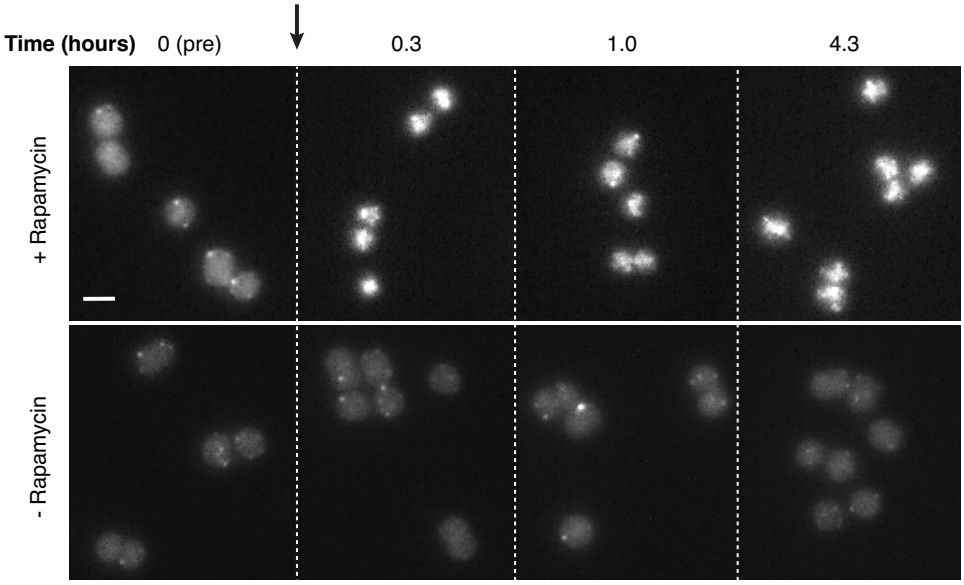

Figure 33: Time-course of rapamycin-induced Vipp1 relocalization. Maximum Z-Intensity Projection images of cells expressing all the components of the anchor-away system and taken from a log-phase culture bubbled with 1% air-CO2 mix at 100 μE light intensity at different times before and after addition of 20 μM rapamycin or solvent (dimethyl sulfoxide) alone showing that relocalization of Vipp1 to the nucleoid at the cell centers is rapid, effective and stable for at least 4-5 hours. The same grayscale contrast was applied to all panels within each condition.

periphery.

To reveal the functional role of native Vipp1 localization, we chose to apply the anchor-away perturbation in three different physiological conditions that could enable us to disentangle interactions between thylakoid formation and photosynthesis *Dark Growth*, *Dark-to-Light Transition*, and *Light Growth* (Fig. 31C). During *Dark Growth* (Anderson and McIntosh, 1991), cells rely on glucose from the media to grow and divide in the absence of photosynthetic light while still multiplying their thylakoid membranes and associated pigments. However, dark-grown cells are not photosynthetically

competent (Barthel et al., 2013), meaning they are incapable of immediate oxygen production upon application of light. In the *Dark-to-Light Transition*, photosynthetic competency is induced when *dark-grown* cells are shifted to light conditions for at least 6 - 8 hours (Barthel et al., 2013). During *Light Growth*, cells rely on photosynthesis to grow and divide in the absence of glucose and also continuously multiply their thylakoids and photosynthetic components. We confirmed in our own settings that dark-grown cells, even after multiple cell divisions, still maintain and produce thylakoids and photosynthetic pigments as measured by electron and fluorescence microscopy, and spectrophotometry respectively (Fig. 32D-E). We note that in electron microscopy images, the thylakoid sheets in dark-grown cells appeared less stacked than in light-grown cells, however they still maintained a cell peripheral arrangement that was also observed in fluorescence images of live cells, which later grew and divided when shifted to *Light Growth* (Fig. 32D). We also confirmed that dark-grown cells are photosynthetically incompetent as evidenced by their inability to evolve oxygen in the presence of light (Fig. 31F). When these dark-grown cells are shifted to *Light Growth* in glucose free media, the photosynthetic competency is restored within 24 h (Fig. 31F) as was previously established (Barthel et al., 2013).

We asked whether the rapid drug-induced relocalization of Vipp1 to the nucleoid in the three conditions (*Dark, Dark-to-Light* and *Light*) would elicit either a growth or thylakoid morphology defect. For each of the conditions, we added rapamycin to induce Vipp1 relocalization and used time-lapse fluorescent microscopy to monitor thylakoid content, Vipp1 localization and cell growth for at least half of the cell cycle (Fig. 34A-C). For longer timescales,

we measured growth by monitoring turbidity in bulk culture (Fig. 35A-B) and examined thylakoid morphology by both fluorescence and electron microscopy (Fig. 35C) We find that Vipp1 relocalization causes no growth or thylakoid morphology (Fig. 34A-B, Fig. 35A-C, Fig. 36) defects in *Dark* or *Light* conditions, but causes a severe growth defect when Vipp1 is relocalized during the *Dark-to-Light Transition* (Fig. 34C, Fig. 35A). The growth defect observed during the *Dark-to-Light Transition* is not due to nucleoid-related changes, but to Vipp1 relocalization, since a control strain in which heterologously expressed GFP is relocalized to the nucleoid exhibits no growth phenotype (Fig. 35B). The severity of the growth defect correlates with the timing of Vipp1 relocalization during the *Dark-to-Light Transition*, with less of a growth defect observed when the perturbation occurs later in the transition (Fig. 37A). Intriguingly, the most severe growth defect is observed when relocalization is induced at the start of the *Dark-to-Light Transition*, which coincides with the peak burst in Vipp1 puncta formation in normal conditions (Fig. 37B).

Based on the following observations, we conclude that the native localization of Vipp1 is not necessary for thylakoid membrane growth, but is necessary for enabling acquisition of photosynthetic competency. First, since perturbing Vipp1 in the *Dark* condition (in which there is active thylakoid growth) triggers no defect in thylakoid growth or morphology (Fig. 35C, Fig. 36A-B), we conclude that Vipp1 localization is not required for thylakoid membrane growth. However, since perturbing Vipp1 in the *Dark-to-Light* condition (in which cells must enable photosynthesis in order to grow) elicits a severe growth defect (Fig. 34C, Fig. 35A), we infer that native Vipp1

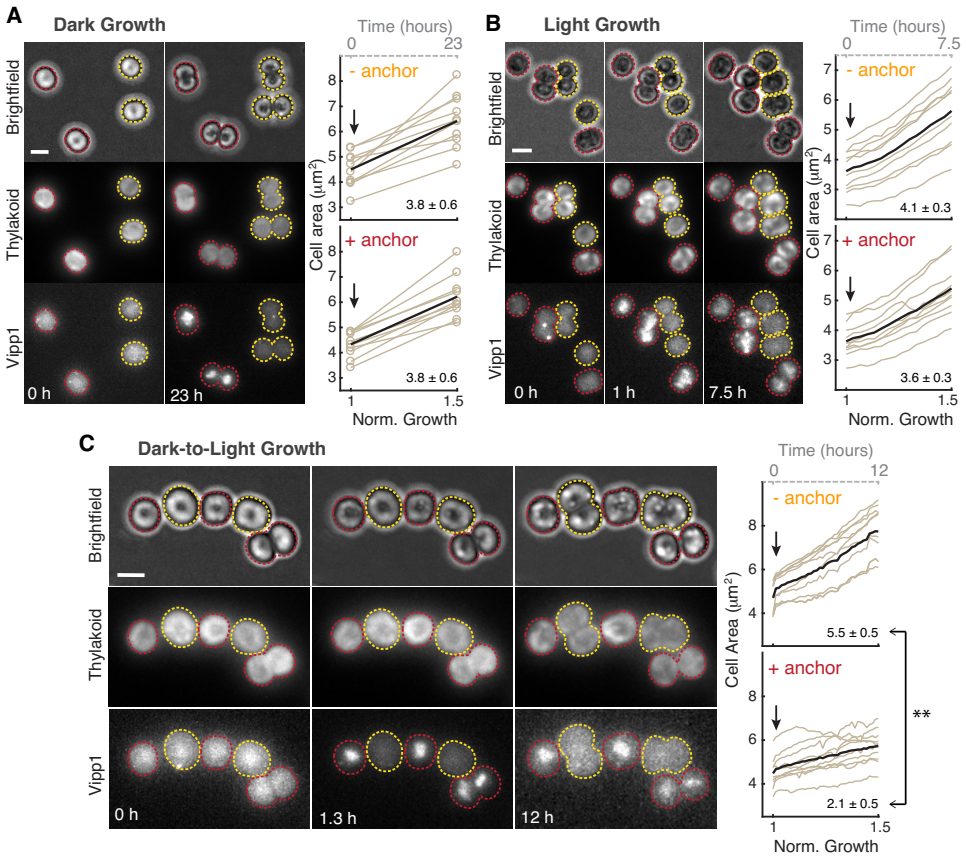

Figure 34: Relocalization of Vipp1 to the nucleoid has no significant effect on Dark or *Light Growth* but elicits a severe growth defect during *Dark-to-Light Transition*.

(Continued) A) Relocalization of Vipp1 induces no growth defect during *Dark Growth*. Left: Example images of dark-grown cells (Maximum Z-Intensity Projections, 12 Z-steps) in brightfield, thylakoid and Vipp1 channels before addition of rapamycin (0 h) and after 23 hours of growth in darkness. A mix of two strains are shown, one (yellow outlines) in which the Vipp1-GFP-FKBP12 construct is expressed alone, and another (red outlines) in which the HU-FRB anchor construct is also expressed in trans (i.e. responsive to rapamycin-based relocalization). 20 μM rapamycin was added to the agarose pad holding the cells on the microscope stage immediately after time 0 h. Dark-grown cells were obtained from a culture growing in darkness in the presence of glucose for at least 144 h with a s. Right: quantification of cell growth as areas (extracted from automatic segmentations in the brightfield channel - open circles) over time of ten random cells obtained from each strain (with or without HU-FRB anchor) in the same field of view. On the bottom x-axis, normalized time is shown as the time it took for -anchor cells to increase the average cell area by 50%. On the top x-axis (dashed grey line), time is shown in hours. Arrows indicate the addition of rapamycin to the agarose pad. Average growth trace is shown as a black bold curve. The mean and standard error of the mean of area expansions rates of the 10 cells are shown in the lower right corners. We found no significant difference between the means of +anchor and - anchor cells as calculated with the two-sample two-tailed t-test at 5% significance level. B) Relocalization of Vipp1 induces no significant growth defect during *Light Growth*. Similar to panel A - on the left: selected images of light-growing cells obtained from a timelapse movie (8 Z-steps, every 30 min for 7.5 hours total). Right: quantifications of cell growth as areas over time of ten random cells obtained from each strain from the same field of view. No significant differences between the means of +anchor and -anchor cells rates (shown at the lower right corner) as calculated with the two sample two-tailed t-test at 5% significance level were found. C) Relocalization of Vipp1 during *Dark-to-Light Transition* induces a severe growth defect. Similar to panel A - on the left: selected images of dark-grown cells shifted to *Light Growth* obtained from a timelapse movie (10 Z-steps, every 20 min for 12 hours). Photosynthetic light and rapamycin were applied immediately after the first timepoint. On the right: quantification of cell growth as areas over time of ten random cells obtained from each strain from the same field of view. The difference between the means of +anchor and -anchor cells rates (shown in the lower right corner) was found to be significant (** $P < 0.001$) based on the two-sample two-tailed t-test at 5% significance level.

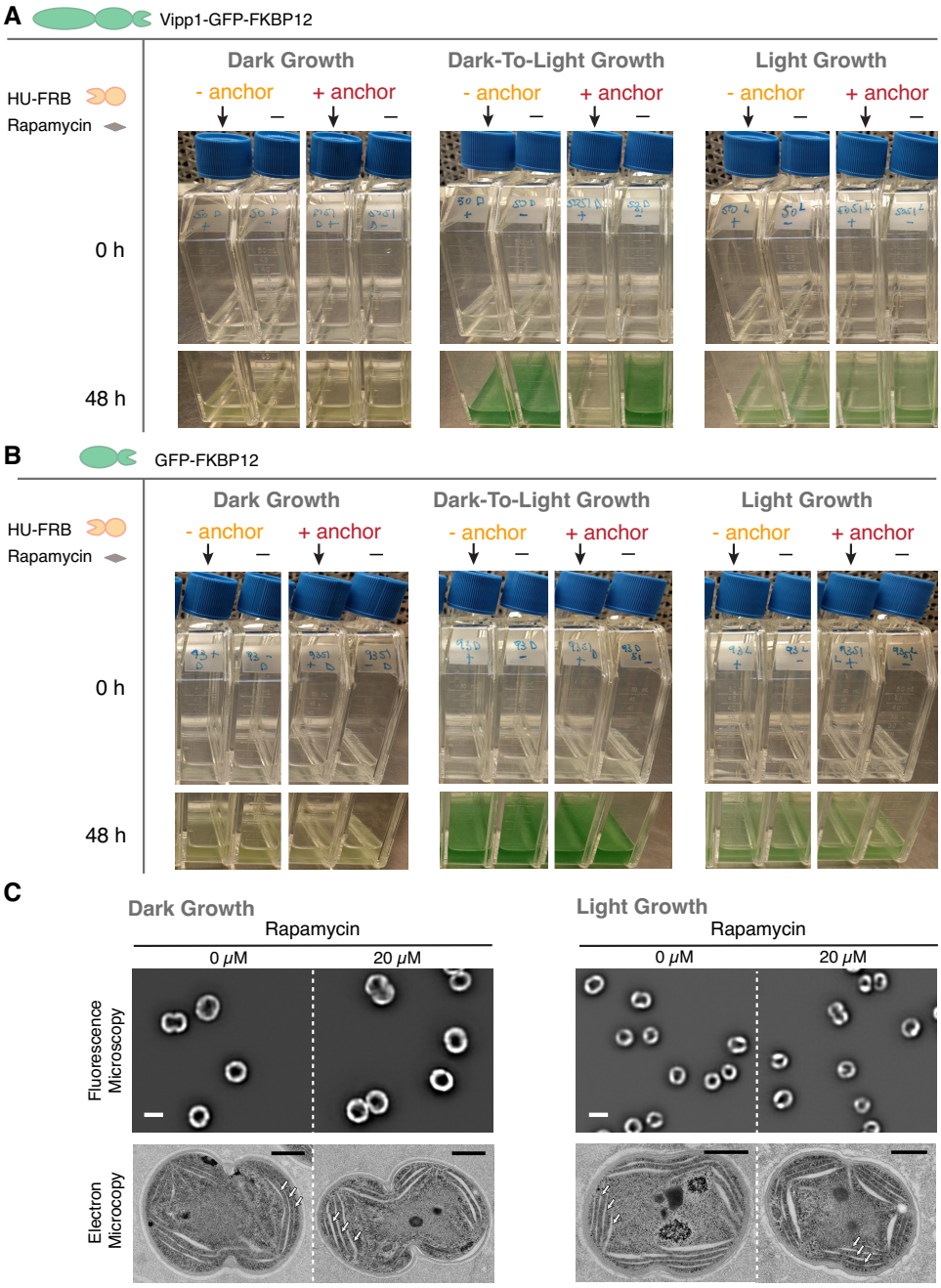

Figure 35: Vipp1 relocalization produces no growth or thylakoid morphology defect in Dark or *Light Growth* cultures but elicits a severe growth defect in the *Dark-to-Light Transition*.

(Continued) Vipp1 relocalization produces no growth or thylakoid morphology defect in Dark or *Light Growth* cultures but elicits a severe growth defect in the *Dark-to-Light Transition*. A) Effect of Vipp1 relocalization on growth in bulk culture. Dark- and light growing vipp1-gfp-fkbp12 cells of the two strains - with or without the HU-FRB anchor construct expressed in trans - were resuspended in fresh media in pairs of flasks at the same starting inoculum and incubated in dark or in light for 48 h. 20 μM rapamycin (arrows) or solvent only (dimethyl sulfoxide, denoted as -) was added at time 0 h. B) Control cultures showing that relocalization of GFP alone at the nucleoid elicits no growth defects. Dark- and light growing gfp-fkbp12 cells of the two strains - with or without the HU-FRB anchor construct expressed in trans - were inoculated in fresh media in pairs of flasks at the same starting inoculum and incubated in dark or in light for 48 h. 20 μM rapamycin (arrows) or solvent only (dimethyl sulfoxide, denoted as -) was added at time 0 h. C) Relocalization of Vipp1 during Dark or *Light Growth* induces no changes in overall thylakoid content or morphology. Top - filtered images in the thylakoid channel (mid-cell plane) obtained from cells expressing all the components of Vipp1 anchor-away in the presence or absence of rapamycin after 48 h of growth in darkness (left) or in light (right). The images are displayed at the same grayscale contrast for rapamycin and no rapamycin treated cells. The overall thylakoid morphology (i.e. peripherally localized signal with occasional gaps in fluorescence intensity that correspond to regions of high curvature of thylakoid membranes) in the rapamycin-treated cells is similar to that of the control cells. Bottom - corresponding electron microscopy images of representative cells from each of the conditions tested. Stacked thylakoids (arrows) are present in both dark- and light-grown cells and their morphology and arrangement remain unaffected by Vipp1 perturbation.

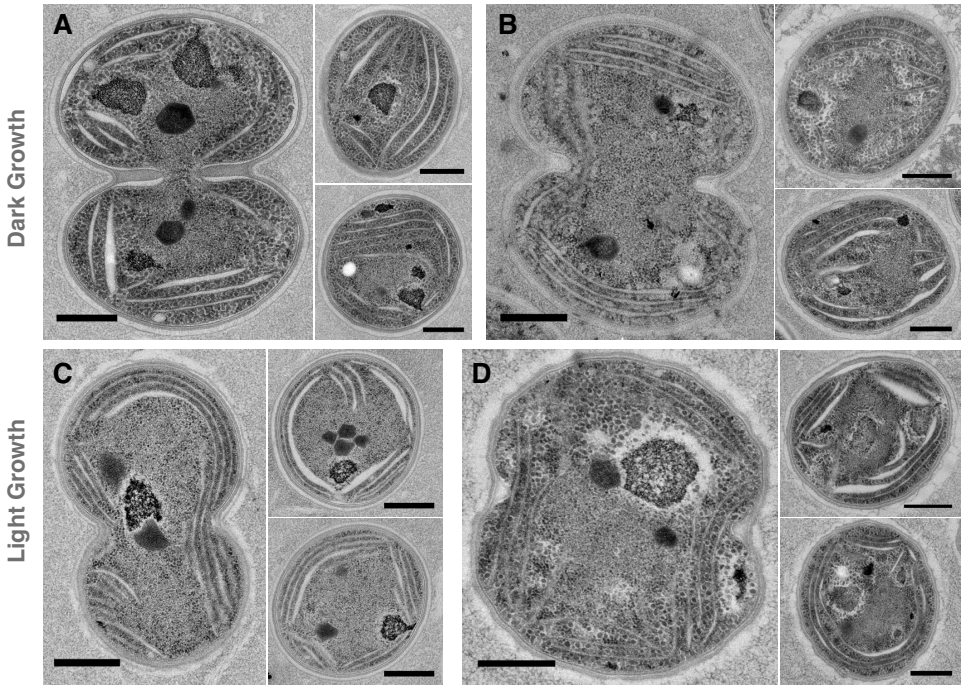

Figure 36: Representative transmission electron microscopy images of dark- and light-grown cells expressing all the components of the anchor-away system after 48 hours of growth in darkness or in light in the absence or presence of 20 μM rapamycin. Dark-grown cells obtained by culturing the cells for 144 hours in darkness in the presence of glucose with brief exposures to activating light (see Methods) were diluted to an OD750 of 0.1 in glucose-containing media and to which either dimethyl sulfoxide solvent (A) or 20 μM rapamycin dissolved in dimethyl sulfoxide was added (B) and further grown in complete darkness for 48 hours. Light-grown cells obtained from a log-growing culture were similarly diluted to an OD750 of 0.1 in media with no glucose to which dimethyl sulfoxide alone (C) or 20 μM rapamycin in dimethyl sulfoxide (D) was added and grown in light for 48 hours. The overall morphology of the thylakoid stacks in rapamycin-treated cells is similar to that of the control cells. Scale bars = 0.5 μm.

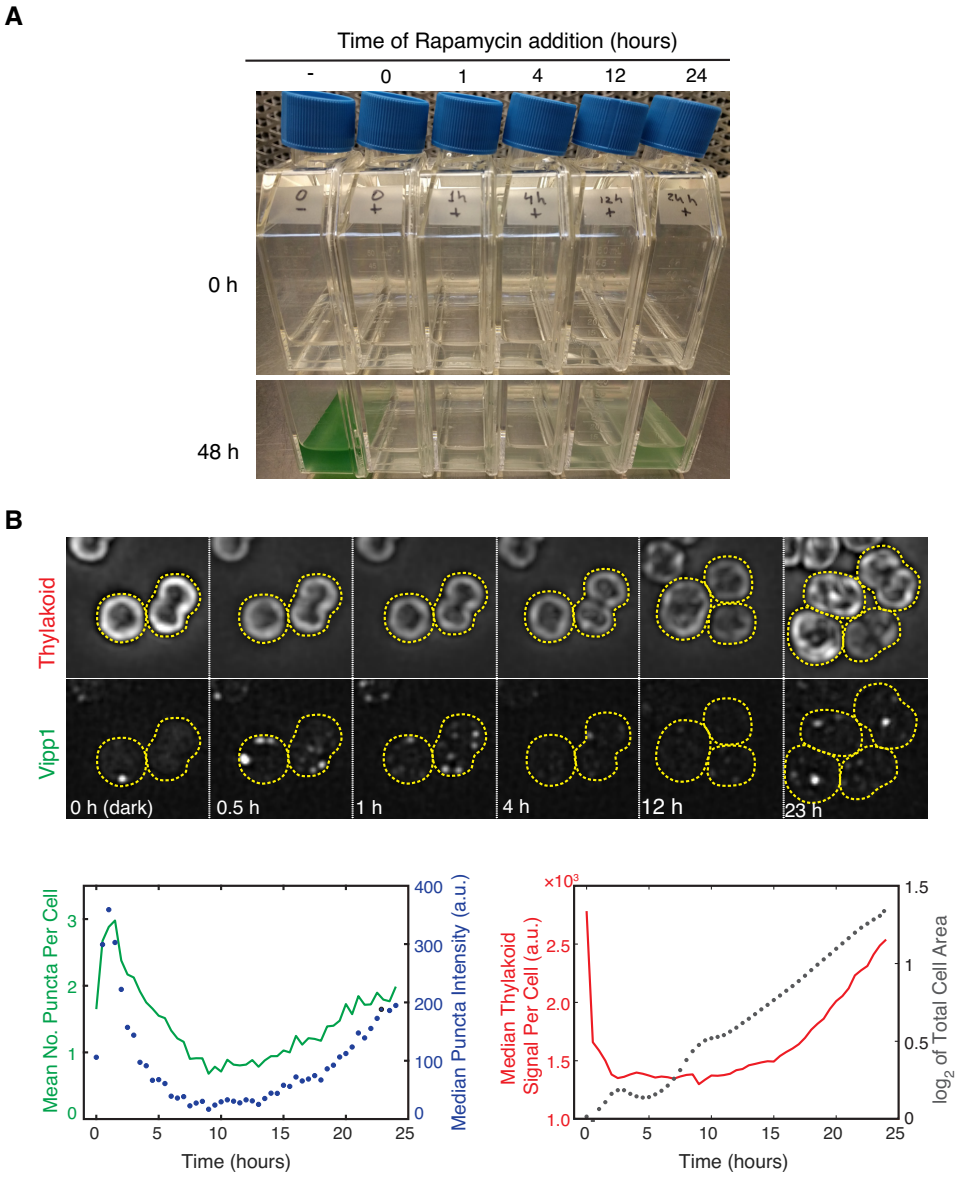

Figure 37: The early stages of *Dark-to-Light Transition* are more prone to the effect of Vipp1 perturbation.

(Continued) A) Delaying Vipp1 relocalization to a later time during the *Dark-to-Light Transition* lessens the impact of the perturbation on growth as evidenced by the increased pigmentation (i.e. growth) of cultures labeled 12 and 24 h. B) Vipp1 undergoes a burst in the rate of puncta formation in the early stages of the *Dark-to-Light Transition*. Top - selected images from a time-lapse montage of dark-grown vipp1-gfp cells transitioning to *Light Growth*. Maximum Z-Intensity Projections of 3D filtered stacks in the thylakoid and Vipp1 channels are shown. The same grayscale contrast was applied to the entire montage within each channel. Bottom: quantification of Vipp1 puncta number per cell, Vipp1 puncta intensities, total thylakoid signal per cell and total cell area extracted from brightfield segmentation masks over time (every 30 min for 24 h) from a field 260 starting vipp1-gfp cells. At time 0 h, before light is turned on, thylakoid auto-fluorescence per cell is high, likely because the existing photosynthetic pigments residing in the thylakoid membranes are uncoupled from photochemical reactions (hence more fluorescent). In growing photosynthetic light, total thylakoid signal per cell is relatively flat early on and starts accumulating after 12 hours as the cells increase their thylakoid content and acclimate fully to the photosynthetic lifestyle. The initial dip in total cell area reflects the shrinkage of a small fraction of cells that fail to grow. For full movie of a representative group of cells.

localization is necessary to enable acquisition of photosynthetic competency. As photosynthetic competency is already established during *Light Growth* and Vipp1 perturbation does not trigger a growth or thylakoid morphology defect (Fig. 34B, Fig. 35C, Fig. 36C-D), we infer that Vipp1 localization is not critical for maintaining thylakoid membranes or supporting active and pre-existing photosynthetic capacity.

8.4 Discussion

Punctate or elongated Vipp1 structures have been previously observed by fluorescence microscopy in both chloroplast (Nordhues et al., 2012; Zhang et al., 2012) and in cyanobacteria (Bryan et al., 2014). However their localization and dynamics in growing cells (non-stress conditions) has not been fully examined. Using live-cell timelapse fluorescence microscopy we

show that Vipp1-GFP continuously exchanges between a diffuse fraction and punctate fraction on a timescale of minutes (Fig. 28C, Fig. 31A-B). This punctate fraction displays limited mobility and localizes to regions of high thylakoid curvature (Fig. 28D-F, Fig. 29), which is consistent with known biochemical properties of membrane binding and oligomerization (Fuhrmann et al., 2009a; Otters et al., 2013; Hennig et al., 2015). Moreover, the localization of Vipp1 puncta to high thylakoid curvature regions is consistent with the previously described in vitro observations in which Vipp1 binds with high affinity to artificial membrane vesicles manifesting high stored-curvature-elastic stress (McDonald et al., 2015).

The ability of *Synechocystis* PCC6803 cells to grow non-photosynthetically, while still maintaining and multiplying their thylakoids (Fig. 32D) - which are also the sites of respiratory activity (Mullineaux, 2014) - allowed us ask whether the localization and dynamics of Vipp1 is important for thylakoid membrane growth alone. By experimentally perturbing the native spatial distribution of Vipp1 in this Dark condition, we show that thylakoid membrane growth is independent of Vipp1 localization (Fig. 35C). This suggest that thylakoids (at least the ones generated in *Dark Growth*) can be made even when Vipp1 is not available at the thylakoids or cell periphery, casting doubt on the lipid (or other components) transport hypothesis from the plasma membrane to thylakoids.

Given that the dark-grown cells are photosynthetically incompetent, yet multiply both their thylakoid membranes (Fig. 32D) and the resident photosynthetic protein components (Barthel et al., 2013) (Fig 32E) as they grow and divide, the *Dark-to-Light Transition* is a sensitized context in which

the cells must either functionally turn on the pre-existing components or synthesize new active components. In other words, Dark-to-Light cells must enable their photosynthetic competency. Vipp1 relocalization in this *Dark-to-Light Transition* causes a severe growth defect (Fig. 34C, Fig. 35A). At this point, the outcome of this relocalization effect can be interpreted in two ways: i) Vipp is involved in enabling photosynthesis or ii) Vipp1 functions as a protective factor from light-induced stress. Given the previously described phenotype of a Vipp1 depletion strain (Gao and Xu, 2009) in which the photosynthetic activity is abrogated without affecting respiration or the thylakoid membranes, we interpret that by altering Vipp1 localization, the dark-grown cells cannot enable the acquisition of photosynthetic competence when light is available. In other words, the light-dependent maturation of photosynthetic machinery occurring during the *Dark-to-Light Transition* is dependent on the localization dynamics of Vipp1.

Finally, Vipp1 relocalization during *Light Growth* must not be rate-limiting for growth (at least over the perturbation time-window), since we do not see a significant effect (Fig. 34B, Fig. 35A). Granted that Vipp1 is not necessary for thylakoid membrane growth (see Dark perturbation), and Vipp1 is necessary for enabling photosynthetic competency (see Dark-to-Light perturbation), we interpret that the effect of Vipp1 relocalization in *Light Growth* is masked because the cells are already photosynthetically competent.

Although relocalization of Vipp1 to the nucleoid alters both the diffuse and the punctate fractions, we hypothesize that it is the punctate fraction enriched at high curvature regions of the thylakoid, that is important for

enabling the acquisition of photosynthetic competency. This is supported by the coincidence in time between the sensitivity to Vipp1 relocalization and the observed burst in Vipp1 puncta formation at the start of the Dark-to-Light condition (Fig. 36), which is also when the need to initiate photosynthetic capacity is high. Furthermore, Vipp1 puncta form dynamically at the cell periphery (Fig. 28B and Fig. 30A) at regions of thylakoids marked by CurT enrichments (Fig. 28D), a specialized zone previously proposed to act as biogenetic centers for new photosynthetic complexes (Heinz et al., 2016). We note that the association between Vipp1 and CurT that that we measured cytologically has not been confirmed in previous biochemical interaction studies, suggesting only an indirect connection between Vipp1 and CurT (Bryan et al., 2014; Heinz et al., 2016).

What could be the molecular mechanisms by which Vipp1 enables acquisition of photosynthetic competency? Combining our analysis of Vipp1 localization and dynamics with its known biochemical properties, we hypothesize that punctuated oligomerization of Vipp1 at regions of high thylakoid curvature mediates functional light-dependent assembly of photosystem(s) complexes. Perhaps due to its membrane-binding and/or lipid rearrangement properties, oligomerized form of Vipp1 could facilitate efficient translation and/or assembly of newly made photosynthetic complexes in the membrane this may be accomplished by alleviating the high curvature membrane stress emerging at thylakoid edges (McDonald et al., 2015), assisting complex assembly by providing lipid cofactors, as previously proposed (Nordhues et al., 2012; McDonald et al., 2017), or mediating putative transient intermembrane contacts (Hennig et al., 2015; Heidrich et al., 2017).

Our conclusions are consistent with the previous analysis of Vipp1 depletion in *Synechocystis*, which showed that when Vipp1 levels are lowered, the photosynthetic function is significantly reduced while respiration and thylakoid morphology remains unchanged (Gao and Xu, 2009). Similarly, deletion of Vipp1 in Synechococcus PCC 7002 implicated Vipp1 primarily in the translation or assembly of photosystem I, not for thylakoid membrane biogenesis per se (Zhang et al., 2014). Additionally, Vipp1 has been shown to interact with specific chaperones, translation factors and core photosynthetic proteins, supporting a role for Vipp1 in the translation or assembly of photosystems (Bryan et al., 2014).

Our results highlight the importance of Vipp1 dynamic behavior at high curvature areas of the thylakoid which could act as sites of activation or synthesis of new photosynthetic complexes. Additionally, we present a new rapid perturbation of Vipp1 and its corresponding synthetic phenotype which opens a new context to biochemically understand which photosynthetic component(s) requires Vipp1 in its functional assembly. Finally, this work highlights how single-cell time lapse imaging and rapid perturbations can complement biochemical and genetic observations in understanding complex processes such photosynthesis that change in space and time.

8.5 Experimental Procedures

8.5.1 Strains and Growth Conditions

Synechocystis sp. PCC 6803 GT strain (kind gift of Dr. Wim Vermaas) (Trautmann et al., 2012) was used for all the work presented. Strains were

propagated on BG11-1.5% GelRite (PlantMedia) containing 10 mM HEPES-KOH pH 8.0 in the presence of the appropriate antibiotic. Photosynthetically grown cultures were grown at 30 °C either in Nalgene plastic flasks tilted at low angle on an orbital shaker (70 rpm) or in 40 ml glass tubes bubbled with 1% CO2-air mix. Philips cool fluorescent tube light bulbs (F40T12/841 Alto) provided the light whose intensity was set to 30 (when flasks were used) or 100 μE m2 s-1 (when air was bubbled) as measured with a LI-COR 190R sensor. Depending on the conditions, the doubling time in the log-phase ranged from 10 to 5 hours (see example of a typical growth curve in Fig. S1B) For all growth or imaging experiments, the cells were prepared and maintained in media with no antibiotics. Prior to imaging, cells were acclimated for 30 min on the microscope stage under a thin agarose pad made with BG-11 medium at 30 °C while illuminated by an external LED white light source (set to 100 $\mu E\ m^2 s^{-1}$ intensity light measured at the objective lens) (See Light Microscopy and Image Analysis Section below). To sustain non-photosynthetic growth (*Dark Growth*), the cultures were supplemented with 27 mM glucose and kept in darkness and stimulated with a daily 5 min exposure of 5 μE m2 s-1 light as was previously established to promote their viability (Anderson and McIntosh, 1991). The cultures were grown in dark up to 144 hours on an orbital shaker (70 rpm) and diluted to an OD750 of 0.1 when their OD750 reached 1 or above (doubling time 16 h). To limit any light effect during *Dark Growth* perturbation of Vipp1, the cultures were kept in continuous darkness for the duration of the experiment.

8.5.2 Genetic Manipulations

To obtain the GFP-tagged allele of Vipp1, we cloned into pBR322 between EcoRI and NcoI using Gibson Assembly (Gibson, 2011) a chimeric construct containing the upstream region and the open reading frame of *sll0617* translationally fused (SGGG linker) at the C-terminus to the codon-optimized variant of mGFPmut3 (Landgraf et al., 2012), the putative native 3 untranslated region of *sll0617* (68 bp), the spectinomycin resistance cassette (aadA) and a 1 kb homology regions downstream of *sll0617*. The resulting plasmid (pAGH42) was used to transform the naturally competent *Synechocystis* according to published procedures (Zang et al., 2007). As for the *vipp1-gfp* construct, the SNAP-tagged version of Vipp1 - also codon optimized - was cloned and transformed into *Synechocystis* cells.

To make the Vipp1-GFP-FKBP12 construct for the anchor-away relocalization system (Liberles et al., 1997) we fused the codon optimized fragment of the human FKBP12 to the C-terminus of GFP (GSGG linker) in the Vipp1-GFP construct described above using the Gibson assembly procedure. The resulting plasmid (pAGH50) was transformed into the appropriate Synechocystis strain to replace the native *vipp1* locus. No differences in growth or Vipp1 localization and dynamics were observed between the *vipp1-gfp-fkbp12* and *vipp1-gfp strains*.

To make the HU-anchor fusion, we designed a construct synthesized by Integrated DNA Technologies (Coralville, IA) in which the codon optimized FRB fragment- was fused to the C-terminus (GSG linker) of HU (encoded by *sll1712*). The DNA construct included the putative native promoter and

5 untranslated region (265 nt) of the *sll1712* gene. At the 3end, the 68 bp 3 untranslated region of *vipp1* was used. The HU-anchor construct was cloned by Gibson assembly downstream of the kanamycin resistance cassette in a pBR322-based plasmid designed to recombine into the *slr0168* neutral site of the *Synechocystis* chromosome (Gao and Xu, 2009).

A GFP-FKBP12 control construct (pAGH93) was made similarly to *vipp1-gfp-fkbp12*, except that the L03 synthetic promoter was used to drive its expression (Huang and Lindblad, 2013).

In all cases, the colonies obtained after transformation were re-streaked onto fresh selection media and the genetic modification was confirmed by PCR amplification with primers flanking the locus of interest. For long-term storage, all *Synechocystis* strains were kept at -80 °C in BG11 with 8% dimethyl sulfoxide. All vectors used in this study are [will be] deposited in the Addgene plasmid bank.

8.5.3 Light microscopy and image analysis

All epifluorescence imaging was performed on a Zeiss AxioObserver.Z1 inverted microscope equipped with an environmental chamber, a Definite Focus module, a hardware-triggered Hamamatsu ORCA Flash4.0 V2+ camera and a Plan-Apochromat 63x/1.40 Oil DIC M27 (NA 1.4) objective. Unless noted, all images were acquired at Nyquist sampling rate as Z-stacks in Zen 2.1 software (Zeiss). Sola SE (Lumencor, Beaverton, OR) white light source was used to image Vipp1 and the thylakoid in the GFP and the far-red fluorescence channels (Zeiss Filter Sets 38HE and 50). All brightfield imaging was done in Khler illumination with light passing through the same filter

cube as the one used for Vipp1 channel.

For superresolution imaging we used Zeiss LSM 880 laser confocal system equipped with an Airyscan unit, environmental control chamber, and 1.4 NA Plan-Apochromat 63x DIC M27 oil objective. Similar to epifluorescence imaging procedures, cells were acclimated on the microscope stage for 30 min, while being illuminated by photosynthetic light. To limit any potential laser-induced stress on cells (and also changes to Vipp1 localization) we identified fields of view with the use 594 nm laser set to the lowest power at maximal detector gain. Image stacks were acquired for 16 Z-planes (512 x 512) with a voxel size 40 x 40 x 180 nm centered at mid-cell where the focus was the sharpest. Laser lines 458 nm, 515 nm and 594 nm were used though the same beamsplitter (MBS 458/514/594) for excitation of CurT-mTurquoise2, Vipp1-GFP (mGFPmut3 variant: excitation maximum - 501 nm, emission maximum - 511 nm) and thylakoid respectively. Emissions from the CurT and Vipp1 channels passed through a dual bandpass filter BP 420-480, BP 495-550). Pixel dwell time (set at the fastest) and laser power were adjusted to avoid saturation and bleaching effects while the detector gain was set to 800. For Airscan processing, Zen Black 2.1 was used to process the image stacks by performing the filtering, deconvolution and pixel reassignments in 3D mode at default settings. This processing confers the increased spatial resolution (by a factor of 1.5 to 1.7x both laterally and axially) and an enhanced signal to noise ratio (Korobchevskaya et al., 2017).

For live-cell imaging, an aliquot of cells (obtained from an OD750 of 0.2 - 0.3 culture) was placed onto the glass of a 35 mm MatTek dish (P35G-0.170-14-C) (MatTek, Ashland, MA) under a thin BG11 1% agarose pad (2

mm thickness) prepared in advance and covered with a coverslip to reduce evaporation. The dish was placed onto the objective in the stage insert and around the agarose pad a basin of water was poured to create a humidified atmosphere. The temperature was maintained at 30 °C while the photosynthetic light was provided by a white light LED ring: RL1360 or DF198 (Advanced Illumination, Rochester, VT) that was positioned on the stage insert and controlled by an analog signal through Zen 2.1. The LED white light was programmed to switch off only during the fluorescence image acquisition. The light intensity reaching the cells at the objective lens was 100 μE which permitted a growth rate comparable to the rate measured in bulk culture. To limit photobleaching and phototoxicity, the imaging light power was limited to 20% by the FL attenuator and the light exposure for image acquisition in the GFP channel (20 to 30 ms) was adjusted to generate a signal-to-background ratio of less than 1.5. Also, the Z-series were limited to 12 steps or fewer. For thylakoid channel, all exposures used were 3.2 ms. Cell doubling time under these conditions was approximately 8 h.

Multispectral datasets were aligned using the affine transformation plugin in Zeiss Zen 2.1 using a standard prepared from TetraSpeck (Invitrogen, Carlsbad, CA) multicolor beads immobilized on a 1.5 coverslip.

To obtain filtered images, a custom MATLAB script (available upon request *Patent Pending*) was written to computationally enhance both diffraction-limited spots and contours of thylakoids by applying a local maxima filter that fit a 3D Gaussian to every position in the dataset and returned the Maximum Likelihood Estimate (MLE) of the amplitude, background and square-error. The amplitude component of the MLEs reports a quantitative

visualization of the diffraction limited signal since it is separated from its background component. See Fig. S1C for an example of the raw and MLE amplitude filtered image. Unless noted, all filtered images presented are MLE visualizations of the signal amplitudes. There are no parameters to adjust for this computational filter.

To automatically segment cells in the brightfield channel a custom MATLAB script was written that implemented correlation imaging as described previously (Julou et al., 2013). In summary, the edge of a cell is uniquely defined by its Z intensity profile given a 3D brightfield Z-stack. Therefore, to computationally find all cell edges, a brightfield Z-stack was filtered in Z by cross-correlating this unique Z-intensity profile using the MATLAB function *convn*. This cross-correlated dataset was then thresholded and all enclosed objects were then segmented by MATLABs function *regionprops*.

Fiji (Schindelin et al., 2012) was used to prepare the kymograms, extract the intensity profiles, and convert the final images. Unless noted, in all images displayed, the grayscale contrast was auto-adjusted linearly before converting to bitmaps.

9 Conclusion

The signal processing module of an imaging apparatus has a disproportionate affect on its total sensitivity because this module is positioned at the end of a series of optical components. The reason for this importance is because the transmission efficiency of each component multiplies, with the total transmission efficiency of the entire optical device exponential to the number of components used (and there are many optical components in even a simple microscope objective). Since signal processing is the last component in this optical path, any improvement in sensitivity and selectivity in this final module will be exponentially magnified by the total efficiency of the optical path.

By deriving and implementing an optimal pattern detection and localization algorithm, we are able to significantly enhance the sensitivity and selectivity of the signal processing used for 3D fluorescent spot detection. Armed with an epi-fluorescent microscope, we enabled long timescale and high temporal resolution 3D molecular imaging of diffraction limited spots in living systems.

A Appendix

A.1 Publications

Gutu, Andrian, Frederick Chang (co-author), and Erin OShea. Dynamical localization of a thylakoid membrane binding protein is required for acquisition of photosynthetic competency.
Content available at:*https://www.biorxiv.org/content/early/2017/10/08/199943*

FC contributed to the design, analysis & writing.

Estrada, Javier, Natalie Andrew, Daniel Gibson, Frederick Chang, Florian Gnad, and Jeremy Gunawardena. "Cellular Interrogation: Exploiting Cell-to-Cell Variability to Discriminate Regulatory Mechanisms in Oscillatory Signalling." *PLoS Comput Biol* 12, no. 7 (2016): e1004995.

FC contributed to the design, microlfuidics & analysis

Liang, Zhangyi, Denise Zickler, Mara Prentiss, Frederick S. Chang, Guillaume Witz, Kazuhiro Maeshima, and Nancy Kleckner. "Chromosomes Progress to Metaphase in Multiple Discrete Steps via Global Compaction/Expansion Cycles." *Cell* 161, no. 5 (2015): 1124-1137.

FC contributed to the imaging experiments

Mirkin, Ekaterina V., Frederick S. Chang, and Nancy Kleckner. "Protein-Mediated Chromosome Pairing of Repetitive Arrays." *Journal of Molecular Biology* 426.3 (2014): 550-557.

FC contributed to the imaging & analysis.

Mirkin, Ekaterina V., Frederick S. Chang, and Nancy Kleckner. "Dynamic Trans Interactions in Yeast Chromosomes." *PloS One* 8.9 (2013): e75895.

FC contributed to the imaging & analysis.

References

Anderson, S. L. and McIntosh, L. (1991). Light-activated heterotrophic growth of the cyanobacterium synechocystis sp. strain PCC 6803: a blue-light-requiring process. *J. Bacteriol.*, 173(9):2761–2767.

Apel, K. and Hirt, H. (2004). Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annu. Rev. Plant Biol.*, 55:373–399.

Aseeva, E., Ossenbühl, F., Sippel, C., Cho, W. K., Stein, B., Eichacker, L. A., Meurer, J., Wanner, G., Westhoff, P., Soll, J., and Vothknecht, U. C. (2007). Vipp1 is required for basic thylakoid membrane formation but not for the assembly of thylakoid protein complexes. *Plant Physiol. Biochem.*, 45(2):119–128.

Balzarotti, F., Eilers, Y., Gwosch, K. C., Gynnå, A. H., Westphal, V., Stefani, F. D., Elf, J., and Hell, S. W. (2017). Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. *Science*, 355(6325):606–612.

Barthel, S., Bernat, G., Seidel, T., Rupprecht, E., Kahmann, U., and Schneider, D. (2013). Thylakoid membrane maturation and PSII activation are linked in greening synechocystis sp. PCC 6803 cells. *Plant Physiol.*, 163(2):1037–1046.

Bossi, M., Fölling, J., Belov, V. N., Boyarskiy, V. P., Medda, R., Egner, A., Eggeling, C., Schönle, A., and Hell, S. W. (2008). Multicolor far-field fluorescence nanoscopy through isolated detection of distinct molecular species. *Nano Lett.*, 8(8):2463–2468.

Bryan, S. J., Burroughs, N. J., Shevela, D., Yu, J., Rupprecht, E., Liu, L.-N., Mastroianni, G., Xue, Q., Llorente-Garcia, I., Leake, M. C., Eichacker, L. A., Schneider, D., Nixon, P. J., and Mullineaux, C. W. (2014). Localisation and interactions of the vipp1 protein in cyanobacteria. *Mol. Microbiol.*

Carlton, P. M., Boulanger, J., Kervrann, C., Sibarita, J.-B., Salamero, J., Gordon-Messer, S., Bressan, D., Haber, J. E., Haase, S., Shao, L., Winoto, L., Matsuda, A., Kner, P., Uzawa, S., Gustafsson, M., Kam, Z., Agard, D. A., and Sedat, J. W. (2010). Fast live simultaneous multiwavelength four-dimensional optical microscopy. *Proc. Natl. Acad. Sci. U. S. A.*, 107(37):16016–16022.

Chen, B.-C., Legant, W. R., Wang, K., Shao, L., Milkie, D. E., Davidson, M. W., Janetopoulos, C., Wu, X. S., Hammer, 3rd, J. A., Liu, Z., English, B. P., Mimori-Kiyosue, Y., Romero, D. P., Ritter, A. T., Lippincott-Schwartz, J., Fritz-Laylin, L., Mullins, R. D., Mitchell, D. M., Bembenek, J. N., Reymann, A.-C., Böhme, R., Grill, S. W., Wang, J. T., Seydoux, G., Tulu, U. S., Kiehart, D. P., and Betzig, E. (2014). Lattice light-sheet microscopy: imaging molecules to embryos at high spatiotemporal resolution. *Science*, 346(6208):1257998.

Cramer, H. (1946). *Mathematical Methods of Statistics. Statistics.*

Cutrale, F., Trivedi, V., Trinh, L. A., Chiu, C.-L., Choi, J. M., Artiga, M. S., and Fraser, S. E. (2017). Hyperspectral phasor analysis enables multiplexed 5D in vivo imaging. *Nat. Methods*, 14(2):149–152.

Dickson, R. M., Cubitt, A. B., Tsien, R. Y., and Moerner, W. E. (1997). On/off blinking and switching behaviour of single molecules of green fluorescent protein. *Nature*, 388(6640):355–358.

Eberhard, S., Finazzi, G., and Wollman, F.-A. (2008). The dynamics of photosynthesis. *Annu. Rev. Genet.*, 42:463–515.

Edwards, A. W. F. (1993). Likelihood, expanded edition. *Biometrics*, 49(4):1285.

Ettinger, A. and Wittmann, T. (2014). Fluorescence live cell imaging. In *Methods in Cell Biology*, pages 77–94.

Fadero, T. C., Gerbich, T. M., Rana, K., Suzuki, A., DiSalvo, M., Schaefer, K. N., Heppert, J. K., Boothby, T. C., Goldstein, B., Peifer, M., Allbritton, N. L., Gladfelter, A. S., Maddox, A. S., and Maddox, P. S. (2017). LITE microscopy: a technique for high numerical aperture, low photobleaching fluorescence imaging.

Finn, E. H., Pegoraro, G., Shachar, S., and Misteli, T. (2017). Comparative analysis of 2D and 3D distance measurements to study spatial genome organization. *Methods*, 123:47–55.

Fossum, E. R., Mendis, S., and Kemeny, S. E. (1995). Active pixel sensor with intra-pixel charge transfer.

Fuhrmann, E., Bultema, J. B., Kahmann, U., Rupprecht, E., Boekema, E. J., and Schneider, D. (2009a). The vesicle-inducing protein 1 from

synechocystis sp. PCC 6803 organizes into diverse higher-ordered ring structures. *Mol. Biol. Cell*, 20(21):4620–4628.

Fuhrmann, E., Gathmann, S., Rupprecht, E., Golecki, J., and Schneider, D. (2009b). Thylakoid membrane reduction affects the photosystem stoichiometry in the cyanobacterium synechocystis sp. PCC 6803. *Plant Physiol.*, 149(2):735–744.

Gao, H. and Xu, X. (2009). Depletion of vipp1 in synechocystis sp. PCC 6803 affects photosynthetic activity before the loss of thylakoid membranes. *FEMS Microbiol. Lett.*, 292(1):63–70.

Grimm, J. B., English, B. P., Chen, J., Slaughter, J. P., Zhang, Z., Revyakin, A., Patel, R., Macklin, J. J., Normanno, D., Singer, R. H., Lionnet, T., and Lavis, L. D. (2015). A general method to improve fluorophores for live-cell and single-molecule microscopy. *Nat. Methods*, 12(3):244–50, 3 p following 250.

Haruki, H., Nishikawa, J., and Laemmli, U. K. (2008). The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. *Mol. Cell*, 31(6):925–932.

Hecht, S., Shlaer, S., and Pirenne, M. H. (1942). ENERGY, QUANTA, AND VISION. *J. Gen. Physiol.*, 25(6):819–840.

Heidrich, J., Thurotte, A., and Schneider, D. (2017). Specific interaction of IM30/Vipp1 with cyanobacterial and chloroplast membranes results in membrane remodeling and eventually in membrane fusion. *Biochim. Biophys. Acta*, 1859(4):537–549.

Heidrich, J., Wulf, V., Hennig, R., Saur, M., Markl, J., Sönnichsen, C., and Schneider, D. (2016). Organization into higher ordered ring structures counteracts membrane binding of IM30, a protein associated with inner membranes in chloroplasts and cyanobacteria. *J. Biol. Chem.*, 291(29):14954–14962.

Heinz, S., Rast, A., Shao, L., Gutu, A., Gügel, I. L., Heyno, E., Labs, M., Rengstl, B., Viola, S., Nowaczyk, M. M., Leister, D., and Nickelsen, J. (2016). Thylakoid membrane architecture in synechocystis depends on CurT, a homolog of the granal CURVATURE THYLAKOID1 proteins. *Plant Cell*.

Hennig, R., Heidrich, J., Saur, M., Schmüser, L., Roeters, S. J., Hellmann, N., Woutersen, S., Bonn, M., Weidner, T., Markl, J., and Schneider, D. (2015). IM30 triggers membrane fusion in cyanobacteria and chloroplasts. *Nat. Commun.*, 6:7018.

Huang, B., Wang, W., Bates, M., and Zhuang, X. (2008). Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science*, 319(5864):810–813.

Huang, F., Hartwich, T. M. P., Rivera-Molina, F. E., Lin, Y., Duim, W. C., Long, J. J., Uchil, P. D., Myers, J. R., Baird, M. A., Mothes, W., Davidson, M. W., Toomre, D., and Bewersdorf, J. (2013a). Video-rate nanoscopy using sCMOS camera–specific single-molecule localization algorithms. *Nat. Methods*, 10(7):653–658.

Huang, F., Hartwich, T. M. P., Rivera-Molina, F. E., Lin, Y., Duim, W. C.,

Long, J. J., Uchil, P. D., Myers, J. R., Baird, M. a., Mothes, W., Davidson, M. W., Toomre, D., and Bewersdorf, J. (2013b). Video-rate nanoscopy using sCMOS camera-specific single-molecule localization algorithms. *Nature methods*, 10(7):653–8.

Jin, C., Zhang, Y., Balakrishnan, S., Wainwright, M. J., and Jordan, M. (2016). Local maxima in the likelihood of gaussian mixture models: Structural results and algorithmic consequences.

Keppler, A., Gendreizig, S., Gronemeyer, T., Pick, H., Vogel, H., and Johnsson, K. (2003). A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.*, 21(1):86–89.

Kopecna, J., Komenda, J., Bucinska, L., and Sobotka, R. (2012). Long-Term acclimation of the cyanobacterium synechocystis sp. PCC 6803 to high light is accompanied by an enhanced production of chlorophyll that is preferentially channeled to trimeric photosystem I. *Plant Physiol.*, 160(4):2239–2250.

Korobchevskaya, K., Lagerholm, B., Colin-York, H., and Fritzsche, M. (2017). Exploring the potential of airyscan microscopy for live cell imaging. *Photonics*, 4(3):41.

Kroll, D., Meierhoff, K., Bechtold, N., Kinoshita, M., Westphal, S., Vothknecht, U. C., Soll, J., and Westhoff, P. (2001). VIPP1, a nuclear gene of arabidopsis thaliana essential for thylakoid membrane formation. *Proc. Natl. Acad. Sci. U. S. A.*, 98(7):4238–4242.

Laissue, P. P., Alghamdi, R. A., Tomancak, P., Reynaud, E. G., and Shroff, H. (2017). Assessing phototoxicity in live fluorescence imaging. *Nat. Methods*, 14(7):657–661.

Landgraf, D., Okumus, B., Chien, P., Baker, T. A., and Paulsson, J. (2012). Segregation of molecules at cell division reveals native protein localization. *Nat. Methods*, 9(5):480–482.

Liberles, S. D., Diver, S. T., Austin, D. J., and Schreiber, S. L. (1997). Inducible gene expression and protein translocation using nontoxic ligands identified by a mammalian three-hybrid screen. *Proc. Natl. Acad. Sci. U. S. A.*, 94(15):7825–7830.

Liu, Z., Legant, W. R., Chen, B.-C., Li, L., Grimm, J. B., Lavis, L. D., Betzig, E., and Tjian, R. (2014). 3D imaging of sox2 enhancer clusters in embryonic stem cells. *Elife*, 3:e04236.

Mathai, A. M. (1993). On noncentral generalized laplacianness of quadratic forms in normal variables. *J. Multivar. Anal.*, 45(2):239–246.

McDonald, C., Jovanovic, G., Ces, O., and Buck, M. (2015). Membrane stored curvature elastic stress modulates recruitment of maintenance proteins PspA and vipp1. *MBio*, 6(5):e01188–15.

McDonald, C., Jovanovic, G., Wallace, B. A., Ces, O., and Buck, M. (2017). Structure and function of PspA and vipp1 n-terminal peptides: Insights into the membrane stress sensing and mitigation. *Biochim. Biophys. Acta*, 1859(1):28–39.

Mortensen, K. I., Churchman, L. S., Spudich, J. a., and Flyvbjerg, H. (2010). Optimized localization analysis for single-molecule tracking and super-resolution microscopy. *Nature methods*, 7(5):377–381.

Mullineaux, C. W. (2014). Co-existence of photosynthetic and respiratory activities in cyanobacterial thylakoid membranes. *Biochim. Biophys. Acta*, 1837(4):503–511.

Neyman, J. and Pearson, E. S. (1933). On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 231(694-706):289–337.

Nickelsen, J. and Rengstl, B. (2013). Photosystem II assembly: from cyanobacteria to plants. *Annu. Rev. Plant Biol.*, 64:609–635.

Nordhues, A., Schöttler, M. A., Unger, A.-K., Geimer, S., Schönfelder, S., Schmollinger, S., Rütgers, M., Finazzi, G., Soppa, B., Sommer, F., Mühlhaus, T., Roach, T., Krieger-Liszkay, A., Lokstein, H., Crespo, J. L., and Schroda, M. (2012). Evidence for a role of VIPP1 in the structural organization of the photosynthetic apparatus in chlamydomonas. *Plant Cell*, 24(2):637–659.

Ollion, J., Cochennec, J., Loll, F., Escudé, C., and Boudier, T. (2013). TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization. *Bioinformatics*, 29(14):1840–1841.

Otters, S., Braun, P., Hubner, J., Wanner, G., Vothknecht, U. C., and Chigri, F. (2013). The first α-helical domain of the vesicle-inducing protein in

plastids 1 promotes oligomerization and lipid binding. *Planta*, 237(2):529–540.

Pavani, S. R. P., Thompson, M. A., Biteen, J. S., Lord, S. J., Liu, N., Twieg, R. J., Piestun, R., and Moerner, W. E. (2009). Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proc. Natl. Acad. Sci. U. S. A.*, 106(9):2995–2999.

Peterson, W., Birdsall, T., and Fox, W. (1954). The theory of signal detectability. *Transactions of the IRE Professional Group on Information Theory*, 4(4):171–212.

Pribil, M., Labs, M., and Leister, D. (2014). Structure and dynamics of thylakoids in land plants. *J. Exp. Bot.*, 65(8):1955–1972.

Protassov, R., van Dyk, D. A., Connors, A., Kashyap, V. L., and Siemiginowska, A. (2002a). Statistics, handle with care: Detecting multiple model components with the likelihood ratio test. *Astrophys. J.*, 571(1):545–559.

Protassov, R., van Dyk, D. A., Connors, A., Kashyap, V. L., and Siemiginowska, A. (2002b). Statistics, handle with care: Detecting multiple model components with the likelihood ratio test. *Apj*, 571(1):545–559.

Rao, C. R. (1945). Information and the accuracy attainable in the estimation of statistical parameters. In *Springer Series in Statistics*, pages 235–247.

Rast, A., Heinz, S., and Nickelsen, J. (2015). Biogenesis of thylakoid

membranes. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 1847(9):821–830.

Rothstein, J. (1954). Probability and information theory, with applications to radar. p. m. woodward. McGraw-Hill, new york; pergamon press, london, 1953. 128 pp. illus. $4.50. *Science*, 119(3103):874–874.

Sage, D., Kirshner, H., Pengo, T., Stuurman, N., Min, J., Manley, S., and Unser, M. (2015a). Quantitative evaluation of software packages for single-molecule localization microscopy. *Nat. Methods*, 12(8):717–724.

Sage, D., Kirshner, H., Pengo, T., Stuurman, N., Min, J., Manley, S., and Unser, M. (2015b). Quantitative evaluation of software packages for single-molecule localization microscopy. *Nature Methods*, (August 2014):1–12.

Smith, C. S., Joseph, N., Rieger, B., and Lidke, K. A. (2010a). Fast, single-molecule localization that achieves theoretically minimum uncertainty. *Nat. Methods*, 7(5):373–375.

Smith, C. S., Joseph, N., Rieger, B., and Lidke, K. a. (2010b). Fast, single-molecule localization that achieves theoretically minimum uncertainty. *Nature methods*, 7(5):373–375.

Turin, G. (1960). An introduction to matched filters. *IEEE Trans. Inf. Theory*, 6(3):311–329.

Vermaas, W. F. J., Timlin, J. A., Jones, H. D. T., Sinclair, M. B., Nieman, L. T., Hamad, S. W., Melgaard, D. K., and Haaland, D. M. (2008). In vivo hyperspectral confocal fluorescence imaging to determine pigment

localization and distribution in cyanobacterial cells. *Proc. Natl. Acad. Sci. U. S. A.*, 105(10):4050–4055.

Voie, A. H., Burns, D. H., and Spelman, F. A. (1993). Orthogonal-plane fluorescence optical sectioning: three-dimensional imaging of macroscopic biological specimens. *J. Microsc.*, 170(Pt 3):229–236.

Vothknecht, U. C., Otters, S., Hennig, R., and Schneider, D. (2012). Vipp1: a very important protein in plastids?! *J. Exp. Bot.*, 63(4):1699–1712.

Wang, W., Li, G.-W., Chen, C., Xie, X. S., and Zhuang, X. (2011). Chromosome organization by a nucleoid-associated protein in live bacteria. *Science*, 333(6048):1445–1449.

Westphal, S., Heins, L., Soll, J., and Vothknecht, U. C. (2001). Vipp1 deletion mutant of synechocystis: a connection between bacterial phage shock and thylakoid biogenesis? *Proc. Natl. Acad. Sci. U. S. A.*, 98(7):4243–4248.

Wilks, S. S. (1938). The Large-Sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math. Stat.*, 9(1):60–62.

Yang, H., Liu, J., Wen, X., and Lu, C. (2015). Molecular mechanism of photosystem I assembly in oxygenic organisms. *Biochim. Biophys. Acta*, 1847(9):838–848.

Zhang, L., Kato, Y., Otters, S., Vothknecht, U. C., and Sakamoto, W. (2012). Essential role of VIPP1 in chloroplast envelope maintenance in arabidopsis. *Plant Cell*, 24(9):3695–3707.

Zhang, S., Shen, G., Li, Z., Golbeck, J. H., and Bryant, D. A. (2014). Vipp1 is essential for the biogenesis of photosystem I but not thylakoid membranes in synechococcus sp. PCC 7002. *J. Biol. Chem.*, 289(23):15904–15914.

Zimmermann, T., Marrison, J., Hogg, K., and O'Toole, P. (2014). Clearing up the signal: spectral imaging and linear unmixing in fluorescence microscopy. *Methods Mol. Biol.*, 1075:129–148.