Computational 3D and Reflectivity Imaging with High Photon Efficiency

by

Dongeek Shin

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of
Master of Science in Computer Science and Engineering
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 2014

ARCHIVES
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
JUN 10 2014
LIBRARIES

© Massachusetts Institute of Technology 2014. All rights reserved.

Signature redacted

Author ....
Department of Electrical Engineering and Computer Science
May 21, 2014

Signature redacted

Certified by..... ....
Vivek K Goyal
Assistant Professor, Boston University
Thesis Supervisor

Signature redacted

Certified by.. ....
Jeffrey H. Shapiro
Julius A. Stratton Professor
Thesis Supervisor

Accepted by......... Signature redacted ....
Leslie Kolodziejski
Chairman, Department Committee on Graduate Theses

Computational 3D and Reflectivity Imaging with High Photon Efficiency

by

Dongeek Shin

Submitted to the Department of Electrical Engineering and Computer Science
on May 21, 2014, in partial fulfillment of the
requirements for the degree of
Master of Science in Computer Science and Engineering

Abstract

Imaging the 3D structure and reflectivity of a scene can be done using photon-counting detectors. Traditional imagers of this type typically require hundreds of detected photons per pixel for accurate 3D and reflectivity imaging. Under low light-level conditions, in which the mean photon count is small, the inverse problem of forming 3D and reflectivity images is difficult due to the Poisson noise inherent in low-flux operation. In this thesis, we propose and study two computational imagers (one passive, one active) that can form accurate images at low light levels. We demonstrate the superior imaging quality of the proposed imagers by comparing them with the state-of-the-art optical imaging techniques.

Thesis Supervisor: Vivek K Goyal
Title: Assistant Professor, Boston University

Thesis Supervisor: Jeffrey H. Shapiro
Title: Julius A. Stratton Professor

Acknowledgments

I am very fortunate to have Professor Vivek Goyal and Professor Jeffrey Shapiro as my research supervisors. I thank Vivek for being supportive ever since I joined the STIR group and guiding me and my research during my first two years at MIT. I would like to thank Jeff for his great mentorship and teaching me how to be a meticulous researcher. Every research meeting with Vivek and Jeff has enlightened me and affected this thesis in positive ways.

I would like to thank Ahmed Kirmani, who motivated me further to look into the topic of time-resolved imaging and gave me invaluable advice on many ideas that went into this thesis.

I would also like to thank my research collaborators Andrea Colaço, Dheera Venkatraman, Dr. Franco Wong, and Hye Soo Yang for creating the synergy that improved the quality of my research as a whole.

Finally, I thank my family for their endless love and support.

This research was supported in part by NSF grant No. 1161413 and a Samsung Scholarship.

Contents

1 Introduction **13**

2 Single-Pixel Reflectivity Imaging with High Photon Efficiency **17**

2.1 Prior Work 17

2.1.1 Classical Passive Reflectivity Imaging 18

2.1.2 Single-Pixel Camera based on Compressive Sensing 19

2.1.3 Single-Pixel Camera based on Multiplexed Sensing 21

2.2 Multiplexed Sensing under Poisson Noise 24

2.3 Non-Negativity of Reflectivity 25

2.4 Novel Image Formation 27

2.4.1 When is the Non-Negativity Constraint Active? 28

2.4.2 Choosing the Multiplexing Pattern 30

2.5 Numerical Experiments 32

2.5.1 Comparison of Mean-Square Errors 32

2.5.2 Natural Images 34

3 Active 3D Imaging with High Photon Efficiency **39**

3.1 Prior Work 39

3.1.1 Classical Active 3D Imaging 39

3.1.2 First-Photon Imaging 40

3.2 Single-Photon Imaging Setup 41

3.2.1 Active Illumination 41

3.2.2 Detection 42

3.2.3 Data Acquisition 43

3.3 Observation Model 43

3.3.1 Poisson Detection Statistics 44

3.3.2 Statistics of Number of Detected Photons 44

3.3.3 Statistics of Single Photon Arrival Times 45

3.4 Novel Image Formation 48

3.5 Experiments 50

3.5.1 Experimental Setup 50

3.5.2 Reflectivity Resolution Test 51

3.5.3 Depth Resolution Test 52

3.5.4 Natural Scenes 53

3.5.5 Pixelwise Root Mean-Square Error Test 56

3.5.6 Effect of System Parameters 56

3.5.7 Limitations 58

3.6 Information Theory for Optical Design 58

4 Conclusions **63**

4.1 Highly Photon-Efficient Reflectivity Imaging 63

4.2 Highly Photon-Efficient 3D Imaging 63

A Proofs **65**

A.1 Multiplexing Failure using Hadamard Matrix under Poisson Noise . . 65

A.2 Multiplexing Failure using Circulant Matrices under Poisson Noise . . 66

A.3 Strict Concavity of Log-Likelihood under Multiplexing 68

A.4 Efficiency of Matrix Inverse Demultiplexing Solution 69

A.5 Violation Probability for Symmetric i.i.d. Matrix Inverse Estimators . 70

A.6 Covariance Matrix of Hadamard Matrix Inverse Solution 71

B Performance Guarantees for Pixelwise Single-Photon Imaging **73**

B.1 Pixelwise Maximum-Likelihood Reflectivity Estimation 74

B.2 Pixelwise Maximum-Likelihood Depth Estimation 75

List of Figures

2-1 Effect of shot noise on image quality for several acquisition times T_a (top) and dark count rate values d (bottom). The maximum reflectivity value of the saturn image is 25. 19

2-2 The multiplexed imaging setup. Measurement $\mathbf{z}_i$ is collected using pattern $\mathbf{w}_i$ for $i = 1, 2, \ldots n$, where n is the total number of image pixels. 21

2-3 The non-negative orthant cone and the polyhedral cone $c(\mathbf{W})$, assuming zero dark count. True reflectivity image (black) lies in the non-negative orthant. From one realization of data, we obtain a feasible reflectivity estimate (blue) using the matrix inverse solution. From another realization of data, the estimate (red) has negative entries and thus is not valid. 26

2-4 Covariance matrices of Hadamard inverse solution $\hat{\mathbf{x}}^{\text{inv}}$ for several sparsity levels of size 63 vector $\mathbf{x}$ (k is the number of non-zero entries). The non-zero support of $\mathbf{x}$ was randomly selected and non-zero entries were set to 100. 30

2-5 Simulated violation probability using Hadamard multiplexing (blue) and the derived theoretical lower bound (red) in Equation (2.16) vs. number of zeros. The reflectivity vector $\mathbf{x}$ has size 15. The non-zero support of signal $\mathbf{x}$ was randomly generated and its non-zero entries were all set to 100. 31

2-6 (Left) MSE vs. number of non-zeros n_0. (Right) RMSE vs. signal strength α, where $\mathbf{x} = \alpha \cdot \mathbb{1}_{n\times1}$ is a constant vector. 33

2-7 Reflectivity estimates of the MIT logo image (top) and the Shepp-Logan image (bottom) from traditional pixelwise ML and the proposed CML that uses multiplexed measurements. All images are sized 127 × 129. 36

2-8 Reflectivity estimates of the Mandrill image (top) and the cameraman image (bottom) using CPML. All images are sized 127 × 129. 37

3-1 Data acquisition model. Rate function $\lambda_{i,j}(t)$ of inhomogeneous Poisson process combining desired scene response and noise sources is shown for pixel (i,j). Here, $N = 3$ and $k_{i,j} = 2$. A noise photon (red) was detected after the second transmitted pulse at $t_{i,j}^{(1)}$, and a signal photon (blue) was detected after the third transmitted pulse at $t_{i,j}^{(2)}$. 42

3-2 Experimental setup for single-photon imaging. 51

3-3 Resolution test experiments. Reflectivity chart imaging (top) was done using $T_a = 300\,\mu$s and mean photon count of 0.48. Depth chart imaging (bottom) was done using $T_a = 6.2\,\mu$s, mean photon count of 1.1, and 33 % of the pixels missing data. The mean photon count was computed simply by averaging the number of photon counts at every pixel over the fixed acquisition time. 52

3-4 Experimental results for imaging natural scenes. 54

3-5 Comparison between our framework and LIDAR technology for 3D imaging. 55

3-6 Sample MSE images with 100 independent trials for two natural scenes. 56

3-7 Effect of dwell time T_a and signal-to-background ratio (SBR) on our 3D recovery method. For acquisition times of $100\,\mu$s and $50\,\mu$s, we calculated the mean photon count $k_{i,j}$ over all pixels to be 1.4 and 0.6, respectively. 57

3-8 Distributions of generalized Gaussian random variables for several values of p with fixed σ. 59

3-9 Comparison of MSE of time-delay recovery using single-photon detections for two different illumination pulses. (Left) Plot of $s_x(t)$, a Gaussian pulse, and $s_y(t)$, an arbitrary bimodal pulse. (Right) MSE of ML estimators vs. number of photon detections. 62

Chapter 1

Introduction

Modern optical imaging systems collect a large number of photon detections in order to suppress optical detector noise and form accurate images of object properties. For example, a commercially available digital camera provides the user with a clean photograph by detecting trillions of photons with the sensor array. Traditional imagers are thus required to operate with long acquisition times or at high light levels such that the total optical flux hitting the detector is sufficiently high.

In this thesis, we propose two computational imagers (one passive, one active) that can accurately recover images of scene properties, such as reflectivity and 3D structure, at low-flux regimes where the mean photon count per pixel can even be less than one. The core principle that allows high-quality imaging for both proposed imagers is the combination of the photodetection physics with the statistical properties of image representations. The proposed imagers are as follows.

1. **A single-pixel camera with high photon efficiency:** We propose a passive single-pixel reflectivity imager that can accurately estimate the scene reflectivity in photon-limited imaging scenarios by using an extension of the theory of multiplexed sensing. The proposed method, which makes use of the physically realistic constraint that reflectivity values are non-negative, disproves the accepted wisdom that reflectivity estimation based on multiplexed measurements always degrades imaging performance under low light-level Poisson noise of the

photodetector. We demonstrate through numerical experiments that the proposed imager outperforms the classical single-pixel camera that takes direct raster observations at low light levels.

2. **An active 3D imaging system with high photon efficiency:** We propose an active imager that will simultaneously recover scene reflectivity and 3D structure with high accuracy using a single-photon detector. The proposed imager allows high-quality 3D imaging by combining the accurate single-photon counting statistics with the prior information that natural scenes are spatially correlated. Experimental results show that the proposed imager outperforms state-of-the-art 3D imaging methods that incorporate denoising algorithms. We form accurate 3D and reflectivity images even when the mean number of photon detections at a pixel is close to 1. Also, we give an information-theoretic framework that can be used to choose optimal design parameters such as the illumination pulse shape, at least for the pixelwise imaging scenario.

Higher photon efficiency of an imaging system directly translates to shorter data acquisition time. For 3D imaging specifically, it can translate to lower illumination power and longer range imaging. Thus, the theory of highly photon-efficient imaging in this thesis opens up many interesting research areas for building **real-time** imaging systems that can robustly operate at **low-power** and under **low light-level** conditions.

The remainder of the thesis is organized as follows. In Chapter 2, we discuss the single-pixel camera and propose a computational imaging method that achieves high photon efficiency in estimating scene reflectivity. Chapter 3 describes the method of 3D imaging using an active illumination source and a single-photon detector, and proposes a highly photon-efficient computational 3D imaging method. Chapter 4 gives the conclusions of the two proposed imaging frameworks.

Bibliographical Note

Parts of Chapter 2 appear in the paper:

- **D. Shin**, A. Kirmani, and V. K. Goyal, "Low-rate Poisson intensity estimation using multiplexed imaging", *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 1364-1368.

Parts of Chapter 3 appear in the papers:

- **D. Shin**, A. Kirmani, V. K. Goyal, and J. H. Shapiro, "Computational 3D and reflectivity imaging with high photon efficiency", *Accepted for publication at IEEE International Conference on Image Processing*, 2014.
- **D. Shin**, A. Kirmani, A. Colaço, and V. K. Goyal, "Parametric Poisson process imaging", *Proceedings of IEEE Global Conference on Signal and Information Processing*, 2013, pp. 1053–1056.
- **D. Shin**, A. Kirmani, V. K. Goyal, and J. H. Shapiro, "Information in a photon: relating entropy and maximum-likelihood range estimation using single-photon counting detectors", *Proceedings of IEEE International Conference on Image Processing*, 2013, pp. 83–87.
- A. Kirmani, D. Venkatraman, **D. Shin**, A. Colaço, F. N. C. Wong, J. H. Shapiro, and V. K. Goyal, "First-photon imaging", *Science*, vol. 343, no. 6166, pp. 58–61, 2014.
- A. Kirmani, A. Colaço, **D. Shin**, and V. K. Goyal, "Spatio-temporal regularization for range imaging with high photon efficiency", *Proceeedings of SPIE*, 2013, vol. 8858, pp. 88581F.

Chapter 2

Single-Pixel Reflectivity Imaging with High Photon Efficiency

2.1 Prior Work

Passive optical imaging relies on an ambient light source to form an image of the scene reflectivity. Commercially available digital cameras are well-known examples of passive optical imagers. Since a two-dimensional sensor array is used for a camera, the reflectivity image is formed by measuring the total optical power hitting the photodetector over a finite acquisition time at every image pixel. The array allows the pixelwise measurements to be made in parallel.

Unlike cameras with a two-dimensional sensor array, the single-pixel camera [1] uses a single photodetector that takes pixelwise measurements not in parallel, but sequentially using a light modulator. The single-pixel camera has its advantages in size, complexity, and cost compared to an array-based camera. However, array based sensing is still preferred over single-pixel sensing when having a short acquisition time is important such as in real-time imaging.

The single-pixel camera is particularly useful in low light-level imaging scenarios, in which one is required to use single-photon avalanche detectors and photomultipliers, that are typically not available as arrays. Accurate formation of intensity images at low light levels is important in engineering applications such as astronomy [2], night vision [3], medical imaging applications such as positron emission tomography (PET), and imaging of light-sensitive biological and chemical samples [4]. In all of these applications, the low-light measurements are collected using single-photon detectors. The main challenge is that the measurements made using such detectors are inherently noisy due to low photon-count levels and are thus corrupted by signal-dependent photon noise that comes from the quantum nature of light detection. In this section, we survey several techniques for low light-level imaging using a single-pixel camera equipped with a photon-number resolving detector.

2.1.1 Classical Passive Reflectivity Imaging

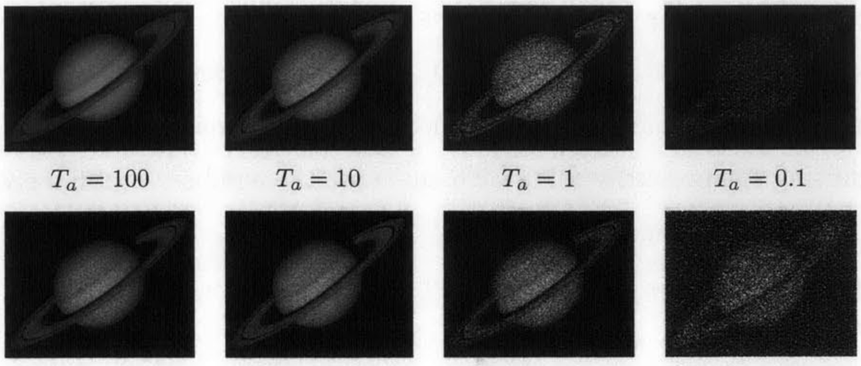
Let $\mathbf{x}$ be the ideal pixelated scene reflectivity of size $n \times 1$ that we are interested in imaging. If the scene is two-dimensional, then $\mathbf{x}$ is its vectorized version. Here, we include the effects of passive illumination and radial fall-off of optical flux in the reflectivity vector $\mathbf{x}$. At every image pixel, we use the photon-counting photodetector to record the number of photons detected in an acquisition time of T_a. Since we have a total of n image pixels in the single-pixel camera, the total image acquisition time using raster scanning then equals nT_a.

Let $\mathbf{y}_i$ be the photon count measurement that the photodetector makes at the i-th pixel. In the absence of illumination, our observations $\mathbf{y}_i$ are simply detector dark counts. We denote the dark count rate as $d \geq 0$. We assume that we know d exactly through a calibration process prior to the imaging experiment. The photon counting noise that corrupts our observations is known as shot noise and is well-modeled as Poisson distributed [5]. Our observation model at pixel i is then

$$\mathbf{y}_i \sim \text{Poisson}(T_a(\mathbf{x}_i + d)). \tag{2.1}$$

The probability mass function of our observation at pixel i is thus

$$\Pr[\mathbf{Y}_i = \mathbf{y}_i \,;\, \mathbf{x}_i, d] = \frac{\exp\{-T_a(\mathbf{x}_i + d)\}\,(T_a(\mathbf{x}_i + d))^{\mathbf{y}_i}}{\mathbf{y}_i\,!}, \tag{2.2}$$



$T_a = 100$ $T_a = 10$ $T_a = 1$ $T_a = 0.1$

$d = 0.1$ $d = 1$ $d = 10$ $d = 100$

Figure 2-1: Effect of shot noise on image quality for several acquisition times T_a (top) and dark count rate values d (bottom). The maximum reflectivity value of the saturn image is 25.

for $\mathbf{y}_i = 0, 1, 2, \ldots$. Thus, the maximum-likelihood (ML) pixelwise reflectivity estimate from our data $\mathbf{y} = [\mathbf{y}_1, \ldots, \mathbf{y}_n]^T$ is simply $\tilde{\mathbf{x}} = \mathbf{y}/T_a - \mathbf{d}$, where $\mathbf{d} = [d, \ldots, d]^T$. We observe that the ML estimate is obtained by normalizing and bias-correcting the raw photon count observations $\mathbf{y}$. The variance of the pixelwise reflectivity estimate $\tilde{\mathbf{x}}$ at the i-th pixel is $\mathrm{Var}(\tilde{\mathbf{x}}_i) = (\mathbf{x}_i + d)/T_a$. As shown in Figure 2-1, we observe that traditional pixelwise ML estimate using the single-pixel camera is limited to having a long acquisition time T_a and low dark count rate d to minimize the variance from shot noise and form high-quality images in photon-limited scenarios.

2.1.2 Single-Pixel Camera based on Compressive Sensing

Recently, the authors of [6] developed an architecture for a photon-efficient single-pixel camera that uses compressed sensing theory. The theory of compressed sensing [7] accurate image formation is guaranteed with high probability even when the number of measurements is less than the number of image pixels. Although it uses only one photodetector like the raster-scanned system considered above, the compressed sensing imaging setup has a major difference from that of traditional pixelwise imaging. Unlike traditional imaging which involves sensing one image pixel at a time and acquiring n measurements in total, compressive imaging acquires a sequence of m measurements by observing light from multiple pixels at the same time using predesigned coded patterns. Typically, the structured pattern of pixels can be achieved using a digital micromirror device (DMD). Experimental results in [8] show the high-quality reflectivity images which are formed using traditional scanning method can be obtained using compressed sensing with one tenth the number of total measurements, and thus with higher photon efficiency.

Let $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ be the set of m vectors, each of size $n \times 1$, describing the coded pattern of pixelwise measurements. By the optical flux conservation law, we see that every entry of $\mathbf{w}_i$ must be between 0 and 1, for $i = 1, 2, \ldots, m$. Then, the observation model for the i-th measurement taken by the compressive single-pixel camera is

$$\mathbf{z}_i \sim \text{Poisson}\left(T_a(\mathbf{w}_i^T\mathbf{x} + d)\right), \tag{2.3}$$

where $\mathbf{w}_i^T$ is the vector transpose of $\mathbf{w}_i$. We can also write the observation model in matrix-vector form:

$$\mathbf{z} \sim \text{Poisson}\left(T_a(\mathbf{W}\mathbf{x} + \mathbf{d})\right), \tag{2.4}$$

where now $\mathbf{W}$ is an $m \times n$ matrix that row-concatenates the m measurement patterns $\{\mathbf{w}_1^T, \ldots, \mathbf{w}_m^T\}$ and $\mathbf{z}$ is the measurement vector of size $m \times 1$. Then, the theory of robust compressed sensing [9] allows us to have $m < n$ and form accurate reflectivity images with high probability, given that the matrix $\mathbf{W}$ satisfies the restricted isometry property (RIP) and the prior information that the image is sparse in the transform-domain.

However, the authors of [10] proved that image estimation based on compressed sensing under low light conditions ($T_a(\mathbf{W}\mathbf{x} + \mathbf{d})_i < 1$ for all $i \in \{1, \ldots, m\}$) fails due to the signal-dependent nature of Poisson noise. Thus, at low light levels, it is instead preferable to rely on traditional pixelwise methods described in the previous section than compressive methods. In other words, although compressive single-pixel imaging achieves higher photon efficiency by making fewer total measurements at high light levels, the method has no performance guarantees when the total flux hitting the detector is low and when the effect of shot noise is significant.

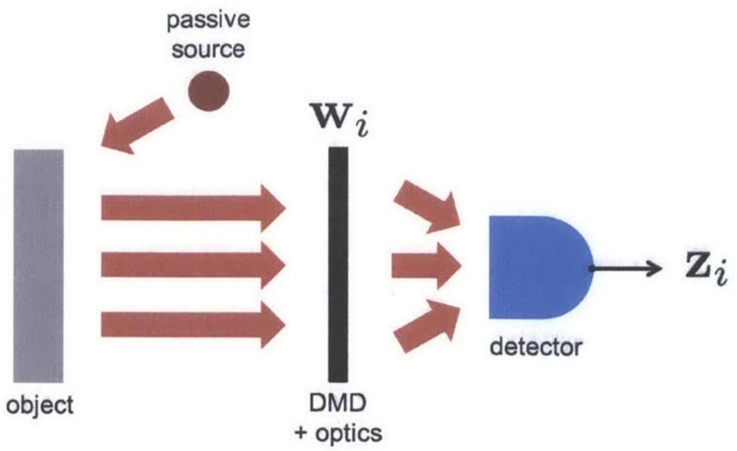


Figure 2-2: The multiplexed imaging setup. Measurement $\mathbf{z}_i$ is collected using pattern $\mathbf{w}_i$ for $i = 1, 2, \ldots n$, where n is the total number of image pixels.

2.1.3 Single-Pixel Camera based on Multiplexed Sensing

Multiplexed imaging is another powerful mechanism used to boost imaging performance when the observation noise is signal independent. Unlike compressed imaging, multiplexed imaging requires the number of measurements of coded patterns to be equal to the number of image pixels ($m = n$). The multiplexed imaging setup is shown in Figure 2-2. Unlike compressed imaging, the method of multiplexed imaging does not require any sparsity assumptions on the scene reflectivity and is thus a non-Bayesian imaging method.

It has been shown [6] that multiplexed imaging outperforms classical pixelwise imaging when the observations are corrupted by signal-independent noise. We emphasize that this is the case when we are operating with a simple photodiode detector at high light levels, where the effect of shot noise and dark current contribution is minimal. Our measurement vector $\mathbf{z}$ of size n can be assumed to be corrupted by additive Gaussian noise:

$$\mathbf{z} = T_a \mathbf{W} \mathbf{x} + \eta, \tag{2.5}$$

where $\mathbf{W} \in [0, 1]^{n \times n}$ is a square multiplexing matrix and η is a vector whose entries are independent and identically distributed (i.i.d.) zero-mean Gaussian random variables with variance σ^2. Note that if $\mathbf{W}$ is the identity matrix, then we have the observation model of traditional pixelwise imaging. Assuming that the multiplexing matrix $\mathbf{W}$ is non-identity and non-singular, we can decode the multiplexed measurements and estimate the reflectivity image by performing a simple matrix inversion. The matrix inversion estimate $\hat{\mathbf{x}}^{\text{inv}} = \mathbf{W}^{-1}(\mathbf{z}/T_a)$ is the traditional demultiplexing solution [11].

Researchers were interested in using multiplexed measurements instead of direct measurements, with the aim of reducing the mean-square error (MSE). For an estimator $\hat{\mathbf{x}}$ of $\mathbf{x}$, the MSE is defined as $\text{MSE}(\mathbf{x}, \hat{\mathbf{x}}) = \text{tr}\left(\mathbb{E}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T]\right)$, where $\text{tr}(\cdot)$ and $\mathbb{E}$ respectively denote the trace and expectation operators. The multiplexing gain $\mathcal{G}(\mathbf{W})$ associated with the multiplexing code $\mathbf{W}$ is defined as the reduction ratio in the root MSE from pixelwise imaging to multiplexed imaging:

$$\mathcal{G}(\mathbf{W}) = \sqrt{\frac{\text{MSE}(\mathbf{x}, \mathbf{y}/T_a)}{\text{MSE}(\mathbf{x}, \hat{\mathbf{x}}^{\text{inv}})}}. \tag{2.6}$$

We see that $\mathcal{G}(\mathbf{W}) > 1$ implies that multiplexed imaging using pattern $\mathbf{W}$ outperforms traditional imaging when the noise is additive Gaussian. In this case, we can write

$$\begin{aligned}
\mathcal{G}(\mathbf{W}) &= \sqrt{\frac{\text{tr}\left(\mathbb{E}[(\mathbf{x} - \mathbf{y}/T_a)(\mathbf{x} - \mathbf{y}/T_a)^T]\right)}{\text{tr}\left(\mathbb{E}[(\mathbf{x} - \hat{\mathbf{x}}^{\text{inv}})(\mathbf{x} - \hat{\mathbf{x}}^{\text{inv}})^T]\right)}} \\
&= \sqrt{\frac{\text{tr}(\mathbb{E}[\eta\eta^T])}{\text{tr}(\mathbb{E}[\mathbf{W}^{-1}\eta\eta^T\mathbf{W}^{-T}])}} \\
&= \sqrt{\frac{\text{tr}(\text{diag}([\sigma^2, \ldots, \sigma^2]))}{\text{tr}(\mathbf{W}^{-1}\text{diag}([\sigma^2, \ldots, \sigma^2])\mathbf{W}^{-T})}} \\
&= \sqrt{\frac{n}{\text{tr}\left((\mathbf{W}^T\mathbf{W})^{-1}\right)}},
\end{aligned} \tag{2.7}$$

where $\text{diag}(\mathbf{v})$ is a matrix that has $\mathbf{v}$ as its diagonal and zeros as its off-diagonal entries. Then, the optimal multiplexing pattern that maximizes the gain is obtained by solving the following optimization problem:

$$\begin{aligned} \max_{\mathbf{W}} \; & \sqrt{\frac{n}{\operatorname{tr}((\mathbf{W}^T\mathbf{W})^{-1})}} \\ \text{s.t. } & \mathbf{W}_{i,j} \in [0,1], \quad i,j = 1,\ldots,n \\ & |\mathbf{W}| \neq 0, \end{aligned} \tag{2.8}$$

where $|\mathbf{W}|$ is the determinant of $\mathbf{W}$. The optimization problem in (2.8) is non-convex in $\mathbf{W}$ and thus is a difficult problem to solve. However, with the extra constraints that $\mathbf{W}_{i,j} \in \{0,1\}$ and that every measurement should combine optical flux from exactly C pixels, previous works [12] proved that Hadamard multiplexing is optimal for certain values of n. The Hadamard multiplexing matrix $\mathbf{H}$ of size $n \times n$ can be constructed by deleting the first row and column of a Hadamard matrix of size $n+1$ and replacing 1's with 0's and -1's with 1's [11]. By construction, the Hadamard multiplexing matrix will have $C = (n+1)/2$ ones and $(n-1)/2$ zeros at every row. Also, because Hadamard matrices can only be constructed when n is a multiple of 4, $\mathbf{H}$ can only be constructed when $n \equiv 3 \pmod 4$. For example, a Hadamard multiplexing matrix of size 7 is

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

It is shown in [13] that the set of eigenvalues of a size-n $\mathbf{H}$ is

$$\lambda(\mathbf{H}) = \left\{ \frac{n+1}{2}, \underbrace{\sqrt{\frac{n+1}{4}}, \ldots \sqrt{\frac{n+1}{4}}}_{\# = (n-1)/2}, \underbrace{-\sqrt{\frac{n+1}{4}}, \ldots, -\sqrt{\frac{n+1}{4}}}_{\# = (n-1)/2} \right\}.$$

Using observation on the set of eigenvalues of a Hadamard multiplexer, the Hadamard multiplexing gain can then be computed to be $\mathcal{G}(\mathbf{H}) = (n+1)/(2\sqrt{n})$ [14]. For high n, we see that Hadamard multiplexed imaging gives an astounding $\sqrt{n}/2$-fold improvement over traditional pixelwise imaging at high light levels when observations are assumed to be corrupted by additive signal-independent noise. Intuitively, the Hadamard multiplexing matrix gives high multiplexing gain because C is a large number such that the signal-to-noise ratio of observations is high, and the rows of $\mathbf{H}$ are almost orthogonal so that the condition number of $\mathbf{H}$ is low.

2.2 Multiplexed Sensing under Poisson Noise

We are mainly interested in photon-limited imaging scenarios, in which the observations are corrupted by signal-dependent Poisson noise. Although Hadamard multiplexing followed by matrix inverse decoding gives a performance boost when the noise is signal independent, we will see that it severely degrades the image quality at low light levels when the shot noise effect is dominant. In the photon-limited imaging setup, the multiplexed observation vector $\mathbf{z}$ of size n obtained from the photon-counting Poisson channel model is

$$\mathbf{z} \sim \text{Poisson}(T_a(\mathbf{W}\mathbf{x} + \mathbf{d})) \tag{2.9}$$

Using the Hadamard multiplexing matrix $\mathbf{H}$, we calculate the multiplexed gain (again, for demultiplexing by code matrix inversion) under Poisson noise as

$$\mathcal{G}(\mathbf{H}) = \sqrt{\frac{\text{tr}\left(\mathbb{E}[(\mathbf{x} - (\mathbf{y}/T_a - \mathbf{d}))(\mathbf{x} - (\mathbf{y}/T_a - \mathbf{d}))^T]\right)}{\text{tr}\left(\mathbb{E}[(\mathbf{x} - \hat{\mathbf{x}}^{\text{inv}})(\mathbf{x} - \hat{\mathbf{x}}^{\text{inv}})^T]\right)}}, \tag{2.10}$$

where $\mathbf{y} \sim \text{Poisson}(T_a(\mathbf{x} + \mathbf{d}))$. Assuming zero dark-count contribution and using the eigenvalue properties of Hadamard multiplexers, we derive that the gain from using Hadamard multiplexing is simply $\mathcal{G}(\mathbf{H}) = \sqrt{(n+1)/(2n)}$ (Appendix A.1). Because $n > 1$ implies $\mathcal{G}(\mathbf{H}) < 1$, the matrix inverse estimate $\hat{\mathbf{x}}^{\text{inv}}$ from Hadamard multiplexed data under Poisson noise will always have MSE higher than that obtained from direct measurements. In fact, it has been numerically demonstrated [15] that general multiplexing methods fail to give MSE reduction in the presence of Poisson noise. Furthermore, in Appendix A.2, we prove that for circulant multiplexing matrices, which are useful for coded aperture imaging applications, multiplexing failure ($\mathcal{G}(\mathbf{W}) \leq 1$) is guaranteed.

One proposed solution to increase the multiplexing gain is to use Bayesian estimators that assume structural properties about the scene [16]. However, many imaging scenarios require one to make no assumptions about the scene and thus it is preferred to use non-Bayesian estimation methods. Hence, it is naturally advised that one should not use conventional multiplexing methods in photon-limited imaging scenarios.

We emphasize that the claims on multiplexing failure ($\mathcal{G}(\mathbf{W}) \leq 1$) under Poisson noise presented in this section only hold when demultiplexing is accomplished using the matrix inverse solution. In the following section, we demonstrate that it is possible to have multiplexing advantage even under Poisson noise, when we enforce the physically-realistic non-negativity constraint of reflectivity images in the estimation process.

2.3 Non-Negativity of Reflectivity

The reflectivity of a scene is always non-negative. However, we will observe that the matrix inverse solution $\hat{\mathbf{x}}^{\text{inv}}$ for demultiplexing may not always be non-negative. First, we see that the matrix inverse solution is equal to the ML estimate using multiplexed measurements under Poisson noise. For notational convenience, we write $\mathbf{x} \geq 0$ to state that every entry of $\mathbf{x}$ is non-negative. The ML estimate is

$$\begin{aligned}
\hat{\mathbf{x}}^{\text{ML}} &= \underset{\mathbf{x}:\ \mathbf{W}\mathbf{x}+\mathbf{d}\geq 0}{\arg\max} \prod_{i=1}^{n} \Pr[\mathbf{Z}_i = \mathbf{z}_i\,; \mathbf{W}, \mathbf{x}, \mathbf{d}], \\
&= \underset{\mathbf{x}:\ \mathbf{W}\mathbf{x}+\mathbf{d}\geq 0}{\arg\max} \prod_{i=1}^{n} \frac{\exp\{-T_a(\mathbf{W}\mathbf{x}+\mathbf{d})_i\}(T_a(\mathbf{W}\mathbf{x}+\mathbf{d})_i)^{\mathbf{z}_i}}{\mathbf{z}_i\,!}.
\end{aligned}$$

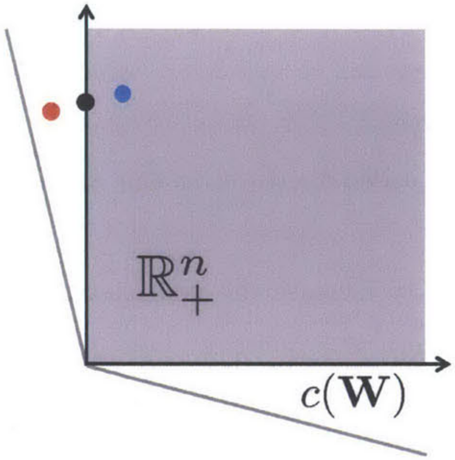


Figure 2-3: The non-negative orthant cone and the polyhedral cone $c(\mathbf{W})$, assuming zero dark count. True reflectivity image (black) lies in the non-negative orthant. From one realization of data, we obtain a feasible reflectivity estimate (blue) using the matrix inverse solution. From another realization of data, the estimate (red) has negative entries and thus is not valid.

We denote the log-likelihood $\mathcal{L}_{\mathbf{x}}(\mathbf{x};\mathbf{z}) = \sum_{i=1}^{n} \log \Pr[\mathbf{Z}_i = \mathbf{z}_i\,;\,\mathbf{W},\mathbf{x},\mathbf{d}]$ so that the ML estimate is equivalent to

$$\underset{\mathbf{x}:\ \mathbf{Wx}+\mathbf{d}\geq 0}{\arg\min} \quad -\mathcal{L}_{\mathbf{x}}(\mathbf{x};\mathbf{z}). \tag{2.11}$$

Due to the non-negativity of Poisson rate parameter $T_a(\mathbf{Wx}+\mathbf{d})$, we observe that the ML solution $\hat{\mathbf{x}}^{\mathrm{ML}}$ is constrained to be in the polyhedral set $c(\mathbf{W}) = \{\mathbf{v} : \mathbf{Wv}+\mathbf{d} \geq 0\}$. Because $\mathcal{L}_{\mathbf{x}}(\mathbf{x};\mathbf{z})$ is strictly convex (see Appendix A.3), if there is a solution $\hat{\mathbf{x}}$ that gives zero gradient ($\nabla_{\mathbf{x}} \mathcal{L}_{\mathbf{x}}(\hat{\mathbf{x}};\mathbf{z}) = 0$) and satisfies $\hat{\mathbf{x}} \in c(\mathbf{W})$, then it is the ML solution.

The gradient of $\mathcal{L}_{\mathbf{x}}(\mathbf{x};\mathbf{z})$ is given by

$$\nabla_{\mathbf{x}} \mathcal{L}_{\mathbf{x}}(\mathbf{x};\mathbf{z}) = -T_a \mathbf{W}^T \mathbb{1} + \sum_{i=1}^{n} \frac{z_i}{e_i^T(\mathbf{Wx}+\mathbf{d})} \mathbf{W}^T e_i, \tag{2.12}$$

where $\mathbb{1}$ is a size-n vector of ones and e_i is a size-n vector that has a single non-zero entry equal to one at i-th index. We observe that the gradient becomes zero at the point defined by the matrix inverse solution $\hat{\mathbf{x}}^{\text{inv}} = \mathbf{W}^{-1}(\mathbf{z}/T_a - \mathbf{d})$. Also, since

$$\mathbf{W}\hat{\mathbf{x}}^{\text{inv}} + \mathbf{d} = \mathbf{W}\mathbf{W}^{-1}(\mathbf{z}/T_a - \mathbf{d}) + \mathbf{d} = \mathbf{z}/T_a$$

and $\mathbf{z}$ is a Poisson random vector, we see that $\hat{\mathbf{x}}^{\text{inv}} \in c(\mathbf{W})$. Thus, the matrix inverse estimate is equal to the ML estimate.

The matrix inverse estimate must then be in the set $c(\mathbf{W})$. Also, because $\mathbf{x} \geq 0$ implies $\mathbf{W}\mathbf{x} + \mathbf{d} \geq 0$, the non-negative orthant cone is always contained in $c(\mathbf{W})$. This implies that there is a non-zero probability of the matrix inverse estimate $\hat{\mathbf{x}}^{\text{inv}}$ being an infeasible reflectivity estimate by having negative entries. Figure 2-3 shows an example of feasible and infeasible reflectivity estimates contained in the polyhedral cone $c(\mathbf{W})$ given $d = 0$.

2.4 Novel Image Formation

In the previous section, we observed that the traditional demultiplexing method based on the matrix inverse solution can give invalid reflectivity estimates that have negative entries. Thus, using the physically realistic constraint that reflectivity values are non-negative, we propose to solve for the non-negatively constrained maximum-likelihood (CML) reflectivity estimate using multiplexed measurements $\mathbf{z}$:

$$\hat{\mathbf{x}}^{\text{CML}} = \underset{\mathbf{x}:\ \mathbf{x} \geq 0}{\arg\min} \sum_{i=1}^{n} \left[(T_a(\mathbf{W}\mathbf{x} + \mathbf{d}))_i - \mathbf{z}_i \log\,(T_a(\mathbf{W}\mathbf{x} + \mathbf{d}))_i \right]. \tag{2.13}$$

Due to the non-negativity constraint, the solution to the constrained maximum-likelihood optimization problem no longer has a closed-form solution. However, because the non-negative orthant is a convex set and the cost function is also a convex function in $\mathbf{x}$, we can use a simple projected gradient algorithm to solve the optimization problem for a global minimum solution [17]. Starting at an initial guess of solution $\mathbf{x}^{(0)}$, we iterate

$$\mathbf{x}^{(k+1)} = \max\left\{\mathbf{x}^{(k)} + \alpha^{(k)}\left(\mathbf{W}^T\mathbb{1} - \sum_{i=1}^{n}\frac{\mathbf{z}_i}{e_i^T(\mathbf{Wx}+\mathbf{d})}\mathbf{W}^T e_i\right),\ 0\right\}, \tag{2.14}$$

where $\alpha^{(k)}$ is the step size chosen at k-th iteration and the maximum operator acts entrywise. The solution at convergence is $\hat{\mathbf{x}}^{\mathrm{CML}}$. It is also possible to use the log-barrier method to enforce the non-negativity constraint, so that we only use a pure descent algorithm to solve for the CML estimate:

$$\min_{\mathbf{x}} \sum_{i=1}^{n} \left[(T_a(\mathbf{Wx}+\mathbf{d}))_i - \mathbf{z}_i \log\,(T_a(\mathbf{Wx}+\mathbf{d}))_i - \lambda \log \mathbf{x}_i\right], \tag{2.15}$$

for a sufficiently small value of $\lambda \geq 0$.

We would like to know when does the proposed CML demultiplexing solution outperform the traditional pixelwise imaging methods and give us a multiplexing advantage. First, we want to understand when does the CML solution diverge away from the traditional matrix inverse demultiplexing solution, which is simply ML. In other words, when does the non-negativity constraint become active in the CML estimation?

2.4.1 When is the Non-Negativity Constraint Active?

The non-negativity constraint becomes active in the constrained solution, only when the unconstrained solution, which is the matrix inverse solution, violates the constraint. We can write the probability mass function of the matrix inverse solution $\hat{\mathbf{x}}^{\mathrm{inv}}$ as

$$\Pr[\hat{\mathbf{X}} = \hat{\mathbf{x}}\,;\, \mathbf{W},\mathbf{x},\mathbf{d}] = \frac{T_a}{|\mathbf{W}^{-1}|}\prod_{i=1}^{n}\frac{(T_a(\mathbf{Wx}+\mathbf{d}))_i^{(T_a(\mathbf{W}\hat{\mathbf{x}}+\mathbf{d}))_i}\exp\{-(T_a(\mathbf{Wx}+\mathbf{d}))_i\}}{(T_a(\mathbf{W}\hat{\mathbf{x}}+\mathbf{d}))_i\,!},$$

for $\hat{\mathbf{x}} \in \{\mathbf{W}^{-1}(\mathbf{z}/T_a - \mathbf{d}) : \mathbf{z} \in \mathbb{Z}_+^n\}$. We are interested in solving for the probability of non-negativity constraint violation (the probability of at least one entry of the matrix inverse solution being negative): $\Pr[\text{Violation}] = 1-\sum_{\hat{\mathbf{x}}\geq 0} \Pr[\hat{\mathbf{X}} = \hat{\mathbf{x}}\,;\, \mathbf{W}, \mathbf{x}, \mathbf{d}]$. If the violation probability is high, then the degree to which the constrained ML estimate diverges away from the unconstrained ML solution by having physically-realistic non-negativity corrections is also high. Thus, we are using the violation probability as a proxy for reduction in MSE.

In Appendix A.4, we show that the matrix inverse solution is in fact an unbiased estimator that achieves minimum mean-square error. Thus, if the CML estimate eventually has lower MSE compared to traditional multiplexed imaging method based on ML, then it must be that the non-negativity constraint introduces bias in the CML estimate and creates a bias-variance tradeoff.

Calculation of the violation probability using the probability mass function can only be done using Monte Carlo methods. Thus, we instead try to derive a lower bound on the violation probability and how it relates to the signal $\mathbf{x}$. In order to construct a lower bound on the violation probability, we first give the following definition.

Definition 1. *A size-n continuous random vector X is symmetric i.i.d. if*

$$p_X(x) = \frac{1}{Z}\prod_{i=1}^{n} f(x_i),$$

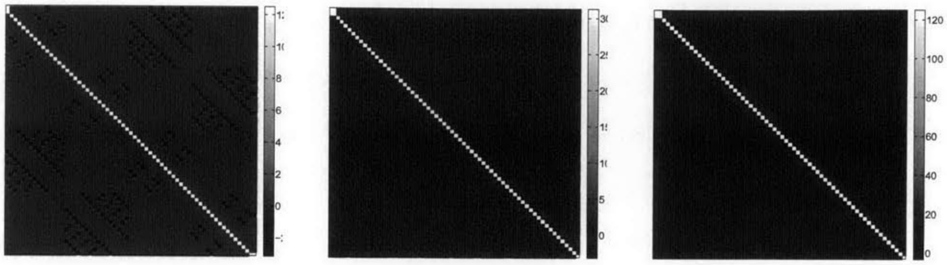
where $f(\cdot)$ is a positive symmetric function and Z is a normalization factor so that $p_X(x)$ is a probability density function of X.

Using the previous definition, we can lower bound the violation probability as follows, assuming zero dark count rate.

Remark 1. *If the matrix inverse solution is unbiased, continuous, and symmetric i.i.d., then*

$$\Pr[\text{violation}] \geq \sum_{k=1}^{n_0} \frac{1}{2^k}, \tag{2.16}$$

where $n_0 \geq 1$ is the number of zeros in the signal vector $\mathbf{x}$.



(a) $k = 5$ (b) $k = 10$ (c) $k = 40$

Figure 2-4: Covariance matrices of Hadamard inverse solution $\hat{\mathbf{x}}^{\text{inv}}$ for several sparsity levels of size 63 vector $\mathbf{x}$ (k is the number of non-zero entries). The non-zero support of $\mathbf{x}$ was randomly selected and non-zero entries were set to 100.

The proof of Remark 1 is given in Appendix A.5. Remark 1 simply states that if the matrix inverse solution is unbiased, continuous, and symmetric i.i.d., then the violation probability increases as the number of zeros in $\mathbf{x}$ increases. Note that it is indeed possible to assume that $\hat{\mathbf{x}}^{\text{inv}}$ is a continuous random vector in certain cases. For example, if we assume that the non-zero entries of $\mathbf{x}$ are arbitrarily large, then the Poisson random vector $\mathbf{z}$ can be approximated as a Gaussian random vector.

2.4.2 Choosing the Multiplexing Pattern

We saw in Remark 1 that the violation probability is related to signal sparsity, when the matrix inverse demultiplexing solution is symmetric i.i.d. Can we construct a multiplexing matrix such that the matrix inverse solution satisfies the symmetric i.i.d. condition? The following remark will imply that the Hadamard matrix inverse solution is approximately symmetric i.i.d. given that the reflectivity signal $\mathbf{x}$ is a constant vector with arbitrarily large entries.

Remark 2. *If the multiplexing pattern is determined by the Hadamard multiplexer* $\mathbf{H}$, *then the matrix inverse solution has a covariance matrix* Σ *with the following properties.*

1. *If* $i = j$, *then* $\Sigma_{i,j} = \frac{1}{T_a}\left(\frac{2}{n+1}\right)^2 \left(\sum_{i=1}^{n}(\mathbf{H}\mathbf{x} + \mathbf{d})_i\right)$,

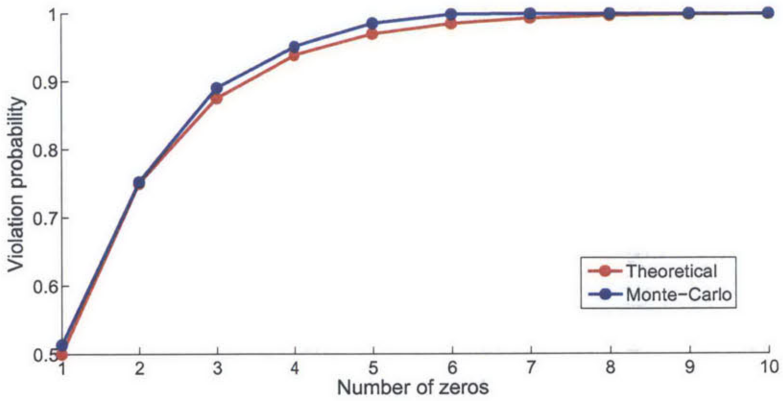


Figure 2-5: Simulated violation probability using Hadamard multiplexing (blue) and the derived theoretical lower bound (red) in Equation (2.16) vs. number of zeros. The reflectivity vector $\mathbf{x}$ has size 15. The non-zero support of signal $\mathbf{x}$ was randomly generated and its non-zero entries were all set to 100.

2. If $i \neq j$, then

$$|\Sigma_{i,j}| \leq \max_{S_0 \in S} \left| \frac{1}{T_a} \left(\frac{2}{n+1} \right)^2 \left(\sum_{k_1 \in S_0} (\mathbf{Hx} + \mathbf{d})_{k_1} - \sum_{k_2 \in \{1,\ldots,n\} \setminus S_0} (\mathbf{Hx} + \mathbf{d})_{k_2} \right) \right|,$$

where $S = \{T \mid T \subset \{1, 2, \ldots, n\}, |T| = (n-1)/2\}$.

We give the proof of Remark 2 in Appendix A.6. Remark 2 tells us that the covariance matrix of Hadamard demultiplexing solution $\hat{\mathbf{x}}^{\text{inv}}$ has large diagonal entries and small off-diagonal entries, given that $\mathbf{x}$ is close to being a constant vector. Figure 2-4 shows that the covariance matrix of $\hat{\mathbf{x}}^{\text{inv}}$ indeed becomes more diagonally dominant as the number of non-zero entries of $\mathbf{x}$ increases.

For example, let $\mathbf{x} = \alpha \mathbb{1}_{n \times 1}$ be a constant vector without any zero entries, where $\alpha > 0$. Assume zero dark counts. Then, $\mathbf{Hx} = \alpha(\frac{n+1}{2})\mathbb{1}_{n \times 1}$ and

1. $\Sigma_{i,j} = \dfrac{\alpha}{T_a} \left(\dfrac{2}{n+1} \right)^2 \left(\dfrac{n(n+1)}{2} \right)$, for $i = j$,

2. $\Sigma_{i,j} = \dfrac{\alpha}{T_a} \left(\dfrac{2}{n+1} \right)^2 \left(\dfrac{n+1}{2} \right)$, for $i \neq j$.

As $n \to +\infty$, we observe that the diagonal entries of the Σ converge to $2\alpha/T_a$ while the off-diagonal entries converge to 0. We then observe that $\hat{\mathbf{x}}^{\text{inv}}$ is symmetric i.i.d., if $n \to +\infty$ so that the limiting covariance matrix is diagonal and $\alpha \to +\infty$ so that $\hat{\mathbf{x}}^{\text{inv}}$ is assumed to be Gaussian distributed.

Given that $\mathbf{x}$ is a large constant vector, we saw that the Hadamard matrix inverse solution can be approximated to be continuous and symmetric i.i.d., and thus obeys the violation probability bound in Equation (2.16). We would like to know how valid is the bound when $\mathbf{x}$ is non-constant and $\hat{\mathbf{x}}^{\text{inv}}$ may not necessarily be symmetric i.i.d. The numerical result in Figure 2-5 shows that the probability of violation when using Hadamard multiplexing is in fact well-described by the lower bound for various signal sparsity levels, at least when the non-zero entries of $\mathbf{x}$ are large and constant.

2.5 Numerical Experiments

2.5.1 Comparison of Mean-Square Errors

We performed numerical experiments to compare the MSE's of traditional pixelwise imaging, multiplexed imaging based on ML matrix inverse solution, and our proposed CML multiplexed imaging. In all our multiplexed imaging experiments, we use the Hadamard multiplexing strategy. Also, we set d to be zero.

The plot on the left of Figure 2-6 shows how the MSE values of estimators depend on the number of zeros in signal $\mathbf{x}$. We set $T_a = 1$. In this experiment, we chose $\mathbf{x}$ to be a vector of size 15. The non-zero support of signal $\mathbf{x}$ was randomly generated and its non-zero entries were all set to 100. As we expected, the ML demultiplexing estimate using matrix inverse performs worse than the pixelwise imaging estimate for all sparsity levels. We also confirm our derivations in Section 2.4.1 that high signal sparsity implies high violation probability, since, as the number of zeros increases, our CML solution diverges away from the ML demultiplexing solution. When the number of non-zero entries in $\mathbf{x}$ is small enough, the CML solution outperforms traditional pixelwise imaging. Contrary to popular belief, we demonstrated that it is possible to achieve multiplexing gain even under Poisson noise in a non-Bayesian way by simply including physically accurate constraints.

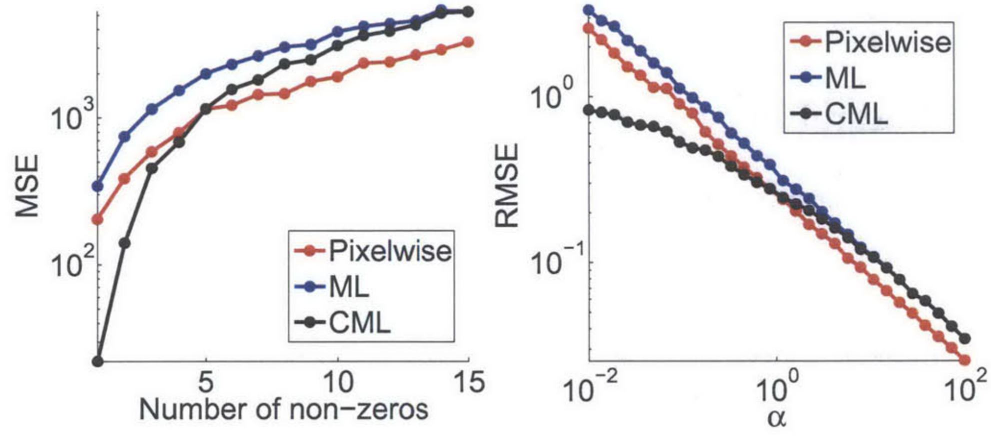


Figure 2-6: (Left) MSE vs. number of non-zeros n_0. (Right) RMSE vs. signal strength α, where $\mathbf{x} = \alpha \cdot \mathbb{1}_{n\times 1}$ is a constant vector.

The plot on the right of Figure 2-6 shows that the signal strength is another factor affecting the violation probability and thus the relative MSE (RMSE), which is defined as

$$\text{RMSE}(\mathbf{x}, \hat{\mathbf{x}}) = \frac{\sqrt{\text{MSE}(\mathbf{x}, \hat{\mathbf{x}})}}{\sum_{i=1}^{n} \mathbf{x}_i}. \tag{2.17}$$

In this experiment, we set the true reflectivity image to be a constant vector $\mathbf{x} = \alpha \cdot \mathbb{1}_{n\times 1}$ of size 15, that is controlled by the parameter $\alpha > 0$. We set $T_a = 1$.

From the definition of RMSE, the RMSE expressions are linear in α, on a log-log scale, for pixelwise imaging and ML demultiplexing methods, and this is observed in the plot. We confirm the failure of traditional ML multiplexing methods compared to pixelwise imaging for all values of α. When $\mathbf{x}$ is a vector with large entries, the violation probability is low and the RMSE of CML is also linear in α on a log-log scale. In this high light-level regime, the CML solution converges to the ML solution and both ML and CML multiplexed imaging methods perform worse than the pixelwise imaging method. However, if $\mathbf{x}$ is described by small entries (low light-level), then we observe that the non-negativity constraint is activated and the CML solution outperforms pixelwise imaging [18]. Thus, in photon-limited imaging scenarios, even when the signal does not have zero entries, multiplexing gain come from the signal having low reflectivity and violating non-negativity conditions.

2.5.2 Natural Images

Figure 2-7 compares the performances of pixelwise imaging and CML multiplexed imaging for two different reflectivity vectors. For all imaging experiments in this section, we set $T_a = 1$ and $d = \sum_{i=1}^{n} \mathbf{x}_i / n$. The quality of image estimate $\hat{\mathbf{x}}$ is quantified by the peak signal-to-ratio (PSNR):

$$\mathrm{PSNR}(\mathbf{x}, \hat{\mathbf{x}}) = 10 \log_{10} \left(\frac{\max\limits_{i=1,\ldots,n} \mathbf{x}_i^2}{\frac{1}{n} \sum_{i=1}^{n} (\mathbf{x}_i - \hat{\mathbf{x}}_i)^2} \right). \tag{2.18}$$

Also, the error image for estimate $\hat{\mathbf{x}}$ of $\mathbf{x}$ is computed as $|\hat{\mathbf{x}} - \mathbf{x}|$, where all operators are entrywise.

We observe that, for the MIT logo image, the CML estimate gives a PSNR boost of 4.7 dB over the classical pixelwise estimate. For the Shepp-logan phantom image, CML estimate boosts PSNR by 4.3 dB. Also, by comparing the error maps for each estimate, we see the superior imaging quality of CML.

So far, we have demonstrated that multiplexing gain comes from using a non-Bayesian demultiplexer that is physically accurate due to the non-negativity constraint. It is possible to improve our multiplexed CML estimation performance by incorporating the prior knowledge that natural scenes are spatially correlated. We propose the constrained and penalized maximum likelihood (CPML) estimate, which is now Bayesian, as the following:

$$\hat{\mathbf{x}}^{\mathrm{CPML}} = \arg\min_{\mathbf{x}:\ \mathbf{x} \geq 0} \sum_{i=1}^{n} \left(T_a(\mathbf{W}\mathbf{x} + \mathbf{d})_i - \mathbf{z}_i \log T_a(\mathbf{W}\mathbf{x} + \mathbf{d})_i \right) + \beta \,\mathrm{pen}(\mathbf{x}), \tag{2.19}$$

where $\mathrm{pen}(\mathbf{x})$ is a function that penalizes the non-smoothness of the image estimate over pixels and β is the regularization parameter controlling the strength of the penalty. If $\mathrm{pen}(\mathbf{x})$ is a convex function in $\mathbf{x}$, then the optimization problem is globally convex and can be solved using computationally efficient first-order gradient methods. Popular penalty functions used for accurate image representations are the ℓ_1-norm of wavelet transform and the total variation semi-norm [19].

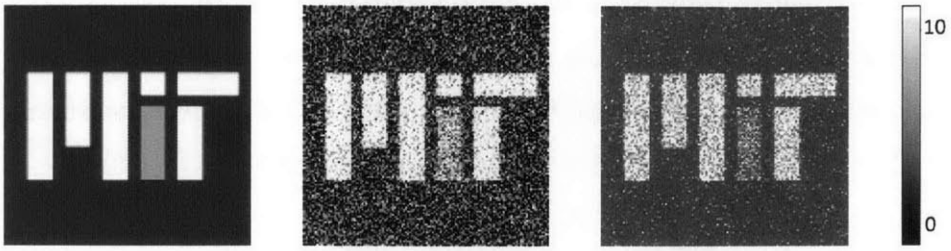
In Figure 2-8, we compare the accuracy of CPML estimation with the state-of-the-art Poisson-denoised pixelwise estimation $\hat{\mathbf{x}}^{\text{den}}$ [20] from direct measurements $\mathbf{y}$.

$$\hat{\mathbf{x}}^{\text{den}} = \arg\min_{\mathbf{x}} \sum_{i=1}^{n} \left(T_a(\mathbf{x}+\mathbf{d})_i - \mathbf{y}_i \log T_a(\mathbf{x}+\mathbf{d})_i \right) + \beta\, \text{pen}(\mathbf{x}). \tag{2.20}$$

In this experiment, for both $\hat{\mathbf{x}}^{\text{den}}$ and $\hat{\mathbf{x}}^{\text{CPML}}$, we use the total variation (TV) semi-norm penalty function

$$\text{pen}(\mathbf{x}) = \|\mathbf{x}\|_{\text{TV}} = \sum_{i=1}^{n} \sum_{j \in N(i)} |\mathbf{x}_i - \mathbf{x}_j|, \tag{2.21}$$

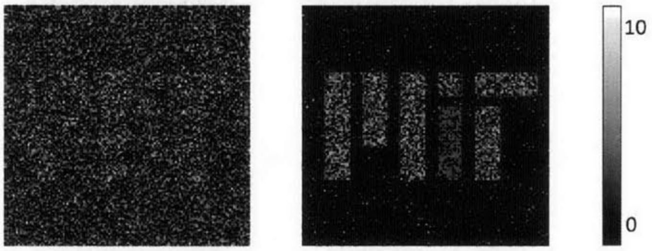
where $N(i)$ is the set of four neighboring pixels of pixel i in the two-dimensional image domain. The TV penalty function allows us to model the spatial smoothness of image and sparsity of edges. Their respective regularization parameters are chosen to maximize PSNR. We see that, for the Mandrill image, the CPML estimate (Figure 2-8 (d)) gives a PSNR boost of 12.9 dB over the baseline pixelwise estimation method (Figure 2-8 (b)) and 6.4 dB over the Poisson denoising method (Figure 2-8 (c)) applied to the pixelwise estimate. For the cameraman image, the PSNR of the CPML estimate (Figure 2-8 (k)) is 17 dB higher than that of the pixelwise estimation (Figure 2-8 (i)) and 5.9 dB higher than that of the Poisson denoised estimate (Figure 2-8 (j)). Also, by comparing the error images, we see the relative high photon efficiency of the CPML estimate.



(a) Ground truth

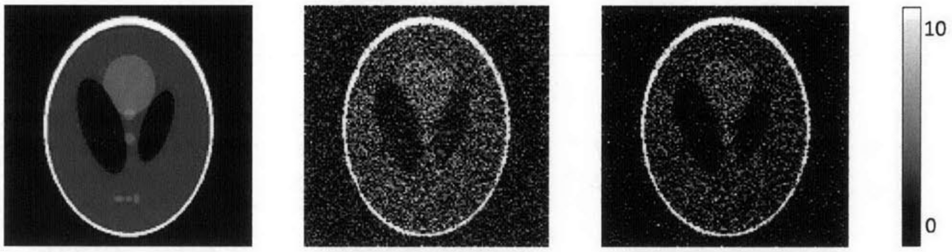
(b) Pixelwise
PSNR = −3.2 dB

(c) CML
PSNR = 1.5 dB



(d) Error map of (b)

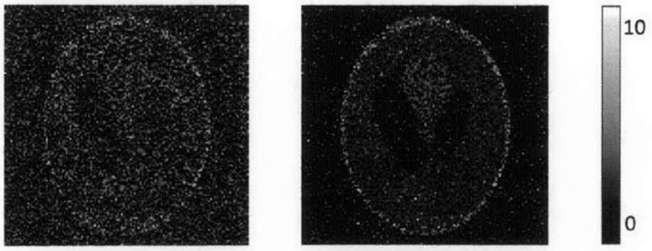
(e) Error map of (c)



(f) Ground truth

(g) Pixelwise
PSNR = 3.5 dB

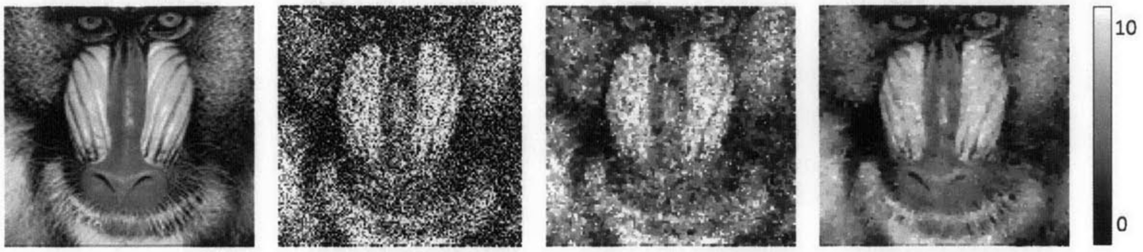
(h) CML
PSNR = 7.8 dB



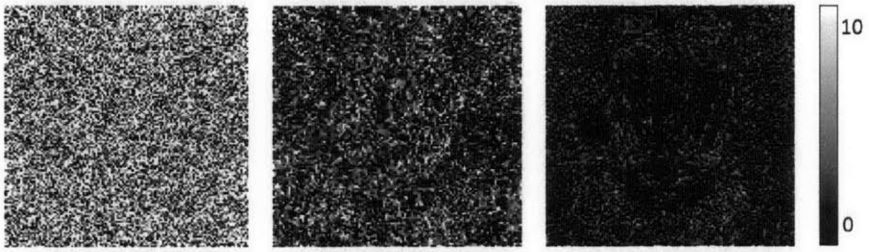
(i) Error map of (g)

(j) Error map of (h)

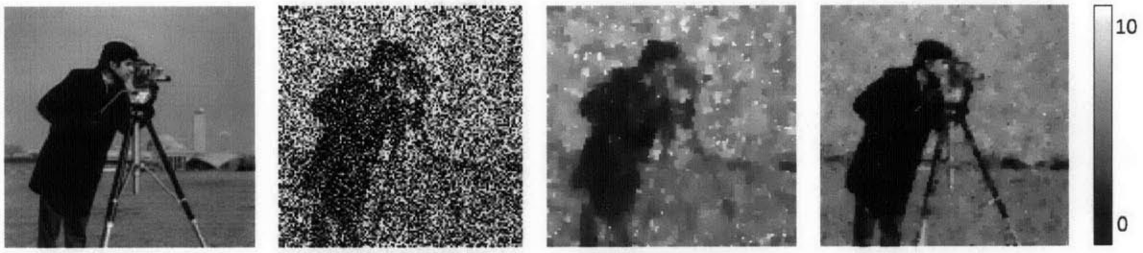
Figure 2-7: Reflectivity estimates of the MIT logo image (top) and the Shepp-Logan image (bottom) from traditional pixelwise ML and the proposed CML that uses multiplexed measurements. All images are sized 127 × 129.



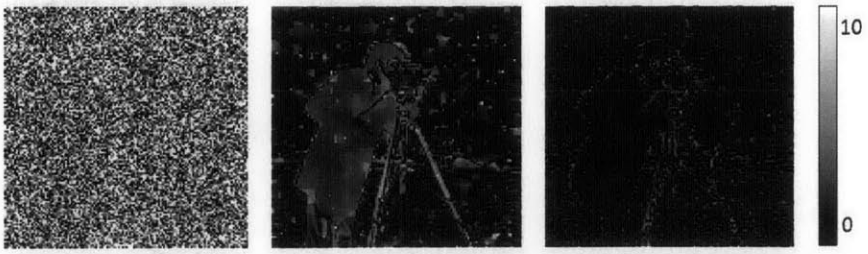
(a) Ground truth (b) Pixelwise PSNR = −8.1 dB (c) Denoised (b) PSNR = −1.6 dB (d) CPML PSNR = 4.8 dB



(e) Error map of (b) (f) Error map of (c) (g) Error map of (d)



(h) Ground truth (i) Pixelwise PSNR = −8.6 dB (j) Denoised (i) PSNR = 2.5 dB (k) CPML PSNR = 8.4 dB



(l) Error map of (i) (m) Error map of (j) (n) Error map of (k)

Figure 2-8: Reflectivity estimates of the Mandrill image (top) and the cameraman image (bottom) using CPML. All images are sized 127 × 129.

Chapter 3

Active 3D Imaging with High Photon Efficiency

3.1 Prior Work

We can acquire 3D structure and reflectivity of a scene using an active imager – one that supplies its own illumination. When the optical flux incident at the detector is high, the shot noise effect is minimal. As the mean count of the flux reaching the detector approaches a few photons, 3D and reflectivity images degrade in quality. In this section, we study how active optical 3D imaging is traditionally done at low-flux regimes.

3.1.1 Classical Active 3D Imaging

Active optical imaging systems differ in how they modulate their illumination. Modulating intensity temporally enables distance measurement by the time-of-flight (ToF) principle. Ordered by increasing modulation bandwidth (shorter pulses), these include: homodyne ToF sensing, pulsed ToF cameras [21], and picosecond laser radar systems [22]. Methods that modulate light spatially include speckle decorrelation imaging, structured light [23], and active stereo imaging [24]. Active 3D imaging methods using spatial light modulation have low photon efficiency because they operate with an always-on optical output. On the other hand, pulsed ToF systems achieve millimeter-accurate sensing using optical output that is activated only for short intervals. Among these, ToF imagers using single-photon avalanche diode (SPAD) detectors have the highest photon efficiency.

Optoelectronic techniques in the low-flux regime: In the low-flux regime, the robustness of imaging technique can be improved using optoelectronic techniques. For example, active ToF imagers use light sources, such as lasers, with narrow spectral bandwidth and spectral filters to suppress ambient background light and dark current. However, optical filtering methods also attenuate signal as well as noise. Range-gated imaging [25] is another popular technique that increases SNR by activating the detector selectively in time. However, range-gating requires a priori knowledge of the approximate object location. A SPAD detector may be replaced with a superconducting nanowire detector (SNSPD) [26], which is much faster, has lower timing jitter, and has lower dark count rates than a typical SPAD. However, SNSPDs have much smaller active areas, and hence have narrower fields-of-view.

Image denoising: When imaging using a SPAD detector in the low-flux regime, it is typical to first obtain a noisy pixelwise estimate of scene depth using photon arrival data, and then apply image denoising methods. This two-step approach usually assumes a Gaussian noise model [20], which is appropriate for high-flux scenarios. At low light levels, denoising is more challenging due to the signal-dependent nature of the noise.

3.1.2 First-Photon Imaging

First-photon imaging (FPI), recently proposed in [27], is a computational imaging framework that allows accurate 3D and reflectivity reconstruction using only the first detected photon at every pixel obtained by raster-scanning the scene. It combines the first-photon arrival statistics with spatial correlations existing in natural scenes for robust low light-level imaging. The statistics of first photon arrival derived in [27, 28, 29] are drastically different than traditional Gaussian noise models in LIDAR. Thus, the FPI framework can allow accurate 3D imaging in the first-photon regime, where traditional denoising algorithms fail due to the inaccuracy in modeling noise.

However, the main limitation of the first-photon imaging framework is that it is limited to a raster-scanning setup, in which the data acquisition time at each pixel is random. Thus, it does not extend naturally to operation using sensor arrays, which employ fixed exposure times, and with which image acquisition time can be greatly reduced in comparison with raster-scanned, single-detector systems. In this chapter, we demonstrate highly photon efficient 3D and reflectivity imaging when the pixelwise dwell time is fixed, thereby opening up the possibility of robust SPAD-array-based imaging under low light-levels and short exposure times. We compare our proposed imaging technique with the state-of-the-art image denoising methods that use sparsity-promoting regularization.

3.2 Single-Photon Imaging Setup

Figure 3-1 shows the signal acquisition model using a pulsed light source and a single SPAD detector. Our aim is to form reflectivity and depth images $\mathbf{x}, \mathbf{z} \in \mathbb{R}_{+}^{n \times n}$ of the scene. We index the scene pixels as (i, j), where $i, j = 1, \ldots, n$. The distance to patch (i, j) is denoted by $\mathbf{z}_{i,j} \geq 0$ and the patch reflectivity is denoted by $\mathbf{x}_{i,j} \geq 0$, including the effect of radial fall-off, view angle, and material properties.

3.2.1 Active Illumination

We use an intensity-modulated light source that illuminates the scene in a raster scanning fashion. This source emits a pulse train with a repetition period of T_r seconds. As shown in Figure 3-1, we reset the clock to 0 at the start of every period of pulse illumination for notational convenience. The photon-flux pulse shape $s(t)$ has units counts/sec (cps). In order to avoid distance aliasing, we assume $T_r > 2\mathbf{z}_{\max}/c$, where $\mathbf{z}_{\max}$ is the maximum scene range and c is the speed of light. With conventional processing, the root mean square (RMS) pulse width T_p governs the achievable depth

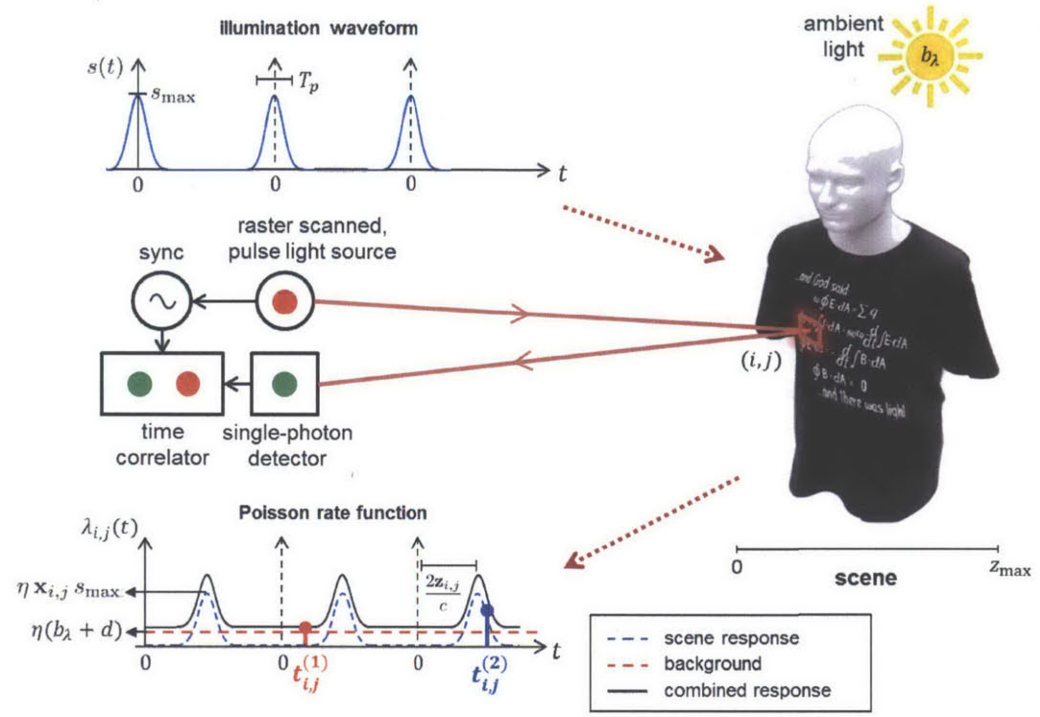

Illumination waveform
$s(t)$
$s_{\max}$
T_p
0
0
0
t
ambient light
b_λ
sync
raster scanned, pulse light source
time correlator
single-photon detector
(i,j)
Poisson rate function
$\lambda_{i,j}(t)$
$\eta\,\mathbf{x}_{i,j}\,s_{\max}$
$\eta(b_\lambda + d)$
$\frac{2\mathbf{z}_{i,j}}{c}$
$t_{i,j}^{(1)}$
$t_{i,j}^{(2)}$
0
scene
$z_{\max}$
scene response
background
combined response

Figure 3-1: Data acquisition model. Rate function $\lambda_{i,j}(t)$ of inhomogeneous Poisson process combining desired scene response and noise sources is shown for pixel (i, j). Here, $N = 3$ and $k_{i,j} = 2$. A noise photon (red) was detected after the second transmitted pulse at $t_{i,j}^{(1)}$, and a signal photon (blue) was detected after the third transmitted pulse at $t_{i,j}^{(2)}$.

resolution in the absence of background light [30]. As typically done in range imaging, we assume that $T_p \ll 2\mathbf{z}_{\max}/c < T_r$.

3.2.2 Detection

A SPAD detector provides time-resolved single-photon detections [31]. Its quantum efficiency η is the fraction of photons passing through the pre-detection optical filter that are detected. Each detected photon is time stamped within a time bin of duration Δ measuring a few picoseconds. Then, as it is typical for a LIDAR system, we have $\Delta \ll T_p \ll 2\,\mathbf{z}_{\max}/c$. For theoretical derivations, we assume that the exact photon detection time is available at each pixel. When the detector records a photon arrival, it becomes inactive for a period of time called the reset time or dead time. We will assume that the detector is active at the start of each illumination period, i.e. immediately after the transmission of each laser pulse, regardless of whether a photon was detected in the previous illumination period.

3.2.3 Data Acquisition

Each patch (i, j) is illuminated with a total of N light pulses. The pixelwise data acquisition time is then $T_a = NT_r$ seconds. We record the total number of observed photon detections $k_{i,j}$, along with their set of photon arrival times $\mathcal{T}_{i,j} = \{t_{i,j}^{(1)}, t_{i,j}^{(2)}, \ldots, t_{i,j}^{(k_{i,j})}\}$ at each pixel. If $k_{i,j} = 0$, then $\mathcal{T}_{i,j} = \varnothing$. Also, modeling realistic imaging scenarios, we assume background light with flux b_λ at the operating optical wavelength λ.

Measurement uncertainty in the photon arrival time results from:

- *Background light*: Ambient light at the operating wavelength causes photon detections unrelated to the scene.

- *Dark counts*: Detection events can occur at times when there is no light incident on the detector.

- *Pulse width*: The illumination pulse has a non-zero width because the modulation bandwidth cannot be infinite. Thus, the timing of a detected photon could correspond to the leading edge of the pulse, the trailing edge, or anywhere in between. This uncertainty translates to error in depth estimation.

Accounting for these characteristics is central to our contribution, as described in the following section.

3.3 Observation Model

Illuminating a scene pixel (i, j) with intensity-modulated light pulse $s(t)$ results in backreflected light signal $r_{i,j}(t) = \mathbf{x}_{i,j}\, s(t - 2\,\mathbf{z}_{i,j}/c) + b_\lambda$ at the detector.

3.3.1 Poisson Detection Statistics

The quantum nature of light in our setup (Figure 3-1) is correctly accounted for by taking the counting process at the SPAD output to be an inhomogeneous Poisson process with rate function

$$\lambda_{i,j}(t) = \eta\, r_{i,j}(t) + d = \eta\, \mathbf{x}_{i,j}\, s(t - 2\,\mathbf{z}_{i,j}/c) + (\eta\, b_\lambda + d). \tag{3.1}$$

For notational convenience, we define $S = \int_0^{T_r} s(t)\, dt$ and $B = (\eta\, b_\lambda + d)T_r$ as the mean signal and background count per period. We assume that both S and B are known, since it is straightforward to measure them before we begin data acquisition. Also, we emphasize that the derivations to follow assume that the total flux is low, i.e., $\eta\, \mathbf{x}_{i,j}\, S + B \to 0^+$, as would be the case in low light-level imaging where photon efficiency is important.

3.3.2 Statistics of Number of Detected Photons

SPAD detectors are not number-resolving photon counters; they only provide us with the knowledge that no photons or one or more photons have been detected. Using Poisson process statistics [32] and the expression of rate function $\lambda_{i,j}(t)$, we have that the probability of the SPAD detector *not* recording a detection from one pulse transmission is

$$P_0(\mathbf{x}_{i,j}) = \frac{\exp\left\{-\left(\int_0^{T_r} \lambda_{i,j}(t)dt\right)\right\}\left(\int_0^{T_r} \lambda_{i,j}(t)dt\right)^0}{0\,!} = \exp\left\{-\left(\eta\, \mathbf{x}_{i,j}\, S + B\right)\right\}. \tag{3.2}$$

Since we illuminate with a total of N pulses, the number of detected photons $K_{i,j}$ is binomially distributed with probability mass function

$$\Pr\left[K_{i,j} = k_{i,j};\, \mathbf{x}_{i,j}\right] = \binom{N}{k_{i,j}} P_0(\mathbf{x}_{i,j})^{N-k_{i,j}} \left(1 - P_0(\mathbf{x}_{i,j})\right)^{k_{i,j}}, \tag{3.3}$$

for $k_{i,j} \in \{0, 1, \ldots, N\}$.

Under the low-flux condition $\eta\, \mathbf{x}_{i,j}\, S + B \to 0^+$ and assuming $N \to +\infty$ such that

$N(1-\exp\{-(\eta\,\mathbf{x}_{i,j}\,S+B)\}) = C$ is a constant, we see that the photon count variable $K_{i,j}$ converges to a Poisson random variable [33] and has the limiting probability mass function

$$\Pr[K_{i,j} = k;\, \mathbf{x}_{i,j}] = \frac{C^k \exp\{-C\}}{k\,!}, \tag{3.4}$$

which is simply a Poisson distribution with mean C.

Pixelwise maximum-likelihood reflectivity estimation: Given the total observed photon count data $k_{i,j}$ at pixel (i,j), we find that the non-negatively constrained maximum-likelihood (ML) reflectivity estimate is

$$\hat{\mathbf{x}}_{i,j}^{\mathrm{ML}} = \underset{\mathbf{x}_{i,j}\geq 0}{\arg\max}\ \Pr\left[K_{i,j} = k_{i,j};\, \mathbf{x}_{i,j}\right] = \max\left\{\frac{1}{\eta S}\left(\log\left(\frac{N}{N-k_{i,j}}\right) - B\right), 0\right\}, \tag{3.5}$$

by checking the zero-gradient condition of the strictly convex log-likelihood function obtained from Equation (3.3). Traditionally, when the number of photon detections is large, the normalized photon-count value is used as the reflectivity estimate [34],

$$\tilde{\mathbf{x}}_{i,j} = \frac{k_{i,j}}{N\eta S}. \tag{3.6}$$

We note that the normalized count value estimate is equal to the ML estimate under the limiting Poisson approximation (Equation (3.4)) of the binomial distribution with the condition $B = 0$. The performance guarantees for the unconstrained ML reflectivity estimation are given in Appendix B.1.

3.3.3 Statistics of Single Photon Arrival Times

At pixel (i,j), the single-photon arrival time $T_{i,j}$ recorded by the SPAD detector is localized down to a bin integration time of Δ. Because the SPAD detector only provides us with the timing information of the first detected photon in a single illumination period, we write the probability of a SPAD firing in a $[t, t+\Delta)$ interval, given that a

SPAD firing must occur in that illumination period, as

$$
\begin{aligned}
&\Pr[\,\text{No fire in } [0,t),\ \text{Fire in } [t,t+\Delta) \mid \text{Fire in } [0,T_r)\,] \\
&= \frac{\Pr[\,\text{No fire in } [0,t)\,] \times \Pr[\,\text{Fire in } [t,t+\Delta)\,]}{\Pr[\,\text{Fire in } [0,T_r)\,]} \\
&\overset{(a)}{=} \frac{1}{1-\exp\{-(\eta\,\mathbf{x}_{i,j}\,S+B)\}}\left(\exp\left\{-\int_0^t (\eta\,\mathbf{x}_{i,j}\,s(\tau-2\,\mathbf{z}_{i,j}/c)+B/T_r)\,d\tau\right\}\right. \\
&\qquad\qquad \left.\times\left(1-\exp\left\{-\int_t^{t+\Delta} (\eta\,\mathbf{x}_{i,j}\,s(\tau-2\,\mathbf{z}_{i,j}/c)+B/T_r)\,d\tau\right\}\right)\right) \\
&= \frac{1}{1-\exp\{-(\eta\,\mathbf{x}_{i,j}\,S+B)\}}\left(\exp\left\{-\int_0^t (\eta\,\mathbf{x}_{i,j}\,s(\tau-2\,\mathbf{z}_{i,j}/c)+B/T_r)\,d\tau\right\}\right. \\
&\qquad\qquad \left.-\exp\left\{-\int_0^{t+\Delta} (\eta\,\mathbf{x}_{i,j}\,s(\tau-2\,\mathbf{z}_{i,j}/c)+B/T_r)\,d\tau\right\}\right)
\end{aligned}
\tag{3.7}
$$

where (a) uses Equation (3.2). Then, assuming infinitesimally small Δ, we can obtain the probability density function of the continuous arrival time random variable $T_{i,j}$,

$$
\begin{aligned}
&f_{T_{i,j}}(t;\ \mathbf{x}_{i,j},\mathbf{z}_{i,j}) \\
&= \frac{1}{1-\exp\{-(\eta\,\mathbf{x}_{i,j}\,S+B)\}} \times \lim_{\Delta\to 0^+}\frac{d}{d\Delta}\left(\exp\left\{-\int_0^t (\eta\,\mathbf{x}_{i,j}\,s(\tau-2\,\mathbf{z}_{i,j}/c)+B/T_r)\,d\tau\right\}\right. \\
&\qquad\qquad \left.-\exp\left\{-\int_0^{t+\Delta} (\eta\,\mathbf{x}_{i,j}\,s(\tau-2\,\mathbf{z}_{i,j}/c)+B/T_r)\,d\tau\right\}\right) \\
&= \frac{\eta\,\mathbf{x}_{i,j}\,s(t-2\,\mathbf{z}_{i,j}/c)+B/T_r}{1-\exp\{-(\eta\,\mathbf{x}_{i,j}\,S+B)\}}\exp\left\{-\int_0^t (\eta\,\mathbf{x}_{i,j}\,s(\tau-2\,\mathbf{z}_{i,j}/c)+B/T_r)\,d\tau\right\}.
\end{aligned}
\tag{3.8}
$$

Then, we observe that the limiting probability density function when $(\eta\,\mathbf{x}_{i,j}\,S+B)\to 0^+$ (low-flux condition) is

$$
\begin{aligned}
f_{T_{i,j}}(t;\ \mathbf{x}_{i,j},\mathbf{z}_{i,j}) &= \frac{\eta\,\mathbf{x}_{i,j}\,s(t-2\,\mathbf{z}_{i,j}/c)+B/T_r}{\int_0^{T_r}(\eta\,\mathbf{x}_{i,j}\,s(t-2\,\mathbf{z}_{i,j}/c)+B/T_r)\,dt} \\
&= \frac{\eta\,\mathbf{x}_{i,j}\,S}{\eta\,\mathbf{x}_{i,j}\,S+B}\underbrace{\left(\frac{s(t-2\,\mathbf{z}_{i,j}/c)}{S}\right)}_{\text{Density 1}}+\frac{B}{\eta\,\mathbf{x}_{i,j}\,S+B}\underbrace{\left(\frac{1}{T_r}\right)}_{\text{Density 2}},
\end{aligned}
\tag{3.9}
$$

for $t \in [0, T_r)$. According to Equation (3.9), it is clear that a photon detection could originate from the backreflected light signal, ambient light, or a dark count. The arrival statistics observed at the detector result from the merging of the Poisson processes corresponding to these sources. Under the low-flux condition, the arrival time of a detected photon originating due to backreflected signal is exactly characterized by the normalized time-shifted pulse shape, which is Density 1. On the other hand, the photon arrival time from background and dark counts is uniformly distributed in the interval $[0, T_r)$ as Density 2. The probability density function in the low-flux limit is a mixture distribution with mixture weights

$$\Pr[\,\text{Detected photon at pixel } (i,j) \text{ is signal}\,] = \frac{\eta\,\mathbf{x}_{i,j}\,S}{\eta\,\mathbf{x}_{i,j}\,S + B},$$
$$\Pr[\,\text{Detected photon at pixel } (i,j) \text{ is noise}\,] = \frac{B}{\eta\,\mathbf{x}_{i,j}\,S + B}.$$

Pixelwise maximum-likelihood depth estimation: Under the low-flux assumption, we use the photon arrival time dataset $\mathcal{T}_{i,j} = \{t_{i,j}^{(1)}, t_{i,j}^{(2)}, \ldots, t_{i,j}^{(k_{i,j})}\}$ to construct the pixelwise constrained ML depth estimate using Equation (3.9):

$$\begin{aligned}\hat{\mathbf{z}}_{i,j}^{\mathrm{ML}} &= \underset{\mathbf{z}_{i,j}\in[0,cT_r/2)}{\arg\max} \prod_{i=1}^{k_{i,j}} f_{T_{i,j}}(t_{i,j}^{(i)}; \mathbf{x}_{i,j}, \mathbf{z}_{i,j}) \\ &= \underset{\mathbf{z}_{i,j}\in[0,cT_r/2)}{\arg\max} \sum_{\ell=1}^{k_{i,j}} \log\left(\eta\,\mathbf{x}_{i,j}\, s(t_{i,j}^{(\ell)} - 2\,\mathbf{z}_{i,j}/c) + B/T_r\right), \qquad (3.10)\end{aligned}$$

assuming that $k_{i,j} \geq 1$. If $k_{i,j} = 0$, then we leave the pixel empty. If $B > 0$, then the ML depth estimate is obtained by solving a non-convex optimization problem. Also, in that case, the ML estimation requires the knowledge of the true reflectivity $\mathbf{x}_{i,j}$, which is not available. Thus, the log-matched filter [32] is instead traditionally used for depth estimation using $k_{i,j}$ photon detections:

$$\tilde{\mathbf{z}}_{i,j} = \underset{\mathbf{z}_{i,j}\in[0,cT_r/2)}{\arg\max} \sum_{\ell=1}^{k_{i,j}} \log s(t_{i,j}^{(\ell)} - 2\,\mathbf{z}_{i,j}/c). \qquad (3.11)$$

We note that the log-matched filter solution can be obtained by assuming $B = 0$ in constrained ML estimate. The performance guarantees for ML depth estimation are given in Appendix B.2.

3.4 Novel Image Formation

In the limit of large sample size or high signal-to-noise ratio (SNR), ML estimation converges to the true parameter value [35]. However, when the data is limited or SNR is low – such as in our problem – pixelwise ML solutions yield inaccurate estimates. We compare our 3D imaging method with the baseline normalized-count reflectivity estimate $\tilde{\mathbf{x}}_{i,j}$ and the log-matched filter depth estimate $\tilde{\mathbf{z}}_{i,j}$, which are ML estimates asymptotically. Along with using the single-photon detection statistics, we exploit the spatial correlation present in real-world scenes by regularizing the ML estimators [36]. Our approach provides significant improvements over pixelwise ML estimators and traditional denoising techniques that may exploit scene sparsity, but assume additive Gaussian noise. Our computational reconstruction proceeds in three steps.

1. **Estimation of reflectivity:** The log-likelihood of scene reflectivity $\mathbf{x}_{i,j}$ given photon count data $k_{i,j}$ is

$$\mathcal{L}_{\mathbf{x}}(\mathbf{x}_{i,j}; k_{i,j}) = -(N - k_{i,j})\, \eta\, \mathbf{x}_{i,j}\, S + k_{i,j} \log\left(1 - \exp\left\{-(\eta\, \mathbf{x}_{i,j}\, S + B)\right\}\right) \tag{3.12}$$

after constants not dependent of $\mathbf{x}_{i,j}$ are dropped. Since $\mathcal{L}_{\mathbf{x}}(\mathbf{x}_{i,j}; k_{i,j})$ is a strictly convex function in $\mathbf{x}_{i,j}$, it is amenable to global minimization using convex optimization [20]. Then, the penalized ML (PML) estimate for scene reflectivity image is obtained from noisy data $\{k_{i,j}\}_{i,j=1}^{n}$ by solving the following convex program:

$$\hat{\mathbf{x}}^{\text{PML}} = \underset{\mathbf{x}:\ \mathbf{x}_{i,j} \geq 0}{\arg\min} \sum_{i=1}^{n} \sum_{j=1}^{n} -\mathcal{L}_{\mathbf{x}}(\mathbf{x}_{i,j}; k_{i,j}) + \beta_{\mathbf{x}}\, \text{pen}(\mathbf{x}), \tag{3.13}$$

where pen($\mathbf{x}$) is a convex function that penalizes the non-smoothness of the estimate and $\beta_{\mathbf{x}}$ controls the degree of penalization.

2. Rejection of background photon detections: Direct application of the similar regularized ML approach to depth estimation using time-of-arrival data is infeasible. This is because the background light contribution to the likelihood function gives rise to a non-convex cost function with locally optimal solutions that may be arbitrarily far from the global optimum. Hence, before estimating depth, the second processing step attempts to censor the photon detections that are due to background light and dark counts.

Detections from background light and dark counts do not contain any scene depth information. Their arrival times are mutually independent over spatial locations with variance scaling with T_r. In contrast, since light pulses have duration $T_p \ll T_r$ and depths $\mathbf{z}_{i,j}$ are correlated over spatial locations, the detection times of signal photons have low conditional variance given data from neighboring positions. Based on this key observation, our method to censor a noisy detection at (i, j) is as follows:

1. Compute the rank-ordered mean (ROM) [37] value $t_{i,j}^{\text{ROM}}$, which is the median value of all the photon arrival times at the 8 neighboring pixels of (i, j). If $t_{i,j}^{\text{ROM}}$ cannot be computed, due to missing data, then we set $t_{i,j}^{\text{ROM}} = +\infty$.

2. Detect and censor out background photons to obtain the index set of signal photons:

$$U_{i,j} = \left\{ \ell \;:\; |t_{i,j}^{(\ell)} - t_{i,j}^{\text{ROM}}| < 2T_p \left(\frac{B}{\eta \, \hat{\mathbf{x}}_{i,j}^{\text{PML}} S + B} \right), \; 1 \leq \ell \leq k_{i,j} \right\} \tag{3.14}$$

It is demonstrated in [37] that the method of rank-ordered means is effective in detecting pixels that are corrupted by high variance uniform noise. Since detections from background light are uniformly distributed, we use the ROM method to reject such detections and only keep signal detections for further processing.

3. **Estimation of 3D structure:** With extraneous background and dark-count detections rejected, the log-likelihood function of depth $\mathbf{z}_{i,j}$ given data $\{t_{i,j}^{(\ell)}\}_{\ell \in U_{i,j}}$ is

$$\mathcal{L}_{\mathbf{z}}(\mathbf{z}_{i,j}; \{t_{i,j}^{(\ell)}\}_{\ell \in U_{i,j}}) = \sum_{\ell \in U_{i,j}} \log s(t_{i,j}^{(\ell)} - 2\,\mathbf{z}_{i,j}/c), \tag{3.15}$$

except when the (i,j)-th pixel does not have uncensored data ($|U_{i,j}| = 0$), in which case we let $\mathcal{L}_{\mathbf{z}}(\mathbf{z}_{i,j}; \{t_{i,j}^{(\ell)}\}_{\ell \in U_{i,j}}) = 0$, so that it has zero contribution to the cost function. Our framework allows the use of arbitrary pulse shapes, but many practical pulse shapes are well approximated as $s(t) = \exp\{-v(t)\}$, where $v(t)$ is a convex function in t. Then, $\mathcal{L}_{\mathbf{z}}(\mathbf{z}_{i,j}; \{t_{i,j}^{(\ell)}\}_{\ell \in U_{i,j}}) = \sum_{\ell \in U_{i,j}} v(t_{i,j}^{(\ell)} - 2\,\mathbf{z}_{i,j}/c)$ is a convex function in $\mathbf{z}_{i,j}$. Our penalized ML estimate for the 3D image is thus obtained using uncensored data and solving the following convex optimization problem:

$$\hat{\mathbf{z}}^{\mathrm{PML}} = \underset{\mathbf{z}:\ \mathbf{z}_{i,j} \in [0, cT_r/2)}{\arg\min} \sum_{i=1}^{n} \sum_{j=1}^{n} -\mathcal{L}_{\mathbf{z}}(\mathbf{z}_{i,j}; \{t_{i,j}^{(\ell)}\}_{\ell \in U_{i,j}}) + \beta_{\mathbf{z}}\,\mathrm{pen}(\mathbf{z}). \tag{3.16}$$

3.5 Experiments

3.5.1 Experimental Setup

The experimental setup used to collect the photon arrival data is shown in the top of Figure 3-2. A pulsed laser diode with pulse width $T_p = 270\,\mathrm{ps}$ and repetition period $T_r = 100\,\mathrm{ns}$ was used as the illumination source. A two-axis galvo was used to raster scan 1000 × 1000 pixels. A lensless SPAD detector with quantum efficiency $\eta = 0.35$ was used for detection. Also, the background light level was set such that the average value of $\eta\,\mathbf{x}_{i,j}\,S$ approximately equaled B. Further details of the experimental setup are given in [27]. Because raster scanning with a fixed pixelwise dwell time is equivalent to using a floodlight illumination source and a detector array, our experimental results are indicative of what can be accomplished in real-time imaging scenarios using SPAD arrays. The experimental dataset was collected by Dheera Venkatraman, a collaborator in this project.

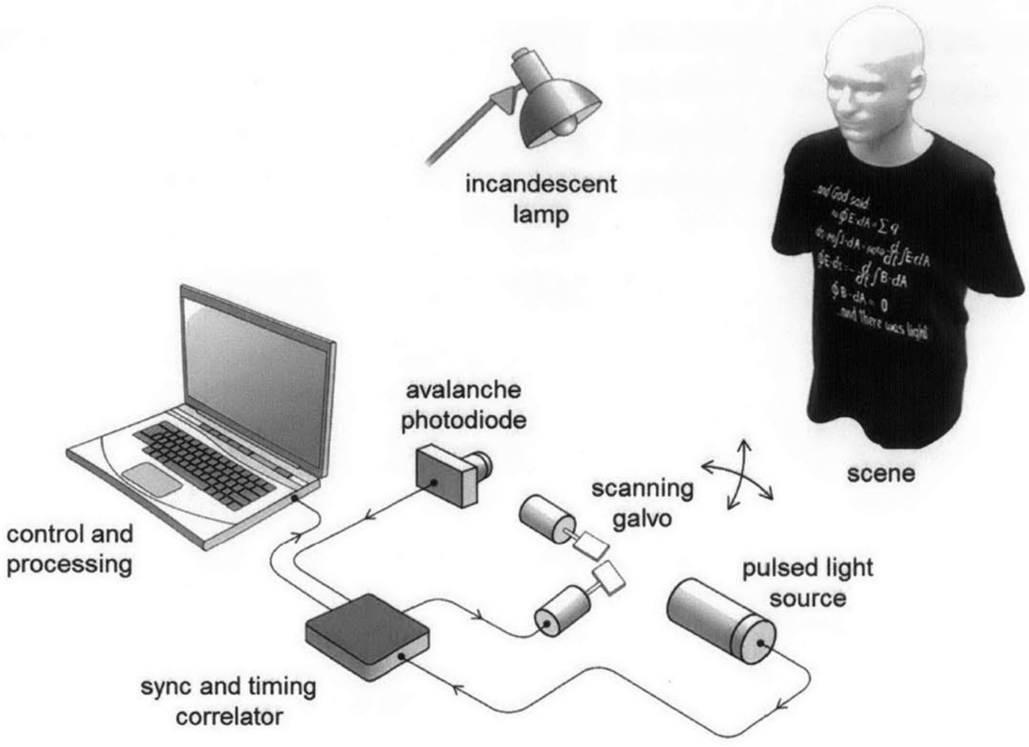

incandescent
lamp
avalanche
photodiode
scanning
galvo
scene
control and
processing
pulsed light
source
sync and timing
correlator

Figure 3-2: Experimental setup for single-photon imaging.

3.5.2 Reflectivity Resolution Test

Reflectivity contrast was tested using a linear grayscale intensity chart. A photograph of the reflectivity chart taken with a DSLR camera is shown in Figure 3-3 (a). As shown in Figure 3-3 (e), our method resolves 16 gray levels using a mean photon count of 0.48. The estimated reflectivity images were rescaled to have visual information in range $[0, 1]$. The performance of our method is comparable to that of the baseline imaging shown in Figure 3-3 (b), which required about 1000 photon detections per pixel.

We quantify the performance of an estimate $\hat{\mathbf{x}}$ of the true reflectivity image $\mathbf{x} \in \mathbb{R}_{+}^{1000\times1000}$ using PSNR defined in Equation (2.18). Our method using the total-variation penalty function outperforms pixelwise ML (Equation (3.6)) by 16.6 dB and denoised pixelwise ML using bilateral filtering [38] by 3.3 dB. The penalty parameter in our method was chosen to maximize PSNR. Also, the bilateral filtering parameters,

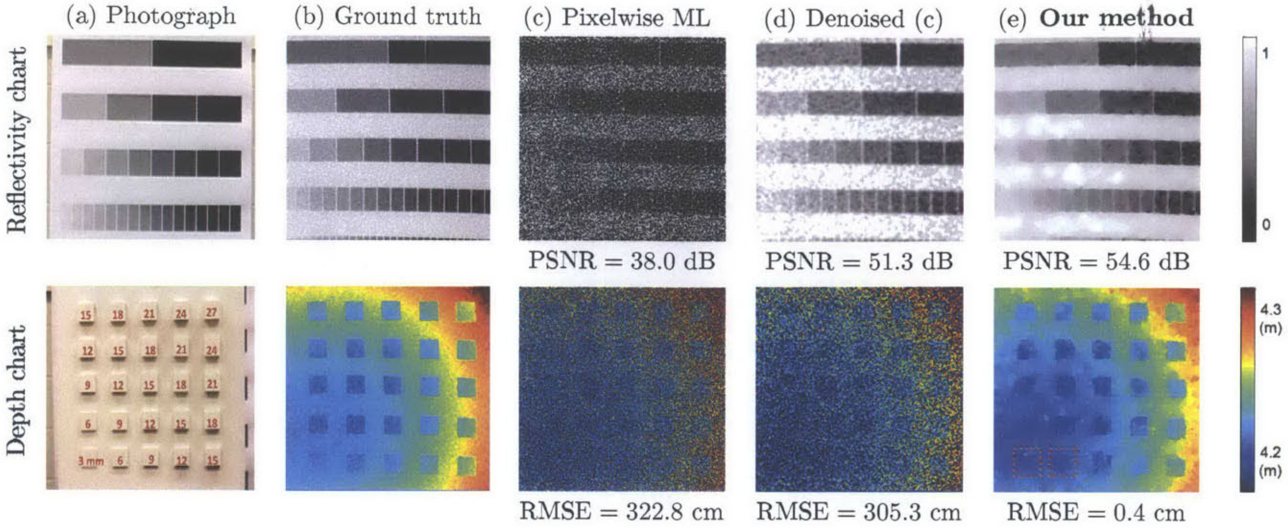


Figure 3-3: Resolution test experiments. Reflectivity chart imaging (top) was done using $T_a = 300\,\mu\text{s}$ and mean photon count of 0.48. Depth chart imaging (bottom) was done using $T_a = 6.2\,\mu\text{s}$, mean photon count of 1.1, and 33 % of the pixels missing data. The mean photon count was computed simply by averaging the number of photon counts at every pixel over the fixed acquisition time.

such as window size, were picked to maximize PSNR.

We note that bilateral filtering is a patch-based denoising method [39] and its computational complexity is linear time in the total number of image pixels. Thus, under certain real-time imaging scenarios, the filtering approach may be preferred over our convex optimization method, which has computational complexity that is polynomial time in the number of pixels.

3.5.3 Depth Resolution Test

Depth resolution was evaluated with a test target comprising 5 cm × 5 cm squares of varying thickness mounted on a flat board. A photograph of the depth chart is shown in the bottom of Figure 3-3 (a). The smallest resolvable height (thickness) above the reference level is an indicator of achievable depth resolution. Our method using total-variation penalty achieves a depth resolution slightly more than 4 mm, and a 3.5 mm depth accuracy using only noisy first photon detections. In contrast, the pixelwise ML depth estimates (Equation (3.11)) are extremely noisy (Figure 3-3 (c)), and the baseline imaging (Figure 3-3 (b)) requires approximately 100 photon detections at each pixel to achieve performance comparable to our method under identical imaging conditions. The penalty parameter in our method was chosen to minimize RMSE.

We quantify the performance of a depth estimator $\hat{\mathbf{z}}$ of the true depth image $\mathbf{z} \in \mathbb{R}_{+}^{1000\times1000}$ using root mean square error (RMSE):

$$\mathrm{RMSE}(\mathbf{z}, \hat{\mathbf{z}}) = \sqrt{\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\mathbf{z}_{i,j} - \hat{\mathbf{z}}_{i,j})^2}. \tag{3.17}$$

As shown in Figure 3-3 (c), the pixelwise ML estimates have an RMSE of at least 3 m. The high error is due to the high variance of background and dark-count noise that is present in our experiment.

Figure 3-3 (d) shows a denoised pixelwise ML depth image. Because 33% of the pixels are empty with no photon arrival data, we first apply bicubic interpolation in the 4-by-4 pixel neighborhood to inpaint the missing pixel values. Then, we performed denoising using median filtering, which is effective in removing high-variance impulse noise in images, of window 3 × 3. We observe that the depth resolution of our method (4 mm) corresponds to approximately 800-fold depth error reduction in RMSE, compared to state-of-the-art denoised pixelwise imaging methods.

3.5.4 Natural Scenes

Imaging results for natural scenes are shown in Figure 3-4, which shows the results of recovering 3D and reflectivity of a life-size mannequin and a scene of a basketball and a soda can using traditional and our imaging methods. In the mannequin dataset, the mean count over all pixels was 1.2 and 55% of pixels were missing data. In the basketball dataset, the mean count over all pixels was 2.1 and 32% of pixels were missing data. Ground truth images were generated using ML estimation from 200 photon detections at each pixel (Figure 3-4 (a)). All reflectivity images were rescaled to have visual information in range $[0, 1]$.

We see, from Figure 3-4 (b), that the pixelwise ML approach gives 3D and reflectiv-

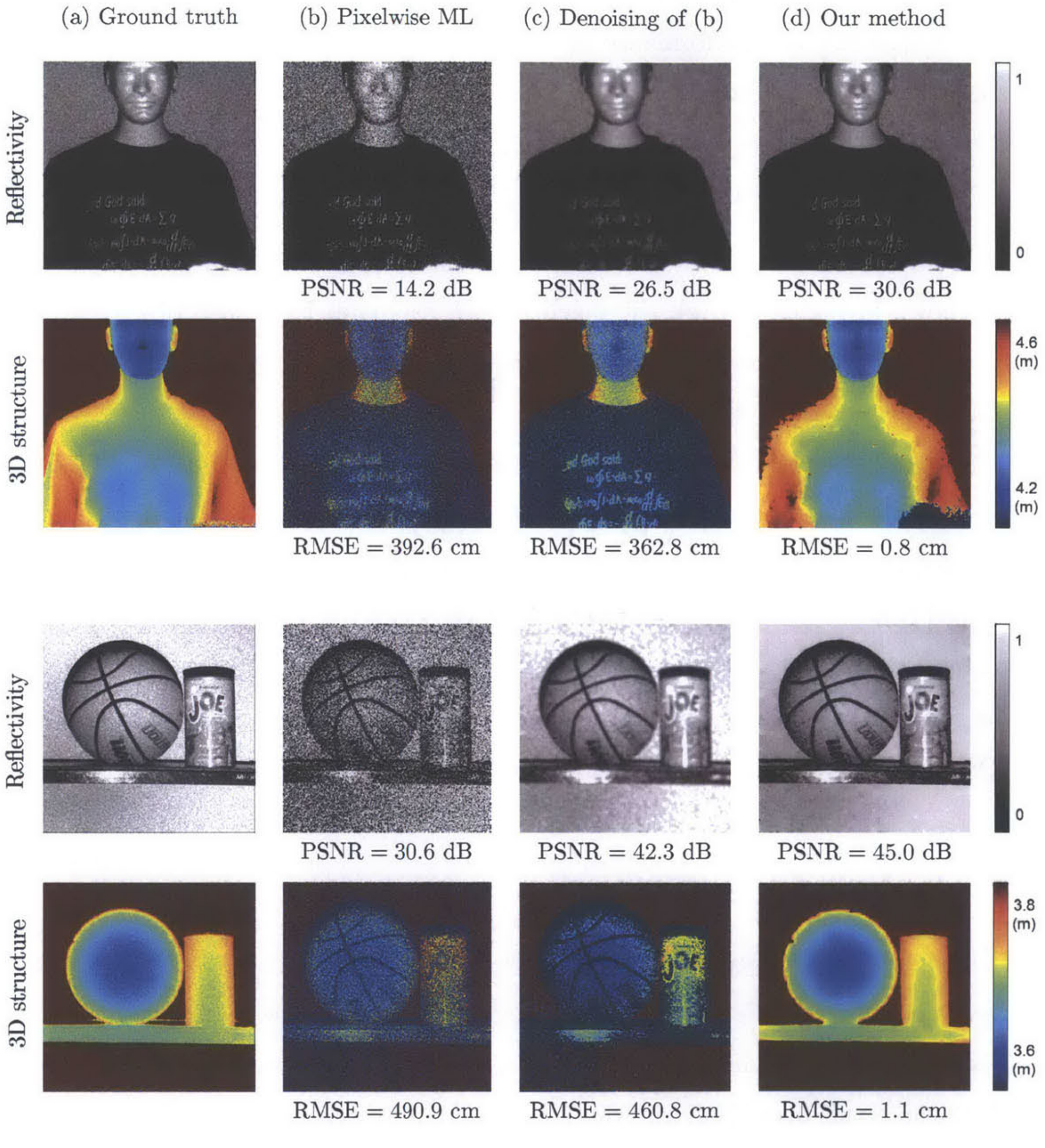
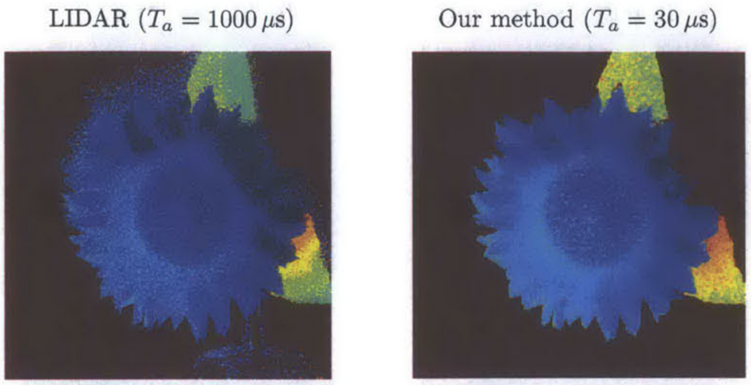

(a) Ground truth
(b) Pixelwise ML
(c) Denoising of (b)
(d) Our method
Reflectivity
PSNR = 14.2 dB
PSNR = 26.5 dB
PSNR = 30.6 dB
1
0
3D structure
RMSE = 392.6 cm
RMSE = 362.8 cm
RMSE = 0.8 cm
4.6 (m)
4.2 (m)
Reflectivity
PSNR = 30.6 dB
PSNR = 42.3 dB
PSNR = 45.0 dB
1
0
3D structure
RMSE = 490.9 cm
RMSE = 460.8 cm
RMSE = 1.1 cm
3.8 (m)
3.6 (m)

Figure 3-4: Experimental results for imaging natural scenes.


LIDAR ($T_a = 1000\,\mu s$)

Our method ($T_a = 30\,\mu s$)

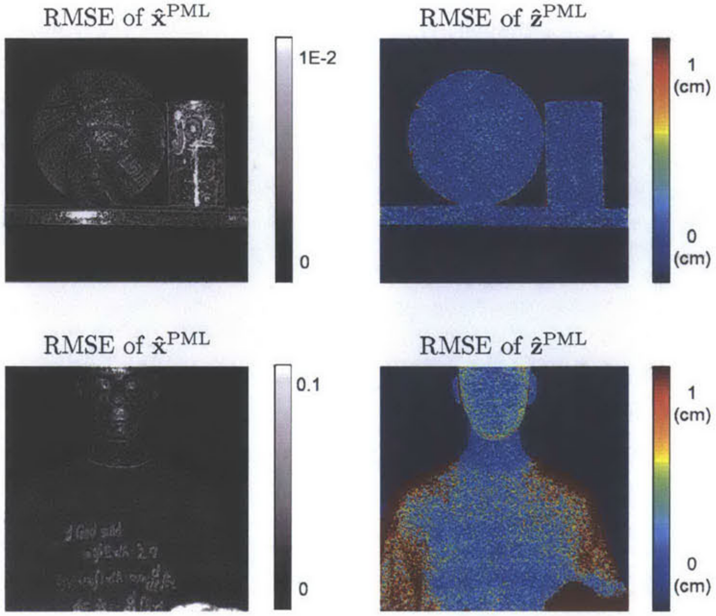
RMSE = 0.59 cm

RMSE = 0.41 cm

Figure 3-5: Comparison between our framework and LIDAR technology for 3D imaging.

ity estimates with high root mean square error (RMSE) and low peak signal-to-noise ratio (PSNR), respectively, due to background light and shot noise at low light-levels. Pixels with missing data were imputed with the average of their neighboring 8 pixelwise ML values. Denoising the ML reflectivity estimate using bilateral filtering [38] and the ML depth estimate using median filtering improves the image qualities (Figure 3-4 (c)). However, we see that denoising the 3D structure of the mannequin shirt fails, since the region has very low reflectivity and many pixels have missing data. On the other hand, our framework, which combines accurate photon arrival statistics with spatial prior information, accurately reconstructs images with RMSE and PSNR values of 0.8 cm and 30.6 dB (Figure 3-4 (d)). We used the total-variation penalty function in our method. The penalty parameters were chosen to minimize RMSE for 3D imaging and maximize PSNR for reflectivity imaging.

Figure 3-5 shows how much photon efficiency we gain over traditional LIDAR-based 3D imaging systems [40] that use the histogramming approach for a scene with a sunflower. The histogramming method, which is a standard pixelwise depth estimation technique in LIDAR systems, simply searches for the location of the peak in the photon arrival time histogram to estimate scene depth at every image pixel. We observe that while the log-matched filter (Equation (3.11)) is asymptotically ML as $B \to 0^+$, the histogramming-based depth estimation method is asymptotically ML as $N \to +\infty$. Thus, when T_a is long, as is the case in traditional LIDAR, it is


RMSE of $\hat{\mathbf{x}}^{\mathrm{PML}}$

RMSE of $\hat{\mathbf{z}}^{\mathrm{PML}}$

RMSE of $\hat{\mathbf{x}}^{\mathrm{PML}}$

RMSE of $\hat{\mathbf{z}}^{\mathrm{PML}}$

Figure 3-6: Sample MSE images with 100 independent trials for two natural scenes.

effective to use the histogramming-based depth estimation method. Based on PSNR and RMSE values, we see that our framework can allow more than 30 × speedup in acquisition, while reconstructing the same high-quality 3D and reflectivity images that a traditional LIDAR system would have formed using long acquisition times.

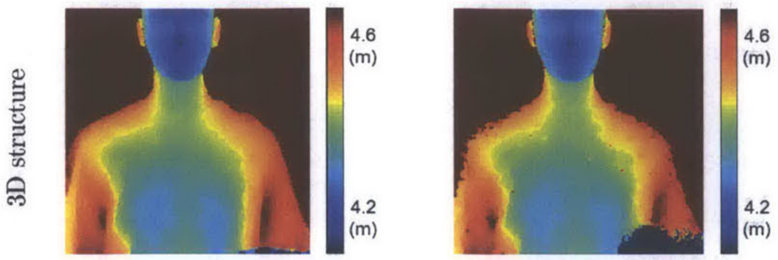
3.5.5 Pixelwise Root Mean-Square Error Test

For the basketball and mannequin scenes, we processed 100 independent photon arrival datasets and obtained the RMSE images of our recovery method by computing $\sqrt{\mathbb{E}[(\mathbf{x}_{i,j} - \hat{\mathbf{x}}_{i,j}^{\mathrm{PML}})^2]}$ and $\sqrt{\mathbb{E}[(\mathbf{z}_{i,j} - \hat{\mathbf{z}}_{i,j}^{\mathrm{PML}})^2]}$ for every pixel (i, j) using sample variance and bias values. The RMSE images in Figure 3-6 show that our computational imager achieves sub-centimeter depth resolution and repeatedly recovers high reflectivity information.

3.5.6 Effect of System Parameters

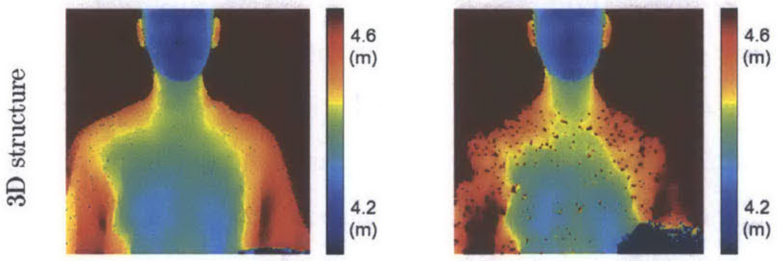
Figure 3-7 shows how the performances of traditional ML and our image formation methods are affected by changing acquisition time T_a and the signal-to-background

(a) $T_a = 100\,\mu\text{s}$, SBR = 7 (b) $T_a = 100\,\mu\text{s}$, SBR = 1



RMSE = 0.51 cm RMSE = 0.88 cm

(c) $T_a = 50\,\mu\text{s}$, SBR = 7 (d) $T_a = 50\,\mu\text{s}$, SBR = 1



RMSE = 5.80 cm RMSE = 11.91 cm

Figure 3-7: Effect of dwell time T_a and signal-to-background ratio (SBR) on our 3D recovery method. For acquisition times of $100\,\mu\text{s}$ and $50\,\mu\text{s}$, we calculated the mean photon count $k_{i,j}$ over all pixels to be 1.4 and 0.6, respectively.

ratio (SBR). We define SBR as the ratio between the mean counts of signal and noise:

$$\text{SBR} = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\eta\,\mathbf{x}_{i,j}\,S}{B}. \tag{3.18}$$

In our experiment, SBR is modified by changing T_r such that $B = (\eta b_\lambda + d)T_r$ changes while $\eta\,\mathbf{x}_{i,j}\,S$ remains the same. We see in Figure 3-7 that RMSE of our recovered 3D image increases monotonically with decreasing T_a and SBR. Even when the acquisition time is short and SBR is low as shown in Figure 3-7 (d), our 3D recovery method demonstrates near pulsewidth RMSE of 11.9 cm (Pulsewidth is $cT_p \approx 9$ cm).

3.5.7 Limitations

Results of 3D imaging using our method have high error near the edges of scene objects. The surface normals at these locations are nearly perpendicular to the line of sight and these regions incur more background noise detections, relative to signal detections. Although our method censors noisy detections near edges, it estimates the missing depth values using spatial correlations, leading to a loss of details.

Also, a detected photon may have originated from an indirect bounce, causing estimation inaccuracy. However, as it was in our experiments, diffuse scattering in quasi-Lambertian scenes causes the light multipath bounces to be considerably weaker than the direct reflection. Combined with Poisson statistics, this implies an exponentially diminishing probability of photon detections from indirect reflections. Finally, our method of estimating reflectivity fails if background noise is sufficient to provide a detection in each pulse-repetition period, with high probability. Hence, in our experiments, we employed a suitably narrowband spectral filter so that $B \ll 1$.

3.6 Information Theory for Optical Design

So far, we described our proposed framework for robust active 3D imaging and compared it against the state-of-the-art LIDAR technology to demonstrate its effectiveness. However, we have not yet commented on how optical design parameters, such as shape of the pulse waveform, can affect depth recovery performance. In this section, for pixelwise imaging scenarios, we study which optical pulse shapes give high recovery performance and how information theory can be used to compare the performance of two arbitrary pulse waveforms, without the relatively expensive computation of estimation-theoretic quantities.

We focus our analysis on a single image pixel. Let z be the depth value and $\tau = 2z/c$ be the time-delay parameter that we are interested in recovering. Also, under the low-flux condition, let t be the arrival time of the single photon that we detect and $f_T(t;\tau)$ be the probability density function given that we detect exactly one photon from the backscatter pulse waveform. According to our previous derivation in

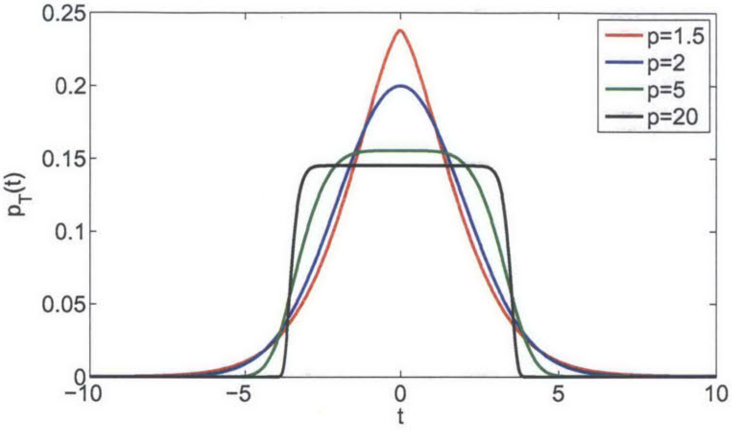


Figure 3-8: Distributions of generalized Gaussian random variables for several values of p with fixed σ.

Equation (3.9), we observe that $f_T(t;\tau) = s(t-\tau)/S$. We see that the Cramér-Rao lower bound (CRLB) of the time-delay parameter given only a single photon detection can then be shown to be

$$\mathrm{CRLB}(\tau) = -\left(\mathbb{E}\left[\frac{d^2}{d\tau^2}\log\left(\frac{s(t-\tau)}{S}\right)\right]\right)^{-1} = \left(\int_0^{T_r}\frac{\dot{f}_T(t;\tau)^2}{f_T(t;\tau)}\,dt\right)^{-1}, \qquad (3.19)$$

where the expectation is over T and $\dot{f}_T(t;\tau) = df_T(t;\tau)/dt$. For example, if our pulse $s(t)$ belongs to the generalized Gaussian family, the likelihood function from a single-photon observation is

$$f_T(t;\tau) = \frac{1}{2a\Gamma(1+1/p)}\exp\left\{-\left(\frac{|t-\tau|}{a}\right)^p\right\}, \qquad (3.20)$$

where $a > 0$ is the scale parameter, $p \geq 1$ is the order parameter, and $\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}\,dt$ is the gamma function. Parameter a controls the pulse width and p controls the concavity of pulse waveform. The generalized Gaussian distribution (see Figure 3-8) includes pulses frequently used in practical depth imaging systems such as the Gaussian pulse ($p = 2$) and near-uniform pulse ($p \gg 1$). We note that the variance of the generalized Gaussian random variable is $\sigma^2 = a^2\Gamma(3/p)/\Gamma(1/p)$.

We have shown in [41] that the CRLB of depth parameter τ given M independent single photon detections using generalized Gaussian pulses is

$$\mathrm{CRLB}(\tau) = \left(\frac{\Gamma(1/p)^2}{p(p-1)\Gamma(3/p)\Gamma(1-1/p)} \right) \frac{\sigma^2}{M}. \tag{3.21}$$

We observe that $\arg\max_{p>1} \mathrm{CRLB}(\tau) = 2$ based on a two-step logic.

1. According to our previous derivation in Equation (3.9), the random variable T of single photon time-of-arrival from the backscattered Gaussian pulse at low light levels is described as $T \sim \mathcal{N}(\tau, \sigma^2)$.

2. Gaussian random variables have minimum Fisher information [42] among all random variables with fixed variance.

Also, for any $p > 1$, the ML estimator using the generalized Gaussian pulse waveform is unbiased, because the generalized Gaussian distribution is always symmetric. Thus, the ML estimator using Gaussian pulse performs worst asymptotically in the generalized Gaussian family of pulses with fixed RMS duration σ. On the other hand, because

$$\lim_{p \to +\infty} \mathrm{CRLB}(\tau) = \lim_{p \to +\infty} \left(\frac{\Gamma(1/p)^2}{p(p-1)\Gamma(3/p)\Gamma(1-1/p)} \right) \frac{\sigma^2}{M} = 0,$$

the ML estimator that uses uniform pulse performs best asymptotically among all the generalized Gaussian pulses with equal σ.

According to the time-bandwidth uncertainty principle [43], if the RMS bandwidth is fixed instead of RMS time duration, then the Gaussian pulse is (approximately) the best performing pulse. The Gaussian pulse does not give the absolute best performance, because the time variance defined for the uncertainty principle is not exactly the same as the mean-square pulse time duration.

In practice, the CRLB using general illumination pulse waveforms can only be computed numerically because we can only observe the photon count histogram of the pulse. Such computations are difficult due to the differentiation operators existing in the CRLB expression. Thus, we study how certain information-theoretic quantities, which require only integrating operations that are numerically stable, can predict error trends without explicitly computing the CRLB.

As we defined earlier, let T be the random variable describing the single photon time-of-arrival. The entropy power of T is defined as [42]

$$N(T) = \frac{1}{2\pi e} \exp\{2\, h(T)\}, \tag{3.22}$$

where $h(T)$ is the differential entropy of T. Based on the results given in [44], we observe the following series of inequalities:

$$\sigma^2 \geq N(T) \geq M \cdot \mathrm{CRLB}(\tau), \tag{3.23}$$

where equalities hold if and only if T is Gaussian. Using Equation (3.23), we show how the information-theoretic quantities can be used to compare performance between the ML estimators of two arbitrary pulse waveforms. Suppose that we have two pulse waveform candidates, $s_x(t)$ and $s_y(t)$, and we would like to choose the pulse that gives rise to a ML estimator with lower MSE. Let the RMS pulse durations of $s_x(t)$ and $s_y(t)$ be σ_x^2 and σ_y^2, respectively. Given M photon detections, can we know whether τ_x^{ML} (ML estimate from using pulse $s_x(t)$ and arrival time data $\{t_x^{(1)}, \ldots, t_x^{(M)}\}$) performs better or worse than τ_y^{ML} (ML estimate from using pulse $s_y(t)$ and arrival time data $\{t_y^{(1)}, \ldots, t_y^{(M)}\}$)? Without loss of generality, if we know the Fisher information of the arrival time variable T_x of pulse $s_x(t)$, we can make the following remark.

Remark 3. *If* $N(T_y) \leq J(T_x)^{-1}$, *then* $\exists\ M_0$ *such that*

$$M > M_0 \ \Rightarrow\ \mathrm{MSE}(\tau, \tau_y^{\mathrm{ML}}) \leq \mathrm{MSE}(\tau, \tau_x^{\mathrm{ML}}).$$

We obtain Remark 3 by combining Equation (3.23) and the asymptotically efficient property of ML estimators. Remark 3 tells us that, using the entropy power of T_y and the Fisher information of T_x, we can compare performances between two pulse waveforms in the large M regime. Note that no computation of estimation-theoretic quantities of T_y is required for the remark to hold.

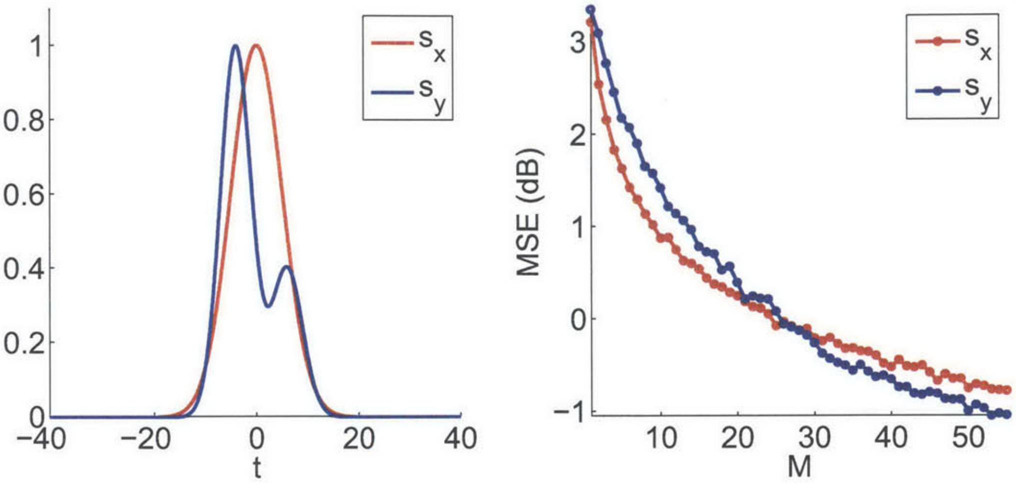


Figure 3-9: Comparison of MSE of time-delay recovery using single-photon detections for two different illumination pulses. (Left) Plot of $s_x(t)$, a Gaussian pulse, and $s_y(t)$, an arbitrary bimodal pulse. (Right) MSE of ML estimators vs. number of photon detections.

Figure 3-9 gives a concrete example of comparing the ML estimation performances of two pulse shapes. In this example, $s_x(t)$ is a Gaussian pulse with $\sigma_x = 5$ and $s_y(t)$ is a bimodal pulse with $\sigma_y = 5.5$. The smaller RMS pulse duration of $s_x(t)$ may lead us to think that τ_x^{ML} is superior to τ_y^{ML} for all M. However, we see that $J(T_x)^{-1} = \sigma^2 = 25$ and $N(T_y)$ is numerically computed to be approximately 23. Thus, the entropy power condition in Remark 3 holds in this example. By Remark 3, even though τ_x^{ML} may perform better than τ_y^{ML} for small M, eventually τ_y^{ML} will outperform τ_x^{ML} for high enough M. We see using numerical simulations of MSE that τ_y^{ML} outperforms τ_x^{ML} after 27 single-photon detections.

Chapter 4

Conclusions

4.1 Highly Photon-Efficient Reflectivity Imaging

We proposed a single-pixel multiplexed method for passive imaging that achieves higher photon efficiency compared to traditional pixelwise imaging methods. Thus, we presented an exception to the view against using multiplexing in the presence of signal-dependent Poisson noise coming from photon-limited imaging scenarios.

Our results entail performance analysis of different multiplexers under signal-dependent Poisson noise. We showed that the multiplexing advantage depends on the strength and sparsity level of the reflectivity image. Based on simulated imaging results, we observed that our method has clear improvements in PSNR over pixelwise methods. Out proposed method is computationally cheap as it involves solving a single convex optimization problem.

It is of future interest to analyze the performance of the proposed multiplexed imaging framework when the reflectivity signal is first degraded through a blurring kernel or a forward imaging operator before being corrupted by Poisson noise.

4.2 Highly Photon-Efficient 3D Imaging

We demonstrated the high accuracy of our proposed active 3D and reflectivity imaging technique, even with significant background light contribution. For the fixed pixelwise dwell-time model, our proposed imager simultaneously recovers high-quality 3D and reflectivity images by combining accurate single-photon counting statistics with the spatial correlation of natural scenes in both depth and reflectivity. Under mild conditions on the pulse shape, our method forms 3D and reflectivity images by solving two convex optimization problems; thus, it is computationally efficient. We observed in our experiments that centimeter accurate 3D imaging is possible, with 1000×1000 spatial resolution, even when the mean photon count per pixel is close to 1, and the signal-to-background ratio is 1. Thus, our computational imager motivates the development of accurate and low-power SPAD array-based 3D and reflectivity imagers.

Our framework can be used in many low light-level imaging applications using photon-counting detectors, such as spatially-resolved fluorescence lifetime imaging (FLIM) [4] and high-resolution LIDAR [31]. Our proposed methods naturally extends to imaging at other wavelengths, making them suitable for practical implementations. Also, several extensions in optoelectronic methods can improve the accuracy of our 3D and reflectivity imager. In particular, our framework can benefit from improved background suppression techniques [40] and range gating methods [25].

Appendix A

Proofs

A.1 Multiplexing Failure using Hadamard Matrix under Poisson Noise

Claim. *Let* $\mathbf{z} \sim \text{Poisson}(T_a \mathbf{W}\mathbf{x})$, *where* $\mathbf{W}$ *is a Hadamard multiplexing matrix* $\mathbf{H}$. *Then,* $\mathcal{G}(\mathbf{W}) = \sqrt{(n+1)/(2n)}$ *and* $n > 1 \Rightarrow \mathcal{G}(\mathbf{W}) < 1$.

Proof. Our pixelwise observation model is simply $\mathbf{y} \sim \text{Poisson}(T_a \mathbf{x})$. Let $T_a = 1$ for notational convenience. Then, by the definition of multiplexing gain,

$$\begin{aligned}
\mathcal{G}(\mathbf{H}) &= \sqrt{\frac{\text{MSE}(\mathbf{x},\mathbf{y})}{\text{MSE}(\mathbf{x},\mathbf{H}^{-1}\mathbf{z})}} \\
&= \sqrt{\frac{\operatorname{tr}\left(\mathbb{E}[(\mathbf{x}-\mathbf{y})(\mathbf{x}-\mathbf{y})^T]\right)}{\operatorname{tr}\left(\mathbb{E}[(\mathbf{x}-\mathbf{H}^{-1}\mathbf{z})(\mathbf{x}-\mathbf{H}^{-1}\mathbf{z})^T]\right)}} \\
&= \sqrt{\frac{\operatorname{tr}(\operatorname{diag}(\mathbf{x}))}{\operatorname{tr}(\mathbf{H}^{-1}\operatorname{diag}(\mathbf{H}\mathbf{x})\mathbf{H}^{-T})}} \\
&\overset{(a)}{=} \sqrt{\left(\frac{\sum_{i=1}^{n}\mathbf{x}_i}{\sum_{i=1}^{n}(\mathbf{H}\mathbf{x})_i}\right)\left(\frac{n}{\operatorname{tr}((\mathbf{H}^T\mathbf{H})^{-1})}\right)},
\end{aligned}$$

where (a) uses the fact that $(\mathbf{H}^T\mathbf{H})^{-1}$ is a circulant matrix by the properties of the Hadamard multiplexing matrix. As shown in [13], the eigenvalue set of the Hadamard multiplexer is

$$\lambda(\mathbf{H}) = \left\{ \frac{n+1}{2}, \underbrace{\sqrt{\frac{n+1}{4}}, \ldots, \sqrt{\frac{n+1}{4}}}_{\#=(n-1)/2}, \underbrace{-\sqrt{\frac{n+1}{4}}, \ldots, -\sqrt{\frac{n+1}{4}}}_{\#=(n-1)/2} \right\},$$

and we observe

$$\begin{aligned} \mathcal{G}(\mathbf{H}) &= \sqrt{\left(\frac{\sum_{i=1}^{n} \mathbf{x}_i}{\sum_{i=1}^{n} (\mathbf{Hx})_i}\right)\left(\frac{n}{\operatorname{tr}((\mathbf{H}^T\mathbf{H})^{-1})}\right)} \\ &= \sqrt{\left(\frac{\sum_{i=1}^{n} \mathbf{x}_i}{((n+1)/2)\sum_{i=1}^{n} \mathbf{x}_i}\right)\left(\frac{n}{(2/(n+1))^2 + 4(n-1)/(n+1)}\right)} \\ &= \sqrt{\frac{n+1}{2n}}. \end{aligned}$$

Since $\mathcal{G}(\mathbf{H}) < 1$ for all $n > 1$, we conclude that Hadamard multiplexing strategy fails under Poisson noise.

□

A.2 Multiplexing Failure using Circulant Matrices under Poisson Noise

Claim. *Let* $\mathbf{z} \sim \text{Poisson}(T_a \mathbf{W}\mathbf{x})$, *where* $\mathbf{W}$ *is a circulant multiplexing matrix. Then,* $\mathcal{G}(\mathbf{W}) \leq 1$.

Proof. A matrix $\mathbf{W}$ is *circulant*, if the $i+1$-th row vector is a right-shifted version of i-th row vector for $i = 1, 2, \ldots, n-1$. A circulant matrix is then necessarily Toeplitz. Let $\mathbf{w}$ be the first column vector of $\mathbf{W}$. Also, for notational convenience, let $T_a = 1$.

Then, the MSE of the matrix inverse solution is

$$\begin{aligned}
\operatorname{tr}\left(\mathbb{E}[(\mathbf{x}-\hat{\mathbf{x}}^{\text{inv}})(\mathbf{x}-\hat{\mathbf{x}}^{\text{inv}})^T]\right) &= \operatorname{tr}\left(\mathbb{E}[(\mathbf{x}-\mathbf{W}^{-1}\mathbf{z})(\mathbf{x}-\mathbf{W}^{-1}\mathbf{z})^T]\right) \\
&= \operatorname{tr}((\mathbf{W}^T\mathbf{W})^{-1}\operatorname{diag}(\mathbf{W}\mathbf{x})) \\
&\overset{(a)}{=} \frac{1}{n}\left(\sum_{i=1}^{n}(\mathbf{W}\mathbf{x})_i\right)\operatorname{tr}((\mathbf{W}^T\mathbf{W})^{-1}) \\
&\overset{(b)}{=} \frac{1}{n}\left(\sum_{i=1}^{n}\mathbf{x}_i\right)\left(\sum_{i=1}^{n}\mathbf{w}_i\right)\operatorname{tr}((\mathbf{W}^T\mathbf{W})^{-1}) \\
&\overset{(c)}{=} \frac{1}{n}\left(\sum_{i=1}^{n}\mathbf{x}_i\right)\left(\sum_{i=1}^{n}\mathbf{w}_i\right)\left(\sum_{i=1}^{n}\frac{1}{|\mathbf{F}\mathbf{w}|_i^2}\right).
\end{aligned}$$

In the above derivation, (a) uses the fact that $(\mathbf{W}^T\mathbf{W})^{-1}$ is circulant since $\mathbf{W}$ is circulant. (b) is true since a circulant matrix has columns with equal norm. Also, (c) holds because of the Fourier diagonalization property of circulant matrices: $\mathbf{W} = \frac{1}{n}\mathbf{F}^*\operatorname{diag}(\mathbf{F}\mathbf{w})\mathbf{F}$, where $\mathbf{F}$ is the $n\times n$ discrete Fourier transform (DFT) matrix. Thus,

$$\begin{aligned}
\mathcal{G}(\mathbf{W}) &= \sqrt{\frac{\sum_{i=1}^{n}\mathbf{x}_i}{\frac{1}{n}\left(\sum_{i=1}^{n}\mathbf{x}_i\right)\left(\sum_{i=1}^{n}\mathbf{w}_i\right)\left(\sum_{i=1}^{n}1/|\mathbf{F}\mathbf{w}|_i^2\right)}} \\
&= \sqrt{\frac{n}{(\sum_{i=1}^{n}\mathbf{w}_i)(\sum_{i=1}^{n}1/|\mathbf{F}\mathbf{w}|_i^2)}}.
\end{aligned}$$

Now, on the contrary to what we aim to prove, suppose that $\mathcal{G}(\mathbf{W}) > 1$. Then, from our above expression, $n/(\sum_{i=1}^{n}1/|\mathbf{F}\mathbf{w}|_i^2) > \sum_{i=1}^{n}\mathbf{w}_i$. However, by the arithmetic-harmonic mean inequality (d) and Parseval's property (e), we observe

$$\frac{n}{\sum_{i=1}^{n}1/|\mathbf{F}\mathbf{w}|_i^2} \overset{(d)}{\leq} \frac{1}{n}\sum_{i=1}^{n}|\mathbf{F}\mathbf{w}|_i^2 \overset{(e)}{=} \sum_{i=1}^{n}\mathbf{w}_i^2 \leq \sum_{i=1}^{n}\mathbf{w}_i,$$

which is a contradiction. □

The claim we have just proved simply states that direct estimation is preferred over circulantly multiplexed estimation when observations are corrupted by Poisson noise. Intuitively, this property of multiplexing failure is due to the signal-dependence of Poisson noise and that mixed observations will thus lead to lower SNR.

A.3 Strict Concavity of Log-Likelihood under Multiplexing

Claim. *The function* $\mathcal{L}_{\mathbf{x}}(\mathbf{x};\mathbf{z})$ *given in Chapter* 2 *is strictly concave in* $\mathbf{x}$.

Proof. It suffices to prove that the negative Hessian $-\nabla^2\mathcal{L}_{\mathbf{x}}(\mathbf{x};\ \mathbf{z})$ is positive definite (has positive eigenvalues). If $\mathbf{z}$ is a zero vector, then trivially the optimal reflectivity solution is also a zero vector. Thus, we assume that $\mathbf{z}$ is not a zero vector. The negative gradient of $\mathcal{L}_{\mathbf{x}}(\mathbf{x};\ \mathbf{z})$ is

$$-\nabla\,\mathcal{L}_{\mathbf{x}}(\mathbf{x};\ \mathbf{z}) = T_a\mathbf{W}^T\mathbb{1} - \sum_{i=1}^{n}\frac{\mathbf{z}_i}{e_i^T(\mathbf{W}\mathbf{x}+\mathbf{d})}\mathbf{W}^T e_i, \qquad \text{(A.1)}$$

where $\mathbb{1}$ is a size-n vector of ones and e_i is a size-n vector with the only non-zero entry, which is at i-th index, equal to one. The negative Hessian of the log-likelihood is

$$-\nabla^2\mathcal{L}_{\mathbf{x}}(\mathbf{x};\ \mathbf{z}) = \mathbf{W}^T\left(\sum_{i=1}^{n}\frac{\mathbf{z}_i}{(e_i^T(\mathbf{W}\mathbf{x}+\mathbf{d}))^2}e_ie_i^T\right)\mathbf{W}. \qquad \text{(A.2)}$$

Then, for any vector $\mathbf{v}\in\mathbf{R}^{n\times 1}$,

$$\begin{aligned}
-\mathbf{v}^T\nabla^2\mathcal{L}_{\mathbf{x}}(\mathbf{x};\ \mathbf{z})\mathbf{v} &= \mathbf{v}^T\left(\mathbf{W}^T\left(\sum_{i=1}^{n}\frac{\mathbf{z}_i}{(e_i^T(\mathbf{W}\mathbf{x}+\mathbf{d}))^2}e_ie_i^T\right)\mathbf{W}\right)\mathbf{v}\\
&= (\mathbf{W}\mathbf{v})^T\left(\underbrace{\sum_{i=1}^{n}\frac{\mathbf{z}_i}{(e_i^T(\mathbf{W}\mathbf{x}+\mathbf{d}))^2}e_ie_i^T}_{=\mathbf{A}}\right)(\mathbf{W}\mathbf{v})\\
&> 0,
\end{aligned}$$

where the last strict inequality is due to $\mathbf{A}$ being a diagonal matrix with positive entries. Thus, the Hessian is positive definite and the log-likelihood function is strictly concave. □

A.4 Efficiency of Matrix Inverse Demultiplexing Solution

Claim. *The matrix inverse solution* $\hat{\mathbf{x}}^{\text{inv}}$ *is an efficient estimator given multiplexed measurements corrupted by Poisson noise.*

Proof. An estimator $\hat{\mathbf{x}}$ is *efficient* if it is unbiased ($\mathbb{E}[\hat{\mathbf{x}}] = \mathbf{x}$) and its mean-square error is equal to the Cramér-Rao lower bound: $\text{CRLB}(\mathbf{x}) = \text{MSE}(\mathbf{x}, \hat{\mathbf{x}})$. An efficient estimator is thus an unbiased estimator with minimum MSE.

Our Poisson observation model is $\mathbf{z} \sim \text{Poisson}(T_a(\mathbf{W}\mathbf{x} + \mathbf{d}))$. Assume $T_a = 1$ for notational convenience. Then, we trivially see that the matrix inverse demultiplexing solution $\hat{\mathbf{x}}^{\text{inv}} = \mathbf{W}^{-1}(\mathbf{z} - \mathbf{d})$ is unbiased, since

$$\mathbb{E}[\hat{\mathbf{x}}^{\text{inv}}] = \mathbb{E}[\mathbf{W}^{-1}(\mathbf{z} - \mathbf{d})] = \mathbf{W}^{-1}\mathbb{E}[\mathbf{z}] - \mathbf{d} = \mathbf{W}^{-1}(\mathbf{W}\mathbf{x} + \mathbf{d}) - \mathbf{d} = \mathbf{x}.$$

In order to derive the CRLB, we derive the Fisher information matrix (FIM) $J(\mathbf{x})$ of the reflectivity signal $\mathbf{x}$ given the Poisson channel after multiplexing.

$$\begin{aligned}
&J(\mathbf{x}) \\
=&\mathbb{E}\left[\left(\frac{\partial}{\partial \mathbf{x}} \log \Pr[\mathbf{z};\ \mathbf{W}, \mathbf{x}, \mathbf{d}]\right)\left(\frac{\partial}{\partial \mathbf{x}} \log \Pr[\mathbf{z};\ \mathbf{W}, \mathbf{x}, \mathbf{d}]\right)^T\right] \\
=&\mathbb{E}\left[\left(\frac{\partial}{\partial \mathbf{x}} \sum_{i=1}^{n} [(\mathbf{W}\mathbf{x} + \mathbf{d})_i - \mathbf{z}_i \log(\mathbf{W}\mathbf{x} + \mathbf{d})_i]\right)\left(\frac{\partial}{\partial \mathbf{x}} \sum_{i=1}^{n} [(\mathbf{W}\mathbf{x} + \mathbf{d})_i - \mathbf{z}_i \log(\mathbf{W}\mathbf{x} + \mathbf{d})_i]\right)^T\right] \\
=&\mathbb{E}\left[\left(\mathbf{W}^T\mathbb{1} - \sum_{i=1}^{n} \frac{\mathbf{z}_i}{e_i^T(\mathbf{W}\mathbf{x} + \mathbf{d})}\mathbf{W}^T e_i\right)\left(\mathbf{W}^T\mathbb{1} - \sum_{i=1}^{n} \frac{\mathbf{z}_i}{e_i^T(\mathbf{W}\mathbf{x} + \mathbf{d})}\mathbf{W}^T e_i\right)^T\right] \\
=&-\mathbf{W}^T\mathbb{1}\mathbb{1}^T\mathbf{W} + \mathbb{E}\left[\left(\sum_{i=1}^{n} \frac{\mathbf{z}_i}{e_i^T(\mathbf{W}\mathbf{x} + \mathbf{d})}\mathbf{W}^T e_i\right)\left(\sum_{i=1}^{n} e_i^T\mathbf{W}\frac{\mathbf{z}_i}{e_i^T(\mathbf{W}\mathbf{x} + \mathbf{d})}\right)\right] \\
=&\sum_{i=1}^{n} \frac{1}{(\mathbf{W}\mathbf{x} + \mathbf{d})_i}\mathbf{W}^T e_i e_i^T \mathbf{W} \\
=&\mathbf{W}^T \text{diag}(\mathbf{W}\mathbf{x} + \mathbf{d})^{-1}\mathbf{W},
\end{aligned}$$

where $\mathbb{1}$ is a $n \times 1$ vector of ones. The CRLB is then computed as

$$\mathrm{CRLB}(\mathbf{x}) = \mathrm{tr}(J(\mathbf{x})^{-1}) = \mathrm{tr}(\mathbf{W}^{-1}\mathrm{diag}(\mathbf{W}\mathbf{x} + \mathbf{d})\mathbf{W}^{-T}).$$

Because $\mathrm{CRLB}(\mathbf{x}) = \mathrm{MSE}(\mathbf{x}, \hat{\mathbf{x}}^{\mathrm{inv}})$, we conclude that the matrix inverse demultiplexer $\hat{\mathbf{x}}^{\mathrm{inv}}$ is an efficient estimator. □

A.5 Violation Probability for Symmetric i.i.d. Matrix Inverse Estimators

Claim. *If the matrix inverse solution is unbiased, continuous, and symmetric i.i.d., then*

$$\Pr[\text{violation}] \geq 1 - \frac{1}{2^{n_0}},$$

where n_0 is the number of zeros in the signal vector $\mathbf{x}$.

Proof. Let S be the index set of $(n - n_0)$ non-zero entries of $\mathbf{x}$. Also, let $\hat{\mathbf{x}}$ be the matrix inverse solution computed using multiplexing matrix $\mathbf{W}$. Then,

$$\begin{aligned}
\Pr[\text{Violation}] &= \Pr[(\hat{\mathbf{x}}_1 < 0) \cup \ldots \cup (\hat{\mathbf{x}}_n < 0);\ \mathbf{W}, \mathbf{x}] \\
&= 1 - \Pr[(\hat{\mathbf{x}}_1 \geq 0) \cap \ldots \cap (\hat{\mathbf{x}}_n \geq 0);\ \mathbf{W}, \mathbf{x}] \\
&\overset{(a)}{=} 1 - \prod_{i=1}^{n} \Pr[\hat{\mathbf{x}}_i \geq 0;\ \mathbf{W}, \mathbf{x}] \\
&\overset{(b)}{=} 1 - \left(\frac{1}{2}\right)^{n_o} \prod_{i \in S} \Pr[\hat{\mathbf{x}}_i \geq 0;\ \mathbf{W}, \mathbf{x}] \\
&\geq 1 - \frac{1}{2^{n_0}}
\end{aligned}$$

where (a) uses the fact that the entries of $\hat{\mathbf{x}}$ are i.i.d. and (b) uses the fact that $\hat{\mathbf{x}}$ is unbiased, continuous and symmetric. □

A.6 Covariance Matrix of Hadamard Matrix Inverse Solution

Claim. *If the multiplexing pattern is determined by the Hadamard multiplexer* $\mathbf{H}$, *then the matrix inverse solution has a covariance matrix* Σ *with the following properties.*

1. *If* $i = j$, *then* $\Sigma_{i,j} = \frac{1}{T_a}\left(\frac{2}{n+1}\right)^2\left(\sum_{i=1}^{n}(\mathbf{Hx}+\mathbf{d})_i\right)$,

2. *If* $i \neq j$, *then*

$$|\Sigma_{i,j}| \leq \max_{S_0 \in S}\left|\frac{1}{T_a}\left(\frac{2}{n+1}\right)^2\left(\sum_{k_1 \in S_0}(\mathbf{Hx}+\mathbf{d})_{k_1} - \sum_{k_2 \in \{1,\ldots,n\}\setminus S_0}(\mathbf{Hx}+\mathbf{d})_{k_2}\right)\right|,$$

   where $S = \{T \,|\, T \subset \{1, 2, \ldots, n\}, |T| = (n-1)/2\}$.

Proof. The covariance matrix of the matrix inverse solution $\hat{\mathbf{x}}$ that is obtained using the Hadamard multiplexing strategy is

$$\Sigma = \frac{1}{T_a}\mathbf{H}^{-1}\mathrm{diag}(\mathbf{Hx}+\mathbf{d})\mathbf{H}^{-T}.$$

The authors in [14] derive a closed-form expression of $\mathbf{H}^{-1}$ in terms of $\mathbf{H}$ as

$$\mathbf{H}^{-1} = \frac{2}{n+1}\left(2\mathbf{H} - \mathbb{1}_{n\times n}\right),$$

where $\mathbb{1}_{n\times n}$ is a $n \times n$ matrix with ones. Since every row of $\mathbf{H}$ has $(n+1)/2$ ones and $(n-1)/2$ zeros, we see that every row of $2\mathbf{H} - \mathbb{1}_{n\times n}$ will have $(n+1)/2$ ones and $(n-1)/2$ minus ones. Let $\mathbf{h}_i$ be the i-th column of $\mathbf{H}^{-1}$. The i-th diagonal entry of the covariance matrix is then,

$$\begin{aligned}\Sigma_{i,i} &= \frac{1}{T_a}\,\mathbf{h}_i^T \mathrm{diag}(\mathbf{Hx}+\mathbf{d})\,\mathbf{h}_i \\ &= \frac{1}{T_a}\sum_{j=1}^{n}\,(\mathbf{Hx}+\mathbf{d})_j\,(\mathbf{h}_i)_j^2 \\ &\overset{(a)}{=} \frac{1}{T_a}\left(\frac{2}{n+1}\right)^2 \sum_{j=1}^{n}(\mathbf{Hx}+\mathbf{d})_j,\end{aligned}$$

where (a) uses the fact that every row of $2\mathbf{H}-\mathbb{1}_{n\times n}$ has $(n+1)/2$ ones and $(n-1)/2$ minus ones.

Before computing the off-diagonal entry of the covariance matrix indexed by (i,j), where $i \neq j$, let $S_{i,j}$ be the index set where entries of $\mathbf{h}_i$ and $\mathbf{h}_j$ have the same signs. Then, we write

$$\begin{aligned}|\Sigma_{i,j}| &= \left|\frac{1}{T_a}\,\mathbf{h}_i^T \mathrm{diag}(\mathbf{Hx}+\mathbf{d})\,\mathbf{h}_j\right| \\ &= \left|\frac{1}{T_a}\left(\frac{2}{n+1}\right)^2\left(\sum_{k_1\in S_{i,j}}(\mathbf{Hx}+\mathbf{d})_{k_1} - \sum_{k_2\in\{1,2,\ldots,n\}\setminus S_{i,j}}(\mathbf{Hx}+\mathbf{d})_{k_2}\right)\right| \\ &\le \max_{S_0\in S}\left|\frac{1}{T_a}\left(\frac{2}{n+1}\right)^2\left(\sum_{k_1\in S_0}(\mathbf{Hx}+\mathbf{d})_{k_1} - \sum_{k_2\in\{1,2,\ldots,n\}\setminus S_0}(\mathbf{Hx}+\mathbf{d})_{k_2}\right)\right|,\end{aligned}$$

where $S = \{T \,|\, T \subset \{1,2,\ldots,n\}, |T| = (n-1)/2\}$.

$\square$

Appendix B

Performance Guarantees for Pixelwise Single-Photon Imaging

The Cramér-Rao lower bound (CRLB) is used to analyze the mean square error (MSE) of an unbiased estimator of a parameter. Let x be a scalar continuous parameter that describes the probability density function $f_Y(y;x)$ of random variable Y. Then, the CRLB is defined as the inverse of the Fisher information $J(x)$ [35]:

$$\mathrm{CRLB}(x) = J(x)^{-1} = \left(\mathbb{E}\left[\frac{d^2}{dx^2}\left(-\log f_Y(y;x)\right)\right]\right)^{-1}. \tag{B.1}$$

For an unbiased estimator $\hat{x}$ of parameter x, the CRLB lower bounds the MSE: $\mathbb{E}[(x-\hat{x})^2] \geq \mathrm{CRLB}(x)$. An unbiased estimator $\hat{x}$ is efficient if $\mathbb{E}[(x-\hat{x})^2] = \mathrm{CRLB}(x)$.

B.1 Pixelwise Maximum-Likelihood Reflectivity Estimation

The CRLB for the reflectivity parameter at pixel (i,j), when data is obtained from setup in Figure 3-1 can be derived as follows.

$$\begin{aligned} \mathrm{CRLB}\,(\mathbf{x}_{i,j}) &= \left(\mathbb{E}\left[\frac{d^2}{d\mathbf{x}_{i,j}^2}\left(-\log \Pr[K_{i,j}=k_{i,j};\mathbf{x}_{i,j}]\right)\right]\right)^{-1} \\ &= \left(\mathbb{E}\left[\frac{k\,\eta^2\,S^2\exp\{\eta\,\mathbf{x}_{i,j}\,S+B\}}{\left(\exp\{\eta\,\mathbf{x}_{i,j}\,S+B\}-1\right)^2}\right]\right)^{-1} \\ &= \frac{\exp\{\eta\,\mathbf{x}_{i,j}\,S+B\}-1}{N\,\eta^2\,S^2}. \end{aligned} \tag{B.2}$$

We see that as the number of pulse repetitions N increases to collect more photons, the CRLB decreases. However, we cannot directly use the CRLB result to lower bound the MSE of the ML reflectivity estimate (without the non-negativity constraint) given by

$$\hat{\mathbf{x}}_{i,j}^{\mathrm{ML}} = \frac{1}{\eta\,S}\left(\log\frac{N}{N-k_{i,j}} - B\right),$$

because it is a biased estimator ($\mathbb{E}\,[\hat{\mathbf{x}}_{i,j}^{\mathrm{ML}}] \neq \mathbf{x}_{i,j}$):

$$\begin{aligned} \mathbb{E}\,\left[\hat{\mathbf{x}}_{i,j}^{\mathrm{ML}}\right] &= \mathbb{E}\left[\frac{1}{\eta S}\log\frac{N}{N-k_{i,j}} - \frac{B}{\eta S}\right] \\ &= \frac{1}{\eta S}\log N - \frac{1}{\eta S}\mathbb{E}\left[\log\left(N-k_{i,j}\right)\right] - \frac{B}{\eta S} \\ &> \frac{1}{\eta S}\log N - \frac{1}{\eta S}\log\left(N-\mathbb{E}[k_{i,j}]\right) - \frac{B}{\eta S} \\ &= \mathbf{x}_{i,j}, \end{aligned}$$

where the strict inequality follows from Jensen's inequality and the fact that the logarithm function is strictly concave.

Under the conditions $N \to +\infty$ and $\eta\,\mathbf{x}_{i,j}\,S + B \to 0^+$, holding $N(\eta\,\mathbf{x}_{i,j}\,S+B)$

constant, the limiting ML estimate is

$$\hat{\mathbf{x}}_{i,j}^{\mathrm{ML}} = \frac{k_{i,j}}{N\eta S} - \frac{B}{\eta S}. \tag{B.3}$$

Then, we observe that the CRLB equals the MSE of the ML reflectivity estimate,

$$\mathrm{CRLB}(\mathbf{x}_{i,j}) = \mathbb{E}\left[(\mathbf{x}_{i,j} - \hat{\mathbf{x}}_{i,j}^{\mathrm{ML}})^2\right] = \frac{1}{N}\left(\frac{\mathbf{x}_{i,j}}{\eta S} + \frac{B}{\eta^2 S^2}\right),$$

We see that the CRLB expression from Poisson likelihood is equal to the first-order Taylor expansion of the CRLB expression of the exact binomial likelihood given by Equation (B.2).

Knowing that the ML solution in the limiting Poisson distribution is unbiased and efficient, we conclude that the ML reflectivity estimate $\hat{\mathbf{x}}_{i,j}^{\mathrm{ML}}$ is efficient asymptotically as $(\eta \mathbf{x}_{i,j} S + B) \to 0^+$ and $N \to +\infty$.

B.2 Pixelwise Maximum-Likelihood Depth Estimation

We again assume the low-flux condition $\eta \mathbf{x}_{i,j} S + B \to 0^+$ and $N \to +\infty$ such that $C = N(1 - \exp\{\eta \mathbf{x}_{i,j} S + B\})$ is a constant at every pixel (i, j). Then, CRLB of depth parameter obeys

$$\begin{aligned}
\mathrm{CRLB}(\mathbf{z}_{i,j}) &= \left(\mathbb{E}\left[\frac{d^2}{d\mathbf{z}_{i,j}^2}\left(-\log f_{T_{i,j}}(\{t_{i,j}^{(\ell)}\}_{\ell=1}^{k_{i,j}}; \mathbf{z}_{i,j})\right)\right]\right)^{-1} \\
&= \left(\mathbb{E}\left[-\sum_{\ell=1}^{k_{i,j}} \frac{d^2}{d\mathbf{z}_{i,j}^2} \log f_{T_{i,j}}(t_{i,j}^{(\ell)}; \mathbf{z}_{i,j})\right]\right)^{-1} \\
&= \frac{1}{C}\left(\int_0^{T_r} \frac{\dot{p}(t; \mathbf{z}_{i,j})^2}{p(t; \mathbf{z}_{i,j})}\, dt\right)^{-1},
\end{aligned} \tag{B.4}$$

where $p(t; \mathbf{z}_{i,j})$ is the distribution obtained by normalizing $\lambda_{i,j}(t)$ to have area 1 over the $[0, T_r)$ interval.

We can exactly compute the MSE for certain pulse waveforms. For example, assume that the illumination waveform is a Gaussian pulse $s(t) \propto \exp\left\{-t^2/(2T_p^2)\right\}$. We saw previously that log-matched filter estimator is ML when $B = 0$. Thus, for the Gaussian pulse $s(t)$, the ML depth estimate (without the $[0, T_r)$ interval constraint) is

$$\hat{\mathbf{z}}_{i,j}^{\mathrm{ML}} = \arg\max_{\mathbf{z}_{i,j}} \sum_{\ell=1}^{k_{i,j}} \log s(t_{i,j}^{(\ell)} - 2\,\mathbf{z}_{i,j}/c) = \frac{c}{2}\left(\frac{\sum_{\ell=1}^{k_{i,j}} t_{i,j}^{(\ell)}}{k_{i,j}}\right),$$

given $k_{i,j} \geq 1$. Let $R = (\mathbf{z}_{i,j} - cT_r/4)^2$ be the mean-square error of the depth estimate that we choose to be the middle of the interval $[0, cT_r/2)$ when $k_{i,j} = 0$. Then, the MSE expression of $\hat{\mathbf{z}}_{i,j}^{\mathrm{ML}}$ can be derived as

$$\begin{aligned}
\mathbb{E}\left[(\mathbf{z}_{i,j} - \hat{\mathbf{z}}_{i,j}^{\mathrm{ML}})^2\right] &= \sum_{k_{i,j}=0}^{+\infty} \Pr[K_{i,j} = k_{i,j}]\,\mathbb{E}\left[(\mathbf{z}_{i,j} - \hat{\mathbf{z}}_{i,j}^{\mathrm{ML}})^2 \mid K_{i,j} = k_{i,j}\right] \\
&= \sum_{k_{i,j}=0}^{+\infty} \frac{\exp\{-C\}C^{k_{i,j}}}{k_{i,j}!}\,\mathbb{E}\left[(\mathbf{z}_{i,j} - \hat{\mathbf{z}}_{i,j}^{\mathrm{ML}})^2 \mid K_{i,j} = k_{i,j}\right] \\
&= \exp\{-C\}\left(R + \sum_{k_{i,j}=1}^{+\infty} \frac{C^{k_{i,j}}}{k_{i,j}!}\,\mathrm{Var}\left(\frac{c}{2} \cdot \frac{\sum_{\ell=1}^{k_{i,j}} t_{i,j}^{(\ell)}}{k_{i,j}}\right)\right) \\
&= \exp\{-C\}\left(R + \sum_{k_{i,j}=1}^{+\infty} \frac{C^{k_{i,j}}}{k_{i,j}!}\,\frac{1}{k_{i,j}}\left(\frac{cT_p}{2}\right)^2\right) \\
&\overset{(a)}{=} \exp\{-C\}\left(R + \left(\frac{cT_p}{2}\right)^2 \int_0^C \frac{\exp\{x\} - 1}{x}\,dx\right),
\end{aligned}$$

where (a) uses an identity given in [45]. We observe that, due to term $\exp\{-C\}$, $\hat{\mathbf{z}}_{i,j}^{\mathrm{ML}}$ achieves zero mean-square error as $C \to +\infty$.

Bibliography

[1] D. Takhar, J. N. Laska, M. B. Wakin, M. F. Duarte, D. Baron, S. Sarvotham, K. F. Kelly, and R. G. Baraniuk, "A new compressive imaging camera architecture using optical-domain compression," *SPIE Electronic Imaging.* International Society for Optics and Photonics, 2006, pp. 606509–606509.

[2] M. J. Eccles, M. E. Sim, and K. P. Tritton, *Low Light Level Detectors in Astronomy*, Cambridge University Press, Cambridge, 2012.

[3] A. M. Waxman, A. N. Gove, M. C. Siebert, D. A. Fay, J. E. Carrick, J. P. Racamato, E. D. Savoye, B. E. Burke, R. K. Reich, W. H. McGonagle, et al., "Progress on color night vision: visible/IR Fusion, perception and search, and llow-light CCD imaging," *Proceedings of SPIE*, 1996, vol. 2736, pp. 96–107.

[4] W. Becker, A. Bergmann, M. Hink, K. König, K. Benndorf, and C. Biskup, "Fluorescence lifetime imaging by time-correlated single-photon counting," *Microscopy Research and Technique*, vol. 63, no. 1, pp. 58–66, 2004.

[5] R. M. Gagliardi and S. Karp, *Optical Communications*, Wiley, New York, 1976.

[6] M. B. Wakin, J. N. Laska, M. F. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. F. Kelly, and R. G. Baraniuk, "An architecture for compressive imaging," *Proceedings of IEEE International Conference on Image Processing*, 2006, pp. 1273–1276.

[7] R. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, 2007.

[8] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 83–91, 2008.

[9] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[10] M. Raginsky, R. M. Willett, Z. T. Harmany, and R. F. Marcia, "Compressed sensing performance bounds under Poisson noise," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 3990–4002, 2010.

[11] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur, "A theory of multiplexed illumination," *Proceedings of IEEE International Conference on Computer Vision*, 2003, pp. 808–815.

[12] N. J. Sloane and M. Harwit, "Masks for Hadamard transform optics, and weighing designs," *Appl. Opt.*, vol. 15, no. 1, pp. 107–114, 1976.

[13] N. Ratner, Y. Y. Schechner, and F. Goldberg, "Optimal multiplexed sensing: bounds, conditions and a graph theory link," *Opt. Express*, vol. 15, no. 25, pp. 17072–17092, 2007.

[14] M. Harwit and N. J. Sloane, *Hadamard Transform Optics*, Elsevier, Amsterdam, 2012.

[15] N. Ratner and Y. Y. Schechner, "Illumination multiplexing within fundamental limits," *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*. IEEE, 2007, pp. 1–8.

[16] A. Mrozack, D. L. Marks, and D. J. Brady, "Coded aperture spectroscopy with denoising through sparsity," *Opt. Express*, vol. 20, no. 3, pp. 2297–2309, 2012.

[17] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, 1999.

[18] D. Shin, A. Kirmani, and V. K. Goyal, "Low-rate Poisson intensity estimation using multiplexed imaging," *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 1364–1368.

[19] M. Elad and A. Feuer, "Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images," *IEEE Trans. on Image Process.*, vol. 6, no. 12, pp. 1646 1658, 1997.

[20] Z. T. Harmany, R. F. Marcia, and R. M. Willett, "This is SPIRAL-TAP: sparse Poisson intensity reconstruction algorithms – theory and practice," *IEEE Trans. on Image Process.*, vol. 21, no. 3, pp. 1084–1096, 2012.

[21] S. Lee, O. Choi, and R. Horaud, *Time-of-Flight Cameras: Principles, Methods and Applications*, Springer, 2013.

[22] A. V. Jelalian, *Laser Radar Systems*, vol. 1, Artech, 1980.

[23] Z. Zhang, "Microsoft Kinect sensor and its effect," *IEEE Multimedia*, vol. 19, no. 2, pp. 4–10, 2012.

[24] D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*, Prentice Hall, 2002.

[25] J. Busck and H. Heiselberg, "Gated viewing and high-accuracy three-dimensional laser radar," *Appl. Opt.*, vol. 43, no. 24, pp. 4705–4710, 2004.

[26] G. Goltsman, O. Okunev, G. Chulkova, A. Lipatov, A. Semenov, K. Smirnov, B. Voronov, A. Dzardanov, C. Williams, and R. Sobolewski, "Picosecond superconducting single-photon optical detector," *Appl. Phys. Lett.*, vol. 79, no. 6, pp. 705–707, 2001.

[27] A. Kirmani, D. Venkatraman, D. Shin, A. Colaço, F. N. Wong, J. H. Shapiro, and V. K. Goyal, "First-photon imaging," *Science*, vol. 343, no. 6166, pp. 58–61, 2014.

[28] A. Kirmani, A. Colaço, D. Shin, and V. K. Goyal, "Spatio-temporal regularization for range imaging with high photon efficiency," *Proceeedings of SPIE*, 2013, vol. 8858, p. 88581F.

[29] D. Shin, A. Kirmani, A. Colaco, and V. K. Goyal, "Parametric Poisson process imaging," *Proceedings of IEEE Global Conference on Signal and Information Processing*, 2013.

[30] B. I. Erkmen and B. Moision, "Maximum likelihood time-of-arrival estimation of optical pulses via photon-counting photodetectors," *Proc. IEEE Int. Symp. Inf. Theory*, 2009, pp. 1909–1913.

[31] B. F. Aull, A. H. Loomis, D. J. Young, R. M. Heinrichs, B. J. Felton, P. J. Daniels, and D. J. Landers, "Geiger-mode avalanche photodiodes for three-dimensional imaging," *Lincoln Laboratory Journal*, vol. 13, no. 2, pp. 335–349, 2002.

[32] D. L. Snyder, *Random Point Processes*, Wiley, New York, 1975.

[33] D. P. Bertsekas and J. Tsitsiklis, *Introduction to Probability*, Athena scientific, 2002.

[34] Y. Chen, J. D. Müller, P. T. So, and E. Gratton, "The photon counting histogram in fluorescence fluctuation spectroscopy," *Biophys. J.*, vol. 77, no. 1, pp. 553–567, 1999.

[35] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 1: Estimation theory*, Prentice Hall, New Jersey, 1998.

[36] D. Shin, A. Kirmani, V. K. Goyal, and J. H. Shapiro, "Computational 3D and reflectivity imaging with high photon efficiency," *Accepted for publication at IEEE International Conference on Image Processing*, 2014.

[37] E. Abreu, M. Lightstone, S. K. Mitra, and K. Arakawa, "A new efficient approach for the removal of impulse noise from highly corrupted images," *IEEE Trans. on Image Process.*, vol. 5, no. 6, pp. 1012–1025, 1996.

[38] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” *Proceedings of IEEE International Conference on Computer Vision*, 1998, pp. 839–846.

[39] K. N. Chaudhury, D. Sage, and M. Unser, “Fast bilateral filtering using trigonometric range kernels,” *Image Processing, IEEE Transactions on*, vol. 20, no. 12, pp. 3376–3382, 2011.

[40] A. McCarthy, R. J. Collins, N. J. Krichel, V. Fernández, A. M. Wallace, and G. S. Buller, “Long-range time-of-flight scanning sensor based on high-speed time-correlated single-photon counting,” *Appl. Opt.*, vol. 48, no. 32, pp. 6241–6251, 2009.

[41] D. Shin, A. Kirmani, V. K. Goyal, and J. H. Shapiro, “Information in a photon: relating entropy and maximum-likelihood range estimation using single-photon counting detectors,” *Proceedings of IEEE International Conference on Image Processing*, 2013, pp. 83–87.

[42] A. Dembo, T. M. Cover, and J. A. Thomas, “Information theoretic inequalities,” *IEEE Transactions on Information Theory*, vol. 37, no. 6, pp. 1501–1518, 1991.

[43] R. E. Blahut, *Theory of Remote Image Formation*, Cambridge University Press, 2004.

[44] A. Stam, “Some inequalities satisfied by the quantities of information of Fisher and Shannon,” *Information and Control*, vol. 2, no. 2, pp. 101–112, 1959.

[45] V. Montazerhodjat, *Photon-limited time of flight depth acquisition: new parametric model and its analysis*, Ph.D. thesis, Massachusetts Institute of Technology, 2013.