

2020

# Probabilistic modeling for single-photon lidar

---

<https://hdl.handle.net/2144/39296>

*Boston University*

BOSTON UNIVERSITY  
COLLEGE OF ENGINEERING

Dissertation

**PROBABILISTIC MODELING FOR  
SINGLE-PHOTON LIDAR**

by

**JOSHUA RAPP**

B.S., Tufts University, 2014  
M.S., Boston University, 2016

Submitted in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

2020

© 2020 by  
JOSHUA RAPP  
All rights reserved

## Approved by

First Reader

---

Vivek K Goyal, Ph.D.  
Associate Professor of Electrical and Computer Engineering

Second Reader

---

W. Clem Karl, Ph.D.  
Professor and Chair of Electrical and Computer Engineering  
Professor of Systems Engineering  
Professor of Biomedical Engineering

Third Reader

---

Lei Tian, Ph.D.  
Assistant Professor of Electrical and Computer Engineering

Fourth Reader

---

Robin M. A. Dawson, Ph.D.  
Principal Member, Technical Staff  
The Charles Stark Draper Laboratory, Inc.

*A human being should be able to change a diaper, plan an invasion, butcher a hog, conn a ship, design a building, write a sonnet, balance accounts, build a wall, set a bone, comfort the dying, take orders, give orders, cooperate, act alone, solve equations, analyze a new problem, pitch manure, program a computer, cook a tasty meal, fight efficiently, die gallantly. Specialization is for insects.*

Robert A. Heinlein, *Time Enough for Love*

## Acknowledgments

Financial support for this work by a Draper Fellowship under Navy contract N00030-16-C-0014, the U.S. National Science Foundation under grant numbers 1422034 and 1815896, and the DARPA REVEAL program under Contract HR0011-16-C-0030 is gratefully acknowledged.

This manuscript has been years in the making, incorporating work that I started in June of 2015 during my first week as an intern at Draper. There have been too many people to count who contributed to my joy, mental health, and intellectual development during that time. I can only express a fraction of the thanks that those people deserve in helping me produce this document. To anyone whom I forget to mention, I offer my gratitude and my apologies.

Above all, I would like to express my deepest appreciation to my advisor, Vivek Goyal. I was extraordinarily fortunate to match with Vivek just before his leave from BU, during which he still made himself available for video chats to give valuable advice and feedback. I cannot mention enough what a pleasure it was to learn from Vivek's writing and editing, his attention to detail, his perceptive questions and innovative project ideas.

I also have to thank Vivek once more for his expert hiring ability. It has been both a joy and a transformative experience learning from and working with my fellow STIR members: John Murray-Bruce, Charlie Saunders, Minxu Peng, Sheila Seidel. Special thanks to Yanting Ma for the mathematical rigor that elevated the work on dead time to one of my favorite projects and for continuing to collaborate even once she had left the group.

My gratitude extends to all those collaborators outside of BU as well. Thanks to J.P. Laine, who initially brought me in to Draper, and especially to Robin Dawson, my Draper Fellowship supervisor, for his years of mentorship and for encouraging me

to learn how single-photon lidar systems actually work. I also appreciate the many PIs (and their students and postdocs) across the Charles River on the MIT/BU DARPA REVEAL team for the inventive ideas and probing questions that have contributed to several of my projects: Bill Freeman, Jeff Shapiro, Antonio Torralba, Greg Wornell, and especially Franco Wong, for his help with several optics experiments and for entrusting me with his lab space although I was not actually an MIT student. Many thanks also to Julián Tachella of Heriot-Watt University for his fruitful visit to BU and the continued effort in collaborating on multiple projects.

I am grateful to BU Professors Clem Karl and Lei Tian for serving on my committee and providing valuable advice and feedback, and to Bobak Nazer for the pleasure of being his TA and learning from him how to think about and teach probability. I would also be remiss not to mention the Tufts University professors who set me up to successfully pursue a PhD, even though I didn't know yet that I wanted one. Thanks to Valencia Koomson for suggesting I attend graduate school and for helping me get there, and to Shuchin Aeron for connecting me with Vivek in April 2014.

Finally, I cannot thank my immediate and extended family enough for their love and support through these years and beyond. Thanks to my parents for teaching me to read, write, count, be curious, and be organized, all of which I've needed to get this thesis to where it is today, and to my sisters Hannah and Becky for being my first teachers and students and for the mutual silliness throughout my life. And most of all, my deepest gratitude goes to my wife, Andie, for being such an exceptional companion and reliable partner during this long academic journey. I appreciate her giving me the space to work when I needed to and simultaneously reminding me of everything in life more important than whatever I was working on. I'm excited for our next adventure.

# PROBABILISTIC MODELING FOR SINGLE-PHOTON LIDAR

JOSHUA RAPP

Boston University, College of Engineering, 2020

Major Professor: Vivek K Goyal, PhD  
Associate Professor of Electrical and Computer  
Engineering

## ABSTRACT

Lidar is an increasingly prevalent technology for depth sensing, with applications including scientific measurement and autonomous navigation systems. While conventional systems require hundreds or thousands of photon detections per pixel to form accurate depth and reflectivity images, recent results for single-photon lidar (SPL) systems using single-photon avalanche diode (SPAD) detectors have shown accurate images formed from as little as one photon detection per pixel, even when half of those detections are due to uninformative ambient light. The keys to such photon-efficient image formation are two-fold: (i) a precise model of the probability distribution of photon detection times, and (ii) prior beliefs about the structure of natural scenes. Reducing the number of photons needed for accurate image formation enables faster, farther, and safer acquisition. Still, such photon-efficient systems are often limited to laboratory conditions more favorable than the real-world settings in which they would be deployed.

This thesis focuses on expanding the photon detection time models to address challenging imaging scenarios and the effects of non-ideal acquisition equipment. The

processing derived from these enhanced models, sometimes modified jointly with the acquisition hardware, surpasses the performance of state-of-the-art photon counting systems.

We first address the problem of high levels of ambient light, which causes traditional depth and reflectivity estimators to fail. We achieve robustness to strong ambient light through a rigorously derived window-based censoring method that separates signal and background light detections. Spatial correlations both within and between depth and reflectivity images are encoded in superpixel constructions, which fill in holes caused by the censoring. Accurate depth and reflectivity images can then be formed with an average of 2 signal photons and 50 background photons per pixel, outperforming methods previously demonstrated at a signal-to-background ratio of 1.

We next approach the problem of coarse temporal resolution for photon detection time measurements, which limits the precision of depth estimates. To achieve sub-bin depth precision, we propose a subtractively-dithered lidar implementation, which uses changing synchronization delays to shift the time-quantization bin edges. We examine the generic noise model resulting from dithering Gaussian-distributed signals and introduce a generalized Gaussian approximation to the noise distribution and simple order statistics-based depth estimators that take advantage of this model. Additional analysis of the generalized Gaussian approximation yields rules of thumb for determining when and how to apply dither to quantized measurements. We implement a dithered SPL system and propose a modification for non-Gaussian pulse shapes that outperforms the Gaussian assumption in practical experiments. The resulting dithered-lidar architecture could be used to design SPAD array detectors that can form precise depth estimates despite relaxed temporal quantization constraints.

Finally, SPAD dead time effects have been considered a major limitation for fast data acquisition in SPL, since a commonly adopted approach for dead time mitigation

is to operate in the low-flux regime where dead time effects can be ignored. We show that the empirical distribution of detection times converges to the stationary distribution of a Markov chain and demonstrate improvements in depth estimation and histogram correction using our Markov chain model. An example simulation shows that correctly compensating for dead times in a high-flux measurement can yield a 20-times speed up of data acquisition. The resulting accuracy at high photon flux could enable real-time applications such as autonomous navigation.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Time-Correlated Single Photon Counting . . . . .	3
1.1.1	Fluorescence Lifetime Imaging . . . . .	3
1.1.2	Single-Photon Lidar . . . . .	5
1.2	Preview of Main Contributions . . . . .	7
<b>2</b>	<b>Background on Single-Photon Lidar</b>	<b>11</b>
2.1	Overview . . . . .	11
2.2	Acquisition System . . . . .	11
2.2.1	Optical Factors . . . . .	13
2.2.2	Detection Electronics . . . . .	16
2.2.3	Acquisition Variations . . . . .	21
2.2.4	Basic Experimental Setup . . . . .	24
2.3	Basic Probabilistic Measurement Model . . . . .	25
2.3.1	Optical Effects . . . . .	25
2.3.2	Non-ideal Electronics Effects . . . . .	29
2.4	Reconstruction Basics . . . . .	31
2.4.1	Depth Map vs. Point Cloud . . . . .	31
2.4.2	Parameter Estimation . . . . .	32
<b>3</b>	<b>Unmixing Signal and Background Light</b>	<b>36</b>
3.1	Overview . . . . .	36
3.2	Unmixing Signal and Background Processes . . . . .	37

3.2.1	Pixelwise Unmixing . . . . .	38
3.2.2	Spatially-Adaptive Unmixing . . . . .	45
3.3	Unmixing Algorithm . . . . .	49
3.3.1	Windowing . . . . .	50
3.3.2	Reflectivity Estimation . . . . .	52
3.3.3	Superpixel Formation . . . . .	53
3.3.4	Depth Estimation . . . . .	54
3.4	Results . . . . .	55
3.4.1	Simulated Results . . . . .	55
3.4.2	Performance Analysis . . . . .	59
3.4.3	Experimental Results . . . . .	63
3.5	Further Performance Evaluation . . . . .	66
3.6	Conclusions . . . . .	68
<b>4</b>	<b>Subtractively-Dithered Ranging</b>	<b>70</b>
4.1	Overview . . . . .	70
4.2	Formulation, Background, and Motivation . . . . .	71
4.2.1	Measurement Model . . . . .	71
4.2.2	Subtractively-Dithered Quantization . . . . .	74
4.3	Generalized Gaussian Approximation and Estimation . . . . .	80
4.3.1	Approximation . . . . .	81
4.3.2	Estimation . . . . .	83
4.4	Estimator Implementations . . . . .	87
4.4.1	ML Estimators . . . . .	87
4.4.2	Order Statistics-Based Estimators . . . . .	87
4.5	Dither Noise Regimes . . . . .	90
4.5.1	Defining $\xi_1$ . . . . .	90

4.5.2	Defining $\xi_2$ . . . . .	91
4.6	Numerical Results . . . . .	93
4.6.1	Normalized MSE vs. $\sigma_Z/\Delta$ . . . . .	94
4.6.2	Order Statistics-Based Estimator Coefficients . . . . .	98
4.6.3	Normalized MSE vs. $K$ . . . . .	100
4.7	Subtractively-Dithered Single-Photon Lidar . . . . .	102
4.7.1	Architecture and Dither Implementation . . . . .	103
4.7.2	Related Work . . . . .	106
4.7.3	Experimental Results for Gaussian Pulse Modeling . . . . .	107
4.7.4	Exponentially-Modified Gaussian Pulse Modeling . . . . .	109
4.7.5	Subtractively-Dithered EMG . . . . .	111
4.7.6	Dithered EMG Mean Estimation . . . . .	112
4.7.7	Experimental Results for EMG Pulse Modeling . . . . .	114
4.8	Conclusions . . . . .	117
4.A	Cramér-Rao Bound . . . . .	119
4.B	Kurtosis Matching . . . . .	120
4.C	Mean Squared Error of $\hat{\mu}_Q$ . . . . .	121
<b>5</b>	<b>Dead Time Compensation</b> . . . . .	<b>123</b>
5.1	Overview . . . . .	123
5.2	Dead Time Characterization . . . . .	125
5.2.1	Paralyzability . . . . .	125
5.2.2	Synchronization . . . . .	125
5.2.3	Dead Time Compensation Approaches . . . . .	126
5.2.4	Other Related Work . . . . .	128
5.3	Empirical Distribution of Detection Times . . . . .	129
5.3.1	Markov Chain Model for Detection Times . . . . .	130

5.3.2	Comparison of Fisher Information . . . . .	136
5.4	Acquisition Parameter Estimation . . . . .	138
5.4.1	Maximum Likelihood Estimator for $\Lambda$ . . . . .	139
5.4.2	Maximum Likelihood Estimator for $B$ . . . . .	140
5.4.3	Estimating $S$ . . . . .	141
5.5	Experimental Verification . . . . .	141
5.6	Arrival Intensity Estimation Algorithm . . . . .	145
5.6.1	Relationship between Arrival and Detection Distributions . . .	145
5.6.2	Nonlinear Inverse Formulation and Algorithm . . . . .	148
5.7	Application to Ranging . . . . .	152
5.7.1	Ranging with True Acquisition Parameters . . . . .	152
5.7.2	Ranging with Estimated Acquisition Parameters . . . . .	158
5.7.3	Signal Quantization for High-Flux Depth Imaging . . . . .	159
5.8	Conclusion . . . . .	163
5.A	Proof of Proposition 1 . . . . .	164
5.B	Derivation of Fisher Information . . . . .	167
5.C	Proof of Proposition 2 . . . . .	167
5.D	Proof of Proposition 3 . . . . .	169
<b>6</b>	<b>Conclusions</b>	<b>171</b>
	<b>References</b>	<b>174</b>
	<b>Curriculum Vitae</b>	<b>194</b>

# List of Tables

3.1	Assumptions about scene, signal, and background properties . . . . .	50
4.1	RMSE performance of the depth estimates in Fig. 4-16 compared to several other methods. The generalized Gaussian-base methods are distinguished by whether the kurtosis matching is performed using a Gaussian (G) or exponentially-modified Gaussian (EMG) pulse shape approximation. The trimmed mean using the EMG-based kurtosis matching, denoted $\hat{\mu}_\alpha$ (EMG), often has the best performance and consistently has lower RMSE than the sample mean. . . . .	117

# List of Figures

1.1	Reflectivity and depth estimation results under photon-starved, high ambient light conditions, with about 4 signal detections per pixel and 25 times as many background detections per pixel on average. We propose a method that censors background detections by searching for clusters of signal detections for each pixel and joins together detection data from similar neighboring pixels. . . . .	8
1.2	Depth estimation results for TCSPC systems with coarse timing resolution. The implementation of subtractive dither and the careful choice of an estimator yields substantial improvement in the ability to precisely estimate fine features. . . . .	9
1.3	Depth estimation results for simulated detections illustrate the effectiveness of using the Markov chain modeling for high-flux acquisition. The proposed method approximately matches the conventional performance with $20\times$ fewer illuminations, greatly speeding up acquisition.	10

2.1	A single-photon lidar system shown in a raster-scanned, confocal configuration. The scene is illuminated by a laser pulse (solid line) passing through a beam-splitter (BS) and raster-scanned via a pair of galvo mirrors. Photons reflected back from the scene (dashed line) pass through a bandpass filter (BPF) tuned to the operating wavelength (a neutral-density filter (ND) providing additional attenuation is optional), before being focused by a lens onto the SPAD. Time correlation between illuminations and detections is performed by the TCSPC module, which streams data to the control computer. . . . .	12
3.1	(a) Detection can be described as an inhomogeneous Poisson process (shown in blue), which is the sum of inhomogeneous signal (green) and homogeneous background (red) processes. As in the example set of detections in (b), signal detections tend to form clusters more readily than the background detections, suggesting windowing as an approach to unmixing signal and background. Pixels with similar transverse position and reflectivity tend to belong to the same object and therefore have similar depth, as in (c). Using this observation leads to borrowing detections from similar neighboring pixels, as illustrated in (d), which can help amplify low signal levels by making signal detection clusters more apparent. . . . .	39
3.2	In comparing the theoretical approximations (solid lines) and Monte Carlo simulation (dashed lines) of the probability of cluster occurrence due to signal (a) or background (b) at various values of $N_{cl}$ , it is apparent that the derivations in the Sections 3.2.1 and 3.2.1 give close approximations to the true clustering probabilities. . . . .	44

3·3	Since only the effective background rate is known, setting a performance threshold $\tau_{\text{FA}}$ as in Fig. 3·2(b) yields a cluster size rule that limits acceptance of background clusters to probabilities less than or equal to $\tau_{\text{FA}}$ . . . . .	45
3·4	Results of a Monte Carlo simulation of ML depth estimates where the background is uniformly distributed over the acquisition interval and $s(t)$ is approximated by a Gaussian pulse of RMS width approximately equal to our measured laser pulse. . . . .	46
3·5	Block diagram of the basic algorithmic sequence. The unmixing procedure continues until enough pixels have reliable depth estimates so that inpainting is needed for only a small fraction with missing entries.	51
3·6	Simulated processing results for Art and Bowling scenes [164] at SBR = 0.04 and 2.0 signal photons per pixel on average. Note that the depth estimates with the PML+ROM method [170] are completely out of range of the actual scene and are instead shown for the range of 6 to 8 meters. . . . .	57
3·7	Absolute error maps for the proposed method applied to simulations with the Art and Bowling scenes [164] at SBR = 0.04 and 2.0 signal photons per pixel on average. The largest depth errors tend to be located at object boundaries. . . . .	58

3·8 Performance evaluation for reflectivity and depth estimation for simulated data sets with 2.0 and 3.0 signal photon detections per pixel on average and a range of SBR values. *Oracle* refers to performance of a penalized ML estimator using only signal detections ( $SBR = \infty$ ). *Shin* refers to PML+ROM, the method of Shin et al. [170]. *Unmixing* refers to the proposed method. Plotted performance is the average error of 10 trials for each value of SBR. . . . . 60

3·9 Performance evaluation for reflectivity and depth estimation for simulated data sets with SBRs of 0.1 and 1.0 and a range of average signal photon detections per pixel. *Oracle* refers to performance of a penalized ML estimator using only signal detections ( $SBR = \infty$ ). *Shin* refers to PML+ROM, the method of Shin et al. [170]. *Unmixing* refers to the proposed method. Plotted performance is the average error of 10 trials for each value of SBR. . . . . 62

3·10 Results of methods on the experimentally-acquired Mannequin data with about 4.05 signal detections per pixel and additional synthetically-generated noise to set the SBR at about 0.04. The error metrics are approximate, since the baseline lidar data is not exactly a ground truth for the scene. Note that the depth estimates with PML+ROM (the method of [170]) are completely out of range of the actual scene and are instead shown for the range of 5 to 8 meters. . . . . 64

3·11 Absolute error maps for the proposed algorithm applied to the experimentally acquired Mannequin data with about 4.05 signal detections per pixel and SBR 0.04. . . . . 65

4.1	Applied to subtractively-dithered measurements, the midrange is approximately optimal only for small $\sigma_Z/\Delta$ (Regime I, red), and the sample mean is approximately optimal only for large $\sigma_Z/\Delta$ (Regime III, blue). For confirmation, in Regime I the midrange approaches $\text{NMSE} = 1/[2(K+1)(K+2)] \approx 3.12 \times 10^{-5}$ and the mean approaches $\text{NMSE} = (1/12)/K \approx 6.67 \times 10^{-4}$ , and in Regime III the mean approaches $\text{NMSE} = (\sigma_Z/\Delta)^2/K$ , which is $8.00 \times 10^{-3}$ at $\sigma_Z/\Delta = 1$ . We seek an estimator simpler than the dithered-sample maximum likelihood that performs at least as well as the mean and midrange for intermediate values of $\sigma_Z/\Delta$ (Regime II, green). . . . .	80
4.2	The value of $\hat{p}_{\tilde{V}}$ goes to infinity as $\sigma_Z/\Delta$ decreases and converges to $\hat{p}_{\tilde{V}} = 2$ as $\sigma_Z/\Delta$ increases, with convergence beginning around $\sigma_Z/\Delta = 1/3$ matching the anticipated behavior. . . . .	82
4.3	The three plots show the noise PDF calculated numerically from the true density (4.17) (solid red) and via the GG approximation (4.22) (dashed black). The close agreement suggests the GGD is a good approximation for the noise. . . . .	83
4.4	As $p$ increases beyond $p = 2$ , $\beta(p)$ becomes much less than 1, implying that $\text{MSE}(\hat{\mu}_{\text{GGML}})$ is much lower than $\text{MSE}(\hat{\mu}_{\text{mean}})$ . . . . .	84
4.5	The value of $\xi_1$ substantially decreases and $\xi_2$ slowly increases as $K$ increases, expanding Regime II. (a) A log-log-cubic fit can be used to compute a close approximation to $\xi_1$ for all $K$ , while a log-log-linear fit suffices for $K > 20$ . (b) The square root of a log-quadratic fit closely approximates $\xi_2$ . . . . .	91

- 4.6 The performance of the estimators is evaluated for  $K =$  (a) 5, (b) 25, and (c) 125 to show the range of behavior as  $\sigma_Z/\Delta$  varies. The ML estimator for dithered measurements  $\hat{\mu}_{\text{DML}}$  and the estimators based on the GGD ( $\hat{\mu}_{\text{GGML}}$ ,  $\hat{\mu}_{\text{NB}}$ ,  $\hat{\mu}_{\text{NL}}$ , and  $\hat{\mu}_\alpha$ ) achieve the lowest NMSE for each  $\sigma_Z/\Delta$  regime (curves are overlapping). Results are shown for 20000 Monte Carlo trials. . . . . 95
- 4.7 The results of our Monte Carlo simulations lead to a simplified decision process for when and how to use dither. If  $\sigma_Z/\Delta > \xi_2$  ( $\approx 1/3$ ), there is no benefit to using anything but  $\hat{\mu}_{\text{Q}}$  applied to the quantized measurements, whereas dither leads to reduced estimation error when  $\sigma_Z/\Delta \leq 1/3$ . If subtractive dithering is not possible, the best performance can be achieved by adding Gaussian noise to set  $\sigma_Z/\Delta \approx 1/3$  and applying  $\hat{\mu}_{\text{QML}}$  (although  $\hat{\mu}_{\text{Q}}$  can be used if simplicity is required). However, larger performance improvements can be achieved with a subtractively-dithered quantizer. For  $K$  subtractively-dithered measurements, compute  $\xi_1$  from either (4.39) or (4.40) to determine whether to use  $\hat{\mu}_{\text{mid}}$  (in Regime I) or  $\hat{\mu}_\alpha$  (in Regime II). . . . . 96
- 4.8 Example dithered measurements are shown in (a-c) for  $\sigma_Z/\Delta = 0.004$ , 0.04, and 0.4 with  $K = 20$ . Plots (d-f) show the resulting coefficient values for the GG estimators given the estimated value of  $\hat{p}_{\hat{v}}$  above. In (g),  $\sigma_Z/\Delta = 0.04$  and  $K = 100$ , highlighting how the coefficients change as  $K$  increases. Note that the coefficients of the order statistics for the NL estimator depend on the measured data sequence shown above. . . . . 99

4.9	The performance of the order statistics estimators is evaluated for $\sigma_Z/\Delta =$ (a) 0.004, (b) 0.04, and (c) 0.4 to show the full range of behavior as the number of measurements $K$ increases. The plots show the MSE normalized to $\Delta = 1$ from 20000 trials per data point. Dashed lines show the theoretical NMSE of the mean and midrange, and the asymptotic variance of the ML estimator of the GGD mean. The ML estimator for dithered measurements $\hat{\mu}_{\text{DML}}$ and the estimators based on the GGD ( $\hat{\mu}_{\text{GGML}}$ , $\hat{\mu}_{\text{NB}}$ , $\hat{\mu}_{\text{NL}}$ , and $\hat{\mu}_\alpha$ ) achieve the lowest NMSE for all $K$ (curves are overlapping).	101
4.10	A photograph of the bistatic, subtractively-dithered single-photon lidar implementation at Draper in Cambridge, MA. The only difference from the conventional setup is the external timing control of a digital delay generator, which inserts 10-picosecond-resolution delays between the laser and SPAD triggers to implement dither.	104
4.11	A diagram showing the conventional SPL setup modified for a subtractively-dithered implementation. Varying delays $d_i$ are inserted by the digital delay generator between the laser and TCSPC triggers to implement dither. The amount of delay is controlled by the computer, which allows the delay to be subtracted from the photon detection times after acquisition.	105
4.12	Photographs showing the 3D-printed mask, “egghead” trophy, and 3D-printed depth resolution chart used for lidar experiments. Note for the depth chart that the heights in millimeters are displayed over each block but are not physically present in the scene.	107

4.13 Depth estimates are presented for the 3D-printed mask and egg-shaped trophy. The ground truth depth in (a) is poorly estimated with coarse quantization (b). Implementing subtractively-dithered measurements significantly improves the depth estimate, although the generalized Gaussian-based estimator (d) does not outperform the sample mean (c) for the experimental results. . . . . 108

4.14 Whereas the Gaussian distribution is a poor approximation to the true pulse shape, the exponentially-modified Gaussian (EMG) distribution captures the exponential tail behavior much more accurately. . . . . 110

4.15 A histogram of simulated dithered EMG data compared to the true distribution and generalized Gaussian approximation. The distribution parameters were chosen to be resemble experimental values, with  $\sigma_Z = 40$  ps,  $\tau_H = 200$  ps, and  $\Delta = 2000$  ps, which were then standardized for the overall distribution to have zero mean and unit variance. While the GG does not match the tail behavior exactly on either side, the value of  $\hat{p}$  determined through kurtosis matching does get the width and overall shape generally correct. . . . . 114

4.16 Point cloud estimates for the (first row) Depth Chart, (second row) Egghead, and (third row) Mask datasets. Additional quantitative results comparing each algorithm to the baseline can be found in Table 4.1. The use of subtractive dither yields a large decrease in the RMSE, and the trimmed-mean estimator produces a small additional improvement. . . . . 115

4.17	Detection counts per pixel for the point clouds in Fig. 4.16. The Depth Chart is fairly uniform, whereas the Egghead and Mask reveal differences in reflectivity and appear to show some shadowing effects due to the bistatic lidar setup at short range. . . . .	116
5.1	Illustration of the effect of dead times on the detection process for $t_r = 100$ ns, $t_d = 75$ ns, $\sigma_p = 5$ ns, $S = 0.5$ , and $B = 1$ . Photon arrival times are generated according to the arrival intensity $\lambda(t)$ . The detection intensity $\mu(t)$ is equal to the arrival intensity $\lambda(t)$ except immediately following a photon detection when $\mu(t) = 0$ , so the detection times are a subset of the arrival times and detection is not a Poisson process. .	132
5.2	Comparisons between histograms of simulated detection times, predicted detection time PDFs, and arrival time PDFs illustrate how dead time affects the detection process. In addition to a shift in the mode toward earlier detection times, the dead time may also cause a ripple in the detection PDF relative to the arrival PDF. Plots are shown for $\sigma_p = 2$ ns, $\Delta = 50$ ps, $n_r = 50000$ illuminations, and $t_d = 75$ ns. The vertical axis scale is constant for each row. An inset with a different vertical scale is included for each plot in the third column to emphasize the ripple that is not easily seen in the original scale. . . . .	135
5.3	The Fisher information ratio $FI_D/FI_A$ indicates the performance improvement that may be gained for the same number of detections when high-flux data is used instead of low-flux data. The plots show for various signal rate $S$ and background rate $B$ and for $\sigma_p = 0.2$ ns, $\Delta = 10$ ps, and $t_d = 75$ ns that when SBR is sufficiently high and $B$ is not too large, the effect of dead time is beneficial for range estimation. . . . .	138

5.4	A photograph of the bistatic single-photon lidar implementation at MIT in Cambridge, MA. The fast-gated SPAD has a variable hold-off time, which enables validation of the probabilistic model for various dead times. . . . .	142
5.5	Results are shown from high-flux experiments with fixed laser power while varying the SPAD hold-off time. Note that because of inaccuracy in estimating the background flux, the values of $\tilde{B}$ were set to minimize the Kolmogorov-Smirnov statistic between the predicted and measured PDFs, with $\hat{S} = \hat{\Lambda}^{\text{HF}} - \tilde{B}$ . The insets show in detail how the change in detection time histogram shape due to dead time is accurately predicted. . . . .	144
5.6	Estimation of arrival intensity (bottom) from detection histogram (top) when $S = B = 3.16$ , $\sigma_p = 2$ ns, $\Delta = 50$ ps, $t_r = 100$ ns, and $t_d = 75$ ns. From left to right: increased number of illuminations ( $n_r$ ), where in the last column, theoretical detection histogram is used as measurement. . . . .	152
5.7	Plots of the MSE for ranging as a function of $n_r$ for $t_r = 100$ ns, $t_d = 75$ ns, $\sigma = 0.2$ ns, $\Delta = 5$ ps and various $S$ and $B$ values. Our proposed methods (MCPDF and MCHC) take advantage of the increased detection rate to perform more accurate ranging than with the low-flux acquisition for all values of $S$ , $B$ , and $n_r$ . . . . .	156
5.8	Plots of the MSE for ranging as a function of the number of detections for $t_r = 100$ ns, $t_d = 75$ ns, $\sigma = 0.2$ ns, $\Delta = 5$ ps and various $S$ and $B$ values. For high $S$ and $B$ not too large, the presence of dead time actually improves ranging accuracy with our MCPDF method relative to the low flux measurements due to the narrowing of the signal pulse. . . . .	157

5.9 Results of ranging simulations performed with acquisition parameters estimated from the detection data, where the true values were  $S = B = 0.562$ ,  $t_r = 100$  ns,  $t_d = 75$  ns,  $\sigma_p = 0.2$  ns, and  $\Delta = 10$  ps. In (a), the estimates  $\hat{B}$ ,  $\hat{\Lambda}$ , and  $\hat{S}$  improve as  $n_r$  increases. The ranging results in (b) using estimated parameters show no degradation in performance compared to the methods with parameters known *a priori*. . . . . 158

5.10 Results for simulated detections from a true 3D scene illustrate the effectiveness of using the Markov chain modeling for high-flux acquisition. Using a 3-bit grayscale reflectivity approximation, our MCPDF method outperforms both low-and high-flux depth estimate for the number of illuminations  $n_r = 100$  and 2000. Our MCPDF method approximately matches the LF performance with  $20\times$  fewer illuminations, greatly speeding up acquisition. . . . . 160

5.11 Comparison of  $\log_{10}(|\hat{z} - z|)$  for HF and MCPDF for the high-flux acquisition with  $n_r = 100$  illuminations. The error for HF in (a) is lower for darker scene patches with detection PDFs less distorted by dead time, whereas the error for MCPDF in (b) is lower for lighter scene patches which reflect back more signal photons. . . . . 161

5.12 The performance of MCPDF improves as more quantization levels are used for the reflectivity estimate. The plot shows the median RMSE values for 100 realizations of detection data for the Bowling scene with  $n_r = 2000$  illuminations. . . . . 162

## List of Abbreviations

ADC	.....	Analog-to-digital converter
AMCW	.....	Amplitude-modulated continuous-wave
AV	.....	Autonomous vehicle
BLUE	.....	Best linear unbiased estimator
BPF	.....	Bandpass filter
CCD	.....	Charge-coupled device
CDF	.....	Cumulative distribution function
CML	.....	Constrained maximum likelihood
CNN	.....	Convolutional neural network
CRB	.....	Cramér-Rao bound
CW	.....	Continuous-wave
DAQ	.....	Data acquisition system
DARPA	.....	Defense Advanced Research Projects Agency
DMD	.....	Digital micromirror device
EMG	.....	Exponentially-modified Gaussian
FLIM	.....	Fluorescence lifetime imaging
FMCW	.....	Frequency-modulated continuous-wave
FPGA	.....	Field-programmable gate array
FPI	.....	First-photon imaging
FRET	.....	Förster resonance energy transfer
FWHM	.....	Full-width at half-maximum
GG	.....	Generalized Gaussian
GGD	.....	Generalized Gaussian distribution
GGRV	.....	Generalized Gaussian random variable
LGV	.....	Laser-gated viewing
LMF	.....	Log-matched filter

MAP	.....	Maximum a posteriori
MCHC	.....	Markov chain-based histogram correction method
MCPDF	.....	Markov chain-based PDF computation method
ML	.....	Maximum likelihood
MRI	.....	Magnetic resonance imaging
MSE	.....	Mean squared error
ND	.....	Neutral density
NMSE	.....	Normalized mean squared error
OD	.....	Optical density
PDF	.....	Probability density function
QIS	.....	Quantum image sensor
RHS	.....	Right-hand side
RJ-MCMC	.....	Reversible-jump Markov chain Monte Carlo
RMS	.....	Root-mean-square
RMSE	.....	Root-mean-square error
ROAD	.....	Rank-ordered absolute deviation
ROM	.....	Rank-ordered mean
SBR	.....	Signal-to-background ratio
SLM	.....	Spatial light modulator
SNSPD	.....	Superconducting nanowire single-photon detector
SPAD	.....	Single-photon avalanche diode
SPL	.....	Single-photon lidar
TAC	.....	Time-to-amplitude converter
TCSPC	.....	Time-correlated single photon counting
TDC	.....	Time-to-digital converter
TV	.....	Total variation

## Chapter 1

# Introduction

Engineering is a philosophical endeavor. Given a problem to solve and the constraints of resources, regulations, and human knowledge, engineers must answer the question, “what matters?” The universe of potential influences on a problem is too large to account for all of them, so any proposed solution implicitly or explicitly determines which factors to consider relevant by choosing a *model*. The model may be derived from mathematical abstraction of physical laws or determined empirically from a set of observations, but every model will fail to capture some level of detail as infinite factors are reduced to a finite number of parameters. While the model itself will always fall short of representing reality with perfect accuracy, its quality may be judged by how well it can be used to solve a particular problem. As George Box succinctly summarized [27], “All models are wrong, but some are useful.” Engineers seek models that account for the most important influences—the factors that measurably affect outcomes—while preferably maintaining the model’s simplicity.

One example of a technology for which the choice of model has a large impact on performance is single-photon lidar (SPL). Lidar (sometimes claimed to be an acronym for “light detection and ranging”) operates with the same principle as radar, using echoes detected from optical-wavelength illuminations to determine the positions and reflectivities of points in the surrounding environment. In SPL, the time-of-flight of individual photons is used to measure the distances to objects. Although range measurements had previously been made using time-resolved single-photon detec-

tion [126, 127], the introduction of “First-Photon Imaging” (FPI) in 2014 revolutionized SPL by detailing how depth and reflectivity maps could be accurately formed from only a single photon detection per pixel location, even when half of the incident photons were due to uninformative ambient light [104]. The two key components enabling the success of FPI were:

1. an acquisition model, used to determine the probabilistic nature of any photon detection time, and
2. a scene model, used to reconstruct depth and reflectivity maps by taking advantage of prior beliefs about the structure of natural scenes.

From the perspective of an optimization problem, these components correspond to a data fidelity term and a regularization term. Both models—of the measurement tool and the subject of the measurements—were more principled than had been used in previous approaches, thus enabling significant strides in recovering depth and reflectivity images from very little detected light.

This thesis focuses primarily on the first component of FPI, exploring practical problems for which existing models are too simple to be useful. We then perform what Box calls “model robustification,” i.e., the identification of the main sources of error and their subsequent incorporation into a more elaborate acquisition model. Unlike many first-order signal processing approaches, we cannot assume that SPL data takes the form of independent measurements from a linear, time-invariant system with additive, white Gaussian noise. Instead, this work accounts for the inherent Poisson randomness of photon counting, the nonlinear effects of quantization, and the statistical dependences incurred from system dead times. Through mathematical analysis, simulation, and experimental validation, we demonstrate how each additional layer of modeling describes the underlying acquisition process accurately enough to better solve a particular problem while keeping our understanding as simple as possible.

## 1.1 Time-Correlated Single Photon Counting

The key to more useful SPL acquisition models is a thorough understanding of time-correlated single photon counting (TCSPC), the operating principle of SPL. Early versions of TCSPC were introduced in 1961 by Bennett [20] and Bollinger and Thomas [24] for measuring the time dependence of the light intensity in scintillators and early lasers (“optical masers”). The basic concept of TCSPC is that of a stopwatch. An illumination source—usually a laser—is pulsed in order to cause a desired optical interaction with an object of interest, e.g., fluorescence, back-reflection, etc. Simultaneous with the illumination is the “start” signal for the timer, either from an electrical trigger signal or a trigger photodiode coupled to the illumination. When the interaction occurs, photons may be generated or reflected back to a single-photon detector such as a single-photon avalanche diode (SPAD), which produces the “stop” signal when a photon detection event occurs. The detection time is then the time difference between the stop and start signals. TCSPC is often used to build up a histogram of detection times and use the empirical distribution to estimate a physical property. The rest of this section introduces some context for two of the most common applications of TCSPC: fluorescence lifetime imaging and single-photon lidar.

### 1.1.1 Fluorescence Lifetime Imaging

Fluorescence is the phenomenon of photon emission from a molecule after an electron excited into a higher energy state returns to the ground state [15, 181]. Intensity measurements from fluorescent molecules called *fluorophores* are commonly used for biological imaging of ion concentrations and protein binding, but intensity measurements are sensitive to fluorophore concentration, photobleaching, and acquisition settings such as focusing and detector gain, all of which may be spatially varying [38]. Additionally, the intensity measurements for fluorophores with similar excitation and

emission wavelengths cannot be spectrally separated.

The time the excited electron spends in the higher energy state is called the *fluorescence lifetime* and has an exponential distribution. The lifetime of a fluorophore depends on the molecular environment, not on the fluorophore concentration, so the measurement of different lifetimes can indicate protein interactions, pH, viscosity, temperature, refractive index, the presence of a chemical or nanoparticle, etc. [16, 112, 181]. Fluorescence lifetimes are thus a useful quantitative measure that can act as an additional dimension for providing contrast, similar to how proton relaxation times provide the contrast in magnetic resonance imaging (MRI) [112]. For example, fluorescence lifetime measurement is the most reliable method of identifying Förster resonance energy transfer (FRET), which can indicate the proximity of proteins below the optical resolution limit [181].

Nanosecond-scale measurements of fluorescence lifetimes have been made since 1926 by Gaviola [66], and lifetime measurement was a main application of early TCSPC systems [21, 113, 137]. Lakowicz [112] first introduced fluorescence lifetime imaging (FLIM) with high spatial resolution from a gated, intensified charge-coupled device (CCD) camera. Previous attempts at FLIM using a scanned TCSPC system had extremely poor spatial resolution (e.g.,  $10 \times 10$  pixels) due to hardware limitations [210]. Modern improvements in laser excitation sources, detectors, timing electronics, and memory capacity have greatly improved scanned TCSPC systems, however, such that they now provide the highest time resolution and sensitivity to the lowest light levels of all FLIM implementations, with capabilities for simultaneous multispectral sensing and multi-exponential estimation [17]. TCSPC thus enables FLIM of weakly-fluorescent samples or the use of a low excitation energy, facilitating new scientific discoveries while reducing damage to a sample [38, 102, 162].

### 1.1.2 Single-Photon Lidar

Non-contact depth measurement has a wide range of uses, from industrial to military or scientific purposes. Active optical methods such as lidar are especially useful due to their high spatial resolution relative to radar or ultrasound methods [5]. Although laser ranging has been under development since the 1960s, mostly for military use, terrain mapping, or atmospheric monitoring [131], commercial lidar development has greatly accelerated since 2005, when all vehicles that completed the DARPA autonomous driving Grand Challenge employed lidar for depth mapping [19].

A number of strategies exist for using light to make range measurements. The most straightforward approach is pulsed lidar, which transmits a short-duration pulse of light at a target and measures the time until the reflected echo is detected. Because the speed of light is well-known, the range can be directly measured from the round-trip time. Aside from its simplicity, the key benefit of pulsed lidar is that the range accuracy depends only weakly on the distance, with the detected pulse amplitude decreasing with the square of the distance [5]. Short-duration pulses can be used to achieve a large peak power while maintaining an eye-safe average power, allowing longer-range measurements [18]. High-bandwidth timing electronics are required to achieve fine range resolution.

Laser ranging can alternatively be performed using continuous, rather than pulsed, illumination. Continuous-wave (CW) lidar, operated with a constant amplitude and frequency, can be used to measure velocity from the Doppler shift [213]. By modulating the light's amplitude or frequency, range measurements can also be performed. The amplitude-modulated continuous-wave (AMCW) method, employed by time-of-flight cameras such as the Microsoft Kinect version 2, measures the phase shift between a typically sinusoidal modulation of the transmitted light and its received echo [62]. The range is encoded in the correlation between transmitted and received

signals, which is ideally also sinusoidal and can therefore be recovered from as few as three measurements [213]. However, AMCW is limited to both moderate precision and range by the modulation bandwidth and the average power output, respectively [18].

Unlike pulsed and AMCW methods, frequency-modulated continuous-wave (FMCW) lidar maintains a constant optical intensity and instead modulates the frequency, typically with a linear chirp. The optical interference of the received signal and a local oscillator creates a beat frequency that is proportional to the distance scaled by the known chirp rate [5, 103]. The range is thus determined by sampling the beat signal and performing Fourier analysis, allowing for sub-micrometer precision despite the use of low-bandwidth electronics [12]. The coherent receiver also makes FMCW more robust to ambient light and interference than pulsed and AMCW approaches [18]. Unfortunately, practical implementation of FMCW lidar is more challenging than amplitude-modulated approaches, as the optical frequency must be tunable, and linear chirping is difficult to achieve [5]. Furthermore, the maximum range for FMCW lidar depends on the coherence length of the laser—the range for which the phase noise is coherent. Modulating the wavelength of a laser while keeping the phase noise low enough for ranging of targets at distances greater than 100 m is considered one of the major challenges of FMCW [103, 220]. Typically this is considered a hardware challenge—requiring a laser with a narrow spectral linewidth—although signal processing approaches have been proposed to perform long-distance ranging despite the lack of coherence [103].

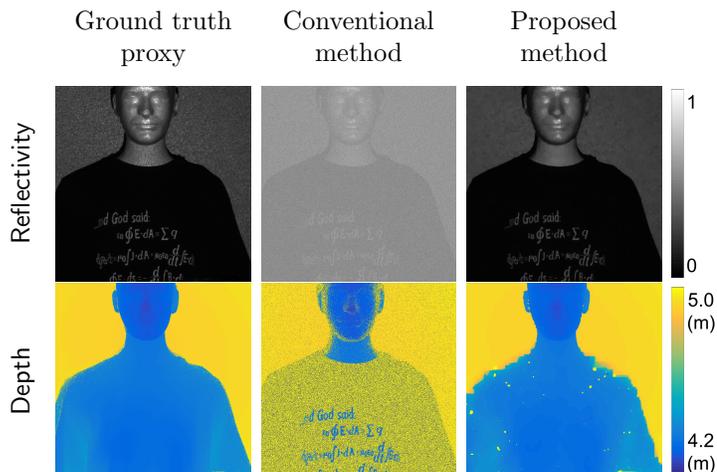
The maximum range is one of the largest remaining challenges for the deployment of lidar in autonomous vehicles (AVs) because it affects driving performance directly: early identification of a potential hazard gives the vehicle more time to make safe driving decisions [85]. At American highway speeds of 65 mph (105 km/h), for in-

stance, a range extension of 30 meters provides an additional second of reaction time. The baseline performance necessary for AVs is generally listed as the ability to detect dark objects with a reflectivity of 10% from a distance of at least 200 m.

The rapidly evolving technology of SPL has potential to overcome the range problem by combining pulsed illumination with single-photon detectors, while achieving high timing precision with TCSPC. Following FPI [104], additional SPL implementations have likewise demonstrated remarkable photon efficiency, forming accurate depth and reflectivity images from about 1.0 detected photon per pixel on average [3, 170, 171]. Such extreme sensitivity means that SPL can handle low-power lasers, fast acquisitions, or large attenuations due to very long distances or fog. For example, experiments with SPL have been used to form point clouds from standoff distances of several hundred meters up to 45 km [116, 143], and even to measure the distance to points on the moon with centimeter precision [54]. By the same principle, airborne SPL systems have been used to measure depth and reflectivity values from greater heights, allowing more efficient scanning of large areas [83, 183]. An extended discussion of SPL system implementation, probabilistic modeling, and basic estimation is the subject of Chapter 2.

## 1.2 Preview of Main Contributions

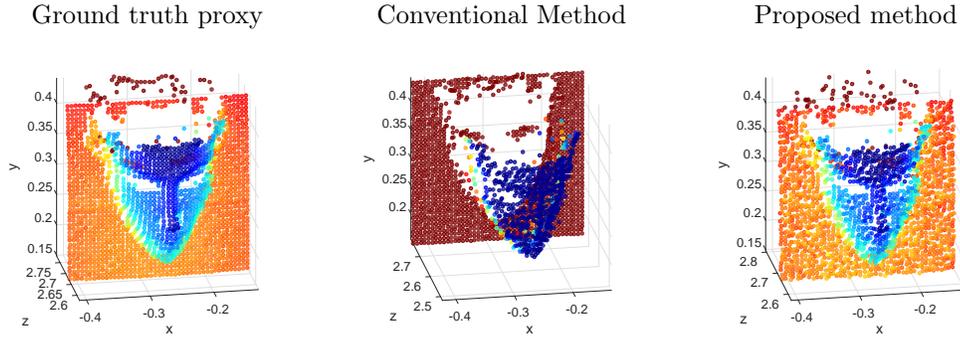
This thesis addresses three challenges to TCSPC acquisition. Although modeling modifications and imaging improvements are shown only for SPL, FLIM is limited by the same constraints of ambient light, temporal quantization, and detector dead times, and some methods developed in this thesis could have direct application in FLIM as well. We now preview some of the improvements that can be achieved by improving the probabilistic modeling in SPL.



**Figure 1-1:** Reflectivity and depth estimation results under photon-starved, high ambient light conditions, with about 4 signal detections per pixel and 25 times as many background detections per pixel on average. We propose a method that censors background detections by searching for clusters of signal detections for each pixel and joins together detection data from similar neighboring pixels.

### Improving Robustness to Strong Ambient Light

Initial photon-efficient lidar experiments were limited to the laboratory, where moderately favorable imaging conditions could be maintained by controlling the ambient light levels. However, practical outdoor uses for SPL must tolerate much stronger ambient light without increasing acquisition times. Chapter 3 describes a method for censoring background detections based on the probability that detections cluster together in time. Spatial correlations between neighboring pixels of similar reflectivity are exploited to extend the censoring to dark regions in which very few pixels have any signal detections. The proposed method is demonstrated on simulated and experimental datasets, such as those shown in Figure 1-1, in which there are only a handful of signal detections and 25 times as many background detections per pixel on average. This material has previously appeared in [147].



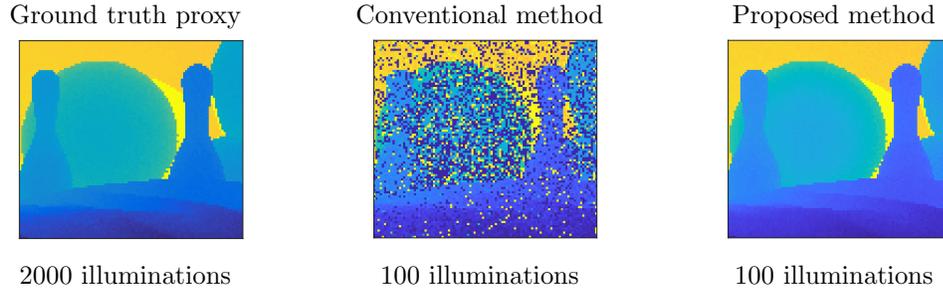
**Figure 1-2:** Depth estimation results for TCSPC systems with coarse timing resolution. The implementation of subtractive dither and the careful choice of an estimator yields substantial improvement in the ability to precisely estimate fine features.

### Mitigating Coarse Temporal Quantization

SPAD array detectors promise fast depth imaging through parallelized acquisition, but current SPAD cameras trade off a higher pixel count for lower timing resolution, which means the precision of measurements using short-duration laser pulses is often limited by the temporal quantization bin size. Chapter 4 demonstrates that implementing subtractive dither within an SPL system can reduce the effects of coarse quantization. Simple estimators are derived and analyzed to efficiently use each photon detection. Experimental results in Figure 1-2 exemplify the improvement in resolving fine features that can be achieved with dithered SPL. Most of this material has previously appeared in [148, 149, 150].

### Compensating for Detector Dead Times

Even in the first TCSPC measurements by Bollinger and Thomas [24], detector dead times were acknowledged as a problem skewing the temporal response histogram if the incident flux was too high. For SPL, dead times can erroneously shift depth estimates, and the standard mitigation approach limits the dynamic range of the scene that can be estimated. Chapter 5 establishes a new way of modeling the absolute



**Figure 1·3:** Depth estimation results for simulated detections illustrate the effectiveness of using the Markov chain modeling for high-flux acquisition. The proposed method approximately matches the conventional performance with  $20\times$  fewer illuminations, greatly speeding up acquisition.

sequence of photon detection times as a Markov chain. Simulated results in Figure 1·3 show an example of how our proposed modeling can avoid the conventional need for attenuation, allowing accurate depth imaging to be performed with far fewer illuminations. Most of this material has previously appeared in [151, 152, 153].

## Chapter 2

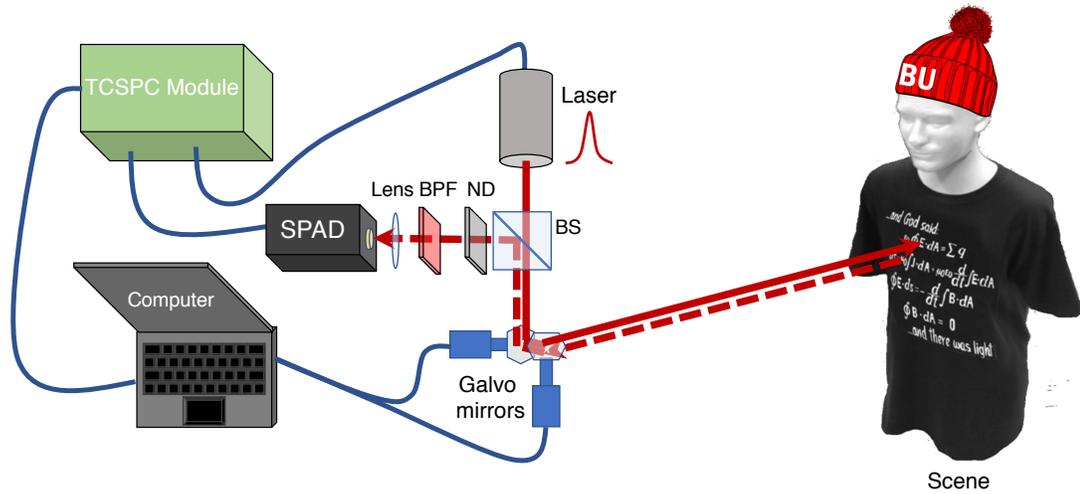
# Background on Single-Photon Lidar

### 2.1 Overview

Following a pulsed laser illumination, whether and when a photon will be detected is a random process. The first step toward forming estimates from SPL data is to understand the various factors affecting the distribution of photon detection times. This chapter begins by discussing the basic components of an SPL system, specifying key details of the laser, detector, and timing electronics, as well as the effects of scene interactions. Next, properties of those components are used to build up a probabilistic model of photon detection times under certain simplifying assumptions. Finally, the basics of reconstructing depth and reflectivity information are discussed, serving as the foundation for the methods introduced in later chapters.

### 2.2 Acquisition System

The key components of a standard SPL system are shown in Figure 2-1. The laser illumination source pulses at megahertz rates and is directed toward one point in a scene by a pair of galvo mirrors. Light reflected back from the scene reaches the SPAD after passing through a bandpass filter at the illumination wavelength and a lens focusing light onto the SPAD's active area. An optional neutral-density filter may also be inserted in the detector path to reduce the flux incident on the detector. Differences between the photon detection times and the most recent laser pulse are recorded by



**Figure 2.1:** A single-photon lidar system shown in a raster-scanned, confocal configuration. The scene is illuminated by a laser pulse (solid line) passing through a beam-splitter (BS) and raster-scanned via a pair of galvo mirrors. Photons reflected back from the scene (dashed line) pass through a bandpass filter (BPF) tuned to the operating wavelength (a neutral-density filter (ND) providing additional attenuation is optional), before being focused by a lens onto the SPAD. Time correlation between illuminations and detections is performed by the TCSPC module, which streams data to the control computer.

timing electronics in a TCSPC module and sent to a computer for processing.

Figure 2.1 highlights that the four main influences on the probabilistic model for SPL—the pulsed laser, single-photon detector, and timing electronics of the SPL system, plus the environmental interactions that occur between light emission and detection—can be divided into two groups based on whether their properties are dominated by *optical* (red) or *electrical* (blue) effects. The following sections describe these effects and emphasize how most probabilistic modeling characterizes only the optical interactions and assumes ideal acquisition equipment. A significant focus of this thesis is to also integrate the effects of non-idealized electronics when relevant.

### 2.2.1 Optical Factors

Optical factors encompass all influences on the light reaching the sensor, including the laser illumination, atmospheric attenuation, surface reflections, and ambient light. We begin with a discussion of the key laser properties and then describe some of the environmental interactions affecting the transient response. We assume throughout a raster-scanned lidar implementation, although the optical model for light reaching a single pixel in a detector array is broadly similar.

#### Laser Properties

The first component in an SPL system is the laser, the source of any useful photon detections. Through excitation of an optical gain medium within a resonant structure, lasers can radiate a beam of light with a narrow spectral width and high power, either continuously or in short pulses [173]. The most important properties of lasers for SPL are the operating wavelength, beam size, pulse duration, peak power, and repetition rate.

The laser wavelength is important for several reasons. The wavelength governs which materials in a scene will reflect, absorb, or even transmit the illumination. The operating wavelength also determines which types of single-photon detector may be used. Silicon SPADs are effective for visible light and some of the ultraviolet spectrum, but their efficiency drops off quickly as the wavelength increases into the infrared range [31]. Active imaging in the near infrared (NIR) and short-wavelength infrared (SWIR) is advantageous because there is little solar contribution in the wavelengths around  $1.3 \mu\text{m}$  and the human eye can tolerate much higher optical power safely, although the available single-photon detectors are less desirable due to manufacturing, temperature, or noise constraints [29]. The experiments presented in this thesis use only visible light illumination for convenience.

A small illumination spot allows for a Dirac transverse sampling model, which means opaque objects have a single depth per pixel and the temporal response function is mostly invariant to the surface normal. A short pulse duration allows measurements with a low depth uncertainty. The peak power is important because it determines in part the number of photons received by the detector. Finally, the repetition rate influences the acquisition speed. As described later in this chapter, typically only a small fraction of illumination pulses lead to photon detections, so it is generally desirable to use high repetition rates to acquire enough detections as fast as possible. However, it is important to note that periodic illumination limits the maximum range that can be unambiguously estimated, so the repetition rate is usually chosen to be the highest rate that can measure the scene extent without depth aliasing.

Diode lasers are usually the illumination source of choice for TCSPC in general and SPL in particular. Compared to other laser technologies, diode lasers can be manufactured in the smallest size package and with the highest differential power efficiency and longest lifetime [165]. A diffraction-limited beam quality is possible, yielding a small spot size. By applying gain-switching, diode lasers can be designed with pulse durations on the order of tens of picoseconds at easily-adjustable repetition rates from kHz to GHz, also allowing for external triggering [10, 117, 142]. The maximum optical power of a gain-switched diode laser is typically in the range of a few tens or hundreds of milliwatts, although fiber amplification is also possible [165]. The pulse’s temporal profile is most symmetric (e.g., closest to a Gaussian shape) at the lowest power, i.e., just above the lasing threshold. Otherwise, secondary “after-pulses” closely following the main peak become noticeable as the pulse energy increases, a phenomenon not to be confused with the SPAD afterpulsing described later in this chapter.

## Environmental Interactions

After the laser illuminates a short-duration pulse, any of the  $Q$  surfaces within the beam width will reflect some light back, delayed by a time dependent on the distance  $z_q$  to surface  $q$ , and attenuated by the surface reflectivity  $\alpha_q$ . The amount of light reflected back is also proportional to the optical intensity of the laser illumination. In addition, ambient light at the same wavelength as the laser, or more generally within the spectral range of the detector, will reach the detector, although the timing of that incident light does not contain any depth information.

Since SPL has been deployed in a wide variety of environments, from controlled laboratory settings to mountings on aircraft or autonomous vehicles, it is important to consider the associated optical interactions and their effects on signal acquisition.

**Number of Photons.** The number of incident photons may be small or even zero for several reasons: the number of illuminations  $n_r$  is kept low for real-time acquisition, the surface reflects very little light because it is weakly-reflective or far away, etc. Beyond the absolute number of photons, an important factor is the relative numbers of informative signal and uninformative ambient photons. For example, estimation is particularly challenging if the ratio between the number of photons due to the laser and ambient illuminations, referred to as the signal-to-background ratio (SBR), is low. Even though optical methods (e.g., confocal configurations, bandpass filters) are used to limit the amount of ambient light that reaches the detector, strong daylight, especially when combined with a weak surface reflection, can result in far more detection events associated with background photons than from signal photons.

**Number of Surfaces.** The most basic 3D reconstruction methods assume a single surface ( $Q = 1$ ) at each pixel location. If a pixel has no object in its line of sight ( $Q = 0$ ), then the histogram contains only background detection events. On the other

hand, there may be reflections from multiple surfaces present at one pixel ( $Q > 1$ ). This may occur because the light passes through a semi-transparent material such as glass. Alternatively, the pixel size or field of view increases with distance (e.g., due to the laser divergence in a scanned setting), so the spot is more likely to cover multiple surfaces. This same principle is often used in foliage-penetrating airborne lidar used for terrain mapping, for example to find an ancient city hidden under dense rainforest [35].

**Pulse Width.** Surfaces are generally assumed to be opaque and approximately normal to the illumination beam so that the reflected temporal response closely resembles the shape of the illumination pulse. However, sub-surface scattering or oblique-angled surfaces due to beam divergence at long distances will return broadened pulse profiles [185].

**Attenuating Media.** Particles in the beam path, such as fog, smoke, rain, or snow, affect the acquired light by scattering photons in different directions after both the illumination (forward path) and reflection (return path). To some extent, the result is similar to that of a signal weakened by additional attenuation and increased background due to scattered photons [188], although the near-range effects of scattering also reshape the temporal distribution of background, with more detections at earlier times [163]. Similar effects are also encountered for lidar in underwater environments [80].

### 2.2.2 Detection Electronics

After light is reflected back toward the detector, the acquisition electronics determine which photon detection events are registered and what time information is stored for each event.

## Single-Photon Avalanche Diodes

Although numerous single photon detection technologies have been employed in SPL, we focus here on SPADs due to their widespread use, including in this work. SPADs are reverse-biased photodiodes biased above the breakdown voltage with basic operation as follows [15, 42, 56, 177]. When a photon hits a SPAD's active area, a charge carrier may be generated via the photoelectric effect. Due to the reverse bias, generation of one carrier will further cause an avalanche of carriers, generating a current that is detectable as a digital signal. In order for the detection circuit to be sensitive to subsequent photon arrivals and to protect the avalanche from damaging the diode, the avalanche is *quenched* by either passively or actively reducing the bias below the breakdown voltage. In actively-quenched detectors, the bias voltage is reset to its initial level after a fixed *hold-off* time  $t_{ho}$ , and the duration of the hold-off plus the reset is considered to be the detector dead time  $t_d$ . This operation lends itself to a nonparalyzable detector model, in which a fixed-duration dead time of insensitivity follows each detection event.

Not all photons incident on the detector actually cause an avalanche of electrons. The *photon detection efficiency*  $\eta$  is the probability that a photon incident on the detector will cause a detection event registered by the TCSPC system. In a SPAD, this value is the product of the *quantum efficiency* (the probability of an incident photon generating a carrier, which depends on material properties), and the *avalanche triggering probability* (the probability that a carrier initiates a detectable avalanche, which depends on the photodiode bias voltage) [67]. Detector efficiency varies greatly depending on the photodetector technology and operating wavelength [56]. Our work assumes the detection efficiency is binary during the detection process, i.e., equal to zero during the dead time and equal to the steady-state value  $\eta$  otherwise. In actual circuits, however, the return of the bias voltage to its initial level after hold-

off is not immediate and is sometimes referred to as the *reset time* [177, 42]. The hold-off time plus the reset time, which is the total time between a photon detection and when the detector is returned to its fully-armed state, is sometimes called the *recovery time*. The *dead time* may then refer to either the hold off time (when the probability of photon detection is zero) or the recovery time (the time until the detector is fully reset); if the reset time is sufficiently short, the definitions are equivalent, but the nomenclature is ambiguous nonetheless. It is also useful for the reset time to be as short as possible because photons arriving during the reset time may cause unpredictable delays in their detection time [67], leading to a perceived arrival rate dependence in the afterpulsing probability [212].

Photons incident on the detector may be due to the desired signal, e.g., back-reflected illumination in lidar or fluorescence from a stimulated molecule in microscopy, or the photons may result from ambient light. Not all avalanches are caused by electrons generated by the photoelectric effect, however. *Dark counts* are false detection events caused in SPADs by thermally-generated carriers that trigger avalanches independent of any incident photons [71]. Dark counts are typically assumed to occur at a constant rate. *Afterpulses* are also false detection events, but unlike dark counts, they are correlated with previous detections [42]. In SPADs, charge carriers can become trapped in semiconductor defect sites during an avalanche; carriers emitted or “detrapped” after the SPAD has been reset can cause subsequent avalanches triggering a detection [47, 96]. Several models have been proposed for the temporal distribution of afterpulses based on the device physics and empirical measurements, including a multi-exponential model [47, 71, 106] and a power-law fit [96], although the best fit model can vary by device [221]. Typically, SPADs are held off long enough for the probability of an afterpulse detection to be sufficiently small and indistinguishable from dark counts [93].

## TCSPC Timing Electronics

While the SPAD provides single-photon sensitivity and a fast response, precise timing is achieved using specialized TCSPC electronics. The classical approach, introduced by Bollinger and Thomas in 1961 and still used by the Becker & Hickl systems popular for FLIM, is based on a stopwatch principle in which one signal starts and another signal stops a timer [15, 24]. The most straightforward implementation has a synchronization signal simultaneous with the illumination to provide the “start,” with a detection event stopping the timer. However, most TCSPC operation occurs with a low-flux input, so that the probability of a detection after any excitation is low, and in most cycles the timer would start without ever receiving a stop signal. Classical systems are thus more commonly operated in reverse start-stop mode instead, in which the timer is started by a detection event and the stop signal is a previous synchronization signal that has been delayed by a known duration. Classical TCSPC systems rely on an analog implementation called a time-to-amplitude converter (TAC), which charges a capacitor in the time between the start and stop signals [99]. The time duration is encoded as the capacitor voltage and can be read out via a high-resolution analog-to-digital converter (ADC). Note that in early TAC implementations, the recording was not digitized but rather captured via other analog means, e.g., a film camera recording an analog oscilloscope, or a mechanical trace [160]. TAC/ADC architectures are incorporated into the highest-resolution commercial systems, such as the SPC-150NXX from Becker & Hickl, which has a minimum bin width of 203 fs. Unfortunately, TAC-based TCSPC is poorly suited to lidar applications due to the slow readout speed, temperature sensitivity, short total measurement range, and inability to record more than one detection event per cycle [99, 207].

An alternative approach to classical TAC-based TCSPC relying on time-to-digital converters (TDCs) has gained popularity since the 1990s. The basic idea of a TDC

is that the time need not first be converted to an analog voltage before conversion to a digital number [158]. Instead, synchronized digital timers can be applied to each input channel (one for the illumination synchronization signal and one for each detector input), and the time differences can be computed digitally. This architecture enables a virtually unlimited measurement range, fast digital readout, the “multi-stop” capability of recording multiple stop signals within one cycle, and monolithic integration [99].

The simplest TDC implementation is a counter driven by a high-speed reference clock; however, the minimum time resolution is then limited by the clock speed, so a 1-THz clock would be required to achieve 1-ps resolution [99, 205]. Instead, most modern TDC-based TCSPC systems combine a coarse, crystal clock-based counter with finer timing that interpolates within clock cycles. To achieve the highest-resolution TDCs, the fine interpolation has typically required an analog component similar to a TAC [205]. A linear voltage ramp or pure sinusoid synchronized with the coarse clock can be sampled and processed to achieve 1-ps resolution, while still maintaining the multi-stop and unlimited range capabilities of the coarse counter. Because such precise measurements require the circuitry to have recovered between event recordings, such TCSPC systems—like the HydraHarp 400 from PicoQuant—enforce a dead time  $t_e$  (of around 80 ns) via a monostable [208, 206]. Purely digital fine interpolation can be performed with a tapped delay line linked to the overall clock through a delay-locked loop [158]. This approach is known as a “Flash TDC” because of the fast digital readout time, but the gate speed of the delay elements usually limits the resolution to be on the order of tens of picoseconds. One commercial example using purely digital timing is the TimeHarp 260 from PicoQuant, which has a short dead time (25 ns) but lower minimum bin resolution (25 ps) [209]. Another all-digital fine interpolator uses a Vernier oscillator approach, but that architecture can lead to much

longer readout times [99]. While these all-digital TDCs are usually implemented in application-specific integrated circuits (ASICs), a number of designs propose direct TDC implementation in a field-programmable gate array (FPGA) [60]. Recently, purely digital TDCs using tapped delay lines implemented in FPGAs, such as the Time Tagger Ultra from Swabian Instruments, have achieved 1-ps resolution with only 2.25-ns dead times [61].

### **2.2.3 Acquisition Variations**

The vast majority of SPL systems use a standard combination of the laser, SPAD, and timing electronics, as described above. However, the particular equipment and the acquisition methodology affect how data is acquired and what information is encoded. This section discusses several of the most notable variations in SPL data acquisition.

#### **Transverse Resolution**

Transverse resolution is achieved through either the illumination or detection or both, each approach presenting different advantages and drawbacks. SPL systems have conventionally employed raster-scanned illumination, as illustrated in Figure 2-1. A laser aimed at one spot in a scene repeatedly pulses for a certain dwell time before being redirected to the next spot by a pair of galvo mirrors [128]. Scanned illumination enables the use of a single-pixel or bucket detector, which is often in a confocal configuration as depicted in Figure 2-1, using a beamsplitter to allow the illumination and detection components to share a common optical axis. The advantage of the confocal arrangement is that the SPAD has a narrow field-of-view (FOV), observing only the region of the scene toward which the laser is aimed and thereby rejecting multipath interference and reducing the contribution of ambient light from other parts of the scene. However, aligning a confocal system can be complicated, so many laboratory setups operating at short range and with controllable ambient light opt

for the simplicity of a non-confocal or *bistatic* configuration, in which the SPAD has a wide FOV and is simply placed next to the galvos to be approximately monostatic. Although the SPAD then detects ambient light from all parts of the scene at each pixel location, an advantage is that the background rate is constant throughout the scene acquisition, allowing calibration with a single simple measurement while the laser is turned off. The background rate for a confocal system must be estimated separately for each pixel. Most of the experimental data presented in this thesis were acquired with a bistatic configuration.

Imaging with a scanned illumination source is sometimes called *pseudoarray* imaging in that it idealizes the array rather than compensating for array-specific non-uniformities. Unfortunately, raster scanning is an inherently slow, serial process. A more recent approach has been to broadly illuminate a swath of the scene and achieve spatial resolution with an array of single-photon sensitive elements [171]. While detector arrays promise faster, parallelized acquisition, existing arrays are still limited in their spatial resolution. One alternative approach is to use a sequence of illuminations patterned by a higher-resolution digital micromirror device (DMD) or other spatial light modulator (SLM) in conjunction with a single-pixel detector. While this technique faces the challenges of nonlinear encoding of depth information and a loss of directional information in the mixed measurements, several implementations have shown that the inherent sparsity of depth maps enables accurate depth and reflectivity recovery via compressed sensing algorithms applied to a reduced number of measurements [45, 92, 182]. Still, broadly diffusing the laser power over a large area reduces the signal strength received at each pixel location. One compromise is the use of a line illumination and line array detector, which reduces the spatial scanning to a single dimension and limits the diffusion of the laser power [119].

## Dwell Time

An important acquisition consideration is determining for how long photons should be acquired at each pixel. This is particularly important for scanned acquisition because the photons are acquired independently for each pixel, so it is easier to introduce acquisition variations. The initial method of photon-efficient SPL proposed in [104] acquired exactly one photon per pixel, which thus required a random number of illuminations before the photon was detected. However, random dwell times are challenging to implement for scanned systems, since real-time feedback is required from the detection electronics to the galvo control, and is impractical for existing detector arrays, which transfer data after fixed numbers of illuminations. Thus the standard SPL acquisition approach uses a fixed dwell time (i.e., a fixed number of illuminations  $n_r$ ) at each pixel, with the number of detected photons randomized by the acquisition and scene properties [170]. Alternative dwell time strategies have also been suggested, such as [129], which proposes an adaptive stopping rule that depends on the number of both illuminations and detections, or [87], which proposes continuous laser scanning (no dwell time at any pixel), with the transverse position of a photon detection determined during the reconstruction.

## Illumination Coding

The standard SPL implementation uses a periodic laser illumination with repetition period  $t_r$ . The periodicity places a hard limit on the maximum scene depth  $z_{\max}$  that can be unambiguously measured without distance aliasing. If instead  $z_{\max}$  is greater than the unambiguous range  $ct_r/2$ , it is possible to avoid aliasing as long as the extent of the scene can be captured within one period. Making absolute range measurements then requires a combination of high and low repetition rate acquisitions [128]. Alternatively, the unambiguous range can be extended arbitrarily

by using non-periodic illumination. For instance, Hiskett et al. [91] and Krichel et al. [108] avoid range ambiguity by triggering the illumination with a pseudo-random pulse sequence that is then cross-correlated with the detection time histogram to identify the correct delay.

#### 2.2.4 Basic Experimental Setup

Later chapters present experimental results acquired with the same or a similar experimental setup and procedure as described in [104, 105]. While we highlight the variations or modifications made when relevant, here we describe the basic setup that informs the derivation of the measurement model.

A picosecond laser (PicoQuant LDH-P-C-640B) at operating wavelength 640 nm is pulsed by the laser driver (PicoQuant PDL-800-D), typically with repetition frequency  $f_r = 10$  MHz (corresponding to a repetition period  $t_r = 100$  ns). The illumination spot is redirected by a pair of galvo mirrors (Thorlabs GVS012), which is controlled by software through a data acquisition (DAQ) interface (such as the National Instruments USB-6363). Simultaneously with the illumination trigger, the laser sends a synchronization signal to the TCSPC electronics (PicoQuant HydraHarp 400), which starts a timer.

The “stop” signal for the timer is a detection event registered by the SPAD detector (Micro Photon Devices PDM-series, with photon detection efficiency  $\approx 35\%$  at 640 nm). Mounted in front of the SPAD is a bandpass filter (Thorlabs FB-640-10) with a transmission efficiency of  $\approx 50\%$  at the operating wavelength and a FWHM bandwidth of 10 nm to reduce the amount of ambient light incident on the detector.

The laser is directed by the galvos to illuminate each point in sequence for a specified dwell time per point. Detected photons are time-stamped by the TCSPC module and streamed to the computer. When the DAQ changes the galvo voltages to change the coordinates of the laser position, it simultaneously sends a marker to

the TCSPC module indicating the position change. The standard configuration is not confocal, so that while the laser is scanned, the detector has a wide field of view encompassing the entire scene. Importantly, this nonconfocal configuration implies that the intensity of background light is uniform over all pixels.

## 2.3 Basic Probabilistic Measurement Model

In this section, we describe the probabilistic model for a periodically-illuminated, raster-scanned SPL system, such as those used for the experiments presented later in this thesis. As before, we structure our discussion by the source of the effect on the model: whether the effect is due to how the illumination interacts with the scene or how the electrical signal is recorded.

### 2.3.1 Optical Effects

The laser is directed toward each scene patch indexed by  $(i, j) \in \{(1, \dots, n_i) \times (1, \dots, n_j)\}$ , and the laser illuminates each patch with  $n_r$  pulses. The arrival times of photons incident on the detector from each  $(i, j)$  are described by a Poisson process with intensity  $r_{i,j}(t)$  [174]. Due to the pseudo-array imaging, which has identical illumination and detection properties for each  $(i, j)$ , we drop the pixel indexing to describe acquisition at one pixel with less tedious notation. The illumination is periodic with period  $t_r$ , and the process intensity in one period is approximately given as

$$r(t) = b_a + \sum_{q=1}^Q \alpha_q \beta s(t - 2z_q/c), \quad \text{for } t \in [0, t_r), \quad (2.1)$$

where  $b_a$  is the intensity of the ambient light at the illumination wavelength, assumed to be constant for the duration of the acquisition;  $\alpha_q \in [0, 1]$  is the *reflectivity* of surface  $q$ , combining attenuation effects due to object reflectance, radial falloff, view angle, etc.;  $\beta \in [0, \infty)$  is the *illumination gain*, corresponding to the expected number

of photon arrivals per illumination from a unit-reflectivity object;  $s(t)$  is the illumination pulse shape normalized to be a valid probability density function (PDF) of arrival times;  $z_q$  is the depth of surface  $q$ ;  $Q$  is the number of surfaces present in the field of view; and  $c$  is the speed of light (around  $3 \times 10^8$  m/s in air). For the majority of this work, we assume the pulse shape is Gaussian  $s(t) = (2\pi\sigma_p^2)^{1/2} \exp(-t^2/2\sigma_p^2)$  with standard deviation  $\sigma_p$ , and  $\beta$  and  $b_a$  are assumed constant for all  $(i, j)$ .

### Detectable Events

At the detector, the photon flux is attenuated by the detection efficiency  $\eta \in [0, 1)$ . Detector dark counts are assumed to be generated as a homogeneous Poisson process with intensity  $b_d$ . Both of these effects maintain the overall nature of *detectable events*—those that could be registered in an ideal system—as a Poisson process with total detection intensity given by

$$\begin{aligned} \lambda(t) &= \eta r(t) + b_d \\ &= \eta\beta \left[ \sum_{q=1}^Q \alpha_q s(t - 2z_q/c) \right] + (\eta b_a + b_d), \quad \text{for } t \in [0, t_r). \end{aligned} \quad (2.2)$$

It is useful to consider the first term as the intensity of an inhomogenous Poisson process  $\lambda^s(t)$  combining all informative *signal* contributions, while the second term is the intensity of a homogenous Poisson process  $\lambda^b$  representing uninformative *background* counts.

The expected number of signal photon arrivals per period at  $(i, j)$  is defined as

$$S_{i,j} = \int_0^{t_r} \lambda_{i,j}^s(t) dt = \alpha_{i,j} \eta \beta, \quad (2.3)$$

where  $\alpha_{i,j} = \sum_{q=1}^Q \alpha_q$ , and the expected number of background arrivals is  $B = \lambda^b t_r$ . The total flux at each point is given as  $\Lambda_{i,j} = S_{i,j} + B$ , and the average signal-to-

background ratio is

$$\text{SBR} = \frac{\eta\beta}{n_i n_j B} \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \alpha_{i,j}. \quad (2.4)$$

### Detection Probability Distributions

SPL systems are typically operated in the low-flux regime, where  $\Lambda \ll 1$ , so the probability of a detection in any given illumination period is small, and the probability of multiple detections in one period is negligible. Little error is thus incurred by ignoring the detector dead time effects at this stage of the model. Each detection time can then be described an independent, identically distributed random variable  $X$  with common probability density [174]

$$p_X(x) = \frac{\lambda(x)}{\Lambda}, \quad x \in [0, \mathbf{t}_r). \quad (2.5)$$

Unlike in [104], which uses a randomized dwell time to collect a fixed number of detections per pixel, we assume the detection process is observed for a fixed, deterministic number of illumination repetition periods; thus, no information is conveyed by the order of the detection times or the identities of the repetition intervals in which the detections occur. The TCSPC histogram acquisition mode exploits the periodicity of  $\lambda(t)$  to wrap detection times (relative to the illumination time) into one period, discarding any indicator of the illumination cycle in which each particular detection occurred. We thus fold time interval  $[0, \mathbf{n}_r \mathbf{t}_r)$  down to  $[0, \mathbf{t}_r)$  to obtain an equivalent model in which all detections occur within one illumination period due to a process with intensity

$$\lambda_{i,j}^{\mathbf{n}_r}(x) = \mathbf{n}_r[\lambda_{i,j}^s(x) + \lambda^b] \quad (2.6)$$

and rate

$$\Lambda_{i,j}^{\mathbf{n}_r} = \mathbf{n}_r(S_{i,j} + B). \quad (2.7)$$

The distribution of photon counts is

$$K_{i,j} \sim \text{Poisson}(\Lambda_{i,j}^{n_r}) \quad (2.8)$$

and the probability density of detection times remains as (2.5). The number of detections due to signal is  $M_{i,j} \sim \text{Poisson}(n_r S_{i,j})$ , and the number of detections due to background is  $N_{i,j} \sim \text{Poisson}(n_r B)$ .

### Binomial vs. Poisson Modeling

In initial photon-efficient SPL experiments [104, 170, 171], the duration of the SPAD dead time was similar to the repetition period  $t_r$ . Thus, initial probabilistic models assumed at most one detection event could be recorded for each pulse-repetition period. As developed in [170], under the simplifying approximation that a dead period ends at the subsequent pulse-repetition boundary, this makes detection within each pulse-repetition period a Bernoulli trial and the total number of detections in  $n_r$  pulse-repetition periods a binomial random variable. More precisely,

$$K_{i,j} \sim \text{binomial}(n_r, 1 - P_0), \quad (2.9a)$$

where

$$P_0(\alpha_{i,j}) = \exp[-(S_{i,j} + B)] \quad (2.9b)$$

is the probability of zero detections in one pulse-repetition period.

Under a low-flux assumption, the models (2.8) and (2.9) for  $K_{i,j}$  are approximately equal; a formal equivalence can be shown through the Poisson limit theorem. A downside of the binomial model is philosophical: it encourages one to discard the detection times when estimating reflectivity, as is done in [104, 170, 171]. The Poisson model instead encourages the separation into signal and background processes, which leads to a separation of  $K_{i,j}$  into its constituents  $(M_{i,j}, N_{i,j})$ ; estimation of reflectivity

from  $M_{i,j}$  is more accurate than from  $K_{i,j}$ , especially when SBR is low, as discussed in Chapter 3.

### 2.3.2 Non-ideal Electronics Effects

Although the occurrence of detectable events may be correctly described by a Poisson process with continuous-valued arrival times, practical TCSPC systems can neither capture every incident photon nor record detection times with infinite precision. We next describe how the acquisition electronics affect the measurement model in two key ways.

#### Temporal Quantization

Most important to the basic acquisition process is the effect of temporal quantization. Each photon detected by the SPAD is assigned a time stamp, where the resolution  $\Delta$  is dictated by the TCSPC electronics. The detection times are recorded in sequence, and we refer to the *absolute detection times* when periodicity is not taken into consideration (i.e., with  $n_r$  illumination cycles, we have  $0 < t_1 \leq t_2, \dots, \leq t_K \leq n_r t_r$ ). Since a large number of bits is required to store picosecond-resolution time tags for acquisitions that can last from seconds to hours, it is more common for periodic illumination to store the *relative detection times*  $x_1, \dots, x_K$ , referring to the time of detection relative to the most recent illumination pulse, which is the absolute detection time modulo  $t_r$ . Rather than processing each of the  $K$  photon detection times separately, especially if the number of detections exceeds the number of time bins, classical approaches to SPL typically form a histogram  $[y_1, \dots, y_{n_b}]$  of  $n_b = \lceil t_r/\Delta \rceil$  time bins, which is constructed using the detection times at each pixel  $(i, j)$ . The observation model for each time bin is then a Poisson random variable, where the

parameter integrates the Poisson process intensity over that bin:

$$y_n \sim \text{Poisson} \left( \int_{(n-1)\Delta}^{n\Delta} \lambda^{nr}(x) dx \right), \quad \text{for } n = 1, \dots, n_b. \quad (2.10)$$

When considered together, forming a histogram for each of the  $n_i \times n_j$  transverse pixel locations yields a 3D data cube.

The ability to accurately resolve transient information depends on the width of the histogram bins. For raster-scanning systems, the bin resolution that can be achieved currently is on the order of picoseconds, which is typically much less than the duration of pulse  $s(t)$ , so quantization effects on the depth estimation can be neglected. However, the timing resolution of detector arrays is usually coarser for each element than for a single-pixel device due to hardware and readout constraints. The single-photon-sensitive elements and timing electronics can easily be constructed as separate elements for a single pixel, whereas in 2D arrays, the timing electronics must be integrated on-chip for each pixel, resulting in a trade-off between the fill factor of the photo-sensitive detector and timing components. This becomes particularly problematic if the bin size  $\Delta$  is larger than the duration of pulse  $s(t)$ . In that case, the depth resolution that can be achieved is quantization-limited and can make object detection and recognition more difficult. The problem of quantization-limited depth estimation is the subject of Chapter 4.

## Dead Times

As previously described, both the detector and the timing electronics have a reset period known as a *dead time* following each detection, during which no further photons can be registered. One of the main implications of these dead times is that the sequence of detection times can no longer be described by a Poisson process: whether a photon will be detected now depends on the time of the most recent detection.

Dead times are thus much more significant in the *high-flux* regime, where  $\Lambda \ll 1$  is no longer valid. The dead time effect, if not handled properly, causes distortions in the detection time histogram that may result in erroneous depth and reflectivity estimates, thereby making accurate localization or object recognition more difficult. Properly modeling and mitigating the effects of dead times is the subject of Chapter 5.

## 2.4 Reconstruction Basics

### 2.4.1 Depth Map vs. Point Cloud

The aim of a reconstruction has an important role both in how the reconstruction is performed and how it can be evaluated. A key distinction is between methods in the number of surfaces assumed per pixel. One category of problem is called target detection [4, 184], in which at most one surface is present per pixel ( $Q \leq 1$ ). More generally, some methods aim to reconstruct a *point cloud*, a set of points with 3D coordinate and reflectivity information. Point clouds are useful when there are multiple surfaces present in some pixels, although reconstruction becomes far more challenging due to multi-modal detection time likelihoods and more complicated priors. One set of strategies for solving the multi-depth problem aims to estimate a 3D volume of reflectivity values, where the volume is sparse (the 3D points correspond to a relatively small number of nonzero reflectivity values). Sparsity is promoted by various convex relaxations: the  $\ell_1$ -norm in [172] promotes sparsity within histograms, the  $\ell_{2,1}$ -norm in [81] promotes correlations across neighboring pixels, and [82] also considers non-local correlations. A second strategy directly estimates a 3D point cloud, where the dimension of the parameter space (i.e., the number of 3D points) is *a priori* unknown. This approach generally uses reversible-jump Markov chain Monte Carlo (RJ-MCMC) algorithms to handle the varying dimensions and a spatial point process prior to model the manifold structure of 3D surfaces [89, 90, 186, 185]. Unfortunately,

having an unknown number of surfaces per pixel also makes reconstruction accuracy difficult to quantify.

In the tradition of [104, 170], this thesis strictly assumes  $Q = 1$ , meaning there is exactly one surface for each scanned pixel. The goal of a reconstruction is then to produce a depth and reflectivity map, with one depth or reflectivity value per pixel. We let  $\mathbf{z}$  and  $\boldsymbol{\alpha}$  denote the full depth and reflectivity images, and  $\hat{\mathbf{z}}$  and  $\hat{\boldsymbol{\alpha}}$  denote their estimates, respectively. The success of a reconstruction can then be easily quantified. The mean-squared error (MSE) used for reflectivity estimation is given as

$$\text{MSE}(\boldsymbol{\alpha}, \hat{\boldsymbol{\alpha}}) = \frac{1}{n_i n_j} \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} (\alpha_{i,j} - \hat{\alpha}_{i,j})^2, \quad (2.11)$$

which is sometimes reported in dB, i.e.,  $10 \log_{10}(\text{MSE})$ . The root mean-square error (RMSE) is used for depth estimates with units of meters:

$$\text{RMSE}(\mathbf{z}, \hat{\mathbf{z}}) = \sqrt{\frac{1}{n_i n_j} \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} (z_{i,j} - \hat{z}_{i,j})^2}. \quad (2.12)$$

## 2.4.2 Parameter Estimation

Basic parameter estimation procedures for reflectivity and depth maps assume one surface per pixel ( $Q = 1$ ), a known background rate  $B$ , and negligible quantization and dead time effects. The form of the reflectivity estimate varies slightly depending on the measurement model (binomial vs. Poisson).

### Binomial Model of Detection

The binomial model (2.9) developed by Shin et al. [170] results in a constrained ML (CML) reflectivity estimate given by

$$\hat{\alpha}_{i,j}^{\text{CML}} = \max \left\{ \frac{1}{\eta\beta} \left[ \log \left( \frac{\mathbf{n}_r}{\mathbf{n}_r - k_{i,j}} \right) - B \right], 0 \right\}. \quad (2.13)$$

Using the approximation

$$\log\left(\frac{n_r}{n_r - k_{i,j}}\right) = \log\left(1 + \frac{k_{i,j}}{n_r - k_{i,j}}\right) \approx \frac{k_{i,j}}{n_r - k_{i,j}} \approx \frac{k_{i,j}}{n_r},$$

the bracketed quantity in (2.13) is essentially the fraction of illumination intervals with a detection, adjusted by  $B$  because  $B$  is the expected number of background detections per illumination period; thus the bracketed quantity is an estimate of the fraction of illumination periods with a signal process detection. Informally, the quality of  $\hat{\alpha}_{i,j}^{\text{CML}}$  as an estimate of  $\alpha_{i,j}$  depends on  $n_r B$  being a good estimate of the true number of background detections  $N_{i,j}$ . Since the variance of  $N_{i,j}$  is proportional to  $B$ , large background rates—especially relative to  $S_{i,j}$ —lead to poor performance of  $\hat{\alpha}_{i,j}^{\text{CML}}$ . This is consistent with Cramér–Rao lower bound computations in [170, App. A].

### Reflectivity Estimation from a Poisson Process

By instead approaching detection entirely as an inhomogeneous Poisson process, one can take advantage of the detection times in both the reflectivity and depth estimates. Estimation of  $\alpha_{i,j}$  then requires the same approach as estimation in amplitude-modulated optical communication as described in [11]. The likelihood function for the set of observed photon detections  $\{x_{i,j}^{(\ell)}\}_{\ell=1}^{k_{i,j}}$  is

$$p\left[\{x_{i,j}^{(\ell)}\}_{\ell=1}^{k_{i,j}}; \alpha_{i,j}, z_{i,j}\right] = e^{-\Lambda_{i,j}^{n_r}} \prod_{\ell=1}^{k_{i,j}} \lambda_{i,j}^{n_r}(x_{i,j}^{(\ell)}),$$

which yields a constrained maximum likelihood (CML) estimate given by

$$\hat{\alpha}_{i,j}^{\text{CML}} = \arg \max_{\alpha_{i,j} \geq 0} \sum_{\ell=1}^{k_{i,j}} \log\left[n_r(\eta\beta\alpha_{i,j}s(x_{i,j}^{(\ell)} - 2z_{i,j}/c) + (\eta b_a + b_d))\right] - \Lambda_{i,j}^{n_r}. \quad (2.14)$$

Differentiating with respect to  $\alpha_{i,j}$ , we have

$$\sum_{\ell=1}^{k_{i,j}} \frac{\eta\beta s(x_{i,j}^{(\ell)} - 2z_{i,j}/c)}{\eta\beta\alpha_{i,j}s(x_{i,j}^{(\ell)} - 2z_{i,j}/c) + (\eta b_a + b_d)} = \mathbf{n}_r\eta\beta. \quad (2.15)$$

Since all terms in (2.15) are nonnegative, the left-hand side is monotonically decreasing in  $\alpha$ , so a unique optimal estimate of  $\alpha_{i,j}$  exists. Unfortunately, this expression requires knowledge of the true depth for the optimal estimate.

At high SBR, an approximate solution is given by

$$\hat{\alpha}_{i,j}^{\text{CML}} = \max \left\{ \frac{k_{i,j} - \mathbf{n}_r B}{\mathbf{n}_r \eta \beta}, 0 \right\}, \quad (2.16)$$

which preserves the non-negativity of  $\alpha$  and simplifies to

$$\hat{\alpha}_{i,j}^{\text{ML,background-free}} = \frac{k_{i,j}}{\mathbf{n}_r \eta \beta} \quad (2.17)$$

if background is completely eliminated. Conveniently, these estimates have a closed form solution, which is simply the normalized photon count. When background is low, all detections are due to signal, so the count is a sufficient statistic for the reflectivity and knowledge of the detection times or true depth is no longer necessary.

## Depth Estimation

The process of depth estimation from the set of detection times is also derived in [11].

The CML depth estimate is given by

$$\hat{z}_{i,j}^{\text{CML}} = \arg \max_{z_{i,j} \in [0, z_{\max})} \sum_{\ell=1}^{k_{i,j}} \log [\eta\alpha_{i,j}s(x_{i,j}^{(\ell)} - 2z_{i,j}/c) + (\eta b_\nu + b_d)]. \quad (2.18)$$

We can see that this requires knowledge of the true  $\alpha_{i,j}$  value, and furthermore that the background term adds a nonconvexity. In practice,  $\hat{z}_{i,j}$  is computed by finding the delay that maximizes the output of a log-matched filter (LMF), equivalent to

cross-correlating the detection time histogram with the logarithm of  $s(t)$ .

Again we remark on the case of zero background, where the depth estimate is given by

$$\hat{z}_{i,j}^{\text{CML}} = \arg \max_{z_{i,j} \in [0, z_{\max})} \sum_{\ell=1}^{k_{i,j}} \log \left[ s(x_{i,j}^{(\ell)} - 2z_{i,j}/c) \right]. \quad (2.19)$$

In this case the background-free solution is also greatly simplified, as it is convex and has no dependence on  $\alpha_{i,j}$ .

## Chapter 3

# Unmixing Signal and Background Light

### 3.1 Overview

The methods of [104, 170] for performing photon-efficient SPL imaging allow the formation of accurate depth and reflectivity maps in low-light scenarios where histogramming techniques perform poorly. However, the imaging accuracy degrades significantly in the presence of high background, which is of particular importance for long-distance or low-power measurements in daylight, when the rate of photon detection from ambient light may be significantly higher than the detection rate from the active illumination. A key contribution of [104, 170] is the use of photon-by-photon processing—using the rank-ordered absolute deviation (ROAD) filter [65] in [104] and the rank-ordered mean (ROM) filter [1] in [170]—that attempts to remove the detections that are likely due to background. While this *censoring* is also an exploitation of spatial structure, it is introduced primarily to remove a nonconvexity inherent to ML estimation of depth in the presence of background. Furthermore, it is applied only to depth estimation—not to reflectivity estimation. The censoring in [171] is also applied only to depth estimation and is based on the depths in the entire field of view being sparse after appropriate discretization.

Building primarily upon [170], this chapter reexamines the model of low-flux detection as an inhomogeneous Poisson mixture process, as laid out in Section 2.3.1 and Section 2.4.2. Given that estimation from few detections has been demon-

---

This chapter includes research conducted jointly with Vivek Goyal [147].

strated when signal and background levels are equal, we aim to use new insights from the model to make accurate imaging possible when background levels are increased twenty-five-fold (with signal levels and other imaging conditions unchanged). The central idea is that by effectively separating the signal and background contributions, estimates can be computed that are almost as good as an oracle that uses only the signal detections. Here we use detection times and intuition from the Poisson process model to approximately unmix signal and background contributions at each pixel. Namely, we observe that signal photons times tend to cluster together, so ML depth estimates are reasonable if the largest cluster of detection times is due to signal photons. Conversely, background detections are uniformly distributed over the range of interest, so the largest cluster of detection times being due to background photons can yield significant errors in the ML depth estimate. We thus search for the largest cluster at each pixel and derive detection count thresholds based on the order statistics of uniform random variables to decide whether clustered detection times are likely to be due to background or signal. For pixels with clusters likely due to background, we introduce spatial adaptivity to add detection data from similar neighboring pixels rather than accept low-reliability depth estimates. While some key concepts are first introduced within a pixelwise ML estimation framework, as in previous works [104, 170, 171], we ultimately apply regularization to improve image formation.

### **3.2 Unmixing Signal and Background Processes**

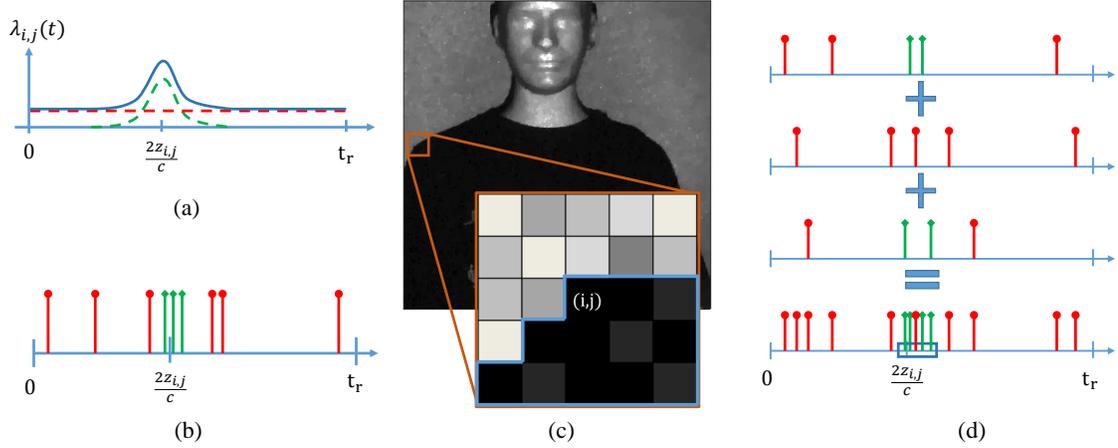
The key observation from the overview of basic parameter estimation in Section 2.4.2 is that the reflectivity and depth estimates are coupled and complicated in the presence of background, but both are greatly simplified if background is removed. Indeed, if we could unmix detection into its component signal and background processes, we

could ignore the background detections and simply apply the background-free estimators. Rather than a conventional approach of forming estimates first and then denoising, this observation is motivation for separating signal from background first and then forming estimates.

### 3.2.1 Pixelwise Unmixing

At an individual pixel, no marker distinguishes between signal and background detections, so no explicit information is available to separate the signal from the background. In order to unmix the processes, the only information we have *a priori* is the different probabilistic models of the detection processes. The signal process rate is related to the short-duration illumination pulse, so signal detection times have a small variance, conditioned on the true depth  $z_{i,j}$ . This suggests that when several signal photons are detected at the same pixel, the detections will be clustered together near the true depth, as illustrated in Figures 3-1(a) and 3-1(b). The background process has a constant rate, meaning no time is more likely than any other to have a background detection. Since the background photons are uniformly distributed in time, we expect them to be fairly spread out in general, unless the background detection rate is very high.

Since signal detections tend to cluster together more readily than background detections, an intuitive approach to identifying signal photons is to search for the largest of those clusters. One way to define a cluster of detections is to choose a window of duration  $t_{\text{wind}}$  and a minimum cluster size  $N_{\text{cl}}$ . The window duration should be chosen such that  $t_{\text{p}} < t_{\text{wind}} \ll t_{\text{r}}$ , so that a well-placed window (one shifted by approximately  $2z_{i,j}/c$ ) is large enough to capture most or all signal detections, without accepting too many background detections. If at  $(i, j)$  there are at least  $N_{\text{cl}}$  detections within some window of duration  $t_{\text{wind}}$ , then we can consider  $(i, j)$  as having a cluster of detections. If there happen to be multiple clusters at  $(i, j)$ , we choose



**Figure 3-1:** (a) Detection can be described as an inhomogeneous Poisson process (shown in blue), which is the sum of inhomogeneous signal (green) and homogeneous background (red) processes. As in the example set of detections in (b), signal detections tend to form clusters more readily than the background detections, suggesting windowing as an approach to unmixing signal and background. Pixels with similar transverse position and reflectivity tend to belong to the same object and therefore have similar depth, as in (c). Using this observation leads to borrowing detections from similar neighboring pixels, as illustrated in (d), which can help amplify low signal levels by making signal detection clusters more apparent.

the window with the most detections  $k_{i,j}^{\max}$  as our signal cluster. From the shift of the window, we have an estimate of the depth  $z_{i,j}$ , and since the mean number of background detections in a short interval  $t_{\text{wind}}$  is close to zero,  $k_{i,j}^{\max}$  yields a rather accurate estimate of the number of signal detections  $m_{i,j}$  analogously to (2.16). As detailed later, the purpose of the  $N_{cl}$  minimum—rather than to seek the largest cluster regardless of its size—is to have a mechanism to produce no depth estimate rather than an unreliable one.

In fact, this intuitive windowing approach falls out of the ML reflectivity and depth estimates in Section 2.4.2. Crudely approximating  $s(t)$  by a square wave of duration  $t_p$  centered at  $2z_{i,j}/c$ , the reflectivity estimate in (2.15) is due only to detections that occur within  $t_p/2$  of the true depth. Even for a more realistic pulse

shape, only detections within a short duration around the true depth contribute non-negligible weight to the reflectivity estimate. Furthermore for the depth estimate, again approximating  $s(t)$  as a square wave, the log-matched filter is maximized at the window containing the largest number of detections.

In the following sections, we derive approximations for probabilities of clusters due to background and due to signal.

### Background Clusters

Since detection of background photons is a homogeneous Poisson process, given  $n$  background detections, the detection times  $\{x_{i,j}^{(\ell)}\}_{\ell=1}^n$  are distributed as the order statistics of  $n$  independent uniform random variables on  $[0, \mathbf{t}_r)$  [174]. Rescaling the set of ordered detections  $0 < x_{(1)} < \dots < x_{(n)} < \mathbf{t}_r$  by  $\mathbf{t}_r$  so they occur in the range  $[0, 1]$ , the  $g$ th order statistic  $X_{(g)}$  has the beta distribution  $\text{Beta}(g, n + 1 - g)$ . According to [51, Sect. 2.3], the time difference  $D_{h-g}^{(g)} = X_{(h)} - X_{(g)}$  between the  $g$ th and  $h$ th detections where  $1 \leq g < h \leq n$  is distributed as  $\text{Beta}((h-g), n+1-(h-g))$ , which depends only on the difference between the  $h$  and  $g$  indices and not on their particular values.

Recall that  $\mathbf{N}_{\text{cl}}$  denotes the minimum number of detections needed in a window of size  $\mathbf{t}_{\text{wind}}$  to consider that window as having a cluster of detections. To have  $\mathbf{N}_{\text{cl}}$  detections in a window beginning at the  $g$ th detection, we must have  $D_{(g+\mathbf{N}_{\text{cl}}-1)-g}^{(g)} = D_{\mathbf{N}_{\text{cl}}-1}^{(g)} < \mathbf{t}_{\text{wind}}/\mathbf{t}_r$ . Now for pixel  $(i, j)$  to not have any clusters, we need all candidate windows to not have clusters. Since there are  $n$  background detections, any of the first  $n - (\mathbf{N}_{\text{cl}} - 1)$  detections may be followed by  $\mathbf{N}_{\text{cl}} - 1$  additional detections within an interval of  $\mathbf{t}_{\text{wind}}$  and thus these are candidates for the beginning of a cluster. Then

the probability of no clusters is

$$\begin{aligned} \Pr[\text{no cluster at } (i, j) \mid N_{i,j} = n] \\ &= \Pr[\{\text{no cluster starting at detection 1}\} \cap \dots \\ &\quad \cap \{\text{no cluster starting at detection } (n - \mathbf{N}_{\text{cl}} + 1)\} \mid N_{i,j} = n]. \end{aligned}$$

The different candidate windows are overlapping and thus the intersected events above are not independent. Making an independence assumption greatly simplifies the computation and gives an approximation that is supported by the numerical evaluations shown in Fig. 3.2(b):

$$\begin{aligned} \Pr[\text{no cluster at } (i, j) \mid N_{i,j} = n] \\ &\approx (\Pr[\text{no cluster starting at detection 1} \mid N_{i,j} = n])^{n - \mathbf{N}_{\text{cl}} + 1} \\ &= (1 - \Pr[D_{\mathbf{N}_{\text{cl}} - 1}^{(1)} < \mathbf{t}_{\text{wind}}/\mathbf{t}_r \mid N_{i,j} = n])^{n - \mathbf{N}_{\text{cl}} + 1}. \end{aligned} \quad (3.1)$$

From this, we have that the conditional probability of a cluster satisfies

$$\begin{aligned} \Pr[\text{cluster at } (i, j) \mid N_{i,j} = n] \\ &\approx 1 - (1 - \Pr[D_{\mathbf{N}_{\text{cl}} - 1}^{(1)} < \mathbf{t}_{\text{wind}}/\mathbf{t}_r \mid N_{i,j} = n])^{n - \mathbf{N}_{\text{cl}} + 1}. \end{aligned} \quad (3.2)$$

Finally, since  $N_{i,j}$  is Poisson-distributed, we can approximate the unconditional probability of a cluster by

$$\begin{aligned} \Pr[\text{cluster at } (i, j)] \\ &\leq \sum_{n=\mathbf{N}_{\text{cl}}}^{\infty} \underbrace{\Pr[N_{i,j} = n]}_{\text{Poisson}(N_{i,j}^{\text{SP}}, n, B)} (1 - \underbrace{(1 - \Pr[D_{\mathbf{N}_{\text{cl}} - 1}^{(1)} < \mathbf{t}_{\text{wind}}/\mathbf{t}_r \mid N_{i,j} = n])^{n - \mathbf{N}_{\text{cl}} + 1}}_{\text{Beta}(\mathbf{N}_{\text{cl}} - 1, n + 1 - (\mathbf{N}_{\text{cl}} - 1))}). \end{aligned} \quad (3.3)$$

Note that  $N_{i,j}^{\text{SP}}$  is defined later in Section 3.2.2 to account for the increased background rate when combining detection data from  $N_{i,j}^{\text{SP}}$  pixels.

## Signal Clusters

We would like to derive the probability of clusters due to signal in a manner similar to that for the background, but the order statistics for many distributions, including our Gaussian pulse shape assumption, are not available in closed form. Instead, we restrict ourselves to consider a cluster present only when  $N_{\text{cl}}$  signal detections occur in a window of length  $t_{\text{wind}}$  centered at the true depth. Since we omit other window positions, we obtain a lower bound for the probability of a signal cluster being present.

The detection time of a signal photon, shifted based on the true depth and divided by  $t_p$ , is given by a standard normal random variable. Denoting the standard normal cumulative distribution function (CDF) by  $\Phi$ , we have the probability of any particular detection landing in the centered window as

$$P_{\text{wind}} = \Phi\left(\frac{t_{\text{wind}}}{t_p}\right) - \Phi\left(-\frac{t_{\text{wind}}}{t_p}\right). \quad (3.4)$$

Given  $m$  signal detections, the probability that exactly  $q$  of them land in the centered window is

$$\begin{aligned} & \Pr[\text{exactly } q \text{ detections in centered window} \mid M_{i,j} = m] \\ &= \binom{m}{q} (P_{\text{wind}})^q (1 - P_{\text{wind}})^{m-q}. \end{aligned} \quad (3.5)$$

The conditional probability of no signal cluster at  $(i, j)$  is the probability of having fewer than  $N_{\text{cl}}$  of the  $m$  detections in the window, which is

$$\begin{aligned} & \Pr[\text{no cluster in centered window} \mid M_{i,j} = m] \\ &= \sum_{q=0}^{N_{\text{cl}}-1} \binom{m}{q} (P_{\text{wind}})^q (1 - P_{\text{wind}})^{m-q}. \end{aligned} \quad (3.6)$$

Finally, since  $M_{i,j}$  is Poisson-distributed, the unconditional probability of a signal

cluster at  $(i, j)$  is bounded as

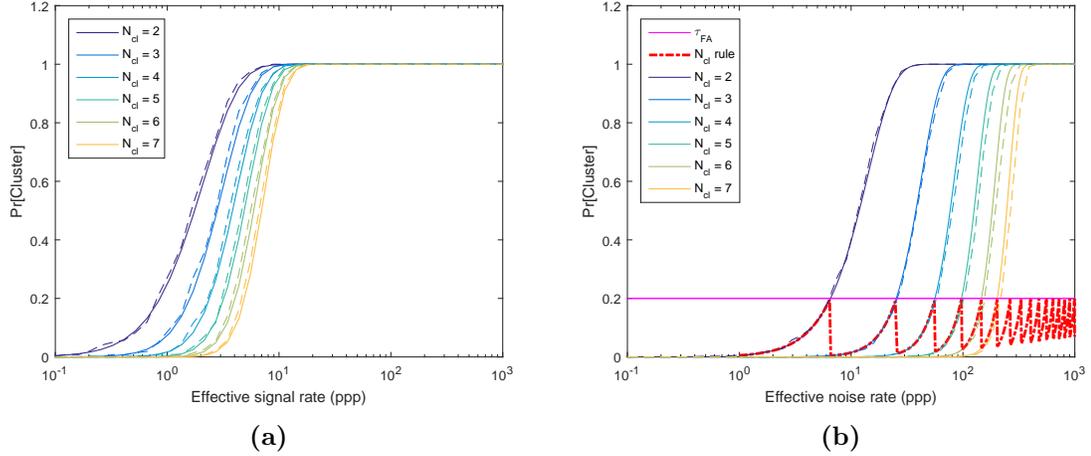
$$\begin{aligned}
& \Pr[\text{cluster at } (i, j)] \\
& \geq \Pr[\text{cluster in centered window}] \\
& = \sum_{m=N_{\text{cl}}}^{\infty} \underbrace{\Pr[M_{i,j} = m]}_{\text{Poisson}(n_r \eta \alpha_{i,j} S)} \left( 1 - \underbrace{\sum_{q=0}^{N_{\text{cl}}-1} \binom{m}{q} (P_{\text{wind}})^q (1 - P_{\text{wind}})^{m-q}}_{\text{binomial}(m, P_{\text{wind}}) \text{ cdf at } N_{\text{cl}}-1} \right). \tag{3.7}
\end{aligned}$$

### Neyman-Pearson Censoring

Figures 3-2(a) and 3-2(b) compare these derivations to Monte Carlo simulations of clustering based on the detection model, confirming that these derivations produce reasonable probability estimates and that the simplifying assumptions are minor. For all experiments and derivations,  $t_{\text{wind}}$  was fixed to  $2t_p$ , where  $t_p = 270$  ps is the measured RMS pulse width of the experiments in [170]. This window size covers more than 95% of the probability mass of signal detection for a Gaussian pulse shape approximation. The pulse repetition period  $t_r = 100$  ns is also used from [170].

Using the plots of these probabilities in Fig. 3-2, we observe that for some rates of signal and background detection, our intuition of finding clusters of detections by windowing is justified. For instance, if the signal and background rates were each 10 photons per pixel (ppp), the probability of observing a cluster of signal detections would be about 1 for any of the minimum cluster sizes shown (see Fig. 3-2(a)), whereas the probability of observing a cluster of background detections would be negligible for  $N_{\text{cl}} > 2$  (see Fig. 3-2(b)). The largest cluster of detections is likely to have more than two detections, and thus the largest cluster could be safely assumed to contain at least one signal detection.

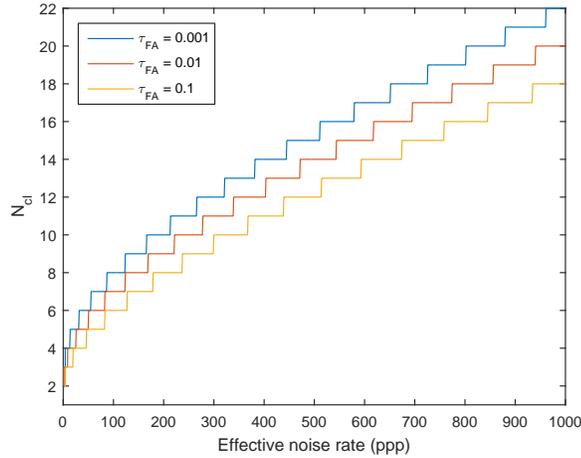
Ideally, we could select optimal  $t_{\text{wind}}$  and  $N_{\text{cl}}$  values for each pixel based on both the local signal and background rates. However, we only know the background rate from



**Figure 3.2:** In comparing the theoretical approximations (solid lines) and Monte Carlo simulation (dashed lines) of the probability of cluster occurrence due to signal (a) or background (b) at various values of  $N_{cl}$ , it is apparent that the derivations in the Sections 3.2.1 and 3.2.1 give close approximations to the true clustering probabilities.

calibration; the signal rate is unknown since it depends on the unknown reflectivity parameter. Our approach from the given data is to fix a reasonable window size and choose  $N_{cl}$  based solely on the background rate, which is known from a calibration acquisition with the laser off, thus capturing only the background input. Given that we assume the largest cluster at a pixel is due to signal, we restrict our minimum cluster size to limit the number of clusters falsely accepted as signal when they are actually due to background. As in Fig. 3.2(b), we set a threshold  $\tau_{FA}$  for the probability of clusters due to background that we will allow. For any given background rate, we can then choose the smallest  $N_{cl}$  that will yield  $\Pr[\text{background cluster}] < \tau_{FA}$ . This method of choosing  $N_{cl}$  as a function of the background rate is illustrated in Fig. 3.3. Since the theoretical derivation tends to slightly overestimate  $\Pr[\text{background cluster}]$ , we are likely to see even fewer clusters due to background than the actual threshold we set.

Now that a reasonable cluster definition is established, we can window the detections at each pixel, and if  $k_{i,j}^{\max} > N_{cl}(n_r B, \tau_{FA}, \mathbf{t}_{\text{window}})$ , we discard all detections except



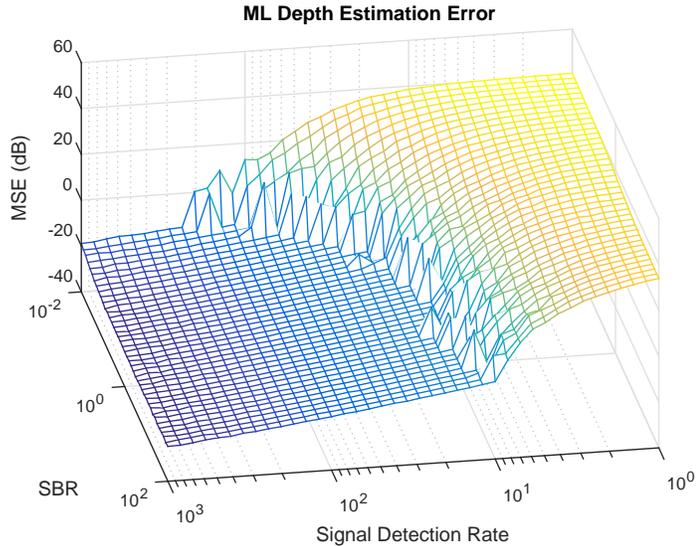
**Figure 3.3:** Since only the effective background rate is known, setting a performance threshold  $\tau_{FA}$  as in Fig. 3-2(b) yields a cluster size rule that limits acceptance of background clusters to probabilities less than or equal to  $\tau_{FA}$ .

those in the best window as background. We dub this *Neyman–Pearson censoring* because of the optimization under bounded false acceptance rate.

### 3.2.2 Spatially-Adaptive Unmixing

While setting a low  $\tau_{FA}$  is a good approach for limiting the number of accepted background clusters, the resulting  $N_{cl}$  may be too high at some pixels for any cluster to be found. Moreover, for photon-efficient imaging, it is common for large regions of scenes to have very few signal detections, such as in [170], where some scenes were reported as having 54% of pixels with no detections. Even with no background present, it would be impossible to estimate the depth purely from windowing, since there are no detections from which to identify clusters. As a result, relying on windowing for low- $\alpha$ , low-SBR data often yields too many pixels with no depth estimate.

A key insight into solving this problem comes from analyzing the behavior of ML depth estimates in background. Fig. 3-4 shows the results of Monte Carlo simulations of depth estimation for various  $\alpha$  and SBR values, where we see a thresholding be-



**Figure 3-4:** Results of a Monte Carlo simulation of ML depth estimates where the background is uniformly distributed over the acquisition interval and  $s(t)$  is approximated by a Gaussian pulse of RMS width approximately equal to our measured laser pulse.

havior that is common to nonlinear estimators [57]. In particular, the plot suggests two conditions for high-accuracy ML depth estimates. First, for any fixed SBR, the ML depth estimate has a low mean-squared error (MSE), so long as enough signal detections are present. This phenomenon is due to the small variance of the signal process relative to the background, resulting in a strong peak in the log-matched filter output where large numbers of signal detections cluster together, even if the number of signal detections  $m_{i,j}$  is much less than the total number of background detections  $n_{i,j}$ . When faced with the problem of making accurate depth estimates, the obvious solution is then to repeat the measurements enough times that sufficient signal photons are detected. This is the approach of conventional photon-counting lidar systems, which make hundreds or thousands of detections per pixel to ensure enough signal is present [30].

The second key observation from Fig. 3-4 is that for a fixed number of signal detections, the accuracy improves as the SBR increases (i.e., the background level

diminishes). This observation is particularly relevant for photon-efficient imaging, since collecting additional data may be impossible due to time constraints or simply undesirable due to the resulting decrease in photon efficiency. While increasing the SBR may be possible, it is practically a more difficult approach to reducing error than increasing the acquisition time. Either prior information about the scene depth must be known, so that range gating can be applied pre- or post-detection, or a method like windowing is needed to identify background detections and censor them.

As previously discussed, reducing the background level alone is insufficient for low- $\alpha$ , low-SBR data. In such cases, we are motivated to combine both approaches—increasing the signal level and the SBR—to achieve an acceptable level of error. More specifically, in combination with windowing, we are tasked with finding an approach that mimics collecting more data, without actually increasing the acquisition time. A few observations about natural scenes suggest a useful solution. First, the depth maps of natural scenes are generally smooth except at object boundaries, so neighboring pixels often have approximately equal depth. This was the justification for the total variation (TV) regularization used in [170], since TV regularization tends to smooth out background while still preserving jump discontinuities [121]. Secondly, edges in reflectivity and object boundaries in depth tend to be co-located, so scene patches that are similar in both reflectivity and transverse position are likely similar in longitudinal position (depth) as well.

These observations can be codified through the construction of *superpixels*, over-segmentations of an image into small regions of similar pixels, which are a common tool in computer vision applications. Superpixels were originally introduced in [154] with the idea that pixels are arbitrary elementary units of digital images, and that breaking images into more natural building blocks could improve and speed up further processing such as larger-scale image segmentation and object detection. The general

reasoning is that pixels that are similar in both some color space (e.g., Lab) and in transverse position have a high probability of belonging to the same object. As a result, superpixels have been used as a preprocessing tool to provide noise robustness and to fill in gaps of depth maps for stereo [222, 34], RGB-d [194], and lidar [124] systems.

Our approach is to use a variant of superpixels to artificially extend acquisition times, which facilitates depth estimation. Consider a small neighborhood of pixels in a scene, such as the one illustrated in Fig. 3-1(c). Assuming that the scene has been sampled with adequate transverse spatial resolution, pixels within this neighborhood will have similar depth values, unless the neighborhood crosses a boundary between objects. If  $(a, b)$  is in the neighborhood of  $(i, j)$ , then  $p_{X_{i,j}}(x) \approx p_{X_{a,b}}(x)$  for all  $x \in [0, t_r)$ . Thus, combining detections from  $(a, b)$  into the  $(i, j)$  vector is almost equivalent to doubling the acquisition time at  $(i, j)$ . This borrowing will maintain SBR but increase  $m_{i,j}$ , helping to reduce the estimation error, as we observed in Fig. 3-4. Borrowing creates some smoothing in the transverse directions, and the ideal trade-off between background reduction and this smoothing probably occurs just to the left of the estimation threshold illustrated in Fig. 3-4.

For practical purposes, reinforcing the signal by borrowing detections from neighboring pixels will enhance the size of signal clusters and make windowing more reliable and useful, as illustrated in Fig. 3-1(d). When superpixels are formed, the background rate is effectively amplified by  $N_{i,j}^{\text{SP}}$ , the number of pixels that contributed to the enhanced detection vector at  $(i, j)$ , so we update our cluster size requirement for windowing to  $\mathbf{N}_{\text{cl}}(N_{i,j}^{\text{SP}}\mathbf{n}_r B, \tau_{\text{FA}}, \mathbf{t}_{\text{wind}})$  to avoid falsely accepting background clusters. This formulation in fact describes the generic windowing procedure, where  $N_{i,j}^{\text{SP}} = 1$  if the detections at only a single pixel are used.

There are many existing superpixel definitions and implementations, each designed

to meet particular performance criteria [2]. In principle, any definition could be used within our algorithm to select groups of similar pixels from which to borrow photon detections. A key difference of our approach to that of the other depth-estimation applications of superpixels in [222, 34, 194, 124] is that the existing methods all incorporate a conventional digital camera. Superpixels are formed using the high-quality color images, and then the assumed redundancy of these regions is applied to corresponding regions of a lower-quality depth map to fill in gaps or filter out background. In our system, reflectivity is a grayscale value estimated from the same active illumination data used to estimate the depth. Due to the low signal counts and high background levels, the intensity data is much less reliable than that from a conventional camera. As a result, we use a simple definition of selecting the subset of pixels in a square region that meet a reflectivity tolerance compared to  $(i, j)$ . Since pixels are chosen within a fixed distance of  $(i, j)$ , the set of candidate pixels changes slightly from one pixel of interest to the next. We use this particular definition in order to promote a high degree of localization, which helps preserve small changes in reflectivity and depth. Other superpixel definitions that consider each region to be homogeneous would smooth over these small changes.

### 3.3 Unmixing Algorithm

Our method for forming depth and reflectivity images from the raw detection data builds off the image formation procedure of Shin et al. [170] (which we denote as PML+ROM because it combines penalized ML estimation with a ROM filter censoring step), adding in the windowing and spatial adaptivity introduced in Section 3.2. The raw data input to the algorithm is the set of photon detections  $\{x_{i,j}^{(\ell)}\}_{\ell=1}^{k_{i,j}}$  for each patch  $(i, j)$ , which implicitly includes the values  $k_{i,j}$ . It is also assumed that  $B$ , the mean background count per patch per pulse-repetition period, and  $\eta\beta$ , the mean

**Table 3.1:** Assumptions about scene, signal, and background properties

Assumptions
One depth per pixel
Low photon flux (so dead time is negligible)
All depths in $[0, ct_r/2)$ (so no depth aliasing)
Background detection rate constant in time, uniform over the scene
$b_a, b_d, \eta,$ and $s(t)$ known (from calibration)

signal count per pulse in the absence of any attenuation, have been measured through calibration or approximated from environmental conditions and hardware specifications. Assumptions for scene, signal, and background properties are summarized in Table 3.1. The procedure is summarized by the block diagram in Fig. 3.5, and we now detail each component. A small number of algorithm parameters are introduced as needed.

### 3.3.1 Windowing

The process of background censoring at patch  $(i, j)$  by adaptive windowing uses two parameters: a window length  $t_{\text{wind}}$  and a target probability of false acceptance of a background cluster  $\tau_{\text{FA}}$ . It is performed as follows, assuming for the moment no borrowing of detections from neighboring patches:

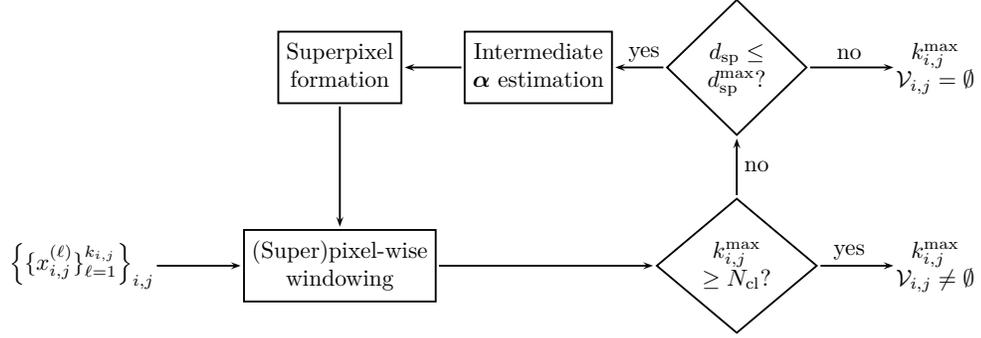
1. For each  $\ell \in \{1, \dots, k_{i,j}\}$ , find the set of detections in the interval of length  $t_{\text{wind}}$  starting at the detection time  $x_{i,j}^{(\ell)}$ :

$$\mathcal{D}_\ell = \left\{ x_{i,j}^{(k)} : x_{i,j}^{(\ell)} \leq x_{i,j}^{(k)} < x_{i,j}^{(\ell)} + t_{\text{wind}} \right\}. \quad (3.8)$$

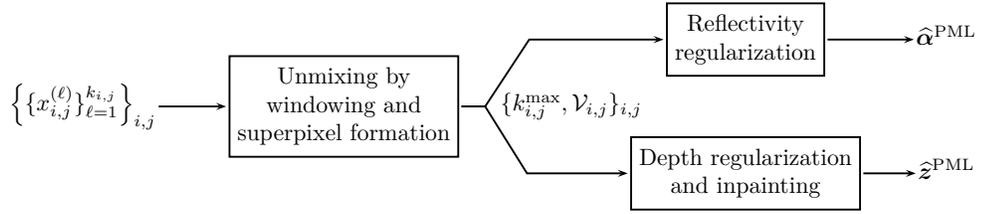
2. Among these sets, select a set  $\mathcal{W}_{i,j}$  with the largest number of detections:

$$\mathcal{W}_{i,j} \text{ satisfies } |\mathcal{W}_{i,j}| = \max_{\ell} |\mathcal{D}_\ell|, \quad (3.9)$$

and define  $k_{i,j}^{\text{max}} = |\mathcal{W}_{i,j}|$ . Resolve ties by choosing uniformly at random among



(a) Unmixing algorithm



(b) Overall algorithm

**Figure 3-5:** Block diagram of the basic algorithmic sequence. The unmixing procedure continues until enough pixels have reliable depth estimates so that inpainting is needed for only a small fraction with missing entries.

the sets with  $k_{i,j}^{\max}$  detections.

- Using false acceptance threshold  $\tau_{\text{FA}}$ , compute minimum cluster size  $\mathbf{N}_{\text{cl}}$  as the smallest integer such that

$$\Pr[\text{background cluster}; \mathbf{N}_{\text{cl}}, N_{i,j}^{\text{sp}} \mathbf{n}_r B, \mathbf{t}_{\text{wind}}] < \tau_{\text{FA}}, \quad (3.10)$$

where  $\Pr[\text{background cluster}]$  is derived in (3.3) and  $N_{i,j}^{\text{sp}} = 1$  for windowing a single pixel. Note that this step does not depend on the detection time data and thus desirable values of  $\mathbf{N}_{\text{cl}}$  may be precomputed.

- If  $k_{i,j}^{\max} \geq \mathbf{N}_{\text{cl}}$ , retain only the detections that fall in the selected window  $\mathcal{W}_{i,j}$

and censor the rest, yielding the set of uncensored detections  $\{x_{i,j}^{(\ell)}\}_{\ell \in \mathcal{V}_{i,j}}$ , where

$$\mathcal{V}_{i,j} = \left\{ \ell \in \{1, \dots, k_{i,j}\} : x_{i,j}^{(\ell)} \in \mathcal{W}_{i,j} \right\}. \quad (3.11)$$

When this windowing is applied with superpixels (i.e.,  $N_{i,j}^{\text{SP}} > 1$ ), the detection time data  $\{x_{i,j}^{(\ell)}\}_{\ell=1}^{k_{i,j}}$  is replaced by augmented detection times  $\{\tilde{x}_{i,j}^{(\ell)}\}_{\ell=1}^{\tilde{k}_{i,j}}$ .

### 3.3.2 Reflectivity Estimation

In the window  $\mathcal{W}_{i,j}$ , the expected number of background detections is  $N_{i,j}^{\text{SP}} \mathbf{n}_r B \mathbf{t}_{\text{wind}} / \mathbf{t}_r$ , which is small even at low SBR and is considerably lower than the number of detections due to background on the entire  $[0, \mathbf{t}_r)$  interval. Since background detection is a homogeneous Poisson process, the variance in the number of background detections in the window is also small, so  $N_{i,j}^{\text{SP}} \mathbf{n}_r B \mathbf{t}_{\text{wind}} / \mathbf{t}_r$  is a good estimate of the number of background detections. Thus, we can modify (2.16) to estimate  $\alpha_{i,j}$  from the window output as

$$\hat{\alpha}_{i,j}^{\text{wind}} = \max \left\{ \frac{k_{i,j}^{\text{max}} - N_{i,j}^{\text{SP}} \mathbf{n}_r B \mathbf{t}_{\text{wind}} / \mathbf{t}_r}{N_{i,j}^{\text{SP}} \mathbf{n}_r \eta \beta}, 0 \right\}. \quad (3.12)$$

For those pixels where  $k_{i,j}^{\text{max}} < N_{\text{cl}}$ , this formula tends to slightly overestimate the reflectivity, since we have likely chosen the window with the largest cluster of background detections. However, the  $\hat{\alpha}_{i,j}^{\text{wind}}$  estimate is temporary for those pixels, since the value of  $k_{i,j}^{\text{max}}$  will be updated after windowing the augmented data from the superpixels.

We form a reflectivity image by regularized ML estimation with a regularization parameter  $\zeta_\alpha \in [0, \infty)$ . Using (2.8), the negative log-likelihood of the scene reflectivity  $\alpha_{i,j}$  given the number of detections in  $\mathcal{W}_{i,j}$  is

$$\mathcal{L}_\alpha(\alpha_{i,j}; k_{i,j}^{\text{max}}) = N_{i,j}^{\text{SP}} \mathbf{n}_r \eta \alpha_{i,j} S - k_{i,j}^{\text{max}} \log [N_{i,j}^{\text{SP}} (\mathbf{n}_r \eta \alpha_{i,j} S + \mathbf{n}_r B \mathbf{t}_{\text{wind}} / \mathbf{t}_r)], \quad (3.13)$$

ignoring terms not dependent on  $\alpha_{i,j}$ . As in [170], we take advantage of spatial

correlations in natural scenes to form a penalized ML (PML) estimate that enforces smoothness:

$$\hat{\boldsymbol{\alpha}}^{\text{PML}} = \arg \min_{\boldsymbol{\alpha}: \alpha_{i,j} \geq 0} \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \mathcal{L}_{\alpha}(\alpha_{i,j}; k_{i,j}^{\max}) + \zeta_{\alpha} \text{pen}_{\alpha}(\boldsymbol{\alpha}). \quad (3.14)$$

### 3.3.3 Superpixel Formation

After windowing, all pixels have a reflectivity estimate, but only those  $(i, j)$  where  $k_{i,j}^{\max} \geq N_{\text{cl}}$  have reliable depth estimates. For those pixels with insufficient signal detection counts, superpixels are formed so that strongly correlated depth data from similar neighboring pixels can be combined to improve the performance of windowing. The key to our superpixel formation is to set bounds for what constitutes a similar pixel. In this chapter, superpixels borrow detections from all neighboring pixels within a fixed distance and a fixed reflectivity tolerance of our pixel of interest. In particular, fix a neighborhood distance  $d_{\text{sp}}$  (typically 1, 2, or 3) and a reflectivity tolerance  $\tau_{\text{sp}}$  (typically around 5% of the full range of  $\hat{\boldsymbol{\alpha}}^{\text{PML}}$  values). The superpixel at  $(i, j)$  is defined as

$$\mathcal{M}_{i,j} = \left\{ (a, b) \in \{1, \dots, n_i\} \times \{1, \dots, n_j\} : \begin{array}{l} |i - a| \leq d_{\text{sp}}, \quad |j - b| \leq d_{\text{sp}}, \quad |\hat{\alpha}_{i,j}^{\text{PML}} - \hat{\alpha}_{a,b}^{\text{PML}}| \leq \tau_{\text{sp}} \end{array} \right\}. \quad (3.15)$$

The set of superpixel detections  $\{\tilde{x}_{i,j}^{(u)}\}_{u=1}^{\tilde{k}_{i,j}}$  is then defined as

$$\{\tilde{x}_{i,j}^{(u)}\}_{u=1}^{\tilde{k}_{i,j}} = \bigcup_{(a,b) \in \mathcal{M}_{i,j}} \{t_{a,b}^{(\ell)}\}_{\ell=1}^{k_{a,b}}, \quad (3.16)$$

where  $\tilde{k}_{i,j}$  is the new detection count for the superpixel at  $(i, j)$ . In this way, the algorithm searches a small local area and adaptively borrows from pixels that are similar in both transverse position and reflectivity.

Once superpixel vectors have been formed, the windowing process of Section 3.3.1

and the reflectivity estimation of Section 3.3.2 are repeated. The windowing is performed on the set of superpixel detections  $\{\tilde{x}_{i,j}^{(u)}\}_{u=1}^{\tilde{k}_{i,j}}$ , resulting in a different (usually larger) value of  $k_{i,j}^{\max}$ . Note that the  $\mathbf{N}_{\text{cl}}$  computation and the reflectivity estimate change to account for the number of pixels  $N_{i,j}^{\text{sp}}$  contributing to the superpixel vector.

Ideally, the smallest possible  $N_{i,j}^{\text{sp}}$  such that  $k_{i,j}^{\max} \geq \mathbf{N}_{\text{cl}}$  would be chosen at each pixel, which would ensure accurate depth estimates with the minimum amount of spatial smoothing. This could be accomplished by incorporating detections from one pixel at a time and re-windowing to check whether the  $\mathbf{N}_{\text{cl}}$  criterion had been met. Unfortunately, this repeated windowing of new detection vectors is too computationally intensive for large images. Instead, we take a coarser approach that gradually increases the candidate neighborhood for forming superpixels by incrementing  $d_{\text{sp}}$ . We cycle through the procedures of windowing, estimating reflectivity, and forming superpixels, gradually increasing  $d_{\text{sp}}$  with each iteration from  $d_{\text{sp}} = 0$  until either  $k_{i,j}^{\max} \geq \mathbf{N}_{\text{cl}}$  for all  $(i, j)$  or some terminal neighborhood size  $d_{\text{sp}}^{\max}$  has been reached. For any remaining pixels without a reliable depth estimate,  $\hat{z}_{i,j}$  is filled in by inpainting during the depth estimation procedure.

### 3.3.4 Depth Estimation

It is assumed that all detections retained in  $\mathcal{V}_{i,j}$  are due to signal, although if too many background clusters are falsely accepted, further rank-ordered mean (ROM) censoring as in [170] can be useful in cleaning up the data. The negative log-likelihood of the depth  $z_{i,j}$  given only signal detections is

$$\mathcal{L}_z(z_{i,j}; \{\tilde{x}_{i,j}^{(\ell)}\}_{\ell \in \mathcal{V}_{i,j}}) = - \sum_{\ell \in \mathcal{V}_{i,j}} \log[s(\tilde{x}_{i,j}^{(\ell)} - 2z_{i,j}/c)]. \quad (3.17)$$

Again applying a smoothness penalization appropriate for depth maps of natural scenes, the PML depth estimate is

$$\hat{\mathbf{z}}^{\text{PML}} = \arg \min_{\mathbf{z} : z_{i,j} \in [0, z_{\max})} \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \mathcal{L}_z \left( z_{i,j}; \{x_{i,j}^{(\ell)}\}_{\ell \in \mathcal{V}_{i,j}} \right) + \zeta_z \text{pen}_z(\mathbf{z}), \quad (3.18)$$

where  $\zeta_z \in [0, \infty)$  controls the amount of penalization.

### 3.4 Results

The experimental setup and procedure follows the basic outline in Section 2.2.4, and a detailed account is given in [104, 105]. The important quantities for our algorithm are the illumination pulse width, measured to be  $t_p = 270$  ps, and the pulse repetition period  $t_r = 100$  ns. The SPAD detector quantum efficiency was  $\eta = 0.35$ .

In [104] and [170], the photon-efficient methods are compared to “ground truth” reconstructions of reflectivity and depth, generated using conventional lidar processing on data from long acquisition times. While these measurements serve as effective baseline comparisons, they still suffer from the same shortcomings as all lidar data. In particular, the conventional processing assumes only one depth exists at each point in the image, and we make this assumption as well. Taking into account multiple depths at a single pixel as in [172] would require adjustments to our algorithm, since superpixels would borrow detections from multiple true depths, only one of which would be registered. Experimental lidar data has effects of shadowing (from nonconfocal configurations) or reflections from multiple depths, so we consider the conventional processing to produce “baseline” estimates but not ground truth.

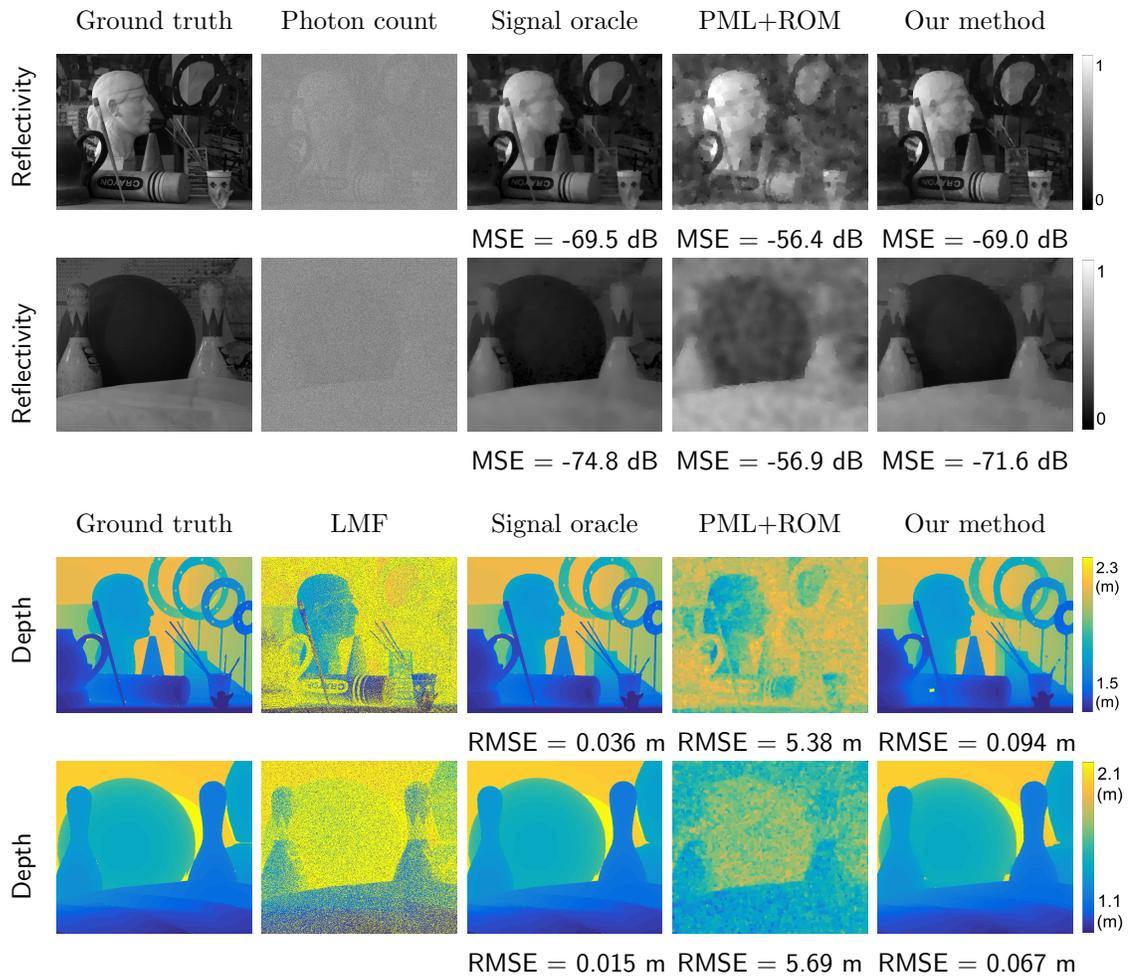
#### 3.4.1 Simulated Results

In order to quantify the algorithm performance compared to an actual ground-truth reference, we first simulated data sets using the model outlined in Section 2.3, where

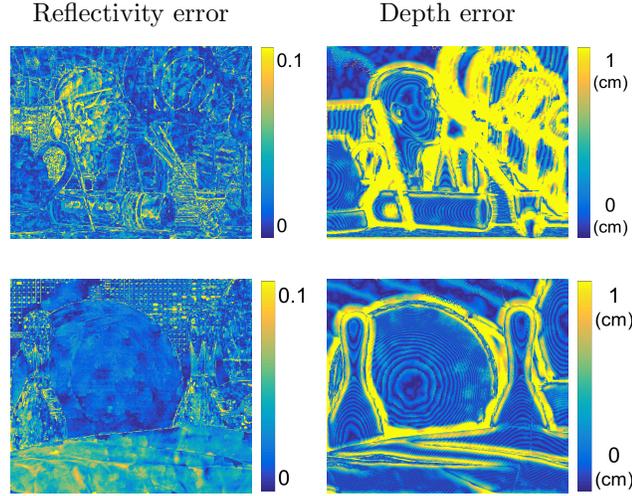
each pixel has only a single true depth. The same parameters from the experiments were used in the simulation in order to maintain consistency. Furthermore, although detections are generated from a model, we use real physical scenes from the Middlebury dataset [164] to form  $\alpha$  and  $z$ . In particular, we chose the Art and Bowling scenes as representative of fairly complex and fairly simple scenes, respectively. The Art scene is  $695 \times 555$  pixels, and the Bowling scene is  $626 \times 555$  pixels.

Signal counts were generated as Poisson random variables with parameters equal to scaled pixel intensities. Signal detection times were generated from a Gaussian pulse shape with mean  $z_{i,j}$  and  $\sigma_p = t_p/2$ . Background detection counts were also generated as Poisson random variables, and given the count at each pixel, the detection times were generated as samples from a uniform distribution over the repetition period  $[0, t_r)$ . In order to meet the low-flux requirement, scenes were simulated so that the average pixel would require 500 illuminations to generate one signal photon. Thus, performance evaluation of scenes with 2.0 and 3.0 signal ppp used 1000 and 1500 illumination periods per pixel, respectively. At the maximum evaluated background level (with  $SBR = 0.04$ ), the average photon detection rate was one detection in approximately 5% of illumination periods, meeting the conventional guideline for operating in the low-flux regime. Performance is quantified as the mean-squared error (MSE) in dB for reflectivity and the root mean-square error (RMSE) with units of meters for depth, as defined in Section 2.4.

Fig. 3-6 shows example simulation results for both scenes at  $SBR = 0.04$  and only 2.0 signal photons per pixel on average. The normalized raw photon count and the log-matched filter (LMF) output are shown as a reference for what conventional methods produce from noisy, low-light data. We compare the results of our proposed method with PML+ROM, which was the state-of-the-art for photon-efficient imaging at lower background intensities. We also show the ideal results from a *signal oracle*, which



**Figure 3-6:** Simulated processing results for Art and Bowling scenes [164] at  $\text{SBR} = 0.04$  and 2.0 signal photons per pixel on average. Note that the depth estimates with the PML+ROM method [170] are completely out of range of the actual scene and are instead shown for the range of 6 to 8 meters.



**Figure 3-7:** Absolute error maps for the proposed method applied to simulations with the Art and Bowling scenes [164] at  $\text{SBR} = 0.04$  and 2.0 signal photons per pixel on average. The largest depth errors tend to be located at object boundaries.

represents the ideal case of perfect unmixing and uses only the signal detections for estimation (equivalently,  $\text{SBR} = \infty$ ). Throughout the simulations, we use  $d_{\text{sp}}^{\text{max}} = 3$ ,  $\tau_{\text{sp}} = 0.05$ , and  $\tau_{\text{FA}} = 0.01$  for our algorithm, which work for a variety of scenes and experimental conditions. These parameters were mainly tuned for very-low SBR data (around 25 times as much background as signal) and could be adjusted to optimize performance for different background conditions or a particular scene.

The results in Fig. 3-6 exemplify the typical performance of the different methods. For reflectivity, it is clear that high levels of background reduce contrast too much for PML+ROM to produce a good estimate from detection counts alone. The unmixing does a much better job at estimating the number of signal detections at each pixel. In particular, the absolute error maps in Figure 3-7 show the smallest errors for the darkest regions, where formation of superpixels allows for precise fractional estimates of signal photon counts.

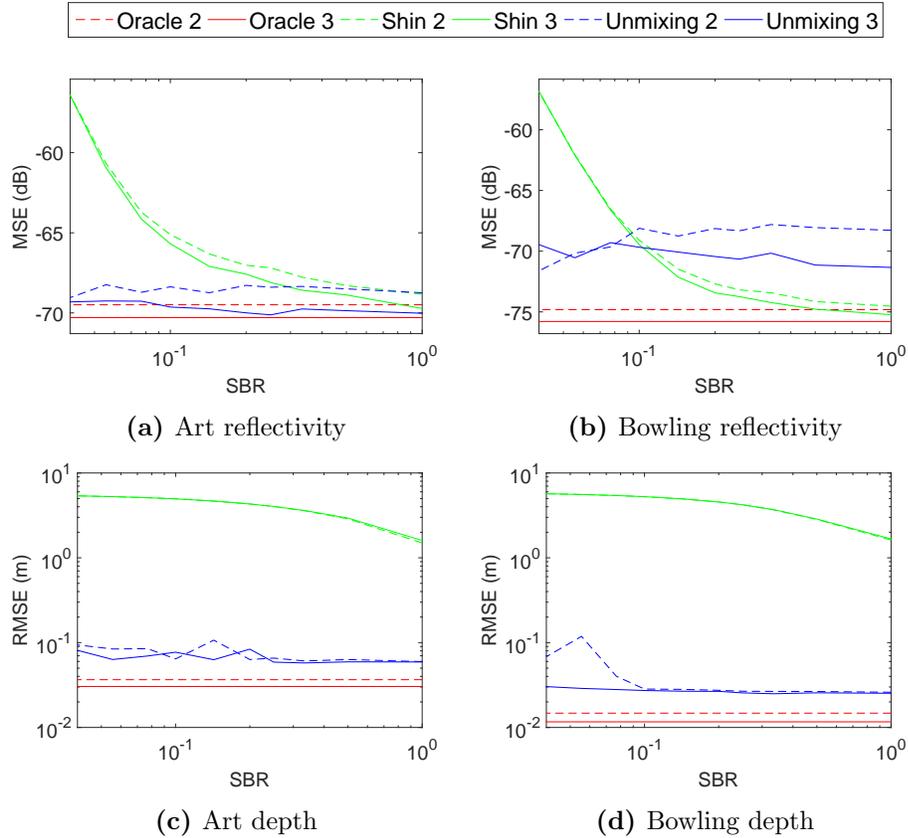
In the case of depth estimation, PML+ROM fails, as background detections pull

the depth estimates towards the mean scene depth (7.5 meters in this case). For our unmixing method, the windowing procedure is much more effective at handling the high-variance background. Figure 3.7 shows that the largest errors that remain in our depth estimation occur in the darkest regions of the scene, particularly at object boundaries. In many of the dark regions, forming superpixels is enough to overcome the low signal photon count. At object boundaries, however,  $N_{i,j}^{\text{SP}}$  decreases since many candidate pixels in the neighborhood fall outside the reflectivity tolerance so the signal clusters are too small, or if the reflectivity contrast between objects at different depths is not sufficient, the superpixels will borrow pixels at multiple depths, causing errors. Nevertheless, the unmixing process produces depth estimates that are almost as good as the signal oracle in many cases.

### 3.4.2 Performance Analysis

Simulating detections from the Art and Bowling scenes allows for a thorough evaluation of the effects of changing signal and background intensities on the unmixing algorithm performance. Fig. 3.8 contains plots comparing the oracle, PML+ROM, and unmixing methods for 2.0 and 3.0 signal detections per pixel at various SBR levels. The MSE and RMSE metrics are shown for the  $\zeta_\alpha$  and  $\zeta_z$  regularization parameter values that produced the best average performance over 10 trials at each value of SBR.

As expected, the best performance is achieved by the oracle estimator with the most signal detections per pixel, since this case has the most signal information available and is not corrupted by background. Estimation of both parameters improves in general for all methods as the signal detection count increases. It is also clear that the reflectivity and depth estimation performance of PML+ROM degrades significantly as SBR decreases. This is due to the shortcomings of the binomial estimator for reflectivity and the limitations of the ROM censoring for removing background



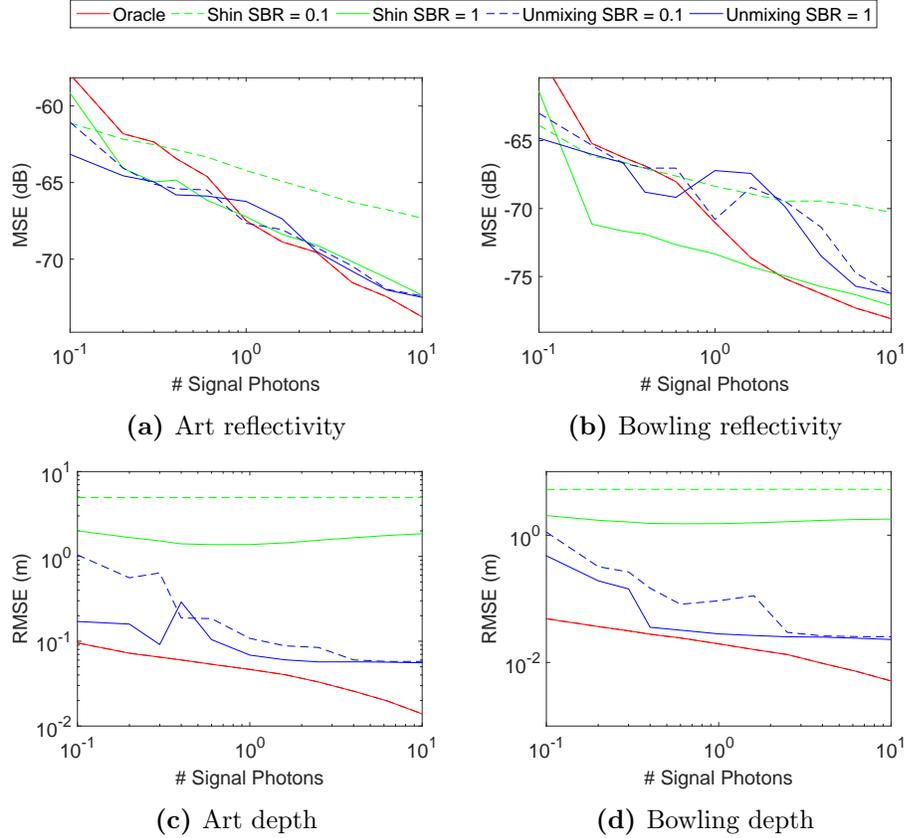
**Figure 3-8:** Performance evaluation for reflectivity and depth estimation for simulated data sets with 2.0 and 3.0 signal photon detections per pixel on average and a range of SBR values. *Oracle* refers to performance of a penalized ML estimator using only signal detections ( $\text{SBR} = \infty$ ). *Shin* refers to PML+ROM, the method of Shin et al. [170]. *Unmixing* refers to the proposed method. Plotted performance is the average error of 10 trials for each value of SBR.

detections at low SBR. For our proposed unmixing method, the parameter estimation accuracy also tends to decrease as SBR decreases, although the change in error is smaller than for PML+ROM, indicating a higher robustness to background. At  $\text{SBR} = 0.04$ , our reflectivity estimate outperforms PML+ROM by about 15 dB. The difference in depth estimation error is even more stark—at  $\text{SBR} = 0.04$ , our method has RMSE almost two orders of magnitude better than PML+ROM.

Fig. 3.9 contains plots showing the dependence of the various methods’ performance on the average number of signal photons per pixel. First, we observe that the oracle performs as expected; since no background is present, the oracle performance depends only on the number of signal detections present, and it is clear that with more signal photons, both reflectivity and depth estimates are improved.

For the unmixing and PML+ROM methods, the reflectivity estimate is likewise improved with higher numbers of signal detections. It is interesting to note that the performance of these methods tends to surpass the oracle for fewer than one signal photon per pixel on average, with the PML+ROM method applied to the Bowling scene being a particularly good example. When the average number of detected signal photons over a scene is very low (especially less than one), many pixels have no signal detections. In the background-free setting, this means that dark patches like the bowling ball likely have no detections over a large area, and the oracle method underestimates the true reflectivity in those areas. On the other hand, when a small amount of background is present, the background detections act like dither for the reflectivity estimate. Occasional background detections in dark regions are smoothed over by regularization, with a resulting estimate that ends up closer to the true reflectivity value for the unmixing and PML+ROM methods.

Like the oracle method, unmixing also produces better depth estimates when more signal is present, although the algorithm appears to suffer from diminishing returns.



**Figure 3-9:** Performance evaluation for reflectivity and depth estimation for simulated data sets with SBRs of 0.1 and 1.0 and a range of average signal photon detections per pixel. *Oracle* refers to performance of a penalized ML estimator using only signal detections (SBR =  $\infty$ ). *Shin* refers to PML+ROM, the method of Shin et al. [170]. *Unmixing* refers to the proposed method. Plotted performance is the average error of 10 trials for each value of SBR.

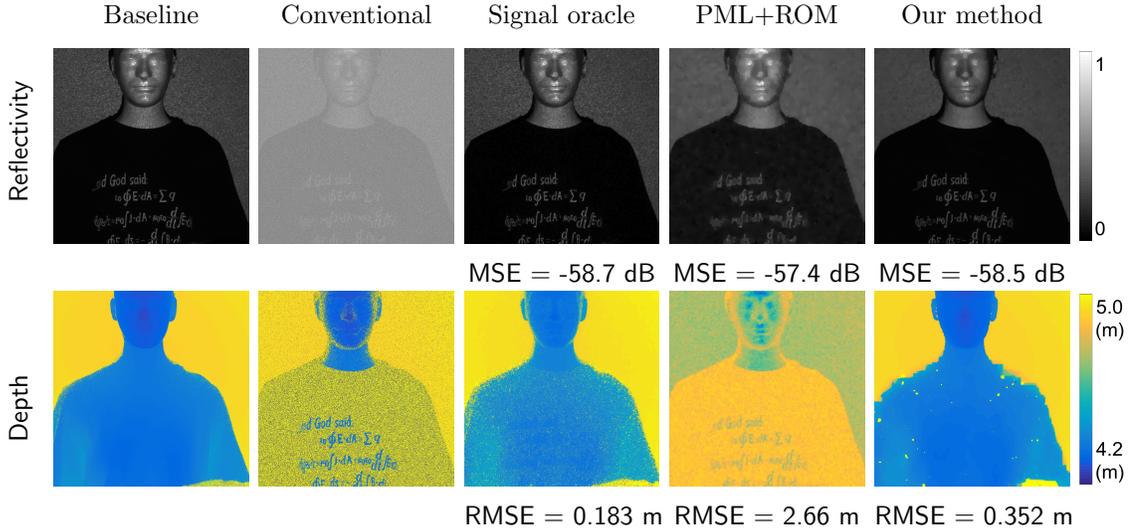
One possible cause is the behavior of the superpixels; although they provide large improvements for background reduction overall, borrowing from neighboring pixels may cause enough smoothing that a few significant errors remain at object boundaries. Unlike the other methods, the results from PML+ROM prove to be fairly constant with respect to the number of signal photons. Since the background increases with the signal for SBR held constant, and since the method of Shin has limited denoising ability, it appears it does not realize the same benefits of additional signal.

The lack of monotonicity in Figures 3-8 and 3-9 is not due primarily to pseudo-random variations in the simulated data; hence, we did not include error bars. The discrete jumps in superpixel sizes and the particular  $\zeta_\alpha$  and  $\zeta_z$  gridding used are the primary contributors to the lack of smoothness of these curves.

### 3.4.3 Experimental Results

We further evaluate the performance of our unmixing algorithm on the  $1000 \times 1000$  pixel dataset of the Mannequin scene from [199], with results shown in Fig. 3-10. Baseline estimates were formed using conventional lidar processing on detection data from long acquisition times under constant conditions at  $\text{SBR} = 1$ . The data was first range-gated to capture the extent of the scene (4.2 to 6 meters), while limiting the influence of background on the baseline estimates. Depth estimates were then formed by applying the log-matched filter to the first 200 detections at each pixel. Reflectivity estimates were formed by scaling the detection count by the number of illumination pulses required to reach 200 detections at each pixel.

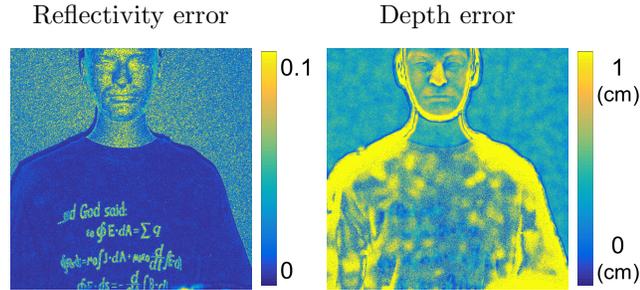
Truncated photon-efficient datasets were created by using only the first 3000 illumination periods ( $300 \mu\text{s}$  per pixel), which resulted in 4.05 signal photons per pixel on average. The signal oracle processing was computed on the truncated data with the same range-gating as the baseline estimates to remove as many background detections as possible. Since the data was collected with some ambient light present in the



**Figure 3-10:** Results of methods on the experimentally-acquired Mannequin data with about 4.05 signal detections per pixel and additional synthetically-generated noise to set the SBR at about 0.04. The error metrics are approximate, since the baseline lidar data is not exactly a ground truth for the scene. Note that the depth estimates with PML+ROM (the method of [170]) are completely out of range of the actual scene and are instead shown for the range of 5 to 8 meters.

scene, the resulting SBR is approximately 8.3, so the oracle data represented favorable (although not entirely background-free) conditions. To evaluate the background performance of the other methods, additional background detections were synthetically generated as uniformly distributed detection times on  $[0, t_r)$ , given a Poisson number of background detections. Including the background detections already present in the data, the background rate was adjusted to set the SBR to 0.04 to match the simulated data. For our algorithm, we use  $d_{sp}^{\max} = 4$ ,  $\tau_{sp} = 0.05$ , and  $\tau_{FA} = 0.01$ . The higher value of  $d_{sp}^{\max}$  is used to account for the larger image size compared to the simulations.

Although there is little distinction when comparing the approximate MSE for all three reflectivity estimates, there is a clear advantage to using our method over that of PML+ROM. The results using PML+ROM are far more smoothed with less contrast, making the text unreadable and the facial features harder to distinguish.



**Figure 3-11:** Absolute error maps for the proposed algorithm applied to the experimentally acquired Mannequin data with about 4.05 signal detections per pixel and SBR 0.04.

Our method instead produces much clearer results, which compare favorably to the oracle and baseline reflectivity estimates.

For the estimate of scene depth, the benefit of unmixing is even more striking. As in the simulations, PML+ROM yields an estimate that is completely out of the range of the true scene. The ROM censoring is insufficient at low SBR, so the entire estimate is dominated by background, which yields an estimate very close to  $z_{\max}/2$ . The resulting RMSE is then mostly an indication of how close the scene subject was to the middle of the imaging range: since the simulated scenes were positioned farther from the center of the scene, the RMSE measures were larger. On the other hand, our unmixing method proves to be considerably more effective at handling the high levels of background.

One important note for this data set is the large discrepancy in reflectivity between the dark and bright regions of the scene. Since the background rate is uniform over the scene, the result is a local SBR that varies greatly with the region being imaged. The depth estimation performance depends on the SBR, so the estimation error varies locally in the same way. Figure 3-11 shows that the brightest regions (the wall, the mannequin’s face, and the shirt around the text) have low absolute depth error. The largest errors occur as in the simulations at object boundaries and in the darkest regions, such as several small patches of the mannequin’s shirt. The effect

of the brightness variation is further apparent in the results from the log-matched filter and PML+ROM; those methods are insufficiently robust to background and thus perform best in the bright regions with less relative background. The effect of the background is a depth estimate biased toward  $t_r/2$ , the midrange point of the acquisition interval. For pixels at the same depth but with different reflectivities (such as the bright text on the dark shirt), the darker pixel depth estimates are biased more towards the midrange, allowing the text to be visible in the depth image. This same phenomenon also occurs in the simulations, such as the crayon text in the Art scene or the bowling pin neck, although the effect is less pronounced there. Finally, the discrepancy in brightness over the mannequin scene required a higher average number of signal detections to achieve acceptable performance in the darkest regions.

### 3.5 Further Performance Evaluation

Due to the large improvement in performance over the previous state of the art, the algorithm described in this chapter has become a standard for benchmarking the performance of other algorithms in photon-starved, high-background acquisitions. We discuss some of the works that compare directly with this algorithm, which we refer to as the unmixing approach.

- Most notably, Lindell et al. [118] aimed to explicitly achieve better performance than the unmixing approach via the use of convolutional neural networks (CNNs), which could be trained to identify clusters of signal detections with more flexibility than the algorithm presented here. From the raw detection data alone, the CNN approach did not outperform the unmixing algorithm for the parameters evaluated in [118]. However, incorporating a ground-truth reflectivity image captured from a coaxial conventional camera did show improved performance with the CNN. Naturally, using a ground-truth reflectivity image

to form more precise superpixels in the unmixing algorithm would likely also lead to improved performance. Lindell et al. also found that the unmixing algorithm performed worse if the SBR value was not correctly known, e.g., if the background calibration was incorrect.

- Tachella et al. [186] introduced an RJ-MCMC approach called ManiPoP for the estimation of point clouds with potentially multiple surfaces per pixel under challenging photon-starved and high-background conditions. In tests with only a single surface per pixel, the unmixing approach performs comparably to ManiPoP for reflectivity estimation but is slightly better at recovering depths in the most photon-starved case, and with significantly shorter computation times. However, the unmixing approach had worse performance for datasets in which not all pixels actually contained surfaces.
- Halimi et al. [82] compared against unmixing in their evaluation of a convex algorithm (NR3D) for multi-depth, multi-spectral, and multi-temporal SPL data. While unmixing is not designed for such scenarios, it did slightly outperform NR3D for depth estimation and some reflectivity estimation in the single-depth, photon-starved, high-background regime, again with significantly shorter computation times. The same algorithm was also tested in [188] for datasets collected through obscurants, including canister smoke, glycol-based smoke, and water-based fog. These obscuring media reduce both the amount of the laser illumination that reaches each pixel and the amount of light reflected back to the detector. The unmixing approach again performed comparably to the proposed NR3D algorithm.
- Li et al. [116] present the longest terrestrial photon-efficient SPL image acquisitions known to date, with depths up to 45 km. At such distances, the laser

beam has substantially diverged to cover multiple depths within one pixel. As a result, Li et al. show improved results from their novel algorithm outperforming the unmixing approach on a dataset with 2.59 signal ppp and an  $\text{SBR} \approx 0.03$ .

- Finally, the unmixing algorithm was shown to perform well in [78] after correction for dead time effects.

### 3.6 Conclusions

The conventional approaches to active imaging in significant ambient light are to increase either the acquisition time or the illumination power. In many situations, neither solution is practically feasible. In the case of autonomous navigation, for instance, vehicle lidar systems need rapid depth acquisition using eye-safe laser intensities and without draining the limited power resources. The only possible approach is a photon-efficient solution, which can make accurate measurements from very little incident signal illumination, even when the ambient light levels are high.

Based on key observations of the probabilistic nature of the signal and background detection processes, a simple windowing approach yields an effective unmixing of the component detection processes. By setting cluster size requirements based on the easily-measured background rate, we ensure that the number of falsely accepted background detections is limited. Remaining gaps where too few signal detections were collected can be effectively filled through the spatially-adaptive process of forming superpixels and aggregating detections within those regions.

While the unmixing algorithm was designed as a general-purpose approach to adaptively handle high levels of background, specific situations that violate some assumptions in Table 3.1 or incorporate additional prior information may benefit more from alternative approaches. For instance, scenes that are highly textured or patterned (many reflectivity edges) may be difficult to image because the depth

and reflectivity edge locations are less correlated. For scenes where an object and the background appear to have the same reflectivity (too few reflectivity edges), the superpixel algorithm will likely borrow detections from objects at different depths, oversmoothing the edges. Other works listed in Section 3.5 have shown that unmixing also does not consistently succeed for scenes with partially transmissive or partially occluding surfaces because the superpixel formation and windowing stages assume a single depth per pixel. Another case where better approaches are available is when all scene patches have similar depths. Forming a histogram of the entire data cube as in [171] can then globally find the cluster of signal detections because the spatial adaptivity of the superpixels is unnecessary.

Nevertheless, a great benefit to our approach is the modularity of the algorithm, which leaves room for improvement and customization with upgrades to the component blocks. For the results presented in this chapter, we perform only a few loops through the algorithm using well-tuned parameters that provide good results at low SBR. While forming superpixels helps fill in values for many pixels with empty depth estimates, we still require some inpainting to fill in the rest. An ideal approach would likely perform more iterations, incrementing  $N_{i,j}^{\text{sp}}$  by one until each pixel has a reasonable depth estimate. A major factor preventing this “Goldilocks” approach for the just-right  $N_{i,j}^{\text{sp}}$  at each pixel is the computational cost of concatenating and windowing many large vectors of detections. Better implementations of our code could take advantage of the embarrassingly parallel problem structure [132] with more distributed or GPU-accelerated computations. Additional approaches to possibly improve results include alternative superpixel definitions, such as the fast SLIC method [2], or regularizers such as Joint Basis Pursuit [189] that take further advantage of correlations between depth and reflectivity images.

## Chapter 4

# Subtractively-Dithered Ranging

### 4.1 Overview

Although depth imaging with SPADs as previously described provides extremely fine depth resolution, the standard acquisition process is often too slow for real-time applications, since the single-pixel detectors require the scene to be raster-scanned with a laser [128]. Arrays of SPADs promise dramatic reductions in acquisition time, as the entire scene can be simultaneously illuminated, with spatial resolution derived from the multiple SPAD elements sensing in parallel [122]. The main problem with current SPAD cameras is that they trade off a higher pixel count for coarser timing resolution [201], which means the precision of range measurements using short-duration laser pulses is often limited by the temporal quantization bin size. Some recent approaches such as [187, 203] have proposed alternative designs for SPAD arrays in which the time-to-digital convertor (TDC) circuitry uses ring oscillators to encode the finer bits; however, such per-pixel TDCs still limit the array fill factor.

Our goal instead is to improve the precision of lidar depth estimates for existing SPAD arrays in a statistically efficient manner, so that the time reduction of parallelized acquisition is not lost by requiring too many photon detections. Such an improvement cannot be achieved by better probabilistic modeling alone but also requires modifications to the standard acquisition procedure. In [171], temporally

---

This chapter includes research conducted jointly with Robin Dawson and Vivek Goyal [148, 149, 150].

spreading a narrow laser pulse, equivalent to adding non-subtractive Gaussian dither, was found to reduce the effects of the detector’s coarse temporal resolution on ranging accuracy. This chapter demonstrates that implementing *subtractive* dither can likewise reduce the effects of coarse quantization.

We begin by examining subtractive dither in generic measurement scenarios not limited to lidar by determining performance limits, optimal estimators, and when one form of dither might be preferable over the other. The estimators developed in this chapter are based on a generalized Gaussian (GG) approximation for the combination of sample variation and quantization noise. We show that our framework is valid for a general set of problems in which quantization of a scalar Gaussian signal occurs. We propose a number of estimators for additive GG noise and find a clear computational advantage with negligible loss in accuracy for simple estimators based on order statistics. We then implement a subtractively-dithered SPL system and show how the GG model performs for lidar data. Finally, we introduce a refined model to compensate for the mismatch between the lidar pulse profile and the Gaussian assumption. We note that since reflectivity estimation can be performed as before, we focus only on recovering depth in this chapter.

## 4.2 Formulation, Background, and Motivation

### 4.2.1 Measurement Model

The basic lidar imaging setup follows the description in Section 2.2.4, with modifications described later in this chapter. We assume the background level is negligible due to spectral filtering or censoring (e.g., as described in Section 3) removing contributions due to ambient light. We further assume that one pixel of a SPAD camera corresponds to a single patch of the scene, such that all back-reflected photons correspond to the single true depth, and that the pulse shape is approximately Gaussian.

The resulting problem of estimating the mean of a Gaussian signal from quantized measurements applies to a broader class of acquisition methods than lidar alone, so we first consider the more generic case of an unknown constant signal  $\mu_X$  corrupted by additive, zero-mean Gaussian noise  $Z \sim \mathcal{N}(0, \sigma_Z^2)$ . Estimation of  $\mu_X$  from  $K$  independent samples

$$X_i = \mu_X + Z_i, \quad i = 1, 2, \dots, K, \quad (4.1)$$

is straightforward, as the sample mean  $\bar{\mu} = (1/K) \sum_{i=1}^K X_i$  is well known to be an efficient estimator of the mean of a Gaussian distribution.

However, all measurement instruments perform some quantization, and the precision of the sample mean estimator strongly depends on the coarseness of the quantization. For instance, consider a uniform midread quantizer  $q(\cdot)$  with bin size  $\Delta$  applied to  $X_i$  when  $\sigma_Z \ll \Delta$ . Except when  $\mu_X$  is close to a quantizer threshold, it will be the case that  $U_i = q(X_i)$  is identical for all  $i$ , so that the “quantized-sample mean” given as

$$\hat{\mu}_Q = \frac{1}{K} \sum_{i=1}^K U_i, \quad (4.2)$$

is no more informative an estimate of  $\mu_X$  than any single measurement. For  $\sigma_Z$  not too small compared to  $\Delta$ , estimation error can be reduced by properly accounting for the quantization and the underlying distribution, e.g., via the maximum likelihood estimator for quantized samples of a Gaussian signal [198, 135]:

$$\hat{\mu}_{\text{QML}} = \arg \max_{\mu_X} \sum_{i=1}^K \log \left[ \Phi \left( \frac{u_i - \mu_X + \frac{\Delta}{2}}{\sigma_Z} \right) - \Phi \left( \frac{u_i - \mu_X - \frac{\Delta}{2}}{\sigma_Z} \right) \right], \quad (4.3)$$

where  $\Phi(\cdot)$  is the cumulative distribution function (CDF) of a standard normal random variable. Still,  $\hat{\mu}_{\text{QML}}$  is no more accurate than  $\hat{\mu}_Q$  when all of the samples have the same value. Because of the coarse quantization mapping every value in

$[j\Delta - \Delta/2, j\Delta + \Delta/2]$  to  $j\Delta$  for  $j \in \mathbb{Z}$ , the resolution of an estimate  $\hat{\mu}_X$  is limited by the bin size  $\Delta$  and the quantization error is signal-dependent.

Statisticians have long recognized that working with rounded data is not the same as working with underlying continuous-valued data. We consider the continuous random variable  $X_{\text{hist}}$  with density constant on intervals  $((j - \frac{1}{2})\Delta, (j + \frac{1}{2})\Delta)$  with  $P(X_{\text{hist}} \in ((j - \frac{1}{2})\Delta, (j + \frac{1}{2})\Delta)) = P(U = j\Delta)$ , for all  $j \in \mathbb{Z}$ . Because of the piecewise-constant form,  $X_{\text{hist}}$  is said to have a *histogram density* [197]. Sheppard [168, 169] introduced widely known corrections that relate the moments of the discrete random variable  $U$  and the moments of  $X_{\text{hist}}$  [101]. However, the moments of  $X_{\text{hist}}$  being close to the moments of the underlying continuous random variable  $X$  depends on continuity arguments and  $\Delta$  being small. In contrast, our interest is in situations like SPL with SPAD arrays where the quantization is coarse relative to the desired precision in estimating  $\mu_X$ .

When quantizing  $X_{\text{hist}}$ , the quantization error  $E_{\text{hist}} = q(X_{\text{hist}}) - X_{\text{hist}}$  is uniformly distributed on  $[-\Delta/2, \Delta/2]$  and independent of  $X_{\text{hist}}$ . In general, however, quantization error being uniformly distributed and independent of the input does not extend to the quantization of  $X$ ; approximating quantization error as such—without regard to whether the input has a histogram density—is often called the “additive-noise model,” “quantization-noise model,” or “white-noise model.” A substantial literature is devoted to understanding the validity of this approximation, e.g. [214, 176, 43, 202, 125].

One approach to improving the precision of estimates from quantization-limited measurements is the use of dither, a small signal introduced before the discretization to produce enough variation in the input such that it spans multiple quantization levels. By combining multiple dithered measurements, estimates can achieve resolution below the least-significant bit and the result may also have desirable statistical and perceptual properties, such as whitened noise. Early applications empirically

demonstrating the benefits of dither include control systems [123, 95], image display [74, 159], and audio [98], with numerous contributions to the statistical theory developed in [64, 97, 214, 166, 176, 75, 211], among others. More recent work has focused on varying the quantizer thresholds primarily for 1-bit measurements in wireless sensing networks, including [140, 161, 157, 49, 58, 59, 68].

#### 4.2.2 Subtractively-Dithered Quantization

When not only the dither distribution but also the exact value is known for each sample, it is possible to achieve an additional reduction in estimation error. Papadopoulos et al. [140] showed that a sawtooth waveform was optimal for known dither, whereas further improvement could be achieved with adaptively-selected dither values. However, a deterministic-length dither sequence is ill-suited to applications such as single-photon lidar, for which the number of measurements is random, and adaptive dither may be infeasible to implement for such high-speed systems. Instead, we consider *subtractively-dithered quantization*, which uses a random dither signal that is measured for every sample. The ML estimator for subtractive dither can be shown to be equivalent to generic quantization schemes with known dither values, e.g., [157, 68]. However, principled selection of the subtractive dither signal distribution makes the quantization error uniformly distributed and independent of the input, which we show can lead to simple, non-iterative estimators.

Define the dither signal  $D_i$ ,  $i = 1, \dots, K$  as a sequence of i.i.d. random variables, independent of the noisy quantizer input  $X_i$ . The output of a subtractively-dithered quantizer is

$$Y_i = q(X_i + D_i) - D_i, \quad (4.4)$$

with the quantization error defined as

$$W_i = Y_i - X_i. \quad (4.5)$$

Define the characteristic function of the dither signal probability density function (PDF) as

$$M_D(ju) = \mathbb{E}[e^{juD}]. \quad (4.6)$$

Then Schuchman's condition [166] is the property of the dither PDF that

$$M_D(j2\pi\ell/\Delta) = 0, \quad \ell \in \mathbb{Z} \setminus 0. \quad (4.7)$$

As long as the quantizer has a sufficient number of levels so that it does not overload, by [176, 75] the Schuchman condition is necessary and sufficient for  $X_i$  to be independent of  $W_j$  for all  $i, j$ , with i.i.d.  $W_i \sim \mathcal{U}[-\Delta/2, \Delta/2]$ . Subtractive dither often uses a uniform dither signal with  $D \sim \mathcal{U}[-\Delta/2, \Delta/2]$  because its characteristic function

$$M_D(ju) = \frac{\sin(u\Delta/2)}{u\Delta/2}$$

meets Schuchman's condition (4.7).

The rest of this chapter considers only when Schuchman's condition is met, with an i.i.d. input signal of the form  $X_i = \mu_X + Z_i$ ,  $Z \sim \mathcal{N}(0, \sigma_Z^2)$ , an i.i.d. dither signal  $D_i \sim \mathcal{U}[-\Delta/2, \Delta/2]$  independent of the input signal, and a non-overloading uniform quantizer. Since Gaussian noise has infinite support, in principle the quantizer must have an infinite number of levels. Then the dithered measurements take the form

$$Y_i = \mu_X + Z_i + W_i, \quad (4.8)$$

and the problem of estimating  $\mu_X$  simply becomes one of mitigating independent additive noise. The sum of the Gaussian and uniform terms can be combined into a single total noise term to obtain

$$Y_i = \mu_X + V_i, \quad (4.9)$$

where  $V_i = Z_i + W_i$  are i.i.d. Then the means and variances simply add so that  $\mu_V = 0$  and  $\sigma_V^2 = \sigma_Z^2 + \Delta^2/12$ .

For convenient shorthand, we refer to measurements from a quantizer without dither as “quantized” and measurements from a subtractively-dithered quantizer as “dithered.” The usual approach to estimating  $\mu_X$  from  $K$  dithered measurements  $Y_i$ ,  $i = 1, 2, \dots, K$ , is via the sample mean

$$\hat{\mu}_{\text{mean}} = \frac{1}{K} \sum_{i=1}^K Y_i. \quad (4.10)$$

The MSE of the sample mean is

$$\text{MSE}(\text{mean}) = \sigma_V^2/K, \quad (4.11)$$

which is  $O(K^{-1})$ . Although using the sample mean is logical when  $\sigma_Z \gg \Delta$  so that the contribution of the uniform noise component is negligible, the sample mean is not in general an efficient estimator. For example, in an alternative case of  $\sigma_Z = 0$ , a maximum likelihood (ML) estimator is the midrange

$$\hat{\mu}_{\text{mid}} = \frac{1}{2} (Y_{(1)} + Y_{(K)}), \quad (4.12)$$

where  $Y_{(1)} \leq Y_{(2)} \leq \dots \leq Y_{(K)}$  are the order statistics of the  $K$  measured samples. We note that any statistic in  $[Y_{(n)} - \Delta/2, Y_{(1)} + \Delta/2]$  is an ML estimator for the mean of a uniform distribution with known variance [133, p. 282]. The midrange is commonly used because it is unbiased and the minimum-variance estimator among linear functions of order statistics [120]. However, no uniformly minimum-variance unbiased estimator exists [133, p. 331]. Whereas the MSE of the sample mean for  $\sigma_Z = 0$  is  $\Delta^2/(12K)$ , the MSE of the midrange is

$$\text{MSE}(\text{mid}) = \Delta^2/[2(K+1)(K+2)], \quad (4.13)$$

which is  $O(K^{-2})$  and hence better than the sample mean by an unbounded factor [120]. Nevertheless, the midrange is not a good estimator in the general case of  $\sigma_Z > 0$ , as it relies on the finite support of the uniform distribution. If instead  $\sigma_Z$  is much larger than  $\Delta$ , rendering the uniform component negligible, then the MSE of the midrange would only improve as  $O(1/\log(K))$  [28]. We observe that normalizing the MSE by  $\Delta^2$  removes the separate dependence on  $\sigma_Z$  and  $\Delta$ , resulting in

$$\text{NMSE}(\text{mean}) = [(\sigma_Z/\Delta)^2 + 1/12]/K, \quad (4.14)$$

and

$$\text{NMSE}(\text{mid}) = 1/[2(K+1)(K+2)]. \quad (4.15)$$

Except in trivial cases ( $\sigma_Z \gg \Delta$  or  $\sigma_Z \ll \Delta$ ),  $V$  has neither Gaussian nor uniform distribution, so the conventional mean and midrange estimators are expected to be suboptimal. Furthermore, existing nonlinear processing schemes for dithered measurements do not adapt to best suit the noise statistics [37].

The key figure of merit for determining estimator performance for quantized Gaussian measurements without dither is then  $\sigma_Z/\Delta$ , a measure of the relative sizes of the noise components [198, 135]. A first approach to finding a better estimator for arbitrary  $\sigma_Z/\Delta$  is to derive the ML estimator for the dithered noise model. From the definitions of the random variables, the PDF of  $W$  is

$$f_W(w) = \begin{cases} 1/\Delta, & w \in [-\Delta/2, \Delta/2] \\ 0, & \text{otherwise,} \end{cases} \quad (4.16)$$

and the PDF of  $Z$  is  $f_Z(z) = \phi(z/\sigma_Z)/\sigma_Z$ , where  $\phi(x)$  is the standard normal PDF. Since the total noise is the sum of independent noise terms, the PDF of the samples

is given by the convolution

$$\begin{aligned} f_V(v) &= f_Z(z) * f_W(w) = \frac{1}{\Delta} \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} f_Z(v - \tau) d\tau \\ &= \frac{1}{\Delta} \left[ \Phi\left(\frac{v + \frac{\Delta}{2}}{\sigma_Z}\right) - \Phi\left(\frac{v - \frac{\Delta}{2}}{\sigma_Z}\right) \right]. \end{aligned} \quad (4.17)$$

For i.i.d. samples from a dithered quantizer, the likelihood function is then

$$\begin{aligned} \mathcal{L}(\{y_i\}_{i=1}^K; \mu_X) &= \prod_{i=1}^K f_V(y_i - \mu_X) \\ &= \prod_{i=1}^K \frac{1}{\Delta} \left[ \Phi\left(\frac{y_i - \mu_X + \frac{\Delta}{2}}{\sigma_Z}\right) - \Phi\left(\frac{y_i - \mu_X - \frac{\Delta}{2}}{\sigma_Z}\right) \right]. \end{aligned} \quad (4.18)$$

From the log-likelihood, the dithered-sample ML estimator of  $\mu_X$  is

$$\hat{\mu}_{\text{DML}} = \arg \max_{\mu_X} \sum_{i=1}^K \log \left[ \Phi\left(\frac{y_i - \mu_X + \frac{\Delta}{2}}{\sigma_Z}\right) - \Phi\left(\frac{y_i - \mu_X - \frac{\Delta}{2}}{\sigma_Z}\right) \right]. \quad (4.19)$$

The ML estimator is notably identical to (4.3), except the dithered measurements are not discrete-valued as are the samples used for  $\hat{\mu}_{\text{QML}}$ .

To determine the efficiency of the mean, midrange, and DML estimators, we derive the Cramér-Rao bound (CRB), which is a limit on the MSE that an unbiased estimator can achieve [195, Chapter 4.2.2]. The CRB normalized by  $\Delta^2$  is derived in Appendix 4.A for one dithered measurement to be

$$\text{NCRB}(\mu_X) = \frac{(\sigma_Z/\Delta)^2}{\int \frac{\left[ \phi\left(\frac{u-1/2}{\sigma_Z/\Delta}\right) - \phi\left(\frac{u+1/2}{\sigma_Z/\Delta}\right) \right]^2}{\Phi\left(\frac{u+1/2}{\sigma_Z/\Delta}\right) - \Phi\left(\frac{u-1/2}{\sigma_Z/\Delta}\right)} du}, \quad (4.20)$$

which can be evaluated via numerical integration. Note that the uniform PDF does not meet the regularity condition required for the CRB to apply, so (4.20) is not

expected to be meaningful for  $\sigma_Z/\Delta = 0$ .

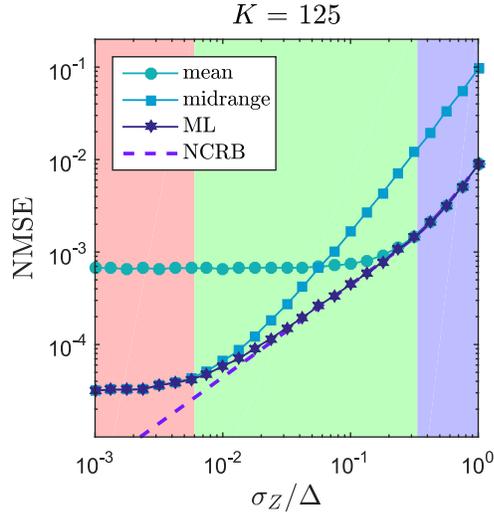
Fig. 4-1 illustrates the suboptimality of the mean and midrange estimators compared to  $\hat{\mu}_{\text{DML}}$  for intermediate values of  $\sigma_Z/\Delta$ . In a Monte Carlo simulation with  $T = 20000$  trials,  $K = 125$  measurements were generated according to (4.4), where both  $\mu_X$  and  $D$  were selected uniformly at random over  $[-\Delta/2, \Delta/2]$ . Computing the normalized MSE of the  $\hat{\mu}_{\text{mean}}$ ,  $\hat{\mu}_{\text{mid}}$ , and  $\hat{\mu}_{\text{DML}}$  estimates as

$$\text{NMSE}(\hat{\mu}_X) = \frac{1}{T} \sum_{t=1}^T \left( \frac{\mu_X - \hat{\mu}_X}{\Delta} \right)^2 \quad (4.21)$$

reveals how the performance of each estimator changes as a function of  $\sigma_Z/\Delta$ .

Fig. 4-1 highlights three distinct regimes of estimator behavior. In Regime I (red), the Gaussian noise component is negligible, so the ML estimator and the midrange are nearly identical and outperform the mean. In Regime III (blue), the uniform noise component is negligible, so the ML estimator and the mean are nearly identical and outperform the midrange. In Regime II (green), neither the uniform nor the Gaussian component dominates, and the DML estimator performs significantly better than both the mean and midrange. Still,  $\hat{\mu}_{\text{DML}}$  does not achieve the CRB for small  $\sigma_Z/\Delta$ , indicating that an efficient estimator of  $\mu_X$  does not exist; however,  $\hat{\mu}_{\text{DML}}$  is asymptotically efficient in  $K$  for  $\sigma_Z/\Delta > 0$  [100, Theorem 7.1].

From the results in Fig. 4-1, it may seem obvious that  $\hat{\mu}_{\text{DML}}$  is a better choice than  $\hat{\mu}_{\text{mean}}$  or  $\hat{\mu}_{\text{mid}}$  for any value of  $\sigma_Z/\Delta$ . However,  $\hat{\mu}_{\text{DML}}$  requires iterative solution, thus making it far more computationally complex than the mean and midrange. One of the primary aims of this chapter is to find a computationally simple estimator that can likewise outperform the mean and midrange in Regime II. We show that a generalized Gaussian approximation to the total noise of a dithered quantizer gives rise to order statistics-based estimators that approach the performance of  $\hat{\mu}_{\text{DML}}$ . In addition, we compare their results to those from quantized measurements without



**Figure 4.1:** Applied to subtractively-dithered measurements, the midrange is approximately optimal only for small  $\sigma_Z/\Delta$  (Regime I, red), and the sample mean is approximately optimal only for large  $\sigma_Z/\Delta$  (Regime III, blue). For confirmation, in Regime I the midrange approaches  $\text{NMSE} = 1/[2(K+1)(K+2)] \approx 3.12 \times 10^{-5}$  and the mean approaches  $\text{NMSE} = (1/12)/K \approx 6.67 \times 10^{-4}$ , and in Regime III the mean approaches  $\text{NMSE} = (\sigma_Z/\Delta)^2/K$ , which is  $8.00 \times 10^{-3}$  at  $\sigma_Z/\Delta = 1$ . We seek an estimator simpler than the dithered-sample maximum likelihood that performs at least as well as the mean and midrange for intermediate values of  $\sigma_Z/\Delta$  (Regime II, green).

dither, leading to design rules for when to use dither and which estimator to apply.

### 4.3 Generalized Gaussian Approximation and Estimation

In order to find a simple estimator for Regime II, we begin by examining the other two regimes and the simple forms of the ML estimator there. We notice that the uniform and Gaussian noise distributions in Regimes I and III are special cases of the generalized Gaussian distribution (GGD), which has PDF [196]

$$f_{\tilde{V}}(v; \mu, \sigma, p) = \frac{1}{2\Gamma(1 + 1/p)A(p)} \exp\left\{-\left(\frac{|v - \mu|}{A(p)}\right)^p\right\}, \quad (4.22)$$

where  $A(p) = \sqrt{\sigma^2 \Gamma(1/p) / \Gamma(3/p)}$  and  $\Gamma(\cdot)$  is the Gamma function. In addition to mean and variance parameters  $\mu$  and  $\sigma^2$ , the GG density has a third parameter  $p$  that controls the exponential decay of its tails. When  $p = 2$  or  $p \rightarrow \infty$ , the GGD simplifies to the Gaussian or uniform distributions, respectively. Another common special case of the GGD is the Laplace distribution for  $p = 1$ .

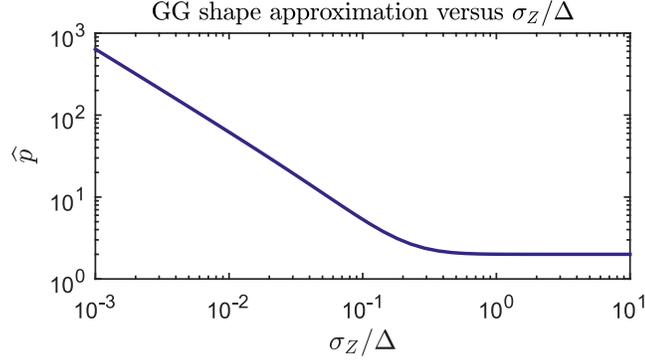
For each of the special cases, we further notice that the ML estimator (median, mean, and midrange for  $p = 1, 2, \infty$ ) is a linear combination of order statistics. When  $p = 1$ , only the middle order statistic has nonzero weight, whereas the reverse is true for  $p \rightarrow \infty$ , with all weight on the two extreme samples. For  $p = 2$ , all of the order statistics are equally weighted. With these two observations in mind, we hypothesize that, if there is a value of  $p$  that approximates intermediate combinations of uniform and Gaussian noise, then there may be a corresponding order statistics-based estimator that approaches the performance of  $\hat{\mu}_{\text{DML}}$ .

### 4.3.1 Approximation

For our stated purpose, it would be ideal if proper selection of  $p$  exactly represented nontrivial sums of Gaussian and uniform terms. Unfortunately the sum of any two independent GG random variables (GGRVs) is another GGRV only when  $p = 2$  for each addend<sup>1</sup> [219]. Nevertheless, the sum of independent GGRVs has many of the same properties as a GGRV, and can be well-approximated as a GGRV through a number of approximation methods. A simple approach from [175] matches the mean, variance, and kurtosis of the GG approximation to the corresponding moments of the true noise distribution as follows. Defining  $\tilde{V}$  as the GG approximation to  $V = Z + W$ , then since the uniform and Gaussian noise components are independent random variables, the mean and variance parameters of the GG noise approximation

---

<sup>1</sup>The limiting distribution of the sum of i.i.d. GGRVs is Gaussian by the Central Limit Theorem [133, Chapter 5.4.2], but the sum of any finite number of GGRVs will only be approximately Gaussian unless each term is Gaussian.



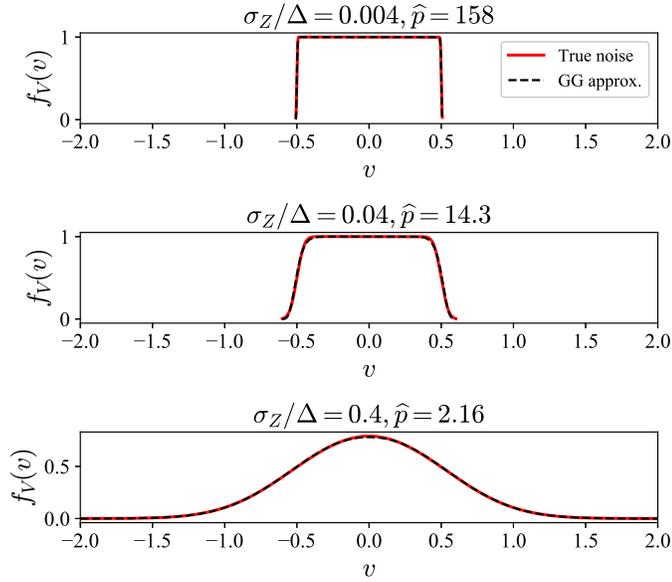
**Figure 4.2:** The value of  $\hat{p}_{\tilde{V}}$  goes to infinity as  $\sigma_Z/\Delta$  decreases and converges to  $\hat{p}_{\tilde{V}} = 2$  as  $\sigma_Z/\Delta$  increases, with convergence beginning around  $\sigma_Z/\Delta = 1/3$  matching the anticipated behavior.

are simply given as  $\mu_{\tilde{V}} = \mu_W + \mu_Z$  and  $\sigma_{\tilde{V}}^2 = \sigma_Z^2 + \sigma_W^2$ . The shape parameter approximation  $\hat{p}_{\tilde{V}}$  for the special case of uniform and Gaussian addends is computed as the unique solution to

$$\frac{\Gamma(1/\hat{p}_{\tilde{V}})\Gamma(5/\hat{p}_{\tilde{V}})}{\Gamma(3/\hat{p}_{\tilde{V}})^2} = 3 - \frac{6}{5} \frac{1}{\left[12 \left(\frac{\sigma_Z}{\Delta}\right)^2 + 1\right]^2} \quad (4.23)$$

(see derivation in Appendix 4.B). We thus see that  $\hat{p}_{\tilde{V}}$  depends on  $\sigma_Z/\Delta$ , with the relationship plotted in Fig. 4.2. Solving (4.23) is fast, and the values of  $\hat{p}_{\tilde{V}}$  for a range of  $\sigma_Z/\Delta$  values could be precomputed and stored in a table if necessary. A rough approximation and good initial value for a solver is  $\hat{p}_{\tilde{V}}^{(0)} = \max\{2, \Delta/\sigma_Z\}$ .

To verify the quality of the generalized Gaussian approximation to the output noise distribution using the kurtosis match, Fig. 4.3 shows comparisons between the true density, computed numerically according to (4.17), and its GG approximation from (4.22). We test  $\sigma_Z/\Delta = 0.004, 0.04,$  and  $0.4$ , maintaining  $\Delta = 1$  for consistency. For  $\sigma_Z \ll \Delta$ , the distribution is close to uniform, and at  $\sigma_Z \approx \Delta$ , the distribution is almost Gaussian. In the intermediate regime, however, the distribution combines attributes of each component, with the flat top of the uniform distribution and exponential tails of the Gaussian distribution. The GGD appears to be a good



**Figure 4.3:** The three plots show the noise PDF calculated numerically from the true density (4.17) (solid red) and via the GG approximation (4.22) (dashed black). The close agreement suggests the GGD is a good approximation for the noise.

approximation of the true noise distribution, almost perfectly matching the shape behavior.

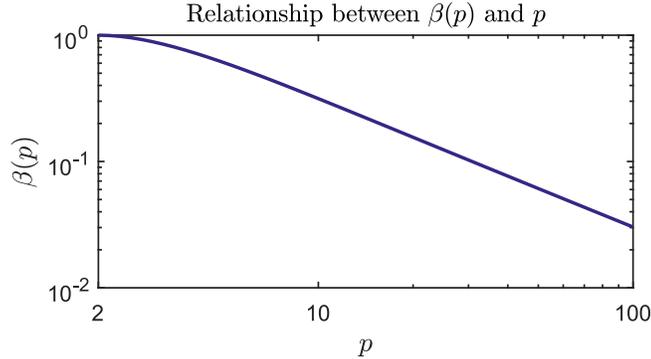
### 4.3.2 Estimation

For i.i.d. samples of a GG distribution, the likelihood function is

$$\mathcal{L}(\{v_i\}_{i=1}^K; \mu, \sigma, p) = \prod_{i=1}^K f_{\tilde{V}}(v; \mu, \sigma, p). \quad (4.24)$$

By differentiating the log of (4.24) with respect to  $\mu$ , the ML estimator  $\hat{\mu}_{\text{GGML}}$  for the mean of a GGRV is given in [196] as the solution to

$$\sum_{i=1}^K \text{sgn}(y_i - \hat{\mu}_{\text{GGML}}) |y_i - \hat{\mu}_{\text{GGML}}|^{p-1} = 0, \quad (4.25)$$



**Figure 4.4:** As  $p$  increases beyond  $p = 2$ ,  $\beta(p)$  becomes much less than 1, implying that  $\text{MSE}(\hat{\mu}_{\text{GGML}})$  is much lower than  $\text{MSE}(\hat{\mu}_{\text{mean}})$ .

and is shown to be asymptotically normal and efficient in  $K$  for  $p \geq 2$ , which is the regime of interest. The asymptotic variance of  $\hat{\mu}_{\text{GGML}}$  normalized by  $\Delta^2$  is given by

$$\text{NVar}(\text{GGML}) = \frac{\beta(p)[(\sigma_Z/\Delta)^2 + 1/12]}{K}, \quad (4.26)$$

where

$$\beta(p) = \frac{\Gamma^2(1/p)}{p^2 \Gamma\left(\frac{2p-1}{p}\right) \Gamma(3/p)}. \quad (4.27)$$

We notice that (4.26) decreases as  $O(K^{-1})$ , but the coefficient  $\beta(p)$ , which is plotted in Fig. 4.4, is much less than 1 for large  $p$ , suggesting that  $\hat{\mu}_{\text{GGML}}$  should outperform  $\hat{\mu}_{\text{mean}}$  for  $\hat{p}_{\tilde{V}} > 2$ . Since the GGD closely approximates the total noise distribution, it would be ideal if  $\hat{\mu}_{\text{GGML}}$  reduced to a computationally simple estimator such as one based on order statistics for all  $p$ . Unfortunately, the ML estimator does not generally have a closed-form expression, except in special cases such as  $p = 1, 2, \infty$  (an explicit expression has also recently been derived for  $p = 4$  [14]), so iterative solution would again be necessary.

We have already observed that the ML estimators for  $p = 1, 2, \infty$  all belong to a class of linear combinations of order statistics called L-estimates [9], which are

attractive because they have closed-form definitions of the form

$$\hat{\mu}_X = \sum_{i=1}^K a_i Y_{(i)}. \quad (4.28)$$

We thus consider how to obtain the coefficients  $a_i$  for L-estimates that perform well for GG noise when  $p$  is not one of the special cases.

An effective L-estimate should weight the order statistics in accordance with the noise distribution. Notable past approaches include that of Lloyd [120], who derived the best linear unbiased estimator (BLUE) of the order statistics. This formulation is impractical, however, as it requires the correlations of the order statistics for a given distribution, which are often not known even for common special cases like the Gaussian distribution. Bovik et al. [25, 26] further specified the minimum variance unbiased L-estimate and then numerically computed results for several values of  $p$  from samples of a GGD with  $K = 3$ .

A number of approximations to Lloyd’s formulation exist to more simply compute near-optimal coefficients for linear combinations of order statistics, including [79, 23]. Öten and de Figueiredo [138] introduced one such method using Taylor expansion approximations to get around the difficulties of knowing distributions of order statistics. This method does still require knowledge of the inverse CDF of the noise distribution, and while there is no closed form expression for the GGD, the necessary values can be pre-computed numerically.

Simpler L-estimates have much longer histories, with consideration of trimming extreme or middle order statistics at least as old as the anonymous suggestion of [6] in 1821 (credited to Gergonne in [180]) The first known mathematical analysis was by Daniell, who called such an estimate the “discard-average” [50, 179]. The method now known as the  $\alpha$ -trimmed mean and popularized by Tukey [192, 193] avoids extensive computation of the weights by trimming a fixed fraction  $\alpha$  from the extremes

of the order statistics. Restrepo and Bovik [156] defined a complementary  $\alpha$ -“outer” trimmed-mean, which retains a fraction  $\alpha$  of the data by trimming the middle order statistics and is suitable for distributions with short tails within the range from Gaussian to uniform distributions. They tabulated several instances of the trimmed mean for GGDs with multiple combinations of  $K$  and  $p$ .

Lastly, Beaulieu and Guo [13] introduced an estimator specifically for the GGD but using nonlinear combinations of the order statistics. The weighting of the order statistics depends on  $p$  via a heuristically-justified function and is shown to perform almost identically to  $\hat{\mu}_{\text{GGML}}$ . This estimator is unbiased and exactly matches the ML estimator for the special cases of  $p = 2$  and  $\infty$ .

In the following section, we consider three of the most computationally-efficient order statistics-based estimators to use for the GG approximation: the nearly-best L-estimate  $\hat{\mu}_{\text{NB}}$  of Öten and de Figueiredo [138], the trimmed-mean estimator  $\hat{\mu}_{\alpha}$  modeled on [156], and the non-linear estimator  $\hat{\mu}_{\text{NL}}$  of Beaulieu and Guo [13]. Each estimator takes the form of (4.28) with different computations of the coefficients  $a_i$ . While  $\hat{\mu}_{\text{NL}}$  is specifically designed for use with GG noise, we modify the more general  $\hat{\mu}_{\text{NB}}$  and  $\hat{\mu}_{\alpha}$  to match the GG approximation. For  $\hat{\mu}_{\text{NB}}$ , we use the PDF and inverse CDF (computed numerically) of the GG approximation to determine the coefficients. One could alternatively compute the coefficients for  $\hat{\mu}_{\text{NB}}$  directly from the true noise distribution in (4.17); however, additional numerical evaluation would be required for the inverse CDF, which we eschew in our search for computationally efficient estimators. There is no explicit distribution assumed by  $\hat{\mu}_{\alpha}$ , but we propose a choice of the trimmed fraction  $\alpha$  based on the estimated  $\hat{p}_{\hat{v}}$  value to implicitly link the estimator to the GGD.

## 4.4 Estimator Implementations

### 4.4.1 ML Estimators

An EM algorithm for obtaining the quantized-sample ML estimate  $\hat{\mu}_{\text{QML}}$  was introduced by Papadopoulos et al. [140, Appendix E]:

$$\hat{\mu}_{\text{QML}}^{(j+1)} = \hat{\mu}_{\text{QML}}^{(j)} + \frac{\sigma_Z}{K\sqrt{2\pi}} \sum_{i=1}^K \frac{\exp\left(-\frac{[u_i - \frac{\Delta}{2} - \hat{\mu}_{\text{QML}}^{(j)}]^2}{2\sigma_Z^2}\right) - \exp\left(-\frac{[u_i + \frac{\Delta}{2} - \hat{\mu}_{\text{QML}}^{(j)}]^2}{2\sigma_Z^2}\right)}{\Phi\left(\frac{u_i + \frac{\Delta}{2} - \hat{\mu}_{\text{QML}}^{(j)}}{\sigma_Z}\right) - \Phi\left(\frac{u_i - \frac{\Delta}{2} - \hat{\mu}_{\text{QML}}^{(j)}}{\sigma_Z}\right)}. \quad (4.29)$$

[223] derived a gradient descent algorithm equivalent to that in (4.29) for the special case of a repeated scalar input and no mixing (i.e., the mixing matrix is a column of 1s). They also showed the negative log-likelihood to be convex, so the EM algorithm converges to the ML estimate. A good initialization is  $\hat{\mu}_{\text{QML}}^{(0)} = \hat{\mu}_{\text{Q}}$ , since the estimators are equal for  $\sigma_Z = 0, \infty$ . Since  $\hat{\mu}_{\text{DML}}$  has the same formulation as  $\hat{\mu}_{\text{QML}}$ , the same algorithm also works for continuous-valued dithered measurements:

$$\hat{\mu}_{\text{DML}}^{(j+1)} = \hat{\mu}_{\text{DML}}^{(j)} + \frac{\sigma_Z}{K\sqrt{2\pi}} \sum_{i=1}^K \frac{\exp\left(-\frac{[y_i - \frac{\Delta}{2} - \hat{\mu}_{\text{QML}}^{(j)}]^2}{2\sigma_Z^2}\right) - \exp\left(-\frac{[y_i + \frac{\Delta}{2} - \hat{\mu}_{\text{QML}}^{(j)}]^2}{2\sigma_Z^2}\right)}{\Phi\left(\frac{y_i + \frac{\Delta}{2} - \hat{\mu}_{\text{QML}}^{(j)}}{\sigma_Z}\right) - \Phi\left(\frac{y_i - \frac{\Delta}{2} - \hat{\mu}_{\text{QML}}^{(j)}}{\sigma_Z}\right)}.$$

We initialize with  $\hat{\mu}_{\text{DML}}^{(0)} = \hat{\mu}_{\text{mid}}$ , since the midrange is known to be the ML estimator for  $\sigma_Z = 0$ . A solver for  $\hat{\mu}_{\text{GGML}}$  was likewise initialized with  $\hat{\mu}_{\text{GGML}}^{(0)} = \hat{\mu}_{\text{mid}}$ .

### 4.4.2 Order Statistics-Based Estimators

To evaluate the GG noise approximation and find the best non-iterative estimator, we compared the three simplest estimators based on the order statistics: the nearly-best L-estimate, the  $\alpha$ -outer mean, and the nonlinear combination from [13]. Since the GGD is symmetric, the coefficients of an unbiased order statistics-based estimator are

defined symmetrically and only half must be uniquely computed. It is thus useful to define  $M = \lfloor K/2 \rfloor$  and  $N = \lceil K/2 \rceil$  using the floor and ceiling functions, respectively.

To derive the nearly-best L-estimate of [138]

$$\hat{\mu}_{\text{NB}} = \sum_{i=1}^K a_i^{\text{NB}} Y_{(i)}, \quad (4.30)$$

we first compute

$$b_1 = f_{\tilde{V}}(c_1)[-2f_{\tilde{V}}(c_1) + f_{\tilde{V}}(c_2)], \quad (4.31a)$$

$$b_i = f_{\tilde{V}}(c_i)[f_{\tilde{V}}(c_{i-1}) - 2f_{\tilde{V}}(c_i) + f_{\tilde{V}}(c_{i+1})], \quad i = 2, \dots, N-1,$$

$$b_N = f_{\tilde{V}}(c_N)[f_{\tilde{V}}(c_{N-1}) - f_{\tilde{V}}(c_N)], \quad (4.31b)$$

where  $c_i = F_{\tilde{V}}^{-1}(i/(K+1))$ , and  $F_{\tilde{V}}^{-1}$  is the inverse of the GG CDF. From this, the weights are derived for  $i = 1, \dots, N$  as

$$a_i^{\text{NB}} = a_{K-i+1}^{\text{NB}} = \begin{cases} b_i / \left( 2 \sum_{i=1}^N b_i \right), & K \text{ even;} \\ b_i / \left( b_N + 2 \sum_{i=1}^M b_i \right), & K \text{ odd.} \end{cases} \quad (4.32)$$

For the simulations in Python, the inverse CDF was numerically computed with the `stats.gennorm.ppf` GGD percentile function in `scipy`, as no closed-form expression exists.

For the  $\alpha$ -outer mean estimate

$$\hat{\mu}_{\alpha} = \sum_{i=1}^K a_i^{\alpha} Y_{(i)}, \quad (4.33)$$

the order statistics' weights  $a_i^{\alpha}$  are only given by Restrepo and Bovik [156] for a

symmetric filter applied to an odd number of samples:

$$a_i^\alpha = a_{K-i+1}^\alpha = \begin{cases} \frac{1}{K\alpha}, & i \leq \lfloor \frac{1}{2}K\alpha \rfloor \\ \frac{\frac{1}{2}K\alpha - \lfloor \frac{1}{2}K\alpha \rfloor}{K\alpha}, & i = \lfloor \frac{1}{2}K\alpha \rfloor + 1, \\ & \alpha \in [0, 1 - 1/K]; \\ \frac{K\alpha - 2\lfloor \frac{1}{2}K\alpha \rfloor}{K\alpha}, & i = \lfloor \frac{1}{2}K\alpha \rfloor + 1, \\ & \alpha \in [1 - 1/K, 1]; \\ 0, & \text{otherwise.} \end{cases} \quad (4.34)$$

Since an even number of measurements is also possible, we similarly define, for all  $\alpha \in [0, 1]$ ,

$$a_i^\alpha = a_{K-i+1}^\alpha = \begin{cases} \frac{1}{K\alpha}, & i \leq \lfloor \frac{1}{2}K\alpha \rfloor; \\ \frac{\frac{1}{2}K\alpha - \lfloor \frac{1}{2}K\alpha \rfloor}{K\alpha}, & i = \lfloor \frac{1}{2}K\alpha \rfloor + 1; \\ 0, & \text{otherwise.} \end{cases} \quad (4.35)$$

Note that the outer mean is equivalent to  $\hat{\mu}_{\text{mean}}$  when  $\alpha = 1$  and reduces to  $\hat{\mu}_{\text{mid}}$  for  $\alpha = 0$ . To match the GGD behavior, we thus propose to define  $\alpha = 2/\hat{p}_{\hat{v}}$ , which yields the ML estimate for both  $\hat{p}_{\hat{v}} = 2$  and  $\hat{p}_{\hat{v}} = \infty$ .

Finally, the nonlinear estimator of [13] is given as

$$\hat{\mu}_{\text{NL}} = \sum_{i=1}^K a_i^{\text{NL}} Y_{(i)}, \quad (4.36)$$

where the data-dependent coefficients are given for  $i = 1, \dots, M$  by

$$a_i^{\text{NL}} = a_{K-i+1}^{\text{NL}} = \frac{1}{2} \frac{[Y_{(K-i+1)} - Y_{(i)}]^{p-2}}{\sum_{j=1}^M [Y_{(K-j+1)} - Y_{(j)}]^{p-2}}. \quad (4.37)$$

Note that if  $K$  is odd, the median term ( $i = N$ ) is ignored, as it would correspond to a numerator of zero.

## 4.5 Dither Noise Regimes

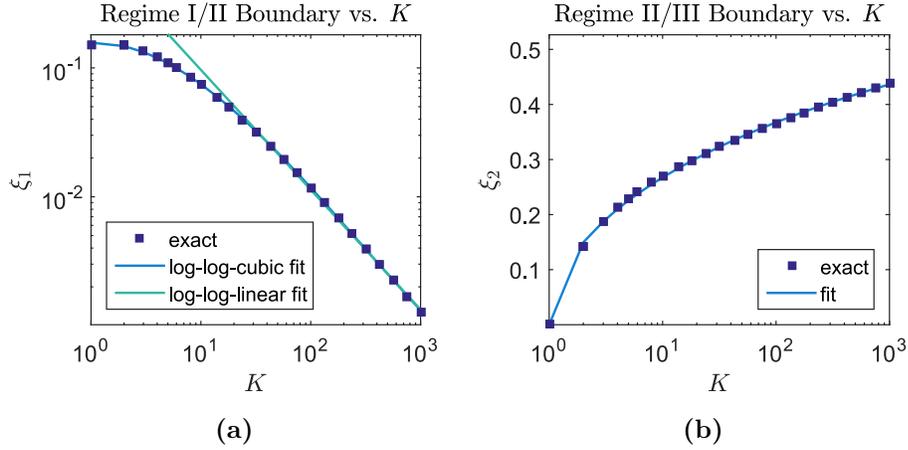
To better understand the dither noise behavior, we have previously described three regimes of the dither noise distribution, with Regimes I and III corresponding to approximately uniform and Gaussian noise, respectively. We have furthermore proposed the GGD with  $p \in (2, \infty)$  as an approximation for the noise distribution in Regime II. However, the boundaries of these regions are imprecise, and we aim to more rigorously define them in this section. We first define  $\xi_1$  and  $\xi_2$  as the values of the ratio  $\sigma_Z/\Delta$  separating the regimes such that the noise distribution is approximately uniform for  $\sigma_Z/\Delta < \xi_1$ , generalized Gaussian for  $\xi_1 \leq \sigma_Z/\Delta < \xi_2$ , and Gaussian for  $\sigma_Z/\Delta \leq \xi_2$ . In each regime, we have an expression for the expected MSE or asymptotic variance of the ML estimator, so we use the intersection or approximate point of convergence of these expressions to define  $\xi_1$  and  $\xi_2$ .

### 4.5.1 Defining $\xi_1$

We define  $\xi_1$ , marking the transition in noise characteristics from uniform to generalized Gaussian ( $p < \infty$ ), as the value of  $\sigma_Z/\Delta$  where  $\text{NMSE}(\text{mid})$  and  $\text{NVar}(\text{GGML})$  intersect, which from (4.15) and (4.26) is the solution to

$$\beta(\hat{p}_{\hat{v}})[(\sigma_Z/\Delta)^2 + 1/12] = \frac{K/2}{K^2 + 3K + 2} \quad (4.38)$$

for a given  $K$ , recalling that  $\hat{p}_{\hat{v}}$  is also dependent on  $\sigma_Z/\Delta$  as shown in (4.23). Fig. 4-5(a) shows that  $\xi_1$  decreases as  $K$  increases, since the probability of observing at least one “outlier” measurement due to the exponential tails increases with  $K$ , so a lower  $\sigma_Z/\Delta$  value (i.e., with shorter tails) is needed for the midrange estimator to achieve nearly-optimal performance. The figure shows the exact values of  $\xi_1$  computed



**Figure 4-5:** The value of  $\xi_1$  substantially decreases and  $\xi_2$  slowly increases as  $K$  increases, expanding Regime II. (a) A log-log-cubic fit can be used to compute a close approximation to  $\xi_1$  for all  $K$ , while a log-log-linear fit suffices for  $K > 20$ . (b) The square root of a log-quadratic fit closely approximates  $\xi_2$ .

by solving (4.38) as well as a log-log-cubic least-squares fit

$$\log \xi_1 \approx 0.0104(\log K)^3 - 0.1760(\log K)^2 + 0.0274(\log K) - 1.8511, \quad (4.39)$$

which can be used quickly to calculate an approximation for a desired value of  $K$ . Since the relationship appears fairly linear for  $K > 20$ , the simple log-log-linear fit

$$\log \xi_1 \approx -0.9301(\log K) - 0.1963, \quad (4.40)$$

which can be rewritten as  $\xi_1 \approx 0.8217/K^{0.9301}$ , is also useful for quick computation. The natural logarithm is used in each case.

#### 4.5.2 Defining $\xi_2$

Defining a useful value of  $\xi_2$ , marking the transition in noise characteristics from generalized Gaussian ( $2 < p < \infty$ ) to Gaussian ( $p = 2$ ), is more subtle. Since  $\text{NMSE}(\text{mean})$  and  $\text{NVar}(\text{GGML})$  both have  $1/K$  factors, they converge where  $\beta(\hat{p}_{\bar{V}}) = 1$ , which is

only the case for  $\widehat{p}_{\widehat{v}} = 2$ . This suggests that equality requires the noise to be exactly Gaussian, which only occurs for  $\sigma_Z/\Delta \rightarrow \infty$ . Instead, we can look for a point where  $\text{NVar}(\text{GGML})$  and  $\text{NMSE}(\text{mean})$  can reasonably be considered to have converged (i.e., the GG is close enough to a Gaussian).

We propose that a reasonable definition of  $\xi_2$  is the value of  $\sigma_Z/\Delta$  that minimizes  $\text{NMSE}(\text{Q})$ , the expected normalized MSE of  $\widehat{\mu}_{\text{Q}}$ . Intuitively, as  $\sigma_Z/\Delta$  increases from  $\xi_2$ , the Gaussian variance will dominate for both quantized and dithered measurements, so that the effect of the quantization error is negligible, whether signal-independent for dithered measurements or signal-dependent without dither. Thus the point at which  $\text{NMSE}(\text{Q})$  is minimized indicates where the Gaussian variance begins to dominate and is a reasonable place to consider a GG approximation to be sufficiently Gaussian. We derive in Appendix 4.C that  $\text{NMSE}(\text{Q})$  is given as

$$\begin{aligned}
\text{NMSE}(\text{Q}) &= \mathbb{E}[(\widehat{\mu}_{\text{Q}} - \mu_X)^2]/\Delta^2 \\
&= \frac{1}{12} + \frac{1}{K} \int_{-1/2}^{1/2} \sum_{m=-M}^M m^2 \Psi(m, \mu_X) d\mu_X \\
&\quad + \frac{K-1}{K} \int_{-1/2}^{1/2} \left( \sum_{m=-M}^M m \Psi(m, \mu_X) \right)^2 d\mu_X \\
&\quad - 2 \int_{-1/2}^{1/2} \mu_X \sum_{m=-M}^M m \Psi(m, \mu_X) d\mu_X, \tag{4.41}
\end{aligned}$$

where

$$\Psi(m, \mu_X) = \Phi\left(\frac{m + 1/2 - \mu_X}{\sigma_Z/\Delta}\right) - \Phi\left(\frac{m - 1/2 - \mu_X}{\sigma_Z/\Delta}\right). \tag{4.42}$$

Defining

$$\xi_2 = \arg \min_{\sigma_Z/\Delta} \mathbb{E}[(\widehat{\mu}_{\text{Q}} - \mu_X)^2]/\Delta^2 \tag{4.43}$$

and solving via a Nelder-Mead algorithm [136] and numerical integration, we show in

Fig. 4-5(b) that the value of  $\xi_2$  changes only slightly as a function of  $K$ . This range of values is notably very close to the value  $\sigma_Z/\Delta = 1/2$  recommended by Vardeman and Lee [198], or the value  $\sigma_Z/\Delta = 1/3$  at which Moschitta et al. [135] suggest that the loss of information from quantizing samples of a Gaussian distribution becomes negligible in estimation of the mean. For quick computation,  $\xi_2$  can be approximated by the square root of a log-quadratic fit:

$$\xi_2 \approx \sqrt{-0.000756(\log K)^2 + 0.328 \log K}. \quad (4.44)$$

We notice that the Regime boundary definitions are inconsistent for  $K < 3$ , as  $\xi_1 > \xi_2$ ; however, the Regime definitions are meaningless for  $K = 1$  or  $2$  anyway, as symmetric order statistics-based estimators (e.g., mean, median, midrange) are all equivalent for such small numbers of measurements, so there is no advantage to distinguishing between noise distributions. We notice also that since  $\xi_1$  decreases monotonically and  $\xi_2$  increases monotonically with  $K$ , Regime II grows as  $K$  increases, since small mismatches between the assumed and true PDFs become easier to observe. Intuitively,  $\xi_1$  decreases much faster than  $\xi_2$  increases because the difference between a PDF with finite support ( $\sigma_Z/\Delta = 0$ ) and one with infinite support ( $\sigma_Z/\Delta > 0$ ) is more significant for large  $K$  than the difference between finite  $\sigma_Z/\Delta$  (e.g., GG approximation with  $\hat{p}_{\tilde{V}} > 2$  and  $\sigma_Z/\Delta \rightarrow \infty$  corresponding to  $\hat{p}_{\tilde{V}} = 2$ ).

## 4.6 Numerical Results

Monte Carlo simulations were performed to compare the NMSE performance of the generalized Gaussian and order statistics-based estimators ( $\hat{\mu}_{\text{NB}}$ ,  $\hat{\mu}_{\text{NL}}$ ,  $\hat{\mu}_{\alpha}$ ) against the ML estimators ( $\hat{\mu}_{\text{DML}}$ ,  $\hat{\mu}_{\text{GGML}}$ ) and the conventional sample mean ( $\hat{\mu}_{\text{mean}}$ ) and midrange ( $\hat{\mu}_{\text{mid}}$ ). Estimates were also computed applying the sample mean ( $\hat{\mu}_{\text{Q}}$ ) and ML estimator ( $\hat{\mu}_{\text{QML}}$ ) to the quantized data to determine under which condi-

tions subtractive dithering actually provides an advantage. As in the motivating example in Section 4.2, for each Monte Carlo trial,  $\mu_X$  was chosen uniformly at random from  $[-\Delta/2, \Delta/2]$ , and  $K$  samples of signal noise  $Z \sim \mathcal{N}(0, \sigma_Z^2)$  and dither  $D \sim \mathcal{U}(-\Delta/2, \Delta/2)$  were generated for (4.4). The quantization bin size was maintained at  $\Delta = 1$  throughout. The normalized MSE was computed for  $T = 20,000$  trials.

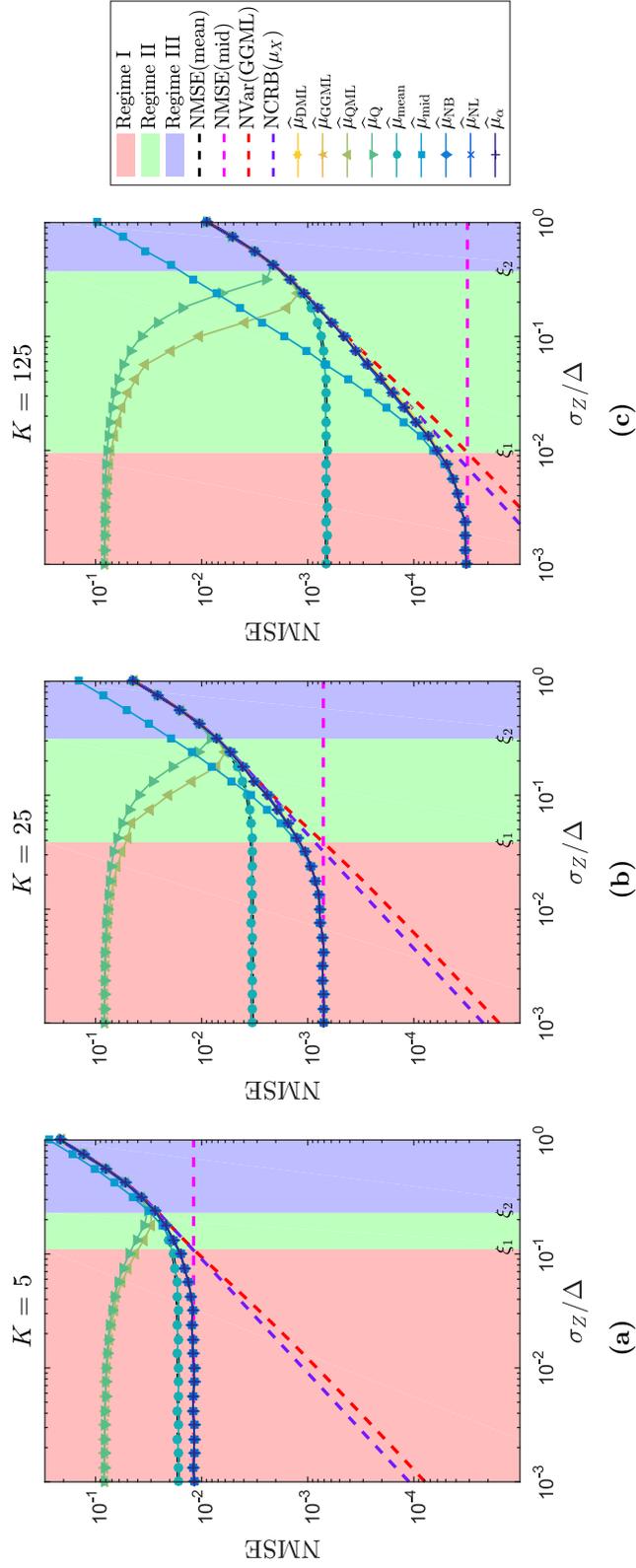
#### 4.6.1 Normalized MSE vs. $\sigma_Z/\Delta$

We begin by discussing the plots in Fig. 4-6 of NMSE as a function of  $\sigma_Z/\Delta$  for  $K = 5, 25, \text{ and } 125$ . The overlapping curves for nine separate estimators and 4 NMSE bounds can make the plots difficult to follow, so a flowchart is included in Fig. 4-7 that summarizes the results and provides a decision-making process for whether to use dither and which estimator to choose.

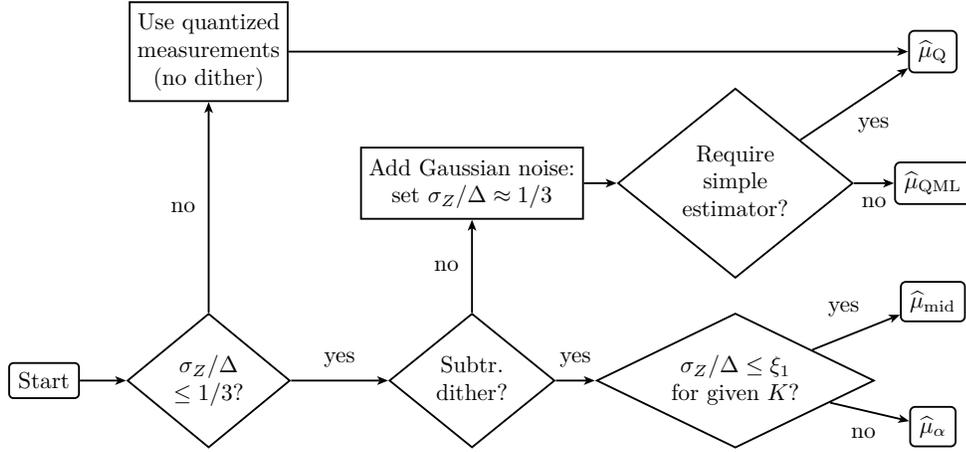
### Generalized Gaussian Estimators

The GG-based estimators ( $\hat{\mu}_{\text{NB}}, \hat{\mu}_{\text{NL}}, \hat{\mu}_\alpha, \hat{\mu}_{\text{GGML}}$ ) have effectively identical performance and match that of  $\hat{\mu}_{\text{DML}}$ . The actual differences in performance vary on the order of a few percent over large ranges of  $K$  and  $\sigma_Z/\Delta$ , compared to the orders of magnitude differences for the mean and midrange. The negligible performance difference further validates approximating the total noise with the GGD. For this same reason, we collectively discuss  $\hat{\mu}_{\text{DML}}$  and the GG-based estimators in the following sections.

The GG-based estimators meet or exceed the performance of all other estimators for all  $\sigma_Z/\Delta$  and for all  $K$ . More specifically, the GG estimators converge to and match the performance of the midrange in Regime I and likewise converge to and match the performance of the mean in Regime III. In Regime II, the GG estimators outperform both the mean and the midrange. Thus, a GG estimator should be the



**Figure 4-6:** The performance of the estimators is evaluated for  $K =$  (a) 5, (b) 25, and (c) 125 to show the range of behavior as  $\sigma_z/\Delta$  varies. The ML estimator for dithered measurements  $\hat{\mu}_{\text{DML}}$  and the estimators based on the GGD ( $\hat{\mu}_{\text{GGML}}$ ,  $\hat{\mu}_{\text{NB}}$ ,  $\hat{\mu}_{\text{NL}}$ , and  $\hat{\mu}_{\alpha}$ ) achieve the lowest NMSE for each  $\sigma_z/\Delta$  regime (curves are overlapping). Results are shown for 20000 Monte Carlo trials.



**Figure 4-7:** The results of our Monte Carlo simulations lead to a simplified decision process for when and how to use dither. If  $\sigma_Z/\Delta > \xi_2$  ( $\approx 1/3$ ), there is no benefit to using anything but  $\hat{\mu}_Q$  applied to the quantized measurements, whereas dither leads to reduced estimation error when  $\sigma_Z/\Delta \leq 1/3$ . If subtractive dithering is not possible, the best performance can be achieved by adding Gaussian noise to set  $\sigma_Z/\Delta \approx 1/3$  and applying  $\hat{\mu}_{QML}$  (although  $\hat{\mu}_Q$  can be used if simplicity is required). However, larger performance improvements can be achieved with a subtractively-dithered quantizer. For  $K$  subtractively-dithered measurements, compute  $\xi_1$  from either (4.39) or (4.40) to determine whether to use  $\hat{\mu}_{mid}$  (in Regime I) or  $\hat{\mu}_\alpha$  (in Regime II).

default estimator choice for any  $\sigma_Z/\Delta$ .

Given the approximate equivalence of the GG estimators, the trimmed-mean  $\hat{\mu}_\alpha$  is the best choice of general-purpose estimator for dithered data. The other estimators either require iterative solvers ( $\hat{\mu}_{DML}$ ,  $\hat{\mu}_{GGML}$ ), rely on numerical computation for the GG inverse CDF ( $\hat{\mu}_{NB}$ ), or are data-dependent ( $\hat{\mu}_{NL}$ ). On the other hand,  $\hat{\mu}_\alpha$  has a simple closed-form solution that can be tabulated if needed.

### Performance by Regime—Dithered Measurements

The plots in Fig. 4-6 validate the concept of three distinct regimes of noise behavior. In the plots, the approximate regime boundaries are computed to be  $\xi_1 = \{0.1098, 3.85 \times$

$10^{-2}, 9.56 \times 10^{-3}$  and  $\xi_2 = \{0.2296, 0.3132, 0.3737\}$  for  $K = \{5, 25, 125\}$ , respectively, confirming that Regime II expands as  $K$  increases. In Regime I, the NMSE performance of all estimators on the dithered data is basically flat and equal to NMSE(mid). This suggests that for a practical system where  $\sigma_Z/\Delta$  can be tuned, once the system is operating in Regime I (dependent on a fixed  $K$ ), there is no benefit from further decreasing  $\sigma_Z/\Delta$ ; performance can only be improved by increasing  $K$ . In Regime II, the GG-based estimators approach NCRB( $\mu_X$ ), especially for large  $K$ . We note that while NVar(GGML) and NCRB( $\mu_X$ ) are close in Regime II, NCRB( $\mu_X$ ) is a tighter bound, as it is based on the true noise distribution, although NVar(GGML) may be easier to compute for a rough estimate of performance. In Regime III, the NMSE performance of all estimators on the dithered data is equal to NMSE(mean). In both Regimes II and III, the NMSE decreases as  $\sigma_Z/\Delta$  decreases. Performance likewise improves with increasing  $K$ .

### Performance by Regime—Quantized Measurements

While the three Regimes were technically defined for dithered measurements in particular, they are also informative of the behavior of estimators applied to quantized measurements. In Regime I,  $\sigma_Z/\Delta$  is so small that, unless  $\mu_X$  lies on the boundary between quantization bins, all measurements are quantized to the same value. As a result, the NMSE of both  $\hat{\mu}_Q$  and  $\hat{\mu}_{QML}$  is dominated by the squared bias term, which is  $1/12$  (the variance is zero). Further decreasing  $\sigma_Z/\Delta$  or increasing  $K$  provides no benefit.

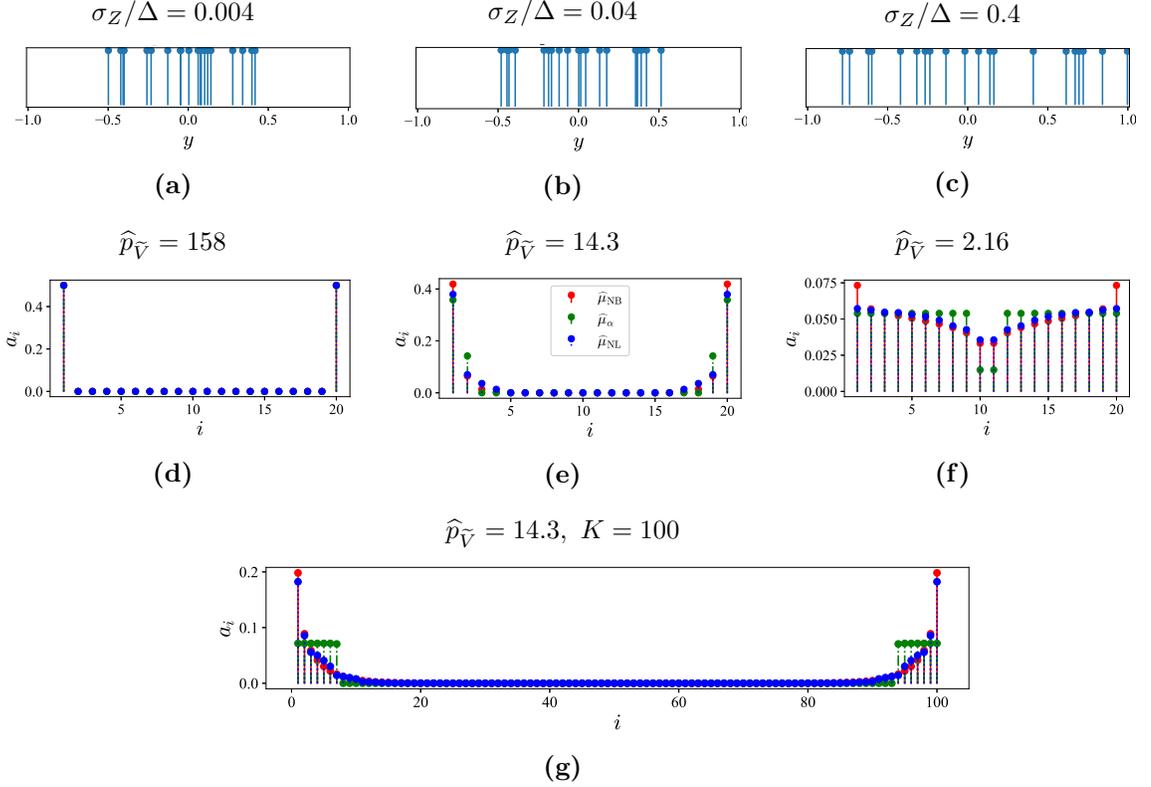
In Regime II,  $\sigma_Z/\Delta$  is large enough that there is often some variation in the measurements due to signal even without the addition of dither. This phenomenon is sometimes referred to in the literature as self-dithering, equivalent to adding nonsubtractive Gaussian dither to a constant signal  $\mu_X$  [36]. Within Regime II, both  $\hat{\mu}_Q$  and  $\hat{\mu}_{QML}$  improve as  $\sigma_Z/\Delta$  increases because the increased signal variation reduces

the bias term of the NMSE faster than the variance increases. The NMSE is minimized for  $\hat{\mu}_Q$  by definition at  $\xi_2$ , and then the NMSE increases as  $\sigma_Z/\Delta$  increases in Regime III. This suggests that if  $\sigma_Z/\Delta$  is small and subtractive dither cannot be used, then quantized measurements benefit from adding nonsubtractive Gaussian dither such that  $\sigma_Z/\Delta = \xi_2$ , which is approximately 1/3. It is in Regime II that  $\hat{\mu}_{QML}$  shows the largest improvement in performance over  $\hat{\mu}_Q$ , with the ML estimator accounting for the form of the signal variation for quantized measurements.

In Regime III, the NMSE of  $\hat{\mu}_Q$  and  $\hat{\mu}_{QML}$  matches that of the best estimators applied to dithered data. Clearly,  $\sigma_Z/\Delta$  is large enough that even the quantized measurements contain sufficient information about the signal variation. This suggests that dither provides no benefit in Regime III, since equal performance can be achieved without dither. Again for both Regimes II and III, the NMSE decreases as  $K$  is increased.

#### 4.6.2 Order Statistics-Based Estimator Coefficients

To better understand why the order statistics-based estimators have essentially identical performance, in Fig. 4-8 we plot the coefficients  $a_i$  from (4.28) for each estimator. The top row shows example measurements for  $K = 20$ ,  $\Delta = 1$ , and  $\sigma_Z/\Delta = 0.004$ , 0.04, and 0.4, respectively, with the samples spreading out as the Gaussian variance increases. The second row of plots depicts the resulting coefficients for  $\hat{\mu}_{NB}$ ,  $\hat{\mu}_{NL}$ , and  $\hat{\mu}_\alpha$  using the estimated value  $\hat{p}_{\hat{V}}$ . Fig. 4-8(d) shows the coefficients are equivalent to those of  $\hat{\mu}_{mid}$  for small  $\sigma_Z/\Delta$ . In Figs. 4-8(e) and 4-8(f), the coefficients of the various estimators are no longer identical. However, the coefficients follow the same trends for each estimator, with zero weight on the middle order statistics for small  $\sigma_Z/\Delta$  and more evenly-distributed weights as the noise model approaches a Gaussian. We note that the coefficients for  $\hat{\mu}_{NL}$  vary depending on the particular set of measurements shown in the top row, and that different sample realizations can result in coefficients



**Figure 4-8:** Example dithered measurements are shown in (a-c) for  $\sigma_Z/\Delta = 0.004, 0.04,$  and  $0.4$  with  $K = 20$ . Plots (d-f) show the resulting coefficient values for the GG estimators given the estimated value of  $\hat{p}_{\hat{V}}$  above. In (g),  $\sigma_Z/\Delta = 0.04$  and  $K = 100$ , highlighting how the coefficients change as  $K$  increases. Note that the coefficients of the order statistics for the NL estimator depend on the measured data sequence shown above.

more or less similar to those of  $\hat{\mu}_{\text{NB}}$  and  $\hat{\mu}_{\alpha}$ . To show the behavior of the coefficients as  $K$  increases, we also plot  $\{a_i\}_{i=1}^K$  for  $K = 100$  and  $\sigma_Z/\Delta = 0.04$  in Fig. 4-8(g). This plot underscores that the coefficients for  $\hat{\mu}_{\alpha}$  are basically indicators of the most significant non-zero coefficients of  $\hat{\mu}_{\text{NB}}$  and  $\hat{\mu}_{\text{NL}}$ . Using the simple formulation of  $\hat{\mu}_{\alpha}$  as a guide, in the limit as  $K \rightarrow \infty$ , only  $K\alpha$  coefficients would have nonzero weight. Since the performance of all three order statistics-based estimators is similar, this further suggests that the selection of which order statistics are used is more important than exactly how much they are weighted.

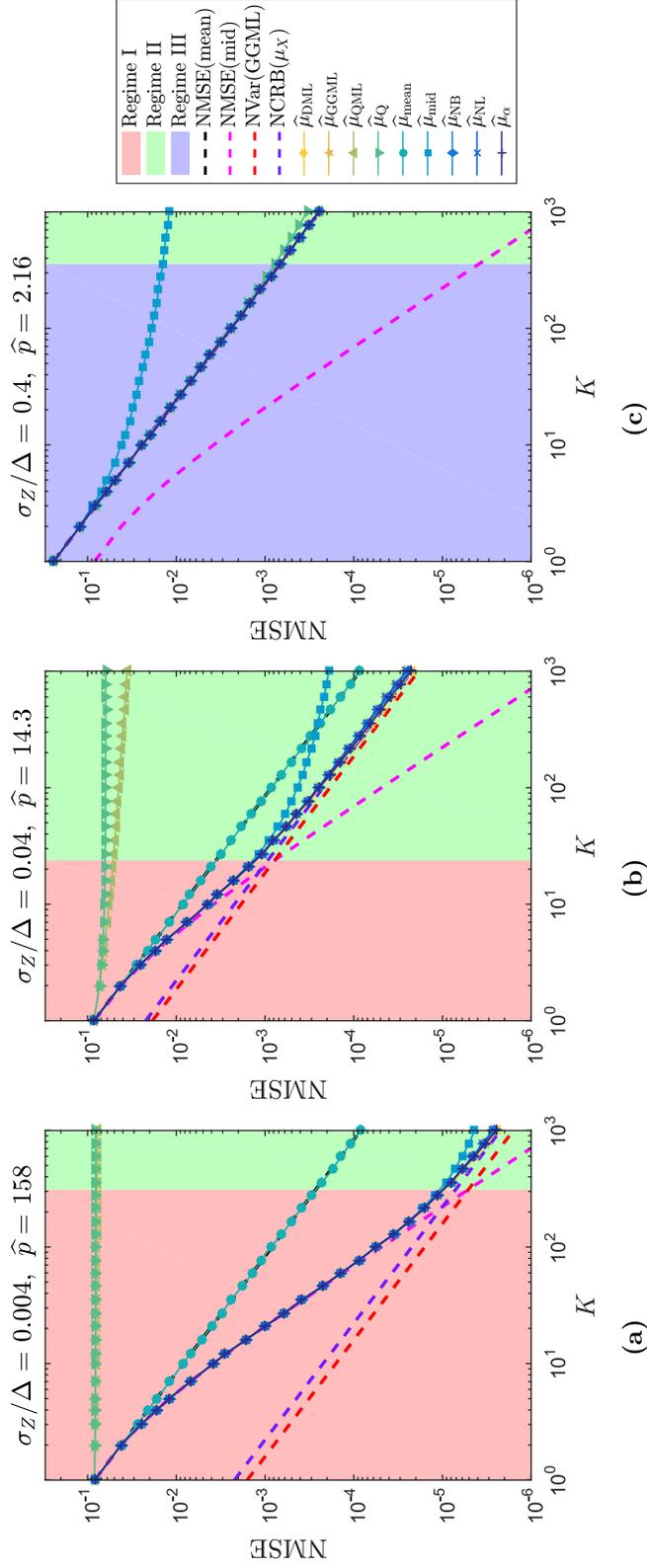
### 4.6.3 Normalized MSE vs. $K$

To better understand how the number of measurements affects the estimators' performance, results are plotted for three fixed values of  $\sigma_Z/\Delta$  in each regime (0.004, 0.04, 0.4) while varying  $K$  in Fig. 4-9.

In Fig. 4-9(a),  $\hat{\mu}_{\text{mid}}$  follows NMSE(mid) as expected for Regime I until  $K \approx 200$ . At that point, the NMSE of the midrange begins to diverge, with slower improvement as  $K$  increases. Similarly, the GG estimators follow NMSE(mid) until  $K \approx 200$  and then switch to NCRB( $\mu_X$ ). This suggests that  $\sigma_Z/\Delta = 0.004$  is in Regime I for  $K < 200$  and in Regime II for  $K > 200$ . This switch between regimes occurs near the intersection of NMSE(mid) and NVar(GGML) as a function of  $K$ , further validating these bounds as useful demarcations of estimator performance. For all  $K$  in the plotted range, the midrange and GG estimators outperform the mean. The quantized estimators show almost no improvement as  $K$  increases.

In Fig. 4-9(b), the midrange performance is similar to that in Fig. 4-9(a), with  $\hat{\mu}_{\text{mid}}$  following NMSE(mid) until the intersection of NMSE(mid) and NVar(GGML) and then improving more slowly as a function of the number of measurements, eventually being outperformed by  $\hat{\mu}_{\text{mean}}$  for large  $K$ . The GG estimators likewise follow NMSE(mid) for small  $K$  and switch to following NCRB( $\mu_X$ ) after the intersection. For large  $K$ , the NMSE of the GG estimators is a constant factor lower than that of  $\hat{\mu}_{\text{mean}}$ , with this factor approximately given by  $\beta(\hat{p}_{\hat{v}})$ . The NMSE of the quantized estimators decreases slowly as  $K$  increases, with marginally better performance for  $\hat{\mu}_{\text{QML}}$  than  $\hat{\mu}_{\text{Q}}$ .

Figures 4-9(a) and 4-9(b) help answer the question of how the order statistics-based estimators “between” the midrange and the mean would perform as a function of  $K$ . The results suggest that these estimators ultimately have  $O(K^{-1})$  NMSE reduction, although this reduction is faster for small values of  $K$ .



**Figure 4-9:** The performance of the order statistics estimators is evaluated for  $\sigma_z/\Delta =$  (a) 0.004, (b) 0.04, and (c) 0.4 to show the full range of behavior as the number of measurements  $K$  increases. The plots show the MSE normalized to  $\Delta = 1$  from 20000 trials per data point. Dashed lines show the theoretical NMSE of the mean and midrange, and the asymptotic variance of the ML estimator of the GGD mean. The ML estimator for dithered measurements  $\hat{\mu}_{\text{DML}}$  and the estimators based on the GGD ( $\hat{\mu}_{\text{GGML}}$ ,  $\hat{\mu}_{\text{NB}}$ ,  $\hat{\mu}_{\text{NL}}$ , and  $\hat{\mu}_{\alpha}$ ) achieve the lowest NMSE for all  $K$  (curves are overlapping).

In Fig. 4.9(c) where  $\sigma_Z/\Delta = 0.4$ , the noise can be sufficiently described as Gaussian for  $K < 359$  because  $\sigma_Z/\Delta > \xi_2$  as shown in Fig. 4.5(b); however,  $\sigma_Z/\Delta < \xi_2$  for  $K \geq 359$ , so the noise distribution transitions from Regime III to Regime II. The midrange has poor performance for all  $K$ , while the other dithered estimators and the quantized estimators have essentially identical performance for  $K < 359$ . Those estimators follow  $\text{NMSE}(\text{mean})$ ,  $\text{NVar}(\text{GGML})$ , and  $\text{NCRB}(\mu_X)$ , which have converged. In Regime III, it is clear that there is no benefit to using dither, as there is minimal improvement in performance even for large  $K$ . In fact, implementing a dithered quantizer is likely more complicated in practice and is discouraged for Regime III.

## 4.7 Subtractively-Dithered Single-Photon Lidar

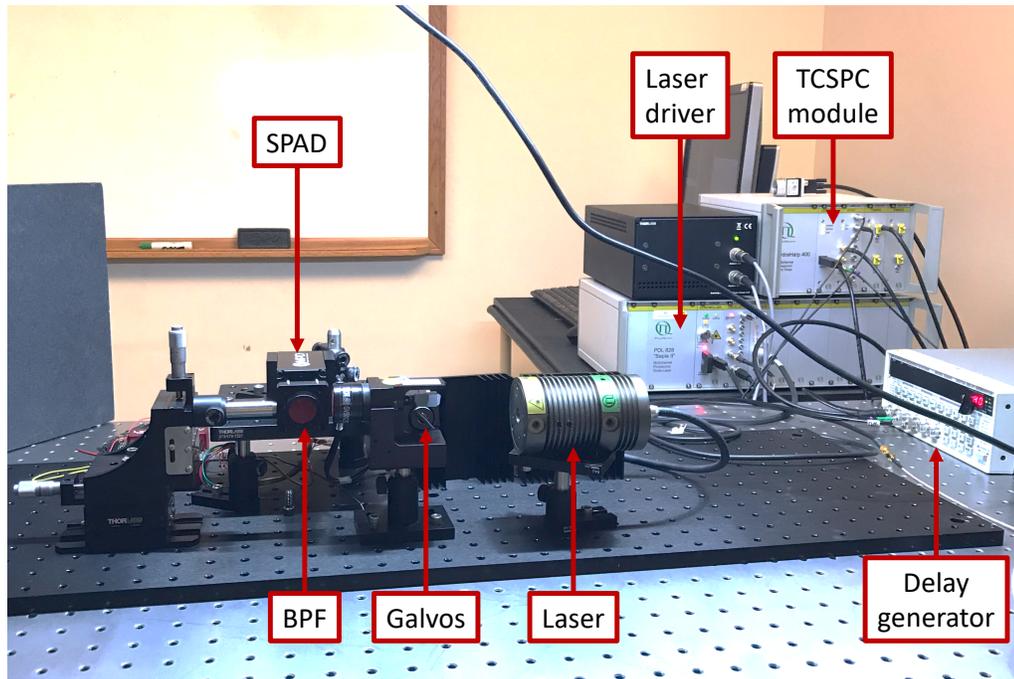
Assuming negligible background, a Gaussian pulse shape, and low-flux acquisition, photons arrival times in SPL have a Gaussian distribution as in (4.1), where  $c\mu_X/2$  corresponds to the true depth. Unfortunately, photon detection times are quantized with bin resolution  $\Delta$ , which may be coarse relative to the pulse width  $\sigma_Z$ , limiting the precision of depth estimates. We have demonstrated that estimation from quantized measurements significantly benefits from the use of both dither in general and generalized Gaussian modeling in particular. We now aim to verify that the improvements in estimation also transfer to actual experimental SPL data. Since a TCSPC system measures time, a dithered system amounts to the insertion of delays in the signal path before quantization occurs. We discuss how these delays can be introduced by controlling the system timing via a digital delay generator. We then show depth imaging results from a dithered lidar system using the generalized Gaussian analysis. Finally, we modify that analysis to better account for the empirical distribution of detection times due to the detector hardware.

### 4.7.1 Architecture and Dither Implementation

Typical TCSPC systems use a control signal from the laser driver to synchronize the illumination with the start of a timer. The SPAD provides the stop signal to the timer when a photon is detected. Dither can be added into the measurement by inserting variable delays into the synchronization signal. Since the dither value must be known so it can be subtracted from the quantized measurement, the delays are most practically implemented as discrete random variables. While discrete dither does not exactly meet the requirements for the quantization error to be independent of the signal, by using a large set of delay values, the behavior of the discrete dither approaches that for a continuous uniform random variable [204].

Because we do not have access to a SPAD camera, we use the system shown in Figure 4-10 to perform pseudoarray imaging by raster-scanning a laser spot over the scene and detecting with a single-pixel SPAD. The laser driver (PicoQuant PDL 828 “Sepia II”) is connected to a 640-nm diode laser (PicoQuant LDH P-C-640B) which has a pulse profile with  $\sigma_Z = 300$  ps. The SPAD (Micro Photon Devices PDM-series) is connected to the TCSPC module (PicoQuant HydraHarp 400). A narrowband bandpass filter centered at the laser wavelength is placed in front of the SPAD to reduce the amount of ambient light incident on the sensor. An electronic signal synchronizing the timing of the illumination and detection is transmitted via a cable connecting the sync output from the laser driver to the sync input of the HydraHarp. A digital acquisition (DAQ) board (a LabJack U3 with an LJTick-DAC card) controls the 2-axis galvo (ThorLabs GVS012) and connects to a Marker input of the HydraHarp to simultaneously indicate when the galvo position changes.

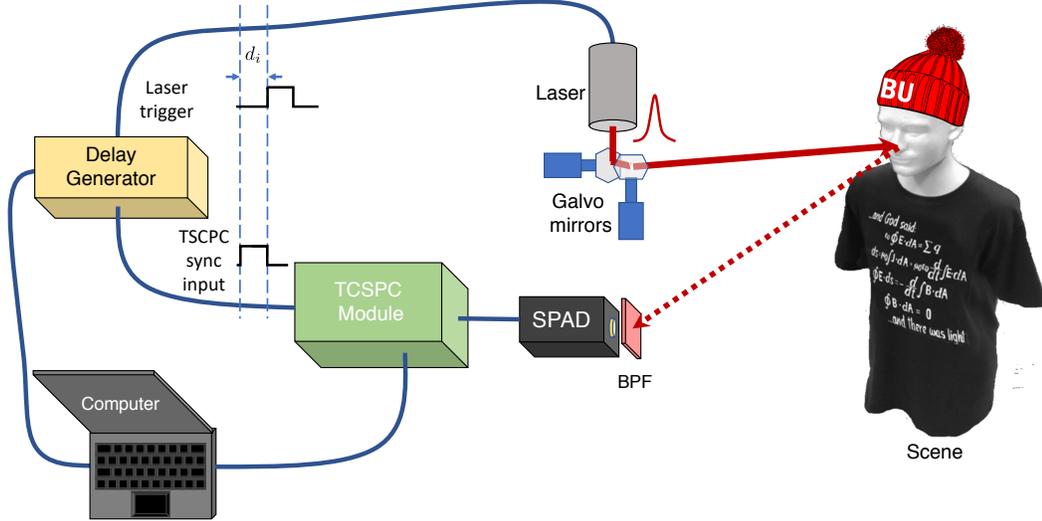
To incorporate dither into the experiment, instead of using the internal clock of the laser driver, we externally control the experiment via a digital delay generator (DDG). As shown in Figure 4-11, we connect one channel of the DDG (T0) to the



**Figure 4.10:** A photograph of the bistatic, subtractively-dithered single-photon lidar implementation at Draper in Cambridge, MA. The only difference from the conventional setup is the external timing control of a digital delay generator, which inserts 10-picosecond-resolution delays between the laser and SPAD triggers to implement dither.

external trigger input of the laser driver, and another DDG channel (AB) connects to the sync input of the HydraHarp. A third channel could be used to control the SPAD gate and further reduce ambient light detection if some depth information is known a priori. Normal lidar operation without dither can proceed by setting no delay between T0 and AB. However, by changing the relative delay between T0 and AB, the perceived time of flight of the illumination can be changed in a predictable and repeatable manner. The DG645 is remotely controlled via a GPIB connection. In order to keep track of the dither step for each detection, an independent DAQ channel is connected to a second Marker input of the HydraHarp.

The experiment is controlled by a MATLAB script, with dithered acquisition



**Figure 4-11:** A diagram showing the conventional SPL setup modified for a subtractively-dithered implementation. Varying delays  $d_i$  are inserted by the digital delay generator between the laser and TCSPC triggers to implement dither. The amount of delay is controlled by the computer, which allows the delay to be subtracted from the photon detection times after acquisition.

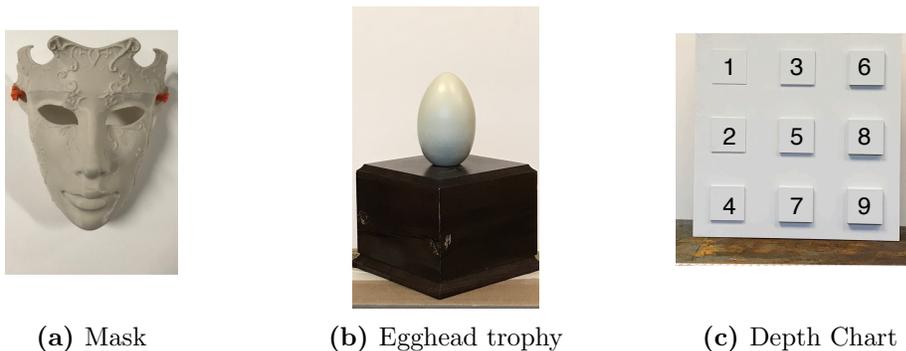
proceeding as follows. A grid of laser points is chosen to cover the desired portion of the scene. A set of  $n_{dv}$  dither delay values is chosen to evenly subdivide  $[0, \Delta)$ . For each point in the grid, MATLAB sets the galvo position and resets the channel delay to zero, recording each change with a marker at the HydraHarp. The laser is pulsed, a MATLAB command increases the delay in channel AB and sends a corresponding marker to the HydraHarp. This process is repeated until all  $n_{dv}$  dither delays have been stepped through in sequence. Then the galvo position is moved, the delay is reset, and the process starts again until all grid points have been addressed. When acquisition is complete, the detection data is sorted into a 3D cell array, with the laser grid position as the first two dimensions and the delay step as the third dimension. Each cell in the array contains a random number of photon detection times.

Since the illumination source has a relatively large  $\sigma_Z = 300$  ps, we demonstrate

the effectiveness of dither by choosing a larger quantization bin size of  $\Delta = 2048$  ps, resulting in  $\sigma_Z/\Delta \approx 0.146$  and  $\hat{p}_{\hat{v}} \approx 3.69$ . (The same  $\sigma_Z/\Delta$  ratio could be achieved for a SPAD array with 350-ps resolution [122] by designing the overall system jitter and laser pulse width to have  $\sigma_Z \approx 51$  ps.) The dither is implemented with a DG645 (Stanford Research Systems) digital delay generator with discrete delay steps  $d_i \in \{0, \dots, 204\} \times 10$  ps to evenly subdivide the quantization bins. The delays are stepped through in sequence at each pixel, and a Poisson number of photons is detected for each delay.

#### 4.7.2 Related Work

Similar ideas using shifted acquisition windows to acquire range information have been used in the past. Busck and Heiselberg [32] implemented a “laser-gated viewing” (LGV) system based off early concepts from Gillespie [69], using a high-speed charge-coupled device (CCD) array with no inherent depth resolution to form depth images. In their work, the scene was repeatedly illuminated by a pulsed laser, and a gate exposed the camera to light for a short duration on the order of a few hundred picoseconds. The light was integrated on the CCD over many illuminations, and once the result was stored, a short delay was added to the gate activation time, and the acquisition process was repeated with the new gate delay. By stepping through a sequence of delays, slices of the transient light response were collected from the scene, and a depth estimate was created by weighting the time of each slice by the grayscale intensity value of the image acquired for that delay. Such an LGV system was used by Christnacher et al. [41] to improve vision through scattering media such as smoke or fog. The same principle was recently used with a Quantum Image Sensor (QIS) to also take advantage of single-photon sensitivity [155]. Crucially, this LGV approach uses detectors with no inherent time resolution. Since light is only detected within the short gate time, all of the light arriving outside this window is discarded, which



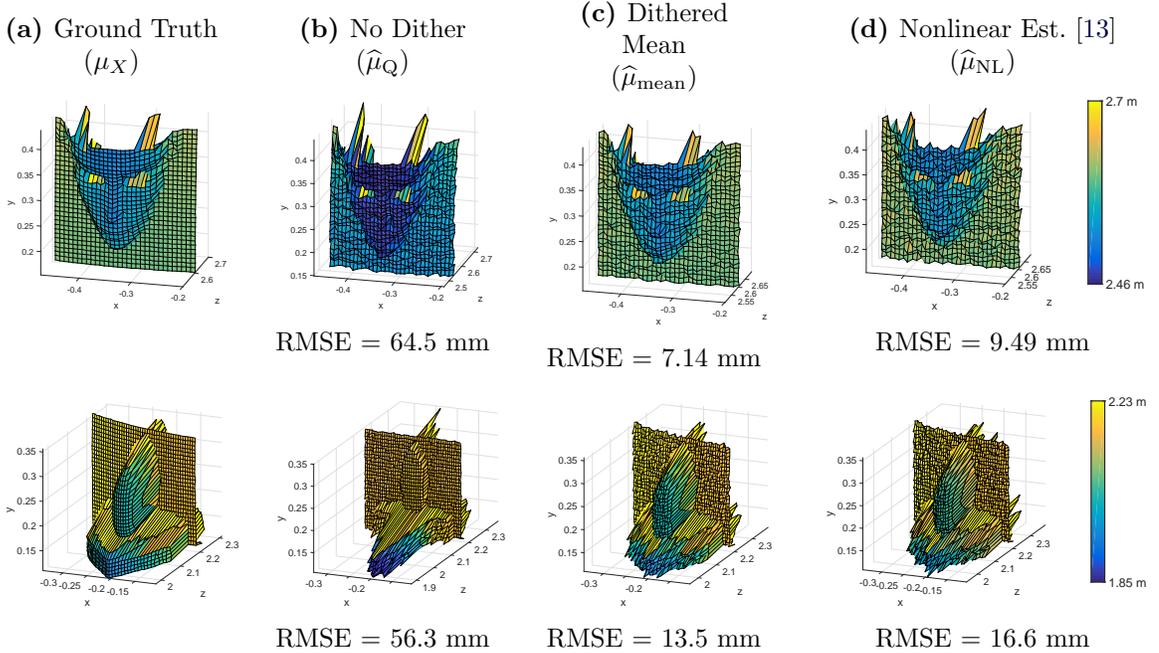
**Figure 4-12:** Photographs showing the 3D-printed mask, “egghead” trophy, and 3D-printed depth resolution chart used for lidar experiments. Note for the depth chart that the heights in millimeters are displayed over each block but are not physically present in the scene.

is an inefficient use of the reflected light and makes the process slow to acquire a full sequence of time slices.

Chen et al. [39] recently proposed a method of dithered SPL using two optical paths with different lengths, such that the longer path length effectively delays laser transmission by a half-bin duration. In simulations, they showed that a simple averaging of signals from different bin shifts reduced the effect of quantization on the depth estimation error. While we share the same basic idea of combining measurements from different sub-bin delays, our electronic delay implementation is far more flexible in terms of the bin sizes and numbers of delays that can be achieved in practice. After the introduction of this work, Raghuram et al. [146] used a similar electronic delay implementation, but their sampling theory approach to processing the measurements does not necessarily assume a single depth per pixel, thus requiring longer acquisition times to capture an entire histogram for each delay.

### 4.7.3 Experimental Results for Gaussian Pulse Modeling

Experimental acquisitions using the dithered SPL setup were performed for the test objects shown in Figure 4-12. Results are shown in Fig. 4-13 for a  $32 \times 32$ -pixel scan of



**Figure 4-13:** Depth estimates are presented for the 3D-printed mask and egg-shaped trophy. The ground truth depth in (a) is poorly estimated with coarse quantization (b). Implementing subtractively-dithered measurements significantly improves the depth estimate, although the generalized Gaussian-based estimator (d) does not outperform the sample mean (c) for the experimental results.

the 3D-printed mask and a  $40 \times 30$ -pixel scan of the egg-shaped trophy. The ground truth depth estimate in Fig. 4-13(a) is formed by averaging thousands of photon detections per pixel measured with 4-ps resolution in order to form a baseline for quantitatively evaluating the results. Coarsely-quantized detection times are acquired for 8.5 ms per pixel for the mask and 10.2 ms per pixel for the egg, yielding averages of 267 and 362 photons per pixel, respectively. The resulting pixelwise depth estimates in Fig. 4-13(b) are strongly biased by the quantization, and the RMS error is on the order of the test objects' large-scale features. Dithered data is acquired by detecting a random number of photons in  $41.5 \mu\text{s}$  (egghead:  $49.8 \mu\text{s}$ ) for each of the 205 dither steps, also yielding 267 (egghead: 362) photons per pixel on average. The sample mean result  $\hat{\mu}_{\text{mean}}$  in Fig. 4-13(c) shows a dramatic improvement over the result

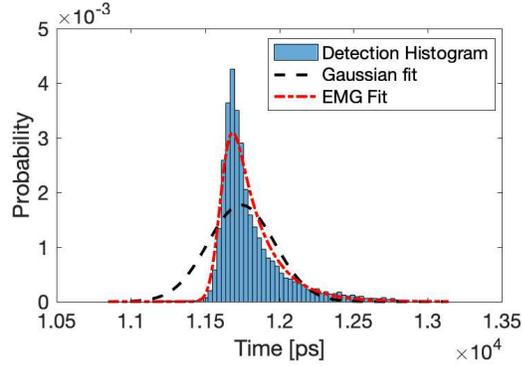
without dither, with the RMS error reduced by a factor of 9.03 for the mask and 4.17 for the egg. The dithering results visually confirm the quantitative improvement, better capturing the shape of the nose and position of the background for the mask, and the right angle of the pedestal and the curvature of the egg than the results without dither. Unfortunately, the benefit of using the GG-based estimator that was observed for simulated measurements does not extend to the experimental data. The RMSE values from applying  $\hat{\mu}_{\text{NL}}$  to both the mask and egg scenes are slightly worse than those from applying the sample mean.

### SPAD Response Function

The worse performance of the GG-based estimator on experimental data is likely due to unmodeled effects, including ambient light as well as the non-Gaussian and asymmetric detection time density shown in Fig. 4-14. In a SPAD, carriers generated as intended in the junction depletion layer are accelerated by the bias voltage, and fluctuations in the buildup of the resulting avalanche do cause a Gaussian response [48]. However, if a photon reaches the neutral region of the semiconductor before generating a carrier, diffusion of the carrier to the depletion layer causes an exponential delay in the avalanche timing [15, 48]. Thus, while other single-photon detectors such as superconducting nanowire single-photon detectors (SNSPDs) tend to have instrument response functions (IRFs) well described by a Gaussian distribution [178], SPADs additionally have a long exponential tail [88].

#### 4.7.4 Exponentially-Modified Gaussian Pulse Modeling

Since the experimental results suggest the GG-based estimator is not robust to this model mismatch, we here propose a modification to the pulse model that accounts for the exponential decay. The sum of Gaussian and exponential random variables, suggested by Hernandez et al. [88] as a good approximation to the SPAD detection time



**Figure 4-14:** Whereas the Gaussian distribution is a poor approximation to the true pulse shape, the exponentially-modified Gaussian (EMG) distribution captures the exponential tail behavior much more accurately.

distribution, has been long been used to describe the asymmetric timing responses in chromatography [72], for which the distribution became known as the exponentially-modified Gaussian (EMG) [76]. The same distribution was also found to be a good fit for reaction times in psychological studies, for which it was called the ex-Gaussian distribution [53].

Let  $H$  be a random variable representing the exponential component of the SPAD response, with PDF

$$f_H(h) = \frac{1}{\tau_H} \exp(-h/\tau_H) = \eta_H \exp(-\eta_H h), \quad (4.45)$$

where  $\eta_H = 1/\tau_H$  with expected value  $E[H] = \tau_H$  and variance  $\sigma_H^2 = \tau_H^2$ . Then the PDF of the total SPAD response time  $R = Z + H$  is the exponentially-modified Gaussian, with PDF given by

$$f_R(r) = f_Z(z) * f_H(h) = \int_{-\infty}^{\infty} f_Z(r-z) f_H(z) dz \quad (4.46)$$

$$= \frac{\eta_H}{\sigma_Z \sqrt{2\pi}} \int_0^{\infty} \exp\left[-\frac{(r-z)^2}{2\sigma_Z^2}\right] \exp(-\eta_H z) dz. \quad (4.47)$$

Through expansion of the integrand and a change of variables, we have

$$\begin{aligned} f_R(r; \mu, \sigma_Z, \eta_H) &= \frac{\eta_H}{2} \exp\left\{-\eta_H \left((r - \mu) - \frac{\sigma_Z^2 \eta_H}{2}\right)\right\} \operatorname{erfc}\left(-\frac{((r - \mu) - \sigma_Z^2 \eta_H)}{\sqrt{2}\sigma_Z}\right) \\ &= \frac{\eta_H}{2} \exp\left\{\eta_H \left(\mu + \frac{\sigma_Z^2 \eta_H}{2} - r\right)\right\} \operatorname{erfc}\left(\frac{\mu + \sigma_Z^2 \eta_H - r}{\sqrt{2}\sigma_Z}\right), \end{aligned} \quad (4.48)$$

where  $\operatorname{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty \exp(-t^2) dt$  is the complementary error function. The CDF of the EMG is not derived here but can be found from the PDF using integration by parts and is given as [134]

$$\begin{aligned} F_R(r) &= \frac{1}{2} \operatorname{erfc}\left(-\frac{r - \mu}{\sqrt{2}\sigma_Z}\right) \\ &\quad - \frac{1}{2} \exp\{-\eta_H [(r - \mu) - \sigma_Z^2 \eta_H]\} \operatorname{erfc}\left[-\frac{(r - \mu) - \sigma_Z^2 \eta_H}{\sqrt{2}\sigma_Z}\right] \\ &= \frac{1}{2} \operatorname{erfc}\left(\frac{\mu - r}{\sqrt{2}\sigma_Z}\right) - \frac{1}{2} \exp\{\eta_H (\mu + \sigma_Z^2 \eta_H - r)\} \operatorname{erfc}\left[\frac{\mu + \sigma_Z^2 \eta_H - r}{\sqrt{2}\sigma_Z}\right]. \end{aligned} \quad (4.49)$$

Figure 4-14 shows that the EMG is a much better fit to the SPL pulse profile than the Gaussian assumption. The maximum likelihood EMG parameters were found using a bounded simplex implementation [218] of the algorithm described in [111]. The parameters were determined to be  $\sigma_Z = 58.4$  ps and  $\tau_H = 191.4$  ps.

#### 4.7.5 Subtractively-Dithered EMG

Now that we have a more accurate model for the photon detection time PDF based on the EMG, we can derive the distribution of photon times in a subtractively-dithered SPL system. We first recall the effect of dither on a generic distribution. Let  $X$  be some random variable with PDF  $f_X(x)$  and CDF  $F_X(x)$ . Let  $W$  be a uniform random variable as defined in (4.16). Then for  $X$  and  $W$  independent, the sum defined as

$Y = X + W$  is a random variable with PDF

$$\begin{aligned}
 f_Y(y) &= f_X(x) * f_W(w) = \int_{-\infty}^{\infty} f_X(x) f_W(y-x) dx \\
 &= \frac{1}{\Delta} \int_{y-\frac{\Delta}{2}}^{y+\frac{\Delta}{2}} f_X(x) dx \\
 &= \frac{1}{\Delta} \left[ F_X\left(y + \frac{\Delta}{2}\right) - F_X\left(y - \frac{\Delta}{2}\right) \right], \tag{4.50}
 \end{aligned}$$

and the CDF is given by

$$F_Y(y) = \int_{-\infty}^y \frac{1}{\Delta} \left[ F_X\left(u + \frac{\Delta}{2}\right) - F_X\left(u - \frac{\Delta}{2}\right) \right] du. \tag{4.51}$$

The distribution of detection times when applying subtractive dither to a pulse with exponentially-modified Gaussian distribution then substitutes the EMG CDF  $F_R(r)$  from (4.49) for the generic CDF  $F_X(x)$  in (4.50) and (4.51).

#### 4.7.6 Dithered EMG Mean Estimation

One of the main challenges in transferring the approach used in the beginning of this chapter to experimental single-photon lidar data is that the Gaussian model of the pulse shape does not adequately capture the asymmetry of the actual SPAD response. Symmetry has often been used to greatly simplify the form of estimators, as the mean, median, midrange, and other symmetric linear combinations of order statistics are all co-located [50]. However, it is still valuable to make estimates of location when that is not the case, although the parameter being estimated must more clearly be identified and the estimation done more carefully [22].

One approach to getting an unbiased estimator that is sensitive to the distribution is to use the BLUE of the order statistics as suggested by Lloyd [120]. Unfortunately,

this requires the ability to compute the expected values and covariance matrix of order statistics, which changes for every change in the number of samples  $K$  and can be complicated depending on the parent distribution. A simplification suggested by Gupta [79] is to assume the covariance matrix is the identity. The location estimator then reweights the coefficients of the order statistics so that the estimate is unbiased. However, this still requires computation of the expected value of each order statistic, which for the dithered EMG requires repeated numerical integration using the substitution of (4.49) into (4.51), which is so unwieldy that we have omitted writing out the full expression. While it would be possible to precompute and tabulate reweighted coefficients for every reasonable number of detections and for fixed system parameters, we find such computation prohibitively expensive for negligible change in the estimator, so we continue without exactly-unbiased estimators here.

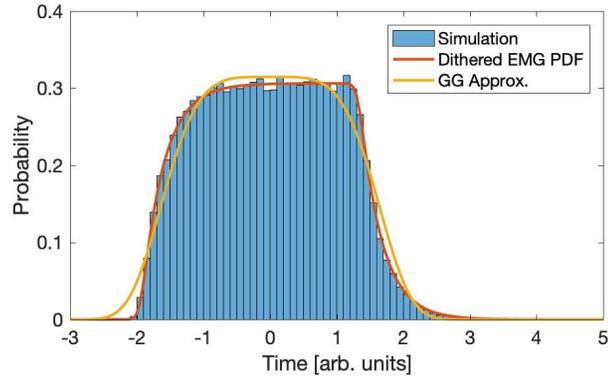
Rather than choosing order statistics and their coefficients to exactly match the dithered detection distribution, we choose to ignore the effect of asymmetry and only modify the GG-based approximation to account for the longer tails incurred from the exponential component. First, we note that the dithered EMG model changes the expected value of the measurements to  $E[Y] = \mu_X + 1/\eta_H$  and the variance to  $\text{Var}[Y] = \sigma_Z^2 + 1/\eta_H^2 + \Delta^2/12$ . Furthermore, the kurtosis of the true distribution changes:

$$\gamma(Y) = \frac{\sigma_Z^4 \gamma(Z) + \sigma_H^4 \gamma(H) + \sigma_W^4 \gamma(W)}{\sigma_Y^4} \quad (4.52)$$

$$= \frac{6\sigma_H^4 - \frac{6}{5}\sigma_W^4}{(\sigma_Z^2 + \sigma_H^2 + \sigma_W^2)^2}, \quad (4.53)$$

where  $\gamma(\cdot)$  refers to the excess kurtosis. Thus, the process of matching the GG kurtosis as described in Appendix 4.B can be updated to solving for  $p_v$  where

$$\frac{\Gamma(1/p_v)\Gamma(5/p_v)}{[\Gamma(3/p_v)]^2} = 3 + \frac{6[1/\eta_H^4 - (\Delta^2/12)^2/5]}{(\sigma_Z^2 + 1/\eta_H^2 + \Delta^2/12)^2}. \quad (4.54)$$

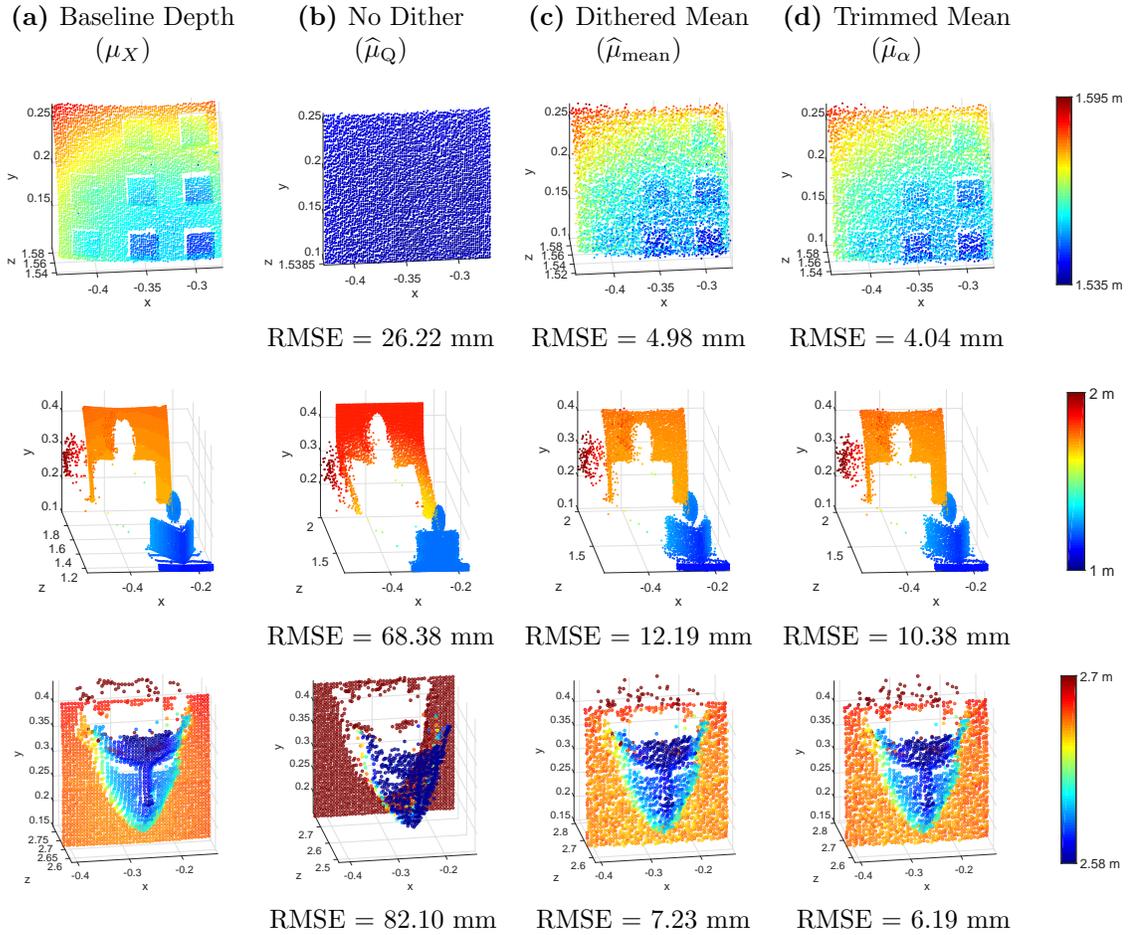


**Figure 4-15:** A histogram of simulated dithered EMG data compared to the true distribution and generalized Gaussian approximation. The distribution parameters were chosen to resemble experimental values, with  $\sigma_Z = 40$  ps,  $\tau_H = 200$  ps, and  $\Delta = 2000$  ps, which were then standardized for the overall distribution to have zero mean and unit variance. While the GG does not match the tail behavior exactly on either side, the value of  $\hat{p}$  determined through kurtosis matching does get the width and overall shape generally correct.

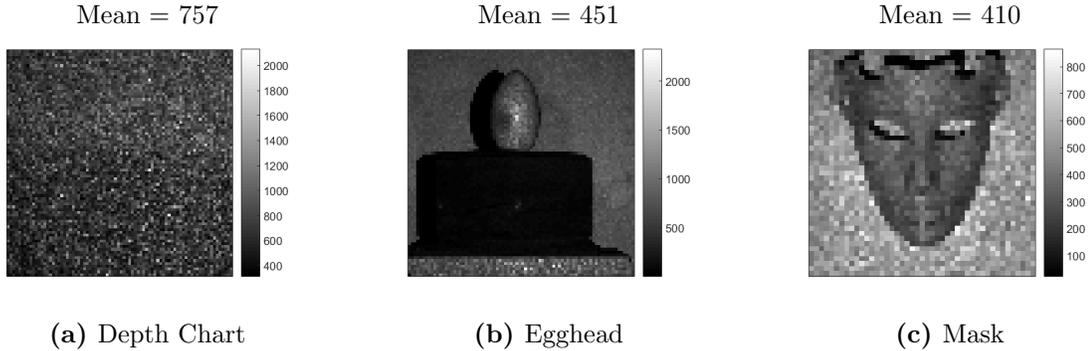
An example comparison between simulated samples of a dithered EMG random variable, the true dithered EMG PDF, and the generalized Gaussian approximation determined solely via kurtosis matching is shown in Fig. 4-15. While the generalized Gaussian is not as close an approximation for the experimental dithered noise distribution, it is still useful for roughly determining the tailed-ness of the noise and which order statistics and weights to use for estimating the mean.

#### 4.7.7 Experimental Results for EMG Pulse Modeling

Results are shown in Figure 4-16 for three additional experimental acquisitions: an  $80 \times 80$ -pixel scan of the depth resolution chart, an  $80 \times 80$ -pixel scan of the egghead trophy, and a  $45 \times 45$ -pixel scan of the 3D-printed mask shown in Figure 4-12. The numbers of detected photons per pixel are shown in Figure 4-17. Baseline depth estimates in Figure 4-16(a) are formed by averaging thousands of photon detections per pixel measured with 4-ps resolution. Coarsely-quantized detection times are ac-



**Figure 4-16:** Point cloud estimates for the (first row) Depth Chart, (second row) Egghead, and (third row) Mask datasets. Additional quantitative results comparing each algorithm to the baseline can be found in Table 4.1. The use of subtractive dither yields a large decrease in the RMSE, and the trimmed-mean estimator produces a small additional improvement.



**Figure 4-17:** Detection counts per pixel for the point clouds in Fig. 4-16. The Depth Chart is fairly uniform, whereas the Egghead and Mask reveal differences in reflectivity and appear to show some shadowing effects due to the bistatic lidar setup at short range.

quired for 6.15 ms per pixel for each of the datasets. Figure 4-17 shows the mean number of photons per pixel in each acquisition. The resulting pixelwise depth estimates in Fig. 4-16(b) are strongly biased by the quantization. In particular, the depth chart is estimated to be a planar surface with no fine structures, the egghead trophy is flattened while its backplane is shifted to the wrong depth, and the features of the mask are lost. Dithered data is acquired by acquiring photons for 30  $\mu\text{s}$  at each of the 205 dither delay steps of 10-ps each. The sample mean result  $\hat{\mu}_{\text{mean}}$  in Fig. 4-16(c) shows a dramatic improvement over the result without dither, recovering many of the fine features that had been lost with the coarse quantization. Finally, Figure 4-16(d) shows the results of the trimmed-mean estimator  $\hat{\mu}_{\alpha}$  defined in Equations (4.33)-(4.35), where with  $p_v$  determined by kurtosis matching with the dithered EMG distribution as in Equation 4.54. Judging from the depth chart, small features on the order of 4 mm are now visible.

Quantitative comparisons to the baseline estimates in Table 4.1 show the  $\hat{\mu}_{\alpha}$  estimator using the EMG distribution always outperforming the sample mean applied to the dithered data and producing the best results more consistently than  $\hat{\mu}_{\text{NL}}$  (defined in (4.36)). Quantitative improvement factors in RMSE over the estimates from

**Table 4.1:** RMSE performance of the depth estimates in Fig. 4.16 compared to several other methods. The generalized Gaussian-base methods are distinguished by whether the kurtosis matching is performed using a Gaussian (G) or exponentially-modified Gaussian (EMG) pulse shape approximation. The trimmed mean using the EMG-based kurtosis matching, denoted  $\hat{\mu}_\alpha$  (EMG), often has the best performance and consistently has lower RMSE than the sample mean.

Dataset	Method RMSE (mm)					
	No Dither	$\hat{\mu}_{\text{mean}}$	$\hat{\mu}_{\text{NL}}$ (G)	$\hat{\mu}_\alpha$ (G)	$\hat{\mu}_{\text{NL}}$ (EMG)	$\hat{\mu}_\alpha$ (EMG)
Depth Chart	26.22	4.98	<b>3.21</b>	4.19	3.63	4.04
Egghead	68.38	12.19	11.12	<b>10.38</b>	11.10	<b>10.38</b>
Mask	82.10	7.23	9.65	6.26	12.23	<b>6.19</b>

coarsely-quantized data range from  $6.5\times$  to  $13\times$ . Thus, we have demonstrated both that subtractive dither can improve depth estimation for SPL with coarse quantization and that modeling of the resulting detection time distribution leads to estimators that yield additional performance improvements.

## 4.8 Conclusions

This chapter introduced the concept of subtractive dither for single-photon lidar, which was studied by looking more generally at the task of estimating the mean of a Gaussian signal from quantized measurements. By applying subtractive dither to the measurement process, the noise becomes signal-independent but no longer has a Gaussian distribution. We showed that the generalized Gaussian distribution is a close and useful approximation for the Gaussian plus uniform total noise distribution. Estimators using the generalized Gaussian approximation effectively match the performance of the ML estimator for the total noise, which is a significant improvement over the more conventional mean and midrange estimators in Regime II, where neither the Gaussian nor uniform noise components dominate. Due to its computational simplicity and efficient performance, we recommend the trimmed mean  $\hat{\mu}_\alpha$ . From further comparison against estimators for quantized measurements, we determined

simple design rules for deciding whether and how to use quantized measurements. In short, there is value in using dither in Regimes I and II, and a GG-based estimator should be used in Regime II.

A subtractively-dithered single-photon lidar system was implemented by sweeping through sub-bin resolution delays between the laser illumination trigger and the TSCPC synchronization “start” signal. The sample mean of dithered detection times yielded significant reductions in RMSE, and additional modeling of the non-Gaussian lidar pulse shape led to further improvements using the generalized Gaussian-based trimmed-mean estimator. The resulting RMS depth error was reduced by up to a factor of 13, compared to the coarsely-quantized data without dither. The estimates could be further improved by incorporating spatial priors such as low total variation through regularization. Future developments can also take into account the discrete nature of the dither as well as the occasional ambient light detections and the pulse asymmetry not accounted for in the model. Finally, these methods would likely also benefit other applications of TCSPC such as fluorescence lifetime imaging (FLIM), especially for lifetimes that are short relative to the timing resolution.

## Appendices

### 4.A Cramér-Rao Bound

The Cramér-Rao Bound is a lower bound on the variance of an unbiased estimator [195], given by

$$\text{CRB}(\mu_X) = 1/I(\mu_X),$$

where  $I(\mu_X)$  is the Fisher information computed as

$$\begin{aligned} I(\mu_X) &= \mathbb{E} \left[ \left( \frac{\partial \log f_Y(y; \mu_X, \sigma_Z, \Delta)}{\partial \mu_X} \right)^2 \right] \\ &\stackrel{(a)}{=} \int \frac{\left( \frac{\partial}{\partial \mu_X} f_Y(y; \mu_X, \sigma_Z, \Delta) \right)^2}{f_Y(y; \mu_X, \sigma_Z, \Delta)} dy \\ &\stackrel{(b)}{=} \frac{1}{\sigma_Z^2 \Delta} \int \frac{\left[ \phi\left(\frac{v-\Delta/2}{\sigma_Z}\right) - \phi\left(\frac{v+\Delta/2}{\sigma_Z}\right) \right]^2}{\Phi\left(\frac{v+\Delta/2}{\sigma_Z}\right) - \Phi\left(\frac{v-\Delta/2}{\sigma_Z}\right)} dv \\ &\stackrel{(c)}{=} \frac{1}{\sigma_Z^2} \int \frac{\left[ \phi\left(\frac{u-1/2}{\sigma_Z/\Delta}\right) - \phi\left(\frac{u+1/2}{\sigma_Z/\Delta}\right) \right]^2}{\Phi\left(\frac{u+1/2}{\sigma_Z/\Delta}\right) - \Phi\left(\frac{u-1/2}{\sigma_Z/\Delta}\right)} du, \end{aligned} \quad (4.55)$$

where step (a) uses the definition of expectation and the chain rule, (b) differentiates (4.17) with respect to  $\mu_X$  for  $v = y - \mu_X$ , and (c) changes variables to  $u = v/\Delta$ . Normalizing by  $\Delta^2$  removes the separate dependence on  $\sigma_Z$  or  $\Delta$ , so we define the normalized CRB as

$$\begin{aligned} \text{NCRB}(\mu_X) &= \text{CRB}(\mu_X)/\Delta^2 \\ &= \frac{(\sigma_Z/\Delta)^2}{\int \frac{\left[ \phi\left(\frac{u-1/2}{\sigma_Z/\Delta}\right) - \phi\left(\frac{u+1/2}{\sigma_Z/\Delta}\right) \right]^2}{\Phi\left(\frac{u+1/2}{\sigma_Z/\Delta}\right) - \Phi\left(\frac{u-1/2}{\sigma_Z/\Delta}\right)} du}. \end{aligned} \quad (4.56)$$

Finally, Fisher information is additive for independent observations, so for  $K$  independent samples, the lower bound on the NCRB is  $1/K$  times that for one observation.

## 4.B Kurtosis Matching

The *kurtosis* of a random variable  $B$  is the standardized fourth central moment [55], defined as

$$\kappa(B) = \frac{\mathbb{E}[(B - \mu_B)^4]}{\{\mathbb{E}[(B - \mu_B)^2]\}^2} = \frac{\mu_4(B)}{\sigma_B^4}. \quad (4.57)$$

The *excess kurtosis*  $\gamma(B) = \kappa(B) - 3$  is often used to simplify computations. Define  $A = B + C$ , where  $B$  and  $C$  are independent random variables. The kurtosis of the sum can be computed by expanding (4.57) as follows:

$$\begin{aligned} \kappa(A) &= \frac{\mathbb{E}[(A - \mu_A)^4]}{\{\mathbb{E}[(A - \mu_A)^2]\}^2} = \frac{\mathbb{E}\{[(B - \mu_B) + (C - \mu_C)]^4\}}{\{\mathbb{E}\{[(B - \mu_B) + (C - \mu_C)]^2\}\}^2} \\ &= \frac{\mu_4(B) + \mu_4(C) + 6\sigma_B^2\sigma_C^2}{(\sigma_B^2 + \sigma_C^2)^2}, \end{aligned}$$

where independence eliminates the odd cross terms. Then the excess kurtosis is

$$\gamma(A) = \frac{\sigma_B^4\gamma(B) + \sigma_C^4\gamma(C)}{\sigma_A^4}. \quad (4.58)$$

The kurtosis of Gaussian and uniform random variables is well-known and straightforward to compute from the definition; the excess kurtosis is 0 for a Gaussian and  $-6/5$  for a uniform distribution. From [175], we have that the excess kurtosis<sup>2</sup> of a GGRV  $V$  with shape parameter  $p_v$  is

$$\gamma(V) = \frac{\Gamma(1/p_v)\Gamma(5/p_v)}{[\Gamma(3/p_v)]^2} - 3. \quad (4.59)$$

---

<sup>2</sup>Note that the definition of kurtosis in [175] corresponds to the excess kurtosis in this work.

To fit the GGD to the sum of uniform and Gaussian random variables, we set the kurtosis of the approximation to match the kurtosis of the sum using (4.58)

$$\begin{aligned} \frac{\Gamma(1/p_v)\Gamma(5/p_v)}{[\Gamma(3/p_v)]^2} &= 3 + \frac{\sigma_Z^4 \cdot 0 + \sigma_W^4(-6/5)}{(\sigma_W^2 + \sigma_Z^2)^2} \\ &= 3 - \frac{6}{5} \frac{1}{\left[1 + 12 \left(\frac{\sigma_Z}{\Delta}\right)^2\right]^2}, \end{aligned} \quad (4.60)$$

where  $\sigma_W^2 = \Delta^2/12$ .

#### 4.C Mean Squared Error of $\hat{\mu}_Q$

We use iterated expectation to compute the MSE of  $\hat{\mu}_Q$  as

$$\mathbb{E}[(\hat{\mu}_Q - \mu_X)^2] = \mathbb{E}[\mathbb{E}[(\hat{\mu}_Q - \mu_X)^2 | \mu_X]], \quad (4.61)$$

with no prior knowledge on the true value so that we assume  $\mu_X \sim \mathcal{U}[-\Delta/2, \Delta/2]$  within a bin. Define a function  $g : \mathbb{R} \rightarrow \mathbb{R}$  as  $g(x) := \mathbb{E}[(\hat{\mu}_Q - \mu_X)^2 | \mu_X = x]$ , then

$$\begin{aligned} g(x) &= \mathbb{E} \left[ \left( \frac{1}{K} \sum_{i=1}^K q(x + Z_i) - x \right)^2 \right] \\ &= x^2 + \frac{1}{K^2} \left( \sum_{i=1}^K \mathbb{E}[(q(x + Z_i))^2] + \sum_{i=1}^K \sum_{j \neq i} \mathbb{E}[q(x + Z_i)] \mathbb{E}[q(x + Z_j)] \right) \\ &\quad - \frac{2x}{K} \sum_{i=1}^K \mathbb{E}[q(x + Z_i)] \\ &= x^2 + \frac{1}{K} \mathbb{E}[(q(x + Z))^2] + \frac{K-1}{K} (\mathbb{E}[q(x + Z)])^2 - 2x \mathbb{E}[q(x + Z)]. \end{aligned} \quad (4.62)$$

Using the definition

$$\Psi(m, x) = \Phi\left(\frac{m + 1/2 - x}{\sigma_Z/\Delta}\right) - \Phi\left(\frac{m - 1/2 - x}{\sigma_Z/\Delta}\right), \quad (4.63)$$

note that

$$\begin{aligned}\mathbb{E}[q(x+Z)] &= \lim_{M \rightarrow \infty} \sum_{m=-M}^M m\Delta \int_{m\Delta-\Delta/2}^{m\Delta+\Delta/2} \frac{1}{\sigma_Z} \phi\left(\frac{z-x}{\sigma_Z}\right) dz \\ &\approx \Delta \sum_{m=-M}^M m\Psi(m, x)\end{aligned}$$

for some large number  $M$ . Similarly,

$$\mathbb{E}[(q(x+Z))^2] \approx \Delta^2 \sum_{m=-M}^M m^2\Psi(m, x).$$

The MSE normalized by  $\Delta^2$  then follows as

$$\begin{aligned}\mathbb{E}[(\hat{\mu}_Q - \mu_X)^2]/\Delta^2 &= \frac{1}{12} + \frac{1}{K} \int_{-1/2}^{1/2} \sum_{m=-M}^M m^2\Psi(m, x) dx \\ &\quad + \frac{K-1}{K} \int_{-1/2}^{1/2} \left( \sum_{m=-M}^M m\Psi(m, x) \right)^2 dx \\ &\quad - 2 \int_{-1/2}^{1/2} x \sum_{m=-M}^M m\Psi(m, x) dx.\end{aligned}\tag{4.64}$$

## Chapter 5

# Dead Time Compensation

### 5.1 Overview

So far in this thesis, we have assumed that every photon detection time can be treated as an independent sample from an identical detection time distribution, which is the normalized Poisson photon arrival process intensity. For low-flux acquisition, this model is accurate enough to be useful. However, the requirement that the photon flux must be low is a limitation on the types of scenes that can be imaged and the applications in which SPL can be deployed. For instance, AVs guided by SPL-generated point clouds would require the depth measurements to be made quickly and over a large range of depths and reflectivities. The conventional approach to imaging bright objects, such as the retro-reflective street signs common to AV environments, is to attenuate the photon flux to the low-flux regime. Otherwise, the dead times in the single photon detectors and the timing electronics may cause distortions in the detection counts acquired detection time histograms, yielding erroneous reflectivity and depth estimates. It is undesirable to attenuate SPL systems in AVs, however, both because potentially useful photons are discarded for bright objects, and because, in combination with the short acquisition times, distant and dark objects such as car tires would reflect too few photons and become essentially invisible.

Allowing for higher incident flux and compensating for the resulting distortions

---

This chapter includes research conducted jointly with Yanting Ma, Robin Dawson, and Vivek Goyal [151, 152, 153].

due to dead time would enable faster acquisition without a loss of accuracy and has been the subject of several recent works [86, 144, 77, 78], although their models of dead time assume a single detection can be recorded per illumination period, an assumption which does not necessarily hold for modern timing electronics. Our aim is to accurately model the effects of dead time on the photon detection process so that photons can be detected at a much higher rate and distortions introduced due to dead time can be predicted and corrected.

This chapter establishes a new way of modeling photon detection times that looks at not only the relative times with respect to the most recent illumination, but also the absolute sequence of times that accounts for time dependence between detections. We explore the empirical distribution of detection times in the presence of dead time and demonstrate that an accurate statistical model can result in reduced ranging error with shorter data acquisition time when operating in the high-flux regime. Specifically, we show that the empirical distribution of detection times converges to the stationary distribution of a Markov chain. Depth estimation can then be performed by passing the empirical distribution through a filter matched to the stationary distribution. Moreover, based on the Markov chain model, we formulate the recovery of arrival distribution from detection distribution as a nonlinear inverse problem and solve it via provably convergent mathematical optimization. By comparing per-detection Fisher information for depth estimation from high- and low-flux detection time distributions, we provide an analytical basis for possible improvement of ranging performance resulting from the presence of dead time. Finally, we demonstrate the effectiveness of our formulation and algorithm via experiments and ranging simulations. Eventually, this approach should lead to the deployment of SPL systems with higher laser powers, shorter acquisition times, and more accurate depth estimation.

## 5.2 Dead Time Characterization

The source of the dead time and resulting behavior of an SPL system may vary greatly depending on the implementation. Specifically, this chapter studies dead time correction for modern TCSPC systems with asynchronous electronics (such as the HydraHarp 400 [208] or TimeHarp 260 [209] from PicoQuant) and a nonparalyzable detector (e.g., a PDM-series [70] or Fast-gated [33] Single Photon Avalanche Diode (SPAD) from Micro Photon Devices, operated in free-running mode), for which the detector dead time  $t_d$  is longer than the electronics dead time  $t_e$ . We refer to Section 2.2.2 for a discussion on how the SPAD design and the architecture of modern TCSPC electronics give rise to dead times. In the following, we formally define paralyzability and synchronization so as to clarify the dead time model we consider.

### 5.2.1 Paralyzability

The dead times of event-counting detectors have been studied since at least the 1940s [73, 107, 114, 109], with Feller [63] first classifying detectors in terms of their paralyzability. *Nonparalyzable (Type I)* detectors are dead for a fixed time  $t_d$  after a detection, regardless of whether additional photons arrive during the dead time. On the other hand, when photons arrive during the dead time of a *paralyzable (Type II)* detector, the dead time restarts and extends for at least another  $t_d$ . We consider only nonparalyzable detectors in our work.

### 5.2.2 Synchronization

An important factor in analyzing the effect of dead times in a TCSPC system is whether the dead times and illuminations are synchronized. We call systems *synchronous* if they ensure that the end of a dead time is synchronized with the start of an illumination period. Synchronous operation is often built into the hardware, such

as in classical TCSPC systems employing TACs or in the gated mode of fast-gated SPADs [33, 144]. In reversed start-stop mode, classical timing electronics may become active in the middle of an illumination period, but that recovery time is consistent, as the dead time is synchronized to the stop signal (either a delayed version of the current pulse or the next pulse). By using digital timing circuitry, modern TCSPC electronics enable *asynchronous* operation, in which there is no enforced synchronization between the dead time and the timing electronics. If a photon is detected towards the end of a cycle and the dead time continues after the next laser excitation, there is no mechanism preventing the detector from becoming active in the middle of that cycle. In other words, the end of the dead time is no longer dependent on the cycle synchronization, but on the arrival time of the most recently detected photon. The synchronous and asynchronous architectures correspond to the “clock-driven” and “event-driven” SPAD recharge mechanisms, which were explored in [7, 8]. While most existing work on the effect of dead time assumes synchronous operation, we consider only asynchronous systems in this work.

### 5.2.3 Dead Time Compensation Approaches

Yu and Fessler [217] outline a number of general strategies for handling the effects of dead time, with the simplest approach being to simply ignore the dead time. Most commonly, the total photon flux at the detector is changed such that the dead time effects are actually negligible and can be ignored. Since the effect of dead time is that photon arrivals within  $t_d$  of a detection are missed, a straightforward approach is to reduce the total photon flux, either by lowering the laser power and ambient light if possible, or by attenuating with a filter at the detector. The conventional wisdom expressed by O’Connor and Phillips [137] is to keep the fraction of excitations causing a detection to be at most 5% to avoid dead time effects, a recommendation that electronics manufacturers have adopted. Reducing the flux inevitably leads to

longer time needed to acquire the same number of photons, which is one of the motivations for recent works in photon-efficient imaging [104, 170, 3], including the work presented in Chapter 3. Other approaches have tried to ignore dead time by changing the hardware setup, such as using multiple detectors so that there is more likely to be a detector not in the reset state when a photon arrives [15].

Rather than attenuate the flux at the detector to avoid dead time effects, another strategy is to correct the distortions in the high-flux data after acquisition. Most algorithm-based attempts at dead time compensation consider synchronous systems due not only to the systems that have historically been available, but also for the convenient property that detection times are statistically independent of each other in different cycles [170]. One of the first methods for dealing with dead time in synchronous systems is that of Coates [44]. Coates's basic algorithm was designed for lifetime measurement, with later work adapting the algorithm to include background subtraction [52]. The basic principle of Coates's algorithm is that for any bin  $n$  in a histogram, the detections in the preceding bins spanning  $t_d$  represent excitation cycles when no photon could have been detected in bin  $n$  because the detector was dead. The number of cycles in which the detector thus must have been dead is used to adjust computation of the photon arrival probability in each bin. Recent work has rederived Coates's expression, which is the ML estimator for the number of photon arrivals in each bin of a histogram in a synchronous system, in order to include priors for maximum a posteriori (MAP) estimation [144]. A few models [215, 200] consider histogram corrections for a hybrid of synchronous and asynchronous systems, which do allow for multiple statistically dependent detections per illumination cycle but without the dependency carrying over into different cycles. A handful of papers address special cases of dead time effects in asynchronous systems: Antolovic et al. [7, 8] consider detection rate estimation for homogeneous arrival processes, whereas Cominelli et al.

[46] explore the special case when  $t_d$  equals an integer multiple of  $t_r$  and no correction is needed. However, these approaches are insufficient to address the typical lidar acquisition mode with inhomogenous arrivals and in which  $t_d$  and  $t_r$  cannot necessarily be arbitrarily adjusted. The only existing work that addresses asynchronous systems generally is that of Isbaner et al. [94], which treats the detection process as a time-dependent attenuation of the arrival process intensity, and uses an iterative procedure to estimate the attenuation and thereby correct the detection histogram. Although both  $t_e$  and  $t_d$  are included in the model for [94], a TCSPC system simplifies to having only one source of dead time when  $t_e < t_d$ , the scenario considered in this chapter.

The last strategy for dealing with dead time is to use the data as acquired but to incorporate dead time into the detection model. In this vein, Heide et al. [86] adjust their parameter estimation procedure to include dead time effects. However, the synchronous system assumption they use is technically only valid for their asynchronous timing electronics (PicoQuant PicoHarp 300, [207]) if zero ambient light is present, which guarantees that the detector will be reset for the next signal pulse.

#### 5.2.4 Other Related Work

In addition to missing photons that arrive during a dead time, events can fail to be registered in the recording of point processes through other forms of “counting loss,” which depends on the measurement system design [15]. A well-documented form of counting loss is “pile-up,” referring to the problem of the rising edge of a pulse overlapping with the tail of a previous pulse, such that the later pulse is not registered by a discriminator. Each piled-up pulse prolongs the duration in which new events cannot be detected, making the discriminator a Type II detector. Pile-up is present in some TCSPC system designs, such as those using passively-recharged SPADs [8] or hybrid photodetectors with negligible dead time [141]. Several approaches correct-

ing for pile-up have recently been proposed for nuclear spectrometry, in which both the pulse times and energies are of interest [190, 167, 191]. Confusingly, the term “pile-up” is also sometimes used to describe the effect of dead time in synchronous TCSPC systems (e.g., [144, 86]) because the effect of earlier detections preventing later detections is similar. Another form of counting loss, named Type III in [217], may also occur in some systems (e.g., [145]) when two pulses occur close together and neither one gets recorded.

In lidar applications, the dead time-affected acquisition results in closer apparent distances, which has a similar effect to the intensity-dependent change in perceived depth known as “range walk error” [139, 110, 40]. Range walk is the result of using a discriminator to trigger in the leading edge of a signal pulse; a stronger signal with a steeper rising edge will be detected earlier than a weaker signal with a smaller slope. Due to the similarity with high-flux ranging, approaches correcting for range walk error could be adapted to compensate for dead time. Several optics-based methods aimed at range walk error correction attempt to experimentally measure and then correct for the bias in depth estimation. He et al. [84] first calibrate the amount of range walk incurred as a function of the detection rate. Then the conventional depth estimation procedure is performed with the dead time-distorted data, and the range walk bias is subtracted off to correct the depth estimate. Ye et al. [216] use a similar method, except they first split the incident light with a 90:10 beamsplitter to two SPADs, using the lower-flux channel for simultaneous bias estimation to subtract off from the lower-variance high-flux estimate.

### 5.3 Empirical Distribution of Detection Times

The challenge of studying the detection time distribution for the asynchronous dead time model is that the detection times are statistically dependent. In this section, we

show that the dependency is Markovian and provide the explicit transition probability density function. From the transition PDF, we can analyze the stationary condition and obtain the stationary distribution, from which our high-flux ranging algorithms are derived.

We established in Section 2.3 that the photon arrival times at a detector are described by a Poisson process with intensity  $\lambda(t)$  given in (2.2). We continue with that formulation for a single pixel, dropping any cumbersome pixel notation for the moment. We especially recall that the photon flux, or the expected number of photon arrivals per illumination period, is denoted as  $S$  and  $B$  for the signal and background processes, respectively. The total flux is given by  $\Lambda := S + B$ , and the signal to background ratio is defined as  $\text{SBR} := S/B$ . Note that for ranging results presented later in this chapter, a single depth is assumed ( $Q = 1$ ), but that is only a strict requirement for some of our dead time compensation methods.

### 5.3.1 Markov Chain Model for Detection Times

If there were no dead time effects, the detection process would be equivalent to the arrival process, which is Poisson with intensity  $\lambda(t)$ . Conditioned on the total number of detections, the absolute detection times (measured with respect to the start of the experiment at time  $t = 0$ ) would be order statistics of i.i.d. random variables with common probability density function  $\propto \lambda(t)$  [174, Section 2.3.3]. However, in the presence of dead time effects, the detection process is no longer Poisson, since the detection intensity, denoted by  $\mu(t)$ , is now a random process depending on the history of the detection process; such a detection process is referred to as a self-exciting process [174]. Specifically, let  $\{K(t) : t \geq 0\}$  denote the detection process with (random) intensity  $\mu(t)$ , where  $\{K(t) : t \geq 0\}$  is characterized by the number of detections  $K(t)$  at time  $t$  and a sequence of absolute detection times  $T_1, \dots, T_{K(t)}$ .

The conditional PDF of  $T_{\ell+1}$  given  $T_1, \dots, T_\ell$  is [174, (6.15) (6.16)]

$$f_{T_{\ell+1}|T_1, \dots, T_\ell}(t|t_1, \dots, t_\ell) = \mu(t) \exp\left(-\int_{t_\ell}^t \mu(\tau) d\tau\right). \quad (5.1)$$

For a general self-exciting process,  $\mu(t)$  can depend on the entire history of the process  $\{K(\tau) : 0 \leq \tau < t\}$ . For the specific detection process considered in this chapter, we have

$$\mu(t) = \begin{cases} \lambda(t), & \text{if } t > T_{K(t)} + \mathbf{t}_d; \\ 0, & \text{if } T_{K(t)} < t \leq T_{K(t)} + \mathbf{t}_d, \end{cases} \quad (5.2)$$

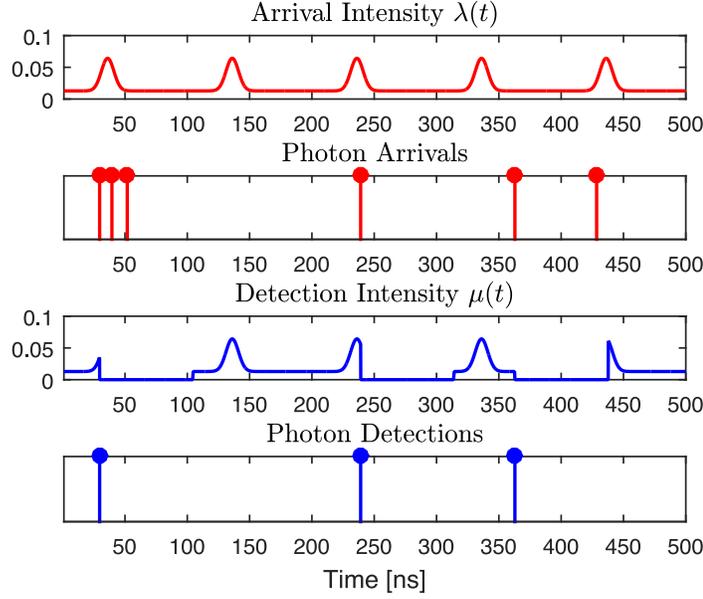
where we introduce the notation  $T_0 := -\infty$ . We can see that  $\mu(t)$  only depends on the latest detection time. Therefore, for the  $\mu(t)$  defined in (5.2), the right-hand side (RHS) of (5.1) depends on  $t_\ell$  but not on  $t_1, \dots, t_{\ell-1}$ . That is, the absolute detection times form a Markov chain with transition PDF

$$f_{T_{\ell+1}|T_\ell}(t|t_\ell) = \lambda(t) \exp\left(-\int_{t_\ell + \mathbf{t}_d}^t \lambda(\tau) d\tau\right) \mathbb{I}\{t > t_\ell + \mathbf{t}_d\}, \quad (5.3)$$

where  $\mathbb{I}$  is the indicator function. An illustration of a realization of the detection process is shown in Fig. 5-1.

Define two sequences of random variables,  $\{N_\ell\}_{\ell \in \mathbb{N}}$  and  $\{X_\ell\}_{\ell \in \mathbb{N}}$ , such that  $N_\ell := \lfloor T_\ell / \mathbf{t}_r \rfloor$ , where  $\lfloor a \rfloor$  is the integer part of  $a \in \mathbb{R}$ , and  $X_\ell := T_\ell \bmod \mathbf{t}_r$ , hence  $T_\ell = N_\ell \mathbf{t}_r + X_\ell$ . That is,  $N_\ell$  is the number of illumination periods before  $T_\ell$  and  $X_\ell$  is the relative location of absolute detection time  $T_\ell$  within illumination period  $N_\ell + 1$ , which is referred to as simply the detection time in this chapter. Note that if there were no dead time effects, the empirical distribution of  $X_\ell$ 's would be identical to the arrival time PDF, given by

$$f_{X_A}(x) = \lambda(x) / \Lambda, \quad \text{for } x \in [0, \mathbf{t}_r). \quad (5.4)$$



**Figure 5.1:** Illustration of the effect of dead times on the detection process for  $t_r = 100$  ns,  $t_d = 75$  ns,  $\sigma_p = 5$  ns,  $S = 0.5$ , and  $B = 1$ . Photon arrival times are generated according to the arrival intensity  $\lambda(t)$ . The detection intensity  $\mu(t)$  is equal to the arrival intensity  $\lambda(t)$  except immediately following a photon detection when  $\mu(t) = 0$ , so the detection times are a subset of the arrival times and detection is not a Poisson process.

The following proposition provides statistical characterization of  $\{X_\ell\}_{\ell \in \mathbb{N}}$  in the presence of dead time.

**Proposition 1.** *Suppose that the photon arrival process is an inhomogeneous Poisson process with periodic intensity function  $\lambda(t)$ , whose period is  $t_r$ , and the detector has dead time  $t_d$ . Define  $x_d := t_d \bmod t_r$ . Let the random sequence  $\{T_\ell\}_{\ell \in \mathbb{N}}$  denote absolute detection times and define detection times as  $X_\ell := T_\ell \bmod t_r$ , for all  $\ell \in \mathbb{N}$ . Then the random sequence  $\{X_\ell\}_{\ell \in \mathbb{N}}$  forms a Markov chain with state space  $[0, t_r)$  and transition PDF*

$$f_{X_{\ell+1}|X_\ell}(x_{\ell+1}|x_\ell) = \frac{\lambda(x_{\ell+1})}{1 - \exp(-\Lambda)} \exp\left(-\int_{x_\ell + x_d}^{\left\lceil \frac{x_\ell + x_d - x_{\ell+1}}{t_r} \right\rceil t_r + x_{\ell+1}} \lambda(\tau) d\tau\right), \quad (5.5)$$

where  $\lceil a \rceil := \lfloor a \rfloor + 1$  and  $\Lambda := \int_0^{\mathbf{t}_r} \lambda(\tau) \, \mathbf{d}\tau$ .

*Proof.* See Appendix 5.A. □

We can check that  $\{X_\ell\}_{\ell \in \mathbb{N}}$  is  $\psi$ -irreducible, recurrent, and aperiodic, and hence it has a unique stationary PDF [130, Proposition 10.4.2]. Denoting the stationary PDF by  $f_{X_D}$ , then for all  $x \in [0, \mathbf{t}_r)$ ,  $f_{X_D}$  satisfies

$$f_{X_D}(x) = \int_0^{\mathbf{t}_r} f_{X_D}(y) f_{X_{\ell+1}|X_\ell}(x|y) \, \mathbf{d}y. \quad (5.6)$$

That is,  $f_{X_D}$  is the eigenfunction corresponding to eigenvalue 1 of the linear operator  $\mathcal{P}$  defined as

$$\mathcal{P}f(x) := \int_0^{\mathbf{t}_r} f(y) f_{X_{\ell+1}|X_\ell}(x|y) \, \mathbf{d}y. \quad (5.7)$$

For the special case where  $\mathbf{x}_d = 0$ , we show in the following that the arrival PDF  $f_{X_A}$  defined in (5.4) satisfies the stationary condition, meaning that dead time does not cause any distortion in detection time distribution; this result has also been noted in [46] with a different derivation. With  $f_{X_D}(x) = \lambda(x)/\Lambda$  and  $\mathbf{x}_d = 0$ , the RHS of (5.6) is

$$\begin{aligned} \int_0^{\mathbf{t}_r} \frac{\lambda(y)}{\Lambda} f_{X_{\ell+1}|X_\ell}(x|y) \, \mathbf{d}y &= \int_0^x \frac{\lambda(y)}{\Lambda} \frac{\lambda(x)}{1 - \exp(-\Lambda)} \exp\left(-\int_y^x \lambda(\tau) \, \mathbf{d}\tau\right) \\ &\quad + \int_x^{\mathbf{t}_r} \frac{\lambda(y)}{\Lambda} \frac{\lambda(x)}{1 - \exp(-\Lambda)} \exp\left(-\int_y^{\mathbf{t}_r+x} \lambda(\tau) \, \mathbf{d}\tau\right) \\ &= \frac{\lambda(x)}{\Lambda} \left[ \frac{\int_0^x \lambda(y) \exp\left(-\int_y^x \lambda(\tau) \, \mathbf{d}\tau\right) \, \mathbf{d}y}{1 - \exp(-\Lambda)} \right. \\ &\quad \left. + \frac{\int_x^{\mathbf{t}_r} \lambda(y) \exp\left(-\int_y^{\mathbf{t}_r+x} \lambda(\tau) \, \mathbf{d}\tau\right) \, \mathbf{d}y}{1 - \exp(-\Lambda)} \right]. \end{aligned}$$

Label the two terms in the square brackets as  $A_1$  and  $A_2$ . Using the chain rule and

the Leibniz rule for differentiation, for any constant  $a$  that does not depend on  $y$ , we have that

$$\frac{d}{dy} \exp \left( - \int_y^a \lambda(\tau) d\tau \right) = \lambda(y) \exp \left( - \int_y^a \lambda(\tau) d\tau \right).$$

Letting  $a = x$ , we have

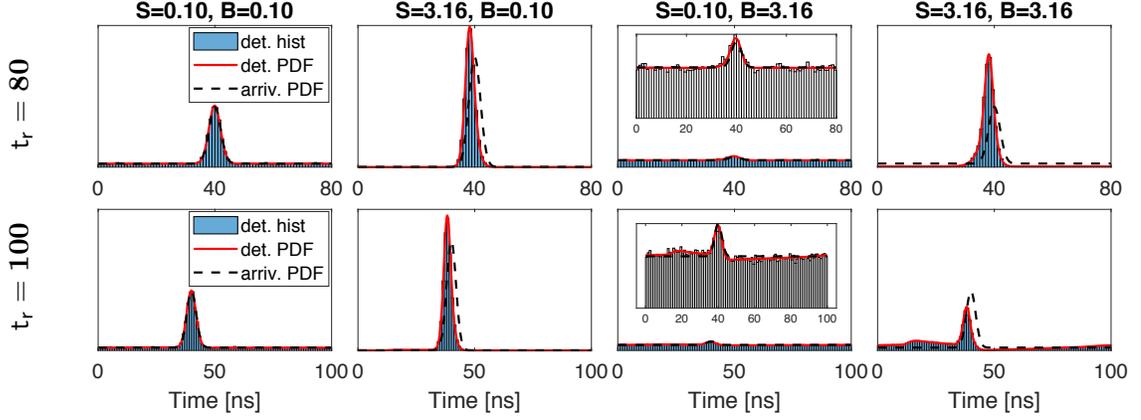
$$A_1 = \frac{\exp \left( - \int_y^x \lambda(\tau) d\tau \right) \Big|_0^x}{1 - \exp(-\Lambda)} = \frac{1 - \exp \left( - \int_0^x \lambda(\tau) d\tau \right)}{1 - \exp(-\Lambda)}.$$

Similarly, let  $a = \mathbf{t}_r + x$ , then

$$A_2 = \frac{\exp \left( - \int_0^x \lambda(\tau) d\tau \right) - \exp(-\Lambda)}{1 - \exp(-\Lambda)}.$$

It follows that  $A_1 + A_2 = 1$ , and so  $\lambda(x)/\Lambda$  is the stationary distribution of the Markov chain when  $\mathbf{x}_d = 0$ .

To numerically demonstrate the correctness of (5.5) for general  $\mathbf{x}_d$ , we partition the state space  $[0, \mathbf{t}_r)$  into  $\mathbf{n}_b$  equally spaced time bins with bin centers  $\{b_m\}_{m=1}^{\mathbf{n}_b}$  and approximate the linear operator  $\mathcal{P}$  defined in (5.7) with an  $\mathbf{n}_b \times \mathbf{n}_b$  matrix  $\mathbf{P}$ , where  $P_{m,n} := f_{X_{\ell+1}|X_\ell}(b_n|b_m)$  with  $f_{X_{\ell+1}|X_\ell}$  defined in (5.5). The matrix  $\mathbf{P}$  is then normalized to have row sum equal to 1 so that it becomes a probability transition matrix  $\tilde{\mathbf{P}}$ . A discrete approximation of  $f_{X_D}$ , denoted by a length- $\mathbf{n}_b$  row vector  $\mathbf{f}_{X_D}$ , is then obtained as the leading left eigenvector of  $\tilde{\mathbf{P}}$ , since  $\mathbf{f}_{X_D}$  should satisfy the Markov chain stationary condition  $\mathbf{f}_{X_D} = \mathbf{f}_{X_D} \tilde{\mathbf{P}}$ . Moreover, if the second largest (in terms of magnitude) eigenvalue of  $\tilde{\mathbf{P}}$  is strictly less than one, in other words,  $\tilde{\mathbf{P}}$  admits a spectral gap, then the corresponding Markov chain converges to its stationary distribution geometrically fast. We have verified that in all parameter settings considered in this paper,  $\tilde{\mathbf{P}}$  admits a spectral gap, thus confirming the convergence of the chain. Finally,  $\mathbf{f}_{X_D}$  is compared with the histogram of a set of simulated detection times, where we expect a close match between the simulated histogram and  $\mathbf{f}_{X_D}$ . Detection



**Figure 5.2:** Comparisons between histograms of simulated detection times, predicted detection time PDFs, and arrival time PDFs illustrate how dead time affects the detection process. In addition to a shift in the mode toward earlier detection times, the dead time may also cause a ripple in the detection PDF relative to the arrival PDF. Plots are shown for  $\sigma_p = 2$  ns,  $\Delta = 50$  ps,  $n_r = 50000$  illuminations, and  $t_d = 75$  ns. The vertical axis scale is constant for each row. An inset with a different vertical scale is included for each plot in the third column to emphasize the ripple that is not easily seen in the original scale.

times are simulated by first generating arrival times according to (2.2). Then starting with detection of the first arrival time generated, subsequent arrivals are culled from the sequence if they are within  $t_d$  of the previous absolute detection time, as in [88]. Note that unlike in [88], both background photons and dark counts are considered to trigger dead times in the same manner as signal detections.

Comparisons between a histogram of detection times collected from simulation and the corresponding  $\mathbf{f}_{X_D}$  are shown in Fig. 5.2. In each simulation, the number of illuminations is  $n_r = 50000$  and the half pulse width is  $\sigma_p = 2$  ns. The close matches between predicted detection PDFs and the simulated histogram results validate the effectiveness of the Markov chain model in deriving the limiting distribution. The figure further illustrates the effect that dead time has on TCSPC. The first column of Fig. 5.2 shows results with  $S = B = 0.1$ , so the total flux  $\Lambda$  is low enough that few photons arrive during the detector dead time, and the arrival and detection

densities are almost identical. If just the signal flux is increased, e.g., by increasing the illumination laser power (second column), the photon detection density narrows and shifts slightly toward earlier detection times (similar to the phenomenon of range walk error), due to early arrivals from the pulse blocking later photons from being detected. When the background flux increases, the distortions in the density due to dead time become more apparent. However, these distortions also depend on the particular values of  $t_r$  and  $t_d$ . When  $t_r$  is slightly larger than  $t_d$  (such as for  $t_r = 80$  in the top row of Fig. 5.2), the dead time triggered after a signal detection will reset just before signal photons from the next pulse arrive at the detector. The dead time thus behaves as a *signal-triggered gate*, blocking detection of many background photons while allowing detection of additional signal. On the other hand, increasing  $t_r$  by just 20 ns (bottom row) causes a significant ripple in the detection PDF a duration  $t_d$  after the main signal peak (modulo  $t_r$ ). The dead time is again often triggered by signal photons when  $S$  is large, but the reset of the detector in the next cycle allows incident background photons to be detected, amplifying the apparent background intensity at that part of the cycle. Note that this pre-pulse ripple could easily be mistaken for optical system inter-reflections or poor electronics thresholding if detector dead time were not taken into account.

### 5.3.2 Comparison of Fisher Information

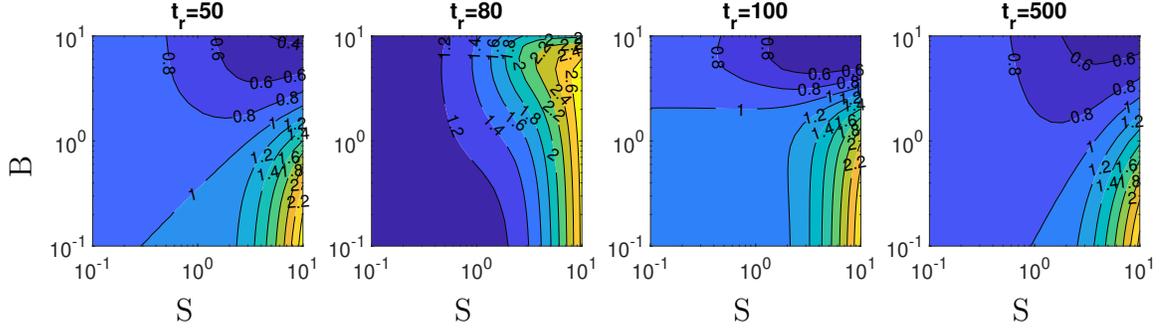
High-flux acquisition enables detection of more photons than low-flux acquisition for a fixed number of illuminations. Although the detection time distribution is distorted in the sense that it is different from the arrival time distribution, our Markov chain model allows us to accurately predict the distortion. Therefore, it is expected that for a fixed number of illuminations, high-flux acquisition with our probabilistic model for detection times can improve ranging performance over the 5% low-flux acquisition rule. Another interesting aspect is to compare estimates from low-flux and high-flux

acquisitions for a fixed number of detections. By comparing the arrival PDF and the detection PDF in Fig. 5-2, we notice that dead time results in a “narrowing” of the pulse, especially for large  $S$ . Since the detection PDF at low flux is identical to the arrival PDF  $f_{X_A}$ , we hypothesize that allowing higher-flux acquisition by forgoing attenuation can incur distortion that may in fact be favorable for depth estimation in some cases.

To verify this somewhat counter-intuitive speculation, we compare the Fisher information per detection for estimating the depth  $z$  from the low-flux PDF  $f_{X_A}$  (5.4) and high-flux PDF  $f_{X_D}$  (the stationary distribution of the Markov chain defined in Proposition 1), which are denoted by  $\text{FI}_A$  and  $\text{FI}_D$ , respectively, and are derived in Appendix 5.B and computed as

$$\begin{aligned}\text{FI}_A &= \int_0^{\mathbf{t}_r} \left( \frac{\partial}{\partial z} f_{X_A}(x; z) \right)^2 \frac{1}{f_{X_A}(x; z)} \mathbf{d}x, \\ \text{FI}_D &= \int_0^{\mathbf{t}_r} \left( \frac{\partial}{\partial z} f_{X_D}(x; z) \right)^2 \frac{1}{f_{X_D}(x; z)} \mathbf{d}x,\end{aligned}$$

where the derivative of  $f_{X_D}$  is computed numerically. Note that while realizations of detection times are not i.i.d. samples of  $f_{X_A}$  or  $f_{X_D}$ , most ranging algorithms only use the empirical distribution of detection times for depth estimation. Therefore, it is reasonable to consider Fisher information of the limiting empirical distributions  $f_{X_A}$  and  $f_{X_D}$  rather than that of the joint distributions. Fig. 5-3 presents the Fisher information ratio  $\text{FI}_D/\text{FI}_A$  for  $\mathbf{t}_d = 75$  ns and with  $\mathbf{t}_r$  varying from 50 to 500 ns. By (5.5), we notice that only  $\mathbf{x}_d := \mathbf{t}_d \bmod \mathbf{t}_r$  affects the detection time distribution. Hence, for the case where  $\mathbf{t}_r = 50$  ns, the effective dead time in terms of detection time distribution is 25 ns. In the regions where the ratio is greater than one,  $f_{X_D}$  is more informative about the depth  $z$  than  $f_{X_A}$  (i.e., the dead time effect is beneficial) in the sense that the per-detection Fisher information is higher. We notice that such



**Figure 5.3:** The Fisher information ratio  $FI_D/FI_A$  indicates the performance improvement that may be gained for the same number of detections when high-flux data is used instead of low-flux data. The plots show for various signal rate  $S$  and background rate  $B$  and for  $\sigma_p = 0.2$  ns,  $\Delta = 10$  ps, and  $t_d = 75$  ns that when SBR is sufficiently high and  $B$  is not too large, the effect of dead time is beneficial for range estimation.

a region usually appears when  $B$  is not too large and the SBR is sufficiently high. A potential reduction in depth error variance was likewise noted by Heide et al. [86] for a synchronous SPL system, which our analysis shows is not a necessary condition for dead time to be beneficial. When  $t_r$  is slightly larger than  $t_d$  (as for  $t_r = 80$  in Fig. 5.3), the signal-triggered gating extends the region in which dead time is beneficial to larger  $B$  compared to the cases where  $t_r$  is much larger than  $t_d$ . Together with the plots in Fig. 5.2, this suggests that the most photon-efficient benefit from dead time is achieved when  $t_r$  is slightly larger than  $t_d$ .<sup>1</sup> This condition may be difficult to achieve in practice as the dead time is not tunable in many devices and adjustment of the illumination period is limited by the required maximum unambiguous range.

## 5.4 Acquisition Parameter Estimation

In this section, we first derive a maximum likelihood (ML) estimator for estimating the total flux  $\Lambda$  from the absolute detection times. We then describe how the background

<sup>1</sup>Note that more photons would be detected with shorter  $t_d$ , but each detection would likely be less informative of the depth.

flux  $B$  can similarly be estimated from a calibration measurement, and the signal flux  $S$  can be determined from the difference of the total and background flux estimates.

#### 5.4.1 Maximum Likelihood Estimator for $\Lambda$

Isbaner et al. [94] note the necessity of estimating  $\Lambda$  in order to correctly reconstruct the histogram of photon arrival times. Define the interdetection period  $R_\ell$  as the number of completed periods after the detector reset at  $T_\ell + \mathbf{t}_d$  before another photon is detected at time  $T_{\ell+1}$ :

$$R_\ell := \left\lfloor \frac{T_{\ell+1} - (T_\ell + \mathbf{t}_d)}{\mathbf{t}_r} \right\rfloor, \quad (5.8)$$

where  $T_\ell$ 's are absolute detection times. Isbaner et al. claim that  $P(R_\ell = r) \propto \exp(-r\Lambda)$  and use weighted least squares to fit an exponential function. In the following proposition, we verify the claim using properties of Poisson processes and the Markov nature of detections with dead time. Moreover, we show that  $R_\ell$ 's are independent, and so an ML estimator for  $\Lambda$  can be easily computed from a realization of  $R_\ell$ 's.

**Proposition 2.** *The random variables  $R_\ell$ 's defined in (5.8) are i.i.d. with the same probability distribution as  $R$ , where*

$$P(R = r) = (1 - \exp(-\Lambda)) \exp(-r\Lambda), \quad r \in \{0\} \cup \mathbb{N}. \quad (5.9)$$

*Proof.* See Appendix 5.C. □

By Proposition 2, given a realization of interdetection periods  $\{r_\ell\}_{1 \leq \ell \leq k}$ , the log-likelihood function is

$$\mathcal{L}(\{r_\ell\}_{\ell=1}^k; \Lambda) = -\Lambda \sum_{\ell=1}^k r_\ell + k \ln(1 - \exp(-\Lambda)).$$

Setting the derivative of  $\mathcal{L}(\{r_\ell\}_{\ell=1}^k; \Lambda)$  with respect to  $\Lambda$  to zero, we obtain the ML

estimator for  $\Lambda$  as

$$\widehat{\Lambda}^{\text{ML}} = -\ln \left( \frac{\sum_{\ell=1}^k r_{\ell}}{k + \sum_{\ell=1}^k r_{\ell}} \right). \quad (5.10)$$

Note that the distribution of  $R$  can be understood as follows. The number of photon arrivals per period is Poisson with parameter  $\Lambda$ , so  $p = 1 - \exp(-\Lambda)$  is the probability of at least one photon arriving in a period. Then  $R$  has a geometric distribution  $P(R = r) = (1 - p)^r p$ , which matches (5.9).

#### 5.4.2 Maximum Likelihood Estimator for $B$

We begin by assuming that background calibration measurements can occasionally be made within the ranging process, for which photons are detected while the laser is turned off. For sequences of ranging measurements such as in raster-scanned 3D imaging, such background-only acquisitions could be made for each laser position or for sets of laser points (e.g., once per row or once per image). Since the background process is homogeneous with  $\lambda(t) = \lambda^b$ , we can rewrite (5.3) as

$$f_{T_{\ell+1}|T_{\ell}}(t|t_{\ell}) = \lambda^b \exp(-\lambda^b(t - (t_{\ell} + \mathbf{t}_d))) \mathbb{I}\{t > t_{\ell} + \mathbf{t}_d\}.$$

Then the conditional distribution of  $T_{\ell}$ 's given  $T_1 = t_1$  is

$$\begin{aligned} f_{T_2, \dots, T_k | T_1}(t_2, \dots, t_k | t_1) &= \prod_{\ell=1}^{k-1} f_{T_{\ell+1}|T_{\ell}}(t_{\ell+1} | t_{\ell}) \\ &= \lambda_b^{k-1} \exp(-\lambda^b(t_k - t_1) + (k-1)\lambda^b \mathbf{t}_d). \end{aligned}$$

Given a set of absolute detection times  $\{t_{\ell}\}_{\ell=1}^k$ , the (conditional) log-likelihood function  $\ln(f_{T_2, \dots, T_k | T_1})$  is

$$\mathcal{L}(\{t_{\ell}\}_{\ell=1}^k; \lambda^b) = (k-1) \ln(\lambda^b) - \lambda^b(t_k - t_1) + (k-1)\lambda^b \mathbf{t}_d.$$

Setting the derivative of  $\mathcal{L}(\{t_\ell\}_{\ell=1}^k; \lambda^b)$  with respect to  $\lambda^b$  to zero, we obtain the (conditional) ML estimator for  $\lambda^b$  as

$$\hat{\lambda}_b^{\text{ML}} = \frac{k-1}{(t_k - t_1) - (k-1)t_d}.$$

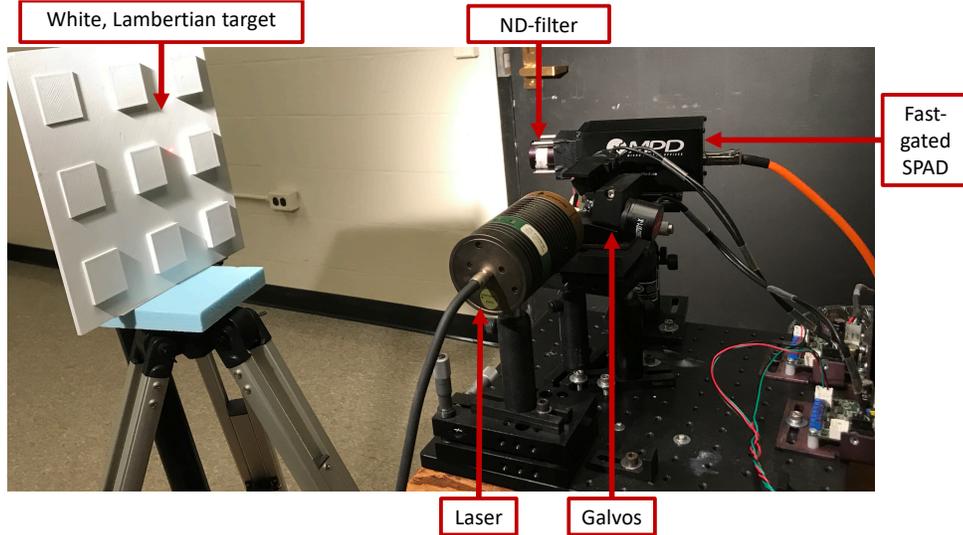
It follows that  $\hat{B}^{\text{ML}} = \hat{\lambda}_b^{\text{ML}} t_r$ .

### 5.4.3 Estimating $S$

From  $\hat{\Lambda}^{\text{ML}}$  and  $\hat{B}^{\text{ML}}$ , we could also compute  $\hat{S} = \max(\hat{\Lambda}^{\text{ML}} - \hat{B}^{\text{ML}}, 0)$  to ensure non-negativity. However, setting  $\hat{S} = 0$  whenever  $\hat{B}^{\text{ML}} > \hat{\Lambda}^{\text{ML}}$  is not informative for depth estimation with the log-matched filter, since  $f_{X_D}$  and  $f_{X_A}$  would be uniform PDFs. Instead, one can assume that there is always at least some small amount of signal and background in the ranging process, so we choose to set minimum values of  $S_{\min} = B_{\min} = 0.01$ . Then  $\hat{B} = \max(\hat{B}^{\text{ML}}, B_{\min})$ ,  $\hat{\Lambda} = \max(\hat{\Lambda}^{\text{ML}}, \hat{B} + S_{\min})$ , and  $\hat{S} = \hat{\Lambda} - \hat{B}$ .

## 5.5 Experimental Verification

We next evaluate the correctness of our model with experimental data from the TC-SPC system shown in Figure 5.4. The SPL system uses the HydraHarp 400 TCSPC module (PicoQuant) with dead time  $t_e \approx 80$  ns, and a fast-gated SPAD detector module (Micro Photon Devices), which has an adjustable hold-off time between 48 ns and 1  $\mu$ s. The illumination source, a pulsed diode laser (PicoQuant LDH-series) at 640 nm and with FWHM pulse duration around 100 ps, was aimed at a Lambertian white target at a fixed distance of around 50 cm. A distortion-free pulse shape calibration was acquired with an OD 3.0 neutral density (ND) filter, and high-flux measurements were acquired with no attenuation (OD 0), while the hold-off time  $t_{ho}$  was varied between 81 and 198 ns. In each case, we assumed that  $t_d \approx t_{ho} + 2$  ns and



**Figure 5.4:** A photograph of the bistatic single-photon lidar implementation at MIT in Cambridge, MA. The fast-gated SPAD has a variable hold-off time, which enables validation of the probabilistic model for various dead times.

$t_e = 80$  ns, thus the dataset included only cases with  $t_d > t_e$ . A high-flux dataset was acquired while the laser was turned off to serve as a background calibration.

The total  $\hat{\Lambda}^{\text{HF}}$  and calibrated background  $\hat{B}^{\text{HF}}$  flux values were estimated using the method proposed in Section 5.4. We attempted to perform the same validation as with simulated data using  $\hat{S}^{\text{HF}} = \hat{\Lambda}^{\text{HF}} - \hat{B}^{\text{HF}}$ . However, the background calibration  $\hat{B}^{\text{HF}}$  does not lead to a good fit between the measured histogram and the predicted effects of dead time for the calibrated low-flux pulse shape. There are several factors not included in the acquisition model that likely contribute to errors in the experimental results:

- **Filter spectral response:** The absorptive ND filters (ThorLabs NEK01) used in the experimental measurements have a fairly limited neutral region of 400-650 nm, so background outside of this passband is attenuated by varying amounts, and the relative attenuation of that background is not consistent for ND filters with different OD values. It is possible that reflective ND filters sets (e.g.,

ThorLabs NDK01) specified as spectrally flat for a greater range of wavelengths would reduce this source of inconsistency.

- **Dark counts:** Our modeling assumes all photons not due to signal can be grouped into a homogeneous “background” process, but attenuation of the incident flux only affects background counts from ambient light, with no effect on dark counts from thermally-generated carriers. While the dark count rate is often quite low for SPADs (on the order of 100 counts per second), for large attenuation factors applied to the ambient light, extrapolation of the high-flux background intensity from a low-flux measurement is not accurately computed.
- **Afterpulsing:** Unlike dark counts, afterpulses are false detection events correlated with previous detection times [42]. SPADs are typically held off long enough for the probability of an afterpulse detection to be sufficiently small and uncorrelated in time, so they simply appear as an increase in the background as the flux increases [93]. For shorter hold-off times, the afterpulse time correlation with respect to the most recent avalanche time is non-negligible, especially at high signal flux due to the periodicity of the signal component.

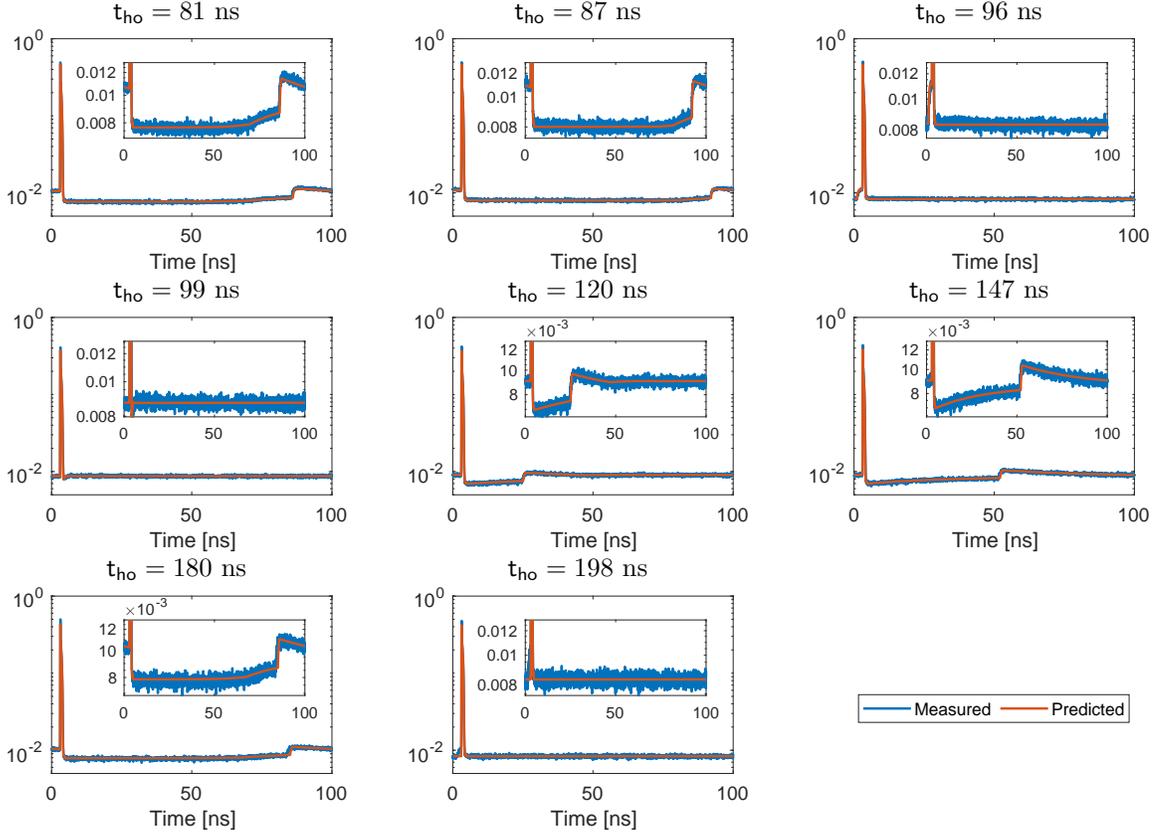
Instead of using poorly-calibrated background values, we choose an approximate background value  $\tilde{B}$  that minimizes the Kolmogorov-Smirnov statistic between the empirical and predicted CDFs [133, Chapter X.2]. We define the empirical CDF from  $k$  detections as

$$F_k(x) = \frac{1}{k} \sum_{\ell=1}^k X_{\ell} \mathbb{I}\{(-\infty, x]\} \quad (5.11)$$

and the predicted CDF

$$F_{X_D}(x; S, B) = \int_{-\infty}^x f_{X_D}(y; S, B) \, dy, \quad (5.12)$$

which uses the predicted detection time PDF given a calibrated signal pulse shape



**Figure 5-5:** Results are shown from high-flux experiments with fixed laser power while varying the SPAD hold-off time. Note that because of inaccuracy in estimating the background flux, the values of  $\tilde{B}$  were set to minimize the Kolmogorov-Smirnov statistic between the predicted and measured PDFs, with  $\hat{S} = \hat{\Lambda}^{\text{HF}} - \tilde{B}$ . The insets show in detail how the change in detection time histogram shape due to dead time is accurately predicted.

with flux  $S$  and a uniform background with flux  $B$ . Then the approximate background flux  $\tilde{B}$  is chosen by performing the following minimization via grid search:

$$\tilde{B} = \arg \min_B \sup_x |F_k(x) - F(x; \hat{\Lambda}^{\text{HF}} - B, B)|. \quad (5.13)$$

Fig. 5-5 shows the close match between the measured and predicted PDFs for a variety of  $t_{\text{ho}}$  values. The insets in the plots highlight how the shape of even the lower-probability regions of the measured PDFs is predicted using our Markov chain

analysis, including the nearly distortion-free special case when  $\mathbf{t}_d$  is an integer multiple of  $\mathbf{t}_r$  as pointed out by [46]. The plots emphasize how the effect of dead times on the detection time distribution is sensitive to many parameters:  $\mathbf{t}_d$ ,  $\mathbf{t}_r$ ,  $S$ , and  $B$ .

## 5.6 Arrival Intensity Estimation Algorithm

Now that we have verified that the Markov chain model indeed represents the behavior of TCSPC systems, we consider several ways of using this modeling to reduce the effects of distortions due to dead time. In Section 5.2.3, we discussed that one common approach to handling dead times is to attempt to correct a histogram so that it represents the photon arrival intensity. In this section, we thus use the Markov chain model to derive an algorithm for estimating the arrival intensity  $\lambda(x)$  for  $x \in [0, \mathbf{t}_r)$  from a histogram of detection times assuming that  $\Lambda$  is known. The algorithm may be implemented with a calibrated  $\Lambda$  when available or with a  $\Lambda$  estimated by, for example, the ML estimator. In what follows, we consider estimating the arrival intensity  $\lambda$  from a detection time histogram assuming that  $\Lambda$  is known. Note that while one could include regularizers to reflect prior knowledge about  $\lambda$ , the goal here is to demonstrate that the proposed method can reconstruct the arrival intensity without any prior knowledge other than the intensity being non-negative. The histogram correction method is thus applicable to a broad class of applications, including FLIM, multi-depth lidar, or non-line-of-sight imaging, in which the arrival intensity is less predictable than SPL with exactly one depth per pixel.

### 5.6.1 Relationship between Arrival and Detection Distributions

We begin by deriving the relationship between  $\lambda(x)$  and  $f_{X_D}(x)$ . Plugging the Markov chain transition PDF  $f_{X_{\ell+1}|X_\ell}(x_{\ell+1}|x_\ell)$  (5.5) into the definition of the stationary PDF

$f_{X_D}(x)$  (5.6), we have

$$f_{X_D}(x) = \lambda(x) \left[ \int_0^{x-x_d} f_{X_D}(y) \frac{\exp\left(-\int_{y+x_d}^x \lambda(\tau) d\tau\right)}{1 - \exp(-\Lambda)} dy + \int_{x-x_d}^{t_r} f_{X_D}(y) \frac{\exp\left(-\int_{y+x_d}^{t_r+x} \lambda(\tau) d\tau\right)}{1 - \exp(-\Lambda)} dy \right] \quad (5.14)$$

for  $x > x_d$ , and

$$f_{X_D}(x) = \lambda(x) \left[ \int_0^{t_r+x-x_d} f_{X_D}(y) \frac{\exp\left(-\int_{y+x_d}^{t_r+x} \lambda(\tau) d\tau\right)}{1 - \exp(-\Lambda)} dy + \int_{t_r+x-x_d}^{t_r} f_{X_D}(y) \frac{\exp\left(-\int_{y+x_d}^{2t_r+x} \lambda(\tau) d\tau\right)}{1 - \exp(-\Lambda)} dy \right] \quad (5.15)$$

for  $x \leq x_d$ . In (5.14) and (5.15), denote the factors in the brackets as  $a(x)$  and we can then write  $f_{X_D}(x) = \lambda(x)a(x)$ , where  $a(x)$  can be interpreted as the attenuation effect on the arrival intensity due to dead time. It is worth mentioning that similar factorization of  $f_{X_D}$  was also used by Isbaner et al. [94] for the derivation of their dead time correction algorithm. However, such a factorization is assumed at the beginning of their derivation, whereas we arrive at this factorization naturally from the stationary condition of a Markov chain.

Plugging  $f_{X_D}(x) = \lambda(x)a(x)$  into (5.14), we have

$$a(x) = \int_0^{x-x_d} \lambda(y)a(y) \frac{\exp\left(-\int_{y+x_d}^x \lambda(\tau) d\tau\right)}{1 - \exp(-\Lambda)} dy + \int_{x-x_d}^{t_r} \lambda(y)a(y) \frac{\exp\left(-\int_{y+x_d}^{t_r+x} \lambda(\tau) d\tau\right)}{1 - \exp(-\Lambda)} dy. \quad (5.16)$$

Differentiating both sides of the above equation with respect to  $x$ :

$$\begin{aligned}
a'(x) &\stackrel{(a)}{=} \frac{1}{1 - \exp(-\Lambda)} \left[ \lambda(x - \mathbf{x}_d) a(x - \mathbf{x}_d) (1 - \exp(-\Lambda)) \right. \\
&\quad - \lambda(x) \left( \int_0^{x - \mathbf{x}_d} \lambda(y) a(y) \exp\left(-\int_{y + \mathbf{x}_d}^x \lambda(\tau) d\tau\right) dy \right. \\
&\quad \left. \left. + \int_{x - \mathbf{x}_d}^{\mathbf{t}_r} \lambda(y) a(y) \exp\left(-\int_{y + \mathbf{x}_d}^{\mathbf{t}_r + x} \lambda(\tau) d\tau\right) dy \right) \right] \\
&\stackrel{(b)}{=} \lambda(x - \mathbf{x}_d) a(x - \mathbf{x}_d) - \lambda(x) a(x), \tag{5.17}
\end{aligned}$$

where step (a) uses the Leibniz rule and the fact that  $\lambda(\mathbf{t}_r + x) = \lambda(x)$  and step (b) follows by noticing from (5.16) that the sum of the two integrals equals  $(1 - \exp(-\Lambda))a(x)$ . Similarly, we can obtain from (5.15) that

$$a'(x) = \lambda(\mathbf{t}_r + x - \mathbf{x}_d) a(\mathbf{t}_r + x - \mathbf{x}_d) - \lambda(x) a(x). \tag{5.18}$$

Note that if we consider periodic extensions of  $a(x)$  and  $f_{X_D}(x)$ , then (5.17) and (5.18) are identical. In the following,  $a(x)$  and  $f_{X_D}(x)$  are considered as their periodic extensions.

Integrating both sides of (5.17), we have that

$$a(x) = - \int_{x - \mathbf{x}_d}^x \lambda(\tau) a(\tau) d\tau + C, \tag{5.19}$$

where  $C$  is a constant. Multiplying both sides of (5.19) by  $\lambda(x)$ , we have that

$$f_{X_D}(x) = -\lambda(x) \int_{x - \mathbf{x}_d}^x f_{X_D}(\tau) d\tau + C\lambda(x). \tag{5.20}$$

Define

$$g(x) := \int_{x-x_d}^x f_{X_D}(\tau) \, d\tau. \quad (5.21)$$

Since  $f_{X_D}(x)$  is a proper probability density function on the interval  $[0, \mathbf{t}_r)$ , it satisfies

$$1 = \int_0^{\mathbf{t}_r} f_{X_D}(x) \, dx = - \int_0^{\mathbf{t}_r} \lambda(x)g(x) \, dx + C \int_0^{\mathbf{t}_r} \lambda(x) \, dx.$$

It follows that

$$C = \frac{1 + \int_0^{\mathbf{t}_r} \lambda(x)g(x) \, dx}{\int_0^{\mathbf{t}_r} \lambda(x) \, dx} = \frac{1 + \int_0^{\mathbf{t}_r} \lambda(x)g(x) \, dx}{\Lambda}. \quad (5.22)$$

Finally, by plugging (5.21) and (5.22) into (5.20), we can establish the following relationship between the arrival intensity function  $\lambda(x)$  and the limiting distribution of the detection times  $f_{X_D}(x)$ :

$$f_{X_D}(x) = -\lambda(x)g(x) + \frac{1 + \int_0^{\mathbf{t}_r} \lambda(x)g(x) \, dx}{\Lambda} \lambda(x). \quad (5.23)$$

### 5.6.2 Nonlinear Inverse Formulation and Algorithm

We next consider inversion of the relationship between  $\lambda(x)$  and  $f_{X_D}(x)$ . Since the measured TCSPC data is in the form of a discrete-time histogram, we attempt only the recovery of a discrete approximation of  $\lambda(t)$ . Consider again that the time interval  $[0, \mathbf{t}_r)$  is partitioned into  $n_b$  equally spaced time bins with bin size  $\Delta$ . Define  $n_d := x_d/\Delta$ . Let the normalized histogram of detection times be denoted by  $\mathbf{h} = (h_1, \dots, h_{n_b})$ , where  $\sum_{m=1}^{n_b} h_m = 1$ . A discrete model for (5.23) is then

$$\mathbf{h} = -\text{diag}(\mathbf{g})\boldsymbol{\lambda} + \Lambda^{-1}\boldsymbol{\lambda} + \Lambda^{-1}(\mathbf{g}^T\boldsymbol{\lambda})\boldsymbol{\lambda} + \boldsymbol{\varepsilon},$$

where  $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_{n_b})$  is a discretization of  $\lambda(t)$ ;  $\Lambda$  is the total flux and is assumed to be known;  $g_n = \sum_{m=n-n_d}^{n-1} h_m$  for  $n > n_d$  and  $g_n = \sum_{m=n-n_d+n_b}^{n_b} h_m + \sum_{m=1}^{n-1} h_m$  for  $n \leq n_d$ , which follows from (5.21);  $\text{diag}(\mathbf{g})$  is a diagonal matrix with  $\mathbf{g}$  on its

diagonal; and  $\varepsilon$  represents the error due to discretization and the difference between the finite-sample empirical distribution and the limiting distribution. For any fixed  $\mathbf{h}$  (hence fixed  $\mathbf{g}$ ), define an operator  $\mathcal{T}(\cdot; \mathbf{h}) : \mathbb{R}^{n_b} \rightarrow \mathbb{R}^{n_b}$  as

$$\boldsymbol{\lambda} \mapsto \mathcal{T}(\boldsymbol{\lambda}; \mathbf{h}) := -\text{diag}(\mathbf{g})\boldsymbol{\lambda} + \Lambda^{-1}\boldsymbol{\lambda} + \Lambda^{-1}(\mathbf{g}^T\boldsymbol{\lambda})\boldsymbol{\lambda}. \quad (5.24)$$

The inverse problem that we need to solve is then to estimate  $\boldsymbol{\lambda}$  from the nonlinear system  $\mathbf{h} = \mathcal{T}(\boldsymbol{\lambda}; \mathbf{h}) + \varepsilon$  given a measurement vector  $\mathbf{h}$ . Define the optimization problem:

$$\min_{\boldsymbol{\lambda}} \{F(\boldsymbol{\lambda}) := D(\boldsymbol{\lambda}) + \delta_{[0, M]^{n_b}}(\boldsymbol{\lambda})\}, \quad (5.25)$$

where  $D(\boldsymbol{\lambda}) := \frac{1}{2}\|\mathbf{h} - \mathcal{T}(\boldsymbol{\lambda}; \mathbf{h})\|^2$  with  $\|\cdot\|$  being the Euclidean norm and  $\delta_{[0, M]^{n_b}}$  the indicator function of the bounded hypercube  $[0, M]^{n_b}$  for some constant  $M$ .

We use a monotone accelerated proximal gradient (APG) algorithm [115] to solve (5.25). Note that the proximal operator for  $\delta_{[0, M]^{n_b}}$  is the orthogonal projector onto  $[0, M]^{n_b}$ , denoted by  $\Pi_{[0, M]^{n_b}}(\cdot)$ , and the gradient of  $D(\boldsymbol{\lambda})$  is computed as follows:

$$\begin{aligned} \nabla D(\boldsymbol{\lambda}) &= \mathbf{J}_{\mathcal{T}}^T(\mathcal{T}(\boldsymbol{\lambda}; \mathbf{h}) - \mathbf{h}) \\ &= \left( \frac{\mathbf{g}\boldsymbol{\lambda}^T}{\Lambda} + \frac{1 + \mathbf{g}^T\boldsymbol{\lambda}}{\Lambda}\mathbf{I} - \text{diag}(\mathbf{g}) \right) (\mathcal{T}(\boldsymbol{\lambda}; \mathbf{h}) - \mathbf{h}), \end{aligned} \quad (5.26)$$

where  $\mathbf{J}_{\mathcal{T}}$  is the Jacobian matrix of  $\mathcal{T}$  and  $\mathbf{I}$  is the identity matrix. We emphasize that  $\mathbf{h}$ ,  $\mathbf{g}$ , and  $\Lambda$  are fixed throughout the algorithm, thus they are treated as constant instead of functions of  $\boldsymbol{\lambda}$  when computing the gradient  $\nabla D(\boldsymbol{\lambda})$  in (5.26). The convergence of the monotone APG algorithm relies on an appropriate choice of the step size  $\gamma$ , which should satisfy  $\gamma < 1/L$ , where  $L$  is the Lipschitz constant of  $\nabla D(\cdot)$  [115]. The following proposition provides an upper-bound  $L_u$  for  $L$ .

**Proposition 3.** *The Lipschitz constant  $L$  of the function  $\nabla D(\cdot)$  defined in (5.26) is*

upper-bounded by  $L_u$  on  $[0, M]^{\mathfrak{n}_b}$ , where  $L_u$  is defined as

$$L_u := 2\Lambda^{-2}\mathfrak{n}_bM^2 + (2\Lambda^{-2} + 2 + 6\Lambda^{-1})\sqrt{\mathfrak{n}_b}M + 4\Lambda^{-1} + 2.$$

*Proof.* See Appendix 5.D. □

Setting the step size  $\gamma = 1/L_u$ , starting with some initialization  $\boldsymbol{\lambda}^0 = \boldsymbol{\lambda}^1 = \mathbf{z}^1 \in [0, M]^{\mathfrak{n}_b}$  and  $q_0 = 0, q_1 = 1$ , for  $k \geq 1$ , the monotone APG algorithm for solving (5.25) proceeds as follows:

$$\begin{aligned} \mathbf{y}^k &= \boldsymbol{\lambda}^k + \frac{q_{k-1}}{q_k}(\mathbf{z}^k - \boldsymbol{\lambda}^k) + \frac{q_{k-1} - 1}{q_k}(\boldsymbol{\lambda}^k - \boldsymbol{\lambda}^{k-1}), \\ \mathbf{z}^{k+1} &= \Pi_{[0, M]^{\mathfrak{n}_b}}(\mathbf{y}^k - \gamma \nabla D(\mathbf{y}^k)), \\ \mathbf{x}^{k+1} &= \Pi_{[0, M]^{\mathfrak{n}_b}}(\boldsymbol{\lambda}^k - \gamma \nabla D(\boldsymbol{\lambda}^k)), \\ q_{k+1} &= \frac{\sqrt{4q_k^2 + 1} + 1}{2}, \\ \boldsymbol{\lambda}^{k+1} &= \begin{cases} \mathbf{z}^{k+1}, & \text{if } F(\mathbf{z}^{k+1}) \leq F(\mathbf{x}^{k+1}); \\ \mathbf{x}^{k+1}, & \text{otherwise.} \end{cases} \end{aligned} \tag{5.27}$$

Since (5.25) is a nonconvex optimization problem, a good initialization is important to avoid converging to local minima that are not global minima. We now introduce an initialization scheme. Let  $C_\lambda$  be a scalar that depends on  $\lambda$  through  $C_\lambda = \int_0^{\text{tr}} \lambda(x)g(x) \, dx$ . Then (5.23) can be written as  $f_{X_D}(x) = -\lambda(x)g(x) + (1 + C_\lambda)\lambda(x)/\Lambda$ , which implies

$$\lambda(x) = \frac{f_{X_D}(x)}{(1 + C_\lambda)/\Lambda - g(x)}. \tag{5.28}$$

Plugging (5.28) back into the definition of  $C_\lambda$ , we obtain a fixed point equation for  $C_\lambda$ :

$$C_\lambda = \int_0^{\text{tr}} \frac{f_{X_D}(x)g(x)}{(1 + C_\lambda)/\Lambda - g(x)} \, dx. \tag{5.29}$$

Notice that by (5.28) the feasible set for  $C_\lambda$  is  $\mathcal{C} = \{C \in \mathbb{R} : (1 + C)/\Lambda - g(x) >$

$0, \forall x \in [0, t_r)\}$ . It follows that the RHS of (5.29) is positive on  $\mathcal{C}$  and monotone decreasing to zero as  $C_\lambda$  goes to infinity. Since the LHS of (5.29) is linearly increasing, a graph will easily show that (5.29) has a unique fixed point on  $\mathcal{C}$ . Therefore, if  $f_{X_D}$  is known perfectly, then one can solve (5.29) for  $C_\lambda$  and plug  $C_\lambda$  into (5.28) to have a perfect reconstruction of  $\lambda$ . However, in practice, we only have a histogram formed by a limited number of measured detection times. Nevertheless, it is plausible to estimate  $C_\lambda$  as the fixed point of

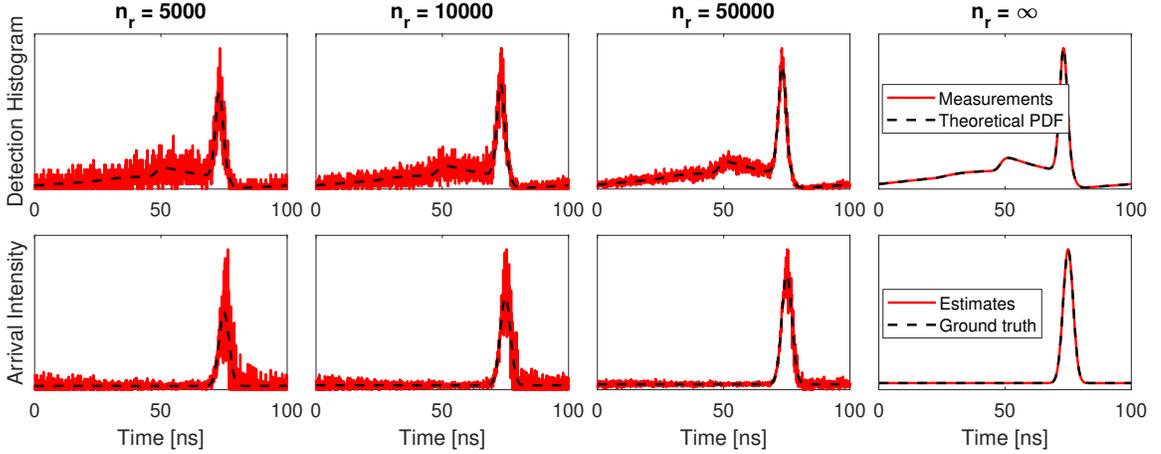
$$\widehat{C}_\lambda = \sum_{m=1}^{n_b} \frac{h_m g_m}{(1 + \widehat{C}_\lambda)/\Lambda - g_m}. \quad (5.30)$$

We can then use  $\boldsymbol{\lambda}^0$  with the  $n^{\text{th}}$  entry being defined as

$$\lambda_m^0 = \frac{h_m}{(1 + \widehat{C}_\lambda)/\Lambda - g_m} \quad (5.31)$$

as the initialization of the nonconvex optimization problem. While we do not have a theoretical guarantee for convergence to the global minimum, simulations with both  $\mathbf{h}$  plus random perturbation and the more principled initialization (5.31) lead to good estimates. Because solving (5.30) is easy and the initial estimate (5.31) is usually close to the solution, using the principled initialization allows the algorithm to converge faster.

Fig. 5-6 presents simulated detection histograms and the corresponding arrival intensity estimates using (5.27), where  $S = B = 3.16$  and  $\sigma_p = 2$  ns. We notice that as  $n_r$  increases, the detection histogram approaches  $f_{X_D}$ . Our estimated arrival intensity likewise approaches the true arrival intensity as  $n_r$  increases. It is interesting to note that while the error in the detection histogram resembles Poisson noise in that the variance increases as the mean increases, the error in the arrival intensity estimate is signal-dependent in a different way. We observe that the error variance is roughly proportional to the pointwise ratio of  $f_{X_D}$  and  $\lambda$ . Although we have no theoretical



**Figure 5.6:** Estimation of arrival intensity (bottom) from detection histogram (top) when  $S = B = 3.16$ ,  $\sigma_p = 2$  ns,  $\Delta = 50$  ps,  $t_r = 100$  ns, and  $t_d = 75$  ns. From left to right: increased number of illuminations ( $n_r$ ), where in the last column, theoretical detection histogram is used as measurement.

results supporting this hypothesis, the observation suggests that the portions of the arrival intensity easiest to reconstruct are those least attenuated by the dead time effects, and vice versa.

## 5.7 Application to Ranging

We now explore how the theory and algorithm developed in Sections 5.3 and 5.6 can be used for depth estimation. In Section 5.7.1, we assume that the acquisition parameters  $S, B, \Lambda$  are known from accurate calibration and compare different methods using the true parameter values. In Section 5.7.2, we use the estimators for  $\Lambda, B$  and  $S$  introduced in Section 5.4 to test our proposed methods using estimated parameters.

### 5.7.1 Ranging with True Acquisition Parameters

The ML depth estimator for the Poisson arrival process passes the set of arrivals through a log-matched filter (LMF) that is matched to the arrival intensity  $\lambda(t)$ , where the LMF is defined as  $v(t) := \log(\lambda(t)) = \log(f_{X_\Lambda}(t)) + \log(\Lambda)$  [11, Eqn. (33)].

Since photon loss due to dead time is negligible for low-flux acquisitions, the filter matched to the arrival PDF  $f_{X_A}$  is thus the ML depth estimator given a set of low-flux relative detection times  $\{x_\ell\}_{\ell=1}^k$ . The LMF for estimating the depth  $z$  is thus defined as

$$\begin{aligned}\widehat{z}(\{x_\ell\}_{\ell=1}^k; f_{X_A}) &:= \arg \max_z \sum_{\ell=1}^k \int_0^{t_r} \delta(t - x_\ell) v(t + 2z/c) dt \\ &= \arg \max_z \sum_{\ell=1}^k \log(f_{X_A}(x_\ell + 2z/c)).\end{aligned}$$

In practical implementations,  $\{x_\ell\}_{\ell=1}^k$  is quantized into  $n_b$  equally spaced time bins over  $[0, t_r)$  with bin centers  $\{b_m\}_{m=1}^{n_b}$ . A histogram  $\mathbf{h} = (h_1, \dots, h_{n_b})$  for the low-flux detection times can then be obtained from the quantized data  $\{\bar{x}_\ell\}_{\ell=1}^k$ . Moreover, instead of estimating the depth  $z$ , we can estimate the time delay  $\tau := 2z/c$ , since the mapping from  $z$  to  $\tau$  is one-to-one. The estimator for  $\tau$  is then

$$\widehat{\tau}(\mathbf{h}; f_{X_A}) := \arg \max_{\tau \in \Gamma} \left\{ \sum_{m=1}^{n_b} h_m \log(f_{X_A}(b_m + \tau)) \right\} + b_{n_b/2},$$

where  $\Gamma := \{-b_{n_b/2}, \dots, -b_1, 0, b_1, \dots, b_{n_b/2}\}$  is a set of on-grid relative time delays, and  $f_{X_A}$  is the arrival PDF assuming  $b_{n_b/2}$  is the true delay.

In Section 5.3, we have derived the limit of the empirical distribution of high-flux detection times  $f_{X_D}$ . Hence, we can similarly define a log-matched filter matched to  $f_{X_D}$  and define an estimator as

$$\widehat{\tau}(\mathbf{h}; f_{X_D}) := \arg \max_{\tau \in \Gamma} \left\{ \sum_{m=1}^{n_b} h_m \log(f_{X_D}(b_m + \tau)) \right\} + b_{n_b/2},$$

which is preferable if  $\mathbf{h}$  is obtained via high-flux acquisition with non-negligible photon loss due to dead time. Note however that  $\widehat{\tau}(\mathbf{h}; f_{X_D})$  is not the ML estimator with dead time effects (even without quantization error), because in this case the joint PDF does

not factorize as product of the marginals. While one can obtain the exact joint PDF from the transition PDF (5.5) and the marginal PDF  $f_{X_D}$ , the true ML estimator is inconvenient to implement. Therefore,  $\hat{\tau}(\mathbf{h}; f_{X_D})$  is used in our simulations.

In Section 5.6, we have derived an algorithm for estimating the arrival time distribution from the detection time distribution. Hence, given a detection histogram, our algorithm can compute an estimate for the arrival histogram  $\hat{\mathbf{h}}^A$ , and then  $\tau(\hat{\mathbf{h}}^A; f_{X_A})$  can be used for depth estimation.

Based on the above discussion, letting  $\mathbf{h}^{\text{LF}}$  and  $\mathbf{h}^{\text{HF}}$  denote the detection time histograms obtained via low-flux and high-flux acquisitions, respectively, we compare six depth estimation methods applicable to asynchronous TCSPC systems. The methods are as follows:

1. **LF**: The low-flux approach first attenuates the incident flux (in practice by applying a neutral density filter) to limit the total flux arriving at the detector to 5% so that dead time effects can be ignored. Since the low-flux detection histogram  $\mathbf{h}^{\text{LF}}$  can then be considered to be the same as the arrival histogram, it then uses  $\hat{\tau}(\mathbf{h}^{\text{LF}}; f_{X_A})$  as the estimator.
2. **HF**: The high-flux method naïvely assumes that dead time has no effect on the acquisition and uses the estimator  $\hat{\tau}(\mathbf{h}^{\text{HF}}; f_{X_A})$ , even when  $\mathbf{h}^{\text{HF}}$  is not a good approximation to the arrival histogram.
3. **SC**: Shift correction assumes that the dead time only adds a bias to the estimate and that the bias can be computed and subsequently subtracted away. In practice, this is equivalent to the optical calibration procedure in [84]; for our simulations, we compute the shift in the mode of  $f_{X_D}$  compared to that of  $f_{X_A}$  and subtract the shift correction from the HF estimate.
4. **Isbaner**: This method, based on the work of Isbaner et al., first estimates the

arrival histogram  $\hat{\mathbf{h}}^A$  from  $\mathbf{h}^{\text{HF}}$  using the algorithm in [94], which has publicly available code,<sup>2</sup> and then applies the estimator  $\hat{\tau}(\hat{\mathbf{h}}^A; f_{X_A})$ . While [94] can estimate  $\Lambda$  from data, we provide the algorithm with the true  $\Lambda$  for fair comparison.

5. **Proposed method 1 – MCPDF:** Our first method computes the Markov chain-based PDF  $f_{X_D}$  to directly apply  $\hat{\tau}(\mathbf{h}^{\text{HF}}; f_{X_D})$ .
6. **Proposed method 2 – MCHC:** Our second method is similar to that of Isbaner et al., except it first estimates the arrival histogram  $\hat{\mathbf{h}}^A$  from  $\mathbf{h}^{\text{HF}}$  using the Markov chain-based histogram correction algorithm introduced in Section 5.6 and then uses  $\hat{\tau}(\hat{\mathbf{h}}^A; f_{X_A})$  as the estimator.

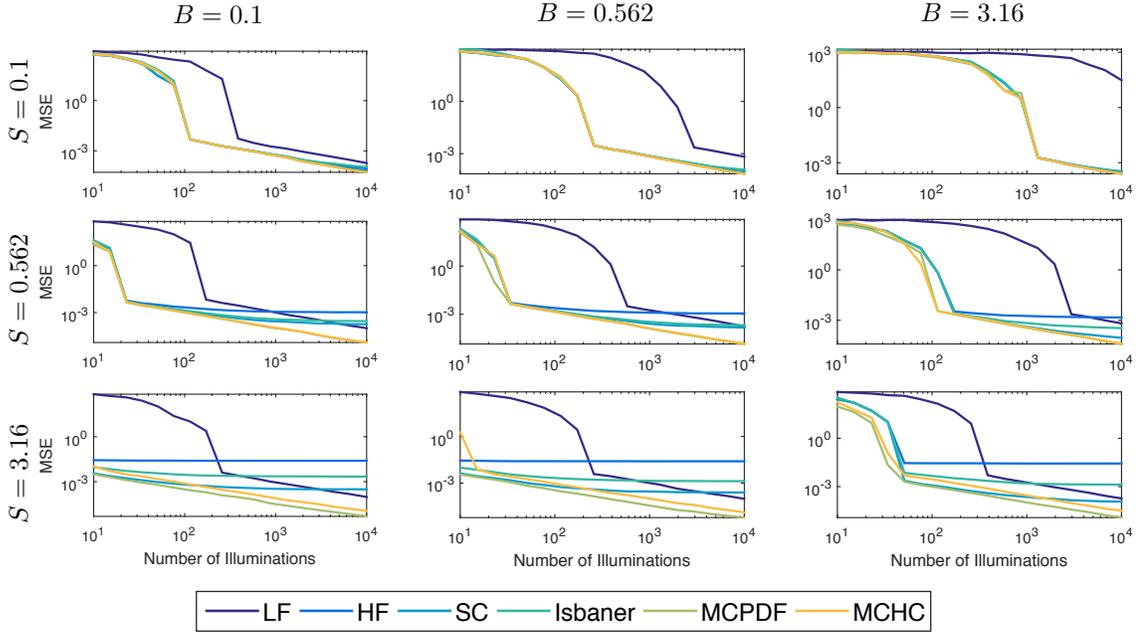
We perform Monte Carlo simulations with  $t_r = 100$  ns,  $t_d = 75$  ns,  $\sigma_p = 0.2$  ns, and bin duration  $\Delta = 5$  ps, which are reasonable experimental parameters for some laboratory settings. For each combination of  $S$  and  $B$ , we generate 600 realizations of the arrival process with  $n_r = 10^4$  illuminations. Starting with the first arrival, the high-flux detection sequence is generated by removing subsequent arrivals if they occur within  $t_d$  of the previous detection. Generation of the corresponding low-flux detection sequence proceeds in the same manner, but the arrival process is first attenuated via Bernoulli thinning, so photons arrive in only 5% of illumination periods on average. For each method, the log-matched filtering is performed via circular cross-correlation (circular convolution of the histogram with the time-reversed PDF). This is due to the asynchronous dead time preserving the shift invariance of the arrival process.

Fig. 5-7 compares the MSE for time delay estimation achieved by the six compared methods as a function of the number of illuminations.

We observe that MCPDF usually achieves the lowest MSE, since it directly performs parameter estimation with the updated detection model. Equivalently, MCPDF

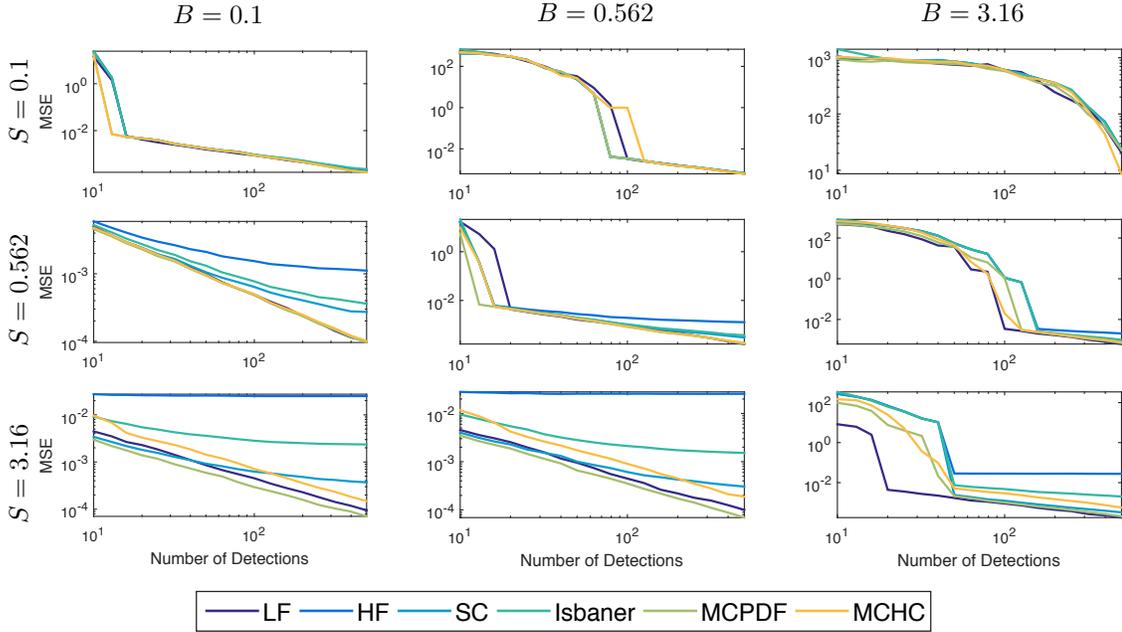
---

<sup>2</sup><http://projects.gwdg.de/projects/deadtimecorrectiontcspe>



**Figure 5-7:** Plots of the MSE for ranging as a function of  $n_r$  for  $t_r = 100$  ns,  $t_d = 75$  ns,  $\sigma = 0.2$  ns,  $\Delta = 5$  ps and various  $S$  and  $B$  values. Our proposed methods (MCPDF and MCHC) take advantage of the increased detection rate to perform more accurate ranging than with the low-flux acquisition for all values of  $S$ ,  $B$ , and  $n_r$ .

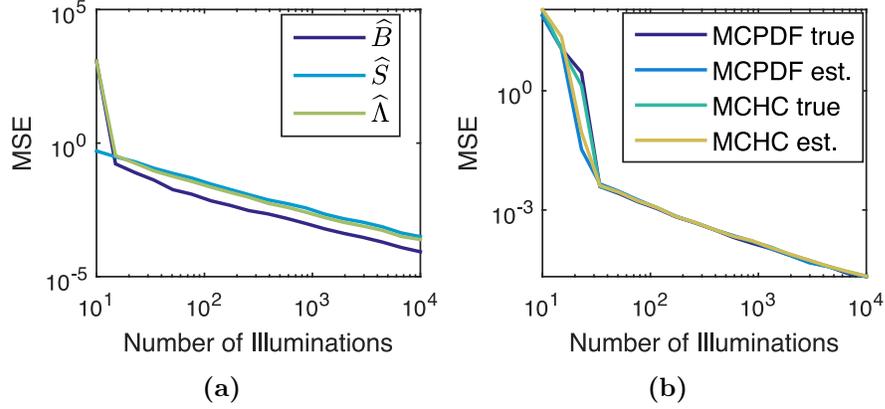
needs the fewest illuminations to achieve a given MSE, hence enabling the fastest acquisition. The MSE of MCHC is comparable to that of MCPDF, limited only in that it must first invert the histogram before estimating the depth. Compared to the LF approach, both MCPDF and MCHC require fewer illuminations to achieve the same MSE, and that time efficiency increases as  $S$  and  $B$  increase and dead time has a more significant impact. Regarding the other approaches, HF is more effective only for low numbers of illuminations but the estimate quickly becomes biased and is therefore not suited to precise depth measurement. Correcting for this bias with SC is quite effective for extending to somewhat higher  $n_r$ , although eventually more accurate modeling is necessary for more precise estimates. The state-of-the-art method for dealing with asynchronous dead-time models by Isbaner et al. [94] achieves low MSE when the total flux is low or moderate, while the accuracy degrades in high-flux



**Figure 5-8:** Plots of the MSE for ranging as a function of the number of detections for  $t_r = 100$  ns,  $t_d = 75$  ns,  $\sigma = 0.2$  ns,  $\Delta = 5$  ps and various  $S$  and  $B$  values. For high  $S$  and  $B$  not too large, the presence of dead time actually improves ranging accuracy with our MCPDF method relative to the low flux measurements due to the narrowing of the signal pulse.

scenarios. The performance degradation is due to their approximation of the detection time distribution being less accurate in high-flux settings (S. Isbaner, personal communication, May 14, 2018).

In addition to enabling faster acquisition, we explored whether dead time could lead to more accurate ranging for an equal number of detected photons. The Fisher information analysis in Section 5.3.2 has provided a theoretical prediction that for sufficiently high SBR, estimating depth from the dead time-distorted detection time distribution can yield lower MSE than that from the arrival time distribution. Although the estimators in our Monte Carlo simulation are not guaranteed to achieve the Cramér-Rao lower bound (i.e., the reciprocal of Fisher information), we would like to see whether the reduction of ranging error due to dead time also exists with



**Figure 5.9:** Results of ranging simulations performed with acquisition parameters estimated from the detection data, where the true values were  $S = B = 0.562$ ,  $t_r = 100$  ns,  $t_d = 75$  ns,  $\sigma_p = 0.2$  ns, and  $\Delta = 10$  ps. In (a), the estimates  $\widehat{B}$ ,  $\widehat{\Lambda}$ , and  $\widehat{S}$  improve as  $n_r$  increases. The ranging results in (b) using estimated parameters show no degradation in performance compared to the methods with parameters known *a priori*.

simple and commonly used estimators. Fig. 5.8 compares the MSE for time delay estimation by the six methods as a function of the number of detections. We notice that for the high SBR cases where  $S = 3.16, B = 0.1$  and  $S = 3.16, B = 0.562$ , MCPDF outperforms LF, which provides numerical evidence that dead time can be beneficial when properly modeled.

### 5.7.2 Ranging with Estimated Acquisition Parameters

The results in Section 5.7.1 use methods that compute  $f_{X_D}$ ,  $f_{X_A}$ , and  $\widehat{\mathbf{h}}^A$  assuming the true values of  $B$ ,  $S$ , and  $\Lambda$  are known. However, in most practical scenarios, this information will not be available *a priori*. We show here results using the parameter estimation strategies outline in 5.4.

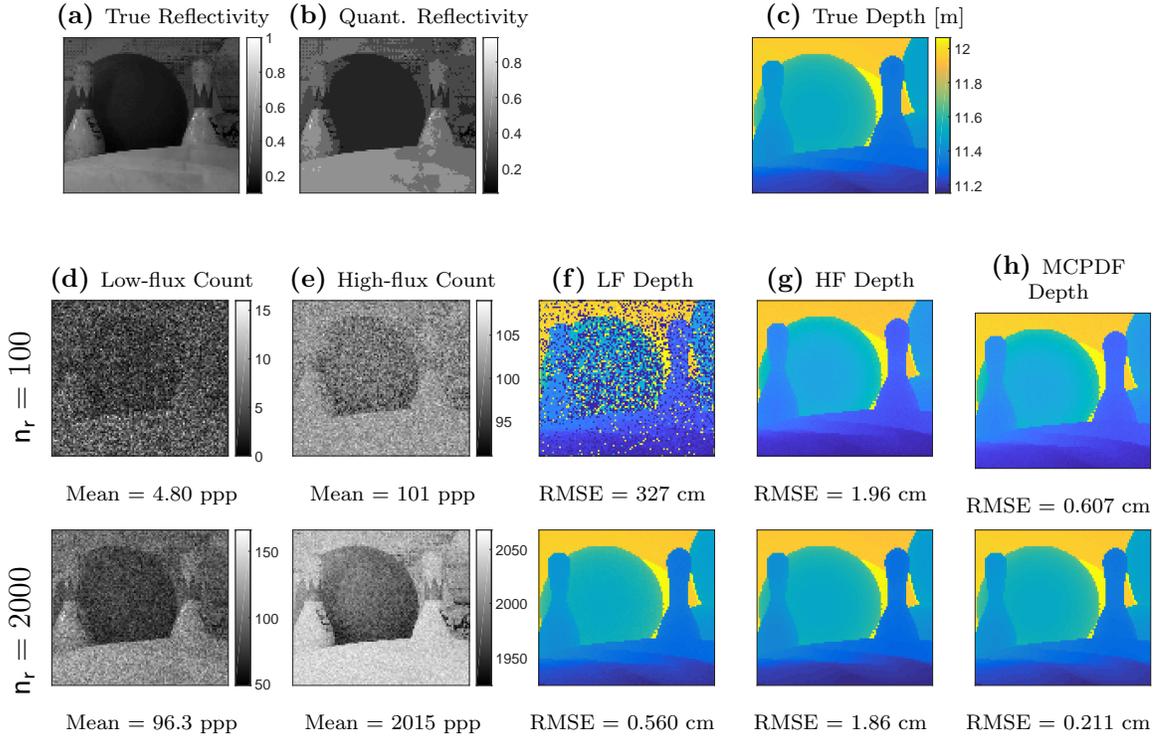
Fig. 5.9 shows one example of estimates using this strategy for 500 Monte Carlo trials with  $S = B = 0.562$ ,  $t_r = 100$  ns,  $t_d = 75$  ns,  $\sigma_p = 0.2$  ns, and  $\Delta = 10$  ps. In Fig. 5.9(a), the  $\widehat{B}$ ,  $\widehat{\Lambda}$ , and  $\widehat{S}$  estimates consistently improve as  $n_r$  increases beyond a

very small number of detections. The resulting depth estimates shown in Fig. 5-9(b) are virtually indistinguishable from the methods using the known parameter values. It is worth noting that, while  $\widehat{B}^{\text{ML}}$  applies for any value of  $B$ ,  $\widehat{\Lambda}^{\text{ML}}$  becomes less reliable for large  $\Lambda$  since  $P(R = 0)$  in (5.9) approaches unity. If the number of illuminations was not fixed in advance, one could pursue an adaptive acquisition strategy as in [129]. Alternatively, for 3D imaging, one could take advantage of spatial correlations to estimate  $\Lambda$ , for example, to borrow measurements from neighboring pixels.

### 5.7.3 Signal Quantization for High-Flux Depth Imaging

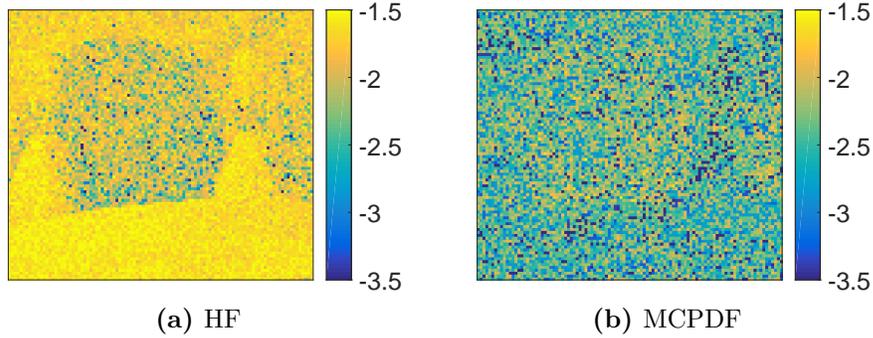
Regardless of whether  $S$  and  $B$  are considered to be known a priori or estimated from the data, scaling up from a single depth measurement to an entire depth map would require recomputing  $\mathbf{f}_{X_D}$  for each pixel. Since recomputation of  $\mathbf{f}_{X_D}$  with each pixel's reflectivity  $\alpha_{i,j}$  would be a slow and inefficient process, this section instead considers the effect of precomputing  $\widetilde{\mathbf{f}}_{X_D}$  for a small set of evenly spaced values of  $\alpha$  over  $[0, 1]$ . Then for each  $(i, j)$  the  $\widetilde{\mathbf{f}}_{X_D}$  for the closest value to  $\alpha_{i,j}$  is used in the log-matched filter to estimate depth. Here we assume a conventional camera coaxially aligned with the lidar and spectrally filtered to accept the same wavelength can acquire a grayscale image that is a sufficient approximation of  $\boldsymbol{\alpha}$ . Then the camera image  $\widetilde{\boldsymbol{\alpha}}$  acquired simultaneously with the lidar data can be used in the depth image reconstruction.

To validate our depth estimation algorithm, we simulate detection data using ground truth depth and reflectivity images from the Middlebury stereo dataset [164]. The color images are first converted to grayscale and then normalized so that  $\alpha_{i,j} \in [0.1, 1.0]$ . The disparity image is converted to a depth map using intrinsic camera properties, and the scene is arbitrarily shifted by 10 m (66.7 ns). Both images are downsampled to  $93 \times 105$  pixels to reduce processing time. For all simulations, we use parameter values of  $\beta = 6$ ,  $B = 3$ ,  $\sigma_p = 0.2$  ns,  $\Delta = 0.02$  ns,  $t_r = 100$  ns, and  $t_d = 75$  ns.



**Figure 5-10:** Results for simulated detections from a true 3D scene illustrate the effectiveness of using the Markov chain modeling for high-flux acquisition. Using a 3-bit grayscale reflectivity approximation, our MCPDF method outperforms both low- and high-flux depth estimate for the number of illuminations  $n_r = 100$  and 2000. Our MCPDF method approximately matches the LF performance with  $20\times$  fewer illuminations, greatly speeding up acquisition.

Fig. 5-10 shows the results of simulated acquisitions and depth estimation for the Bowling scene with the number of illuminations  $n_r = 100$  and 2000. Reflectivity measurement by a 3-bit grayscale camera is emulated. Low-flux acquisition results in 4.80 and 96.3 detected photons per pixel over the scene for the short and long acquisitions, respectively. Without first attenuating the photon flux, the high-flux acquisition detects photons much faster, with an average rate of just more than one photon detection per illumination, emphasizing the multi-stop ability of modern TC-SPC systems. The depth estimation results for the short acquisition demonstrate why the LF approach is insufficient for real-time applications—there are simply too

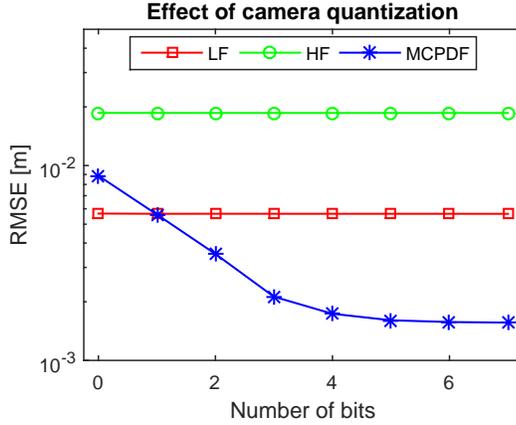


**Figure 5.11:** Comparison of  $\log_{10}(|\hat{z} - z|)$  for HF and MCPDF for the high-flux acquisition with  $n_r = 100$  illuminations. The error for HF in (a) is lower for darker scene patches with detection PDFs less distorted by dead time, whereas the error for MCPDF in (b) is lower for lighter scene patches which reflect back more signal photons.

few signal detections to reliably estimate the depth. Increasing  $n_r$  enables an improvement in root mean squared error (RMSE) of several orders of magnitude for the LF method, with RMSE computed as in (2.12). On the other hand, the results from the HF estimator barely improve as  $n_r$  increases since it does not take dead time into account, so the error is dominated by a bias. The RMSE of our proposed Markov chain-based MCPDF method continues to decrease as more data is acquired. Furthermore, MCPDF achieves nearly the same RMSE for the short acquisition as LF does for the long acquisition, enabling accurate depth imaging to be performed 20 times faster.

A comparison for the high-flux acquisition in Fig. 5.11 further illustrates the advantage of MCPDF. The absolute error for HF and MCPDF is shown on a logarithmic scale. The figure reveals that when the overall arrival rate is high enough, the smallest errors for HF somewhat counterintuitively occur for the darkest pixels, since their detection PDFs are least distorted by the dead time. On the other hand, by correctly anticipating the dead time distortion, errors for MCPDF occur roughly proportionally with the number of detected photons.

We also explore the effect of the quantization of  $\tilde{\alpha}$  on the reconstruction error.



**Figure 5-12:** The performance of MCPDF improves as more quantization levels are used for the reflectivity estimate. The plot shows the median RMSE values for 100 realizations of detection data for the Bowling scene with  $n_r = 2000$  illuminations.

For the same experimental parameters as previously used, 100 realizations of photon detection data were generated for different numbers of bits for the Bowling scene reflectivity. Fig. 5-12 shows the median RMSE results over the 100 trials. The median is plotted since significant outliers occasionally occurred when the RMSE for the LF method was dominated by a small number of pixels with large depth errors. It is clear from the plot that the performance of MCPDF greatly improves with the number of quantization levels, whereas the RMSE of the LF and HF methods does not depend on information about the reflectivity. Fig. 5-2 helps illustrate why the methods depend differently on the quantization of  $\tilde{\alpha}$ . For the LF and HF methods, which use  $\mathbf{f}^A$ , a change in the estimated value of  $\alpha_{i,j}$  only changes the strength of the signal relative to the background, but the position of the peak is unaffected, so the log-matched filter depth estimates are mostly unchanged. On the other hand, the shape of  $\mathbf{f}_{X_D}$ , including the position of its peak, depends strongly on the exact  $\alpha_{i,j}$  value, so a closer approximation of  $\alpha_{i,j}$  from finer quantization yields a more accurate approximation of  $\mathbf{f}_{X_D}$  and thus a better log-matched filter depth estimate. Moreover, Fig. 5-12 implies that for the Bowling scene used in our experiment, 3 bits are sufficient to achieve

small performance degradation due to reflectivity quantization, though the number of bits needed is likely affected by the range of  $\alpha_{i,j}$  values.

## 5.8 Conclusion

This chapter studied dead time compensation for a modern, asynchronous, nonparalyzable detector. By using a Markov chain model for detection times, we obtained the limit of the empirical distribution of detection times as the stationary distribution of the Markov chain. We found that the Fisher information per detection can be higher for this limiting distribution than for the arrival distribution, which suggests that the distortion due to dead time can be beneficial for depth estimation if used properly. Indeed, simulation results showed that our first proposed method MCPDF, which is a log-matched filter matched to the limiting distribution, achieved lower error than the low-flux method for a fixed number of detections when the SBR is sufficiently high. By exploiting the stationary condition for the Markov chain, we derived our second proposed method MCHC, which estimates the arrival distribution from the detection distribution by solving a nonlinear inverse problem with a provably convergent optimization algorithm, and then the corrected histogram is used in a log-matched filter estimator. Depth imaging results show our method can achieve accurate depth images 20 times faster than the conventional method. The use of spatial regularization could allow for accurate imaging with even lower photon counts, further reducing acquisition times. Although we only tested MCHC in the context of ranging, it makes no assumptions about the arrival intensity and should thus be applicable to other dead-time-limited TCSPC applications, including FLIM and NLOS imaging; we leave such extensions as future work.

## Appendices

### 5.A Proof of Proposition 1

First, we show that  $\{X_i\}_{i \in \mathbb{N}}$  is a Markov chain. Define  $\mathcal{B}_i := \{X_i \leq x_i\}, \forall i \in \mathbb{N}$ . We need to establish that

$$P(\mathcal{B}_{i+1} | X_n = x_n, \forall n \leq i) = P(\mathcal{B}_{i+1} | X_i = x_i). \quad (5.32)$$

The following equivalence of events will be useful in the proof, as it relates the sets defined by elements of  $\{X_i\}_{i \in \mathbb{N}}$  to those of  $\{T_i\}_{i \in \mathbb{N}}$ , which has a known transition density (5.3):

$$\begin{aligned} \{X_{i+1} \leq x_{i+1}\} &= \cup_{n=0}^{\infty} \{n\mathbf{t}_r < T_{i+1} \leq n\mathbf{t}_r + x_{i+1}\}, \\ \{T_i = n_i\mathbf{t}_r + x_i\} &= \{N_i = n_i\} \cap \{X_i = x_i\}. \end{aligned} \quad (5.33)$$

Let  $\mathbf{t}_d = \mathbf{k}_d\mathbf{t}_r + \mathbf{x}_d$ , where  $\mathbf{k}_d = \lfloor \mathbf{t}_d/\mathbf{t}_r \rfloor$  and  $\mathbf{x}_d = \mathbf{t}_d \bmod \mathbf{t}_r$ . Moreover, define  $\mathcal{A}_n := \{n\mathbf{t}_r < T_{i+1} \leq n\mathbf{t}_r + x_{i+1}\}, \forall n \in \mathbb{N} \cup \{0\}$ . Then we have

$$\begin{aligned} &P(\mathcal{B}_{i+1} | X_j = x_j, \forall j \leq i) \\ &= \sum_{n_1, \dots, n_i} P(\mathcal{B}_{i+1} | X_j = x_j, N_j = n_j, \forall j \leq i) P(N_j = n_j, \forall j \leq i | X_j = x_j, \forall j \leq i) \\ &= \sum_{n_1, \dots, n_i} P(\cup_{n=0}^{\infty} \mathcal{A}_n | T_j = n_j\mathbf{t}_r + x_j, \forall j \leq i) P(N_j = n_j, \forall j \leq i | X_j = x_j, \forall j \leq i) \end{aligned} \quad (5.34)$$

where the summation is over all  $0 \leq n_1 \leq \dots \leq n_i < \infty$  and the last equality follows by (5.33). In the following, we show that the first probability in (5.34) only depends

on  $x_i$  and  $\mathbf{x}_d$ .

$$\begin{aligned}
& P(\cup_{n=0}^{\infty} \mathcal{A}_n | T_j = n_j \mathbf{t}_r + x_j, \forall j \leq i) \\
& \stackrel{(a)}{=} \sum_{n=0}^{\infty} P(\mathcal{A}_n | T_i = n_i \mathbf{t}_r + x_i) \\
& \stackrel{(b)}{=} \int_{(n_i + \mathbf{k}_d) \mathbf{t}_r + x_i + x_d}^{(n_i + \mathbf{k}_d) \mathbf{t}_r + x_{i+1}} \lambda(t) \exp \left( - \int_{(n_i + \mathbf{k}_d) \mathbf{t}_r + x_i + x_d}^t \lambda(\tau) d\tau \right) dt \\
& + \sum_{n=n_i + \mathbf{k}_d + 1}^{\infty} \int_{n \mathbf{t}_r}^{n \mathbf{t}_r + x_{i+1}} \lambda(t) \exp \left( - \int_{(n_i + \mathbf{k}_d) \mathbf{t}_r + x_i + x_d}^t \lambda(\tau) d\tau \right) dt, \quad (5.35)
\end{aligned}$$

where step (a) follows by the Markov property of  $\{T_i\}_{i=1}^{\infty}$  and  $\{\mathcal{A}_n\}$  being disjoint and in step (b), we have plugged in (5.3) and assumed that  $x_i + \mathbf{x}_d \leq x_{i+1} \leq \mathbf{t}_r$ . Note that other relationships between  $x_i$ ,  $x_{i+1}$ ,  $\mathbf{x}_d$ ,  $\mathbf{t}_r$  may lead to slightly different expression, but the derivation follows similarly. (We will see that the expression does not depend on  $\mathbf{k}_d$ .) Label the two terms in (5.35) as  $S_1$  and  $S_2$ . First, consider  $S_1$ :

$$\begin{aligned}
S_1 &= \int_{x_i + \mathbf{x}_d}^{x_{i+1}} \lambda(t) \exp \left( - \int_{n_i \mathbf{t}_r + x_i + x_d}^{n_i \mathbf{t}_r + t} \lambda(\tau) d\tau \right) dt \\
&= \int_{x_i + \mathbf{x}_d}^{x_{i+1}} \lambda(t) \exp \left( - \int_{x_i + \mathbf{x}_d}^t \lambda(\tau) d\tau \right) dt,
\end{aligned}$$

which follows by change of variable and  $\lambda(t + n_i \mathbf{t}_r) = \lambda(t)$ . Next consider  $S_2$ :

$$\begin{aligned}
S_2 &= \sum_{n=n_i+k_d+1}^{\infty} \int_0^{x_{i+1}} \lambda(t) \exp\left(-\int_{(n_i+k_d)\mathbf{t}_r+x_i+x_d}^{n\mathbf{t}_r+t} \lambda(\tau) d\tau\right) dt \\
&= \sum_{n=0}^{\infty} \int_0^{x_{i+1}} \lambda(t) \exp\left(-\int_{(n_i+k_d)\mathbf{t}_r+x_i+x_d}^{(n+n_i+k_d+1)\mathbf{t}_r+t} \lambda(\tau) d\tau\right) dt \\
&= \sum_{n=0}^{\infty} (\exp(-\Lambda))^n \int_0^{x_{i+1}} \lambda(t) \exp\left(-\int_{x_i+x_d}^{\mathbf{t}_r+t} \lambda(\tau) d\tau\right) dt \\
&= \frac{\int_0^{x_{i+1}} \lambda(t) \exp\left(-\int_{x_i+x_d}^{\mathbf{t}_r+t} \lambda(\tau) d\tau\right) dt}{1 - \exp(-\Lambda)}.
\end{aligned}$$

Notice that neither  $S_1$  nor  $S_2$  depends on  $k_d$ ,  $\{N_j\}_{j \leq i}$ , or  $\{X_j\}_{j < i}$ . Plugging  $S_1$  and  $S_2$  back into (5.34), we have that

$$\begin{aligned}
P(\mathcal{B}_{i+1} | X_j = x_j, \forall j \leq i) &= \sum_{n_1, \dots, n_i} (S_1 + S_2) P(N_j = n_j, \forall j \leq i | X_j = x_j, \forall j \leq i) \\
&= S_1 + S_2,
\end{aligned}$$

where the last equality holds since  $P(\cdot | X_j = x_j, \forall j \leq i)$  is a probability measure and that the summation  $\sum_{n_1, \dots, n_i}$  is over all  $0 \leq n_1 \leq \dots \leq n_i < \infty$ . Hence, we have established (5.32), and therefore proved that  $\{X_i\}_{i \in \mathbb{N}}$  is a Markov chain.

Next, we compute the transition PDF to justify (5.5):

$$\begin{aligned}
f_{X_{i+1}|X_i}(x_{i+1}|x_i) &= \frac{d}{dx_{i+1}} P(X_{i+1} \leq x_{i+1} | X_i = x_i) \\
&= \frac{d}{dx_{i+1}} S_1 + \frac{d}{dx_{i+1}} S_2 = \frac{\lambda(x_{i+1}) \exp\left(-\int_{x_i+x_d}^{x_{i+1}} \lambda(\tau) d\tau\right)}{1 - \exp(-\Lambda)}.
\end{aligned}$$

Recall that we have assumed  $x_i + x_d \leq x_{i+1} \leq \mathbf{t}_r$  in the derivation above, and we can check that it matches (5.5) for this case. Other cases can be derived similarly.

## 5.B Derivation of Fisher Information

We present the derivation for  $\text{FI}_D$ ; the derivation for  $\text{FI}_A$  follows similarly. By definition of Fisher information:

$$\begin{aligned}
\text{FI}_D &= \int_0^{\mathfrak{t}_r} \left( -\frac{\partial^2}{\partial z^2} \log(f_{X_D}(x; z)) \right) f_{X_D}(x; z) \, dx \\
&= \int_0^{\mathfrak{t}_r} \frac{\left( \frac{\partial}{\partial z} f_{X_D}(x; z) \right)^2}{f_{X_D}(x; z)} \, dx - \int_0^{\mathfrak{t}_r} \frac{\partial^2}{\partial z^2} f_{X_D}(x; z) \, dx \\
&\stackrel{(a)}{=} \int_0^{\mathfrak{t}_r} \frac{\left( \frac{\partial}{\partial z} f_{X_D}(x; z) \right)^2}{f_{X_D}(x; z)} \, dx - \frac{\partial^2}{\partial z^2} \left( \int_0^{\mathfrak{t}_r} f_{X_D}(x; z) \, dx \right) \\
&= \int_0^{\mathfrak{t}_r} \frac{\left( \frac{\partial}{\partial z} f_{X_D}(x; z) \right)^2}{f_{X_D}(x; z)} \, dx,
\end{aligned}$$

where the interchange of derivative and integral in step (a) holds trivially, since the range of the integral is finite.

## 5.C Proof of Proposition 2

In the following, we will show that

$$P(R_i = r_i, R_{i-1} = r_{i-1}) = \left(1 - \exp(-\Lambda)\right)^2 \prod_{j=i-1}^i \exp(-r_j \Lambda),$$

which would imply that Proposition 2 is valid for  $R_i$  and  $R_{i-1}$ ; the proof for more than two  $R_i$ 's follows similarly.

Define event  $\mathcal{E}_j$  for  $j = i - 1, i$  as

$$\mathcal{E}_j := \left\{ r_j \mathfrak{t}_r + T_j + \mathfrak{t}_d \leq T_{j+1} < (r_j + 1) \mathfrak{t}_r + T_j + \mathfrak{t}_d \right\}.$$

By definition of  $R_i$  in (5.8),

$$\begin{aligned} P\left(R_i = r_i, R_{i-1} = r_{i-1}\right) &= P\left(\mathcal{E}_i \cap \mathcal{E}_{i-1}\right) \\ &= \mathbb{E}\left[P\left(\mathcal{E}_i \cap \mathcal{E}_{i-1} \mid T_{i-1}\right)\right]. \end{aligned}$$

Note that by the Markov property of absolute detection times as discussed in Section 5.3.1, the joint PDF of  $T_{i+1}, T_i$  given  $T_{i-1} = t_{i-1}$  is

$$f_{T_{i+1}, T_i | T_{i-1}}(t_{i+1}, t_i | t_{i-1}) = f_{T_{i+1} | T_i}(t_{i+1} | t_i) f_{T_i | T_{i-1}}(t_i | t_{i-1}).$$

Let  $a_j := r_j \mathbf{t}_r + t_j + \mathbf{t}_d$  for  $j = i - 1, i$ . Then

$$P\left(\mathcal{E}_1, \mathcal{E}_2 \mid T_{i-1} = t_{i-1}\right) = \int_{a_{i-1}}^{\mathbf{t}_r + a_{i-1}} \int_{a_i}^{\mathbf{t}_r + a_i} f_{T_{i+1} | T_i}(t_{i+1} | t_i) f_{T_i | T_{i-1}}(t_i | t_{i-1}) dt_{i+1} dt_i.$$

First consider the inner integral:

$$\begin{aligned} \int_{a_i}^{\mathbf{t}_r + a_i} f_{T_{i+1} | T_i}(t_{i+1} | t_i) dt_{i+1} &= \int_{a_i}^{\mathbf{t}_r + a_i} \lambda(t_{i+1}) \exp\left(-\int_{t_i + \mathbf{t}_d}^{t_{i+1}} \lambda(\tau) d\tau\right) dt_{i+1} \\ &= -\exp\left(-\int_{t_i + \mathbf{t}_d}^{t_{i+1}} \lambda(\tau) d\tau\right) \Big|_{t_{i+1}=a_i}^{t_{i+1}=\mathbf{t}_r + a_i} \\ &= (1 - \exp(-\Lambda)) \exp(-r_i \Lambda). \end{aligned}$$

Note that the inner integral does not depend on  $t_i$ . Using similar calculation, we have that the outer integral does not depend on  $t_{i-1}$ . Hence,

$$\begin{aligned} P\left(\mathcal{E}_1 \cap \mathcal{E}_2 \mid T_{i-1} = t_{i-1}\right) &= P\left(\mathcal{E}_1 \cap \mathcal{E}_2\right) \\ &= \left(1 - \exp(-\Lambda)\right)^2 \prod_{j=i-1}^i \exp(-r_j \Lambda), \end{aligned}$$

which is the desired result.

## 5.D Proof of Proposition 3

In the following, we will find an upper bound for the Lipschitz constant  $L$  of  $\nabla D(\cdot)$  defined in (5.26). For brevity, we omit the dependence on  $\mathbf{h}$  in the notation for  $\mathcal{T}$ . By (5.26), we have

$$\begin{aligned}\Lambda \nabla D(\boldsymbol{\lambda}) &= \mathbf{g}\boldsymbol{\lambda}^T \mathcal{T}(\boldsymbol{\lambda}) + \mathcal{T}(\boldsymbol{\lambda}) + \mathbf{g}^T \boldsymbol{\lambda} \mathcal{T}(\boldsymbol{\lambda}) - \Lambda \text{diag}(\mathbf{g}) \mathcal{T}(\boldsymbol{\lambda}) \\ &\quad - \mathbf{g}\boldsymbol{\lambda}^T \mathbf{h} - \mathbf{h} - \mathbf{g}^T \boldsymbol{\lambda} \mathbf{h} + \Lambda \text{diag}(\mathbf{g}) \mathbf{h}.\end{aligned}$$

It follows that for any  $\mathbf{u}, \mathbf{v} \in [0, M]^{\mathfrak{n}_b}$ , we have by triangle inequality that

$$\begin{aligned}\Lambda \|\nabla D(\mathbf{u}) - \nabla D(\mathbf{v})\| &\leq \|\mathbf{g}\mathbf{u}^T \mathcal{T}(\mathbf{u}) - \mathbf{g}\mathbf{v}^T \mathcal{T}(\mathbf{v})\| + \|\mathcal{T}(\mathbf{u}) - \mathcal{T}(\mathbf{v})\| \\ &\quad + \|\mathbf{g}^T \mathbf{u} \mathcal{T}(\mathbf{u}) - \mathbf{g}^T \mathbf{v} \mathcal{T}(\mathbf{v})\| \\ &\quad + \Lambda \|\text{diag}(\mathbf{g}) \mathcal{T}(\mathbf{u}) - \text{diag}(\mathbf{g}) \mathcal{T}(\mathbf{v})\| \\ &\quad + \|\mathbf{g}\mathbf{u}^T \mathbf{h} - \mathbf{g}\mathbf{v}^T \mathbf{h}\| + \|\mathbf{g}^T \mathbf{u} \mathbf{h} - \mathbf{g}^T \mathbf{v} \mathbf{h}\|.\end{aligned}$$

Label the six terms on the right hand side as  $T_1, \dots, T_6$ . We will show that there exist constants  $L_1, \dots, L_6 < \infty$  such that  $T_i \leq L_i \|\mathbf{u} - \mathbf{v}\|, \forall i = 1, \dots, 6$ . Then the Lipschitz constant  $L$  of the gradient  $\nabla D$  is upper bounded by  $\Lambda^{-1} \sum_{i=1}^6 L_i$ .

First consider  $T_2$ . Let  $\hat{g} := \max_{i \in [\mathfrak{n}_b]} g_i$ . Then

$$\begin{aligned}T_2 &\stackrel{(a)}{\leq} \|\mathbf{g}^T \mathbf{u} \mathbf{u} - \mathbf{g}^T \mathbf{v} \mathbf{v}\| + \|\mathbf{u} - \mathbf{v}\| + \|\text{diag}(\mathbf{g})(\mathbf{u} - \mathbf{v})\| \\ &\stackrel{(b)}{\leq} \|\mathbf{g}^T \mathbf{u} \mathbf{u} - \mathbf{g}^T \mathbf{u} \mathbf{v}\| + \|\mathbf{g}^T \mathbf{u} \mathbf{v} - \mathbf{g}^T \mathbf{v} \mathbf{v}\| + (1 + \hat{g}) \|\mathbf{u} - \mathbf{v}\| \\ &\stackrel{(c)}{\leq} \|\mathbf{g}\| \|\mathbf{u}\| \|\mathbf{u} - \mathbf{v}\| + \|\mathbf{g}\| \|\mathbf{v}\| \|\mathbf{u} - \mathbf{v}\| + (1 + \hat{g}) \|\mathbf{u} - \mathbf{v}\| \\ &\stackrel{(d)}{\leq} 2\sqrt{\mathfrak{n}_b} M \|\mathbf{u} - \mathbf{v}\| + 2\|\mathbf{u} - \mathbf{v}\| = 2(\sqrt{\mathfrak{n}_b} M + 1) \|\mathbf{u} - \mathbf{v}\|,\end{aligned}$$

where step (a) follows by triangle inequality, step (b) follows by triangle inequality

and the fact that the largest eigenvalue of a diagonal matrix equals to the largest entry on its diagonal, and step (c) follows by Cauchy–Schwarz. To see step (d), notice that  $\|\mathbf{u}\|, \|\mathbf{v}\| \leq \sqrt{n_b}M$ , since  $\mathbf{u}, \mathbf{v} \in [0, M]^{n_b}$  and  $\max_{i \in [n_b]} g_i \leq \|\mathbf{g}\| \leq \|\mathbf{g}\|_1 \leq \|\mathbf{h}\|_1 = 1$  (the second inequality follows by the fact that  $\mathbf{g}$  is non-negative and so  $\|\mathbf{g}\| = \sqrt{\sum_{i=1}^{n_d} g_i^2} \leq \sqrt{(\sum_{i=1}^{n_d} g_i)^2} = \|\mathbf{g}\|_1$ , the third inequality assumed  $\mathbf{t}_d \leq \mathbf{t}_r$  and the last equality follows by  $\mathbf{h}$  being a proper probability density function). Similarly, we can show that  $T_4 \leq 2\Lambda(\sqrt{n_b}M + 1)\|\mathbf{u} - \mathbf{v}\|$ ,  $T_5 \leq \|\mathbf{u} - \mathbf{v}\|$ , and  $T_6 \leq \|\mathbf{u} - \mathbf{v}\|$ .

Next consider  $T_1$ :

$$\begin{aligned}
T_1 &\stackrel{(a)}{\leq} |\mathbf{u}^T \mathcal{T}(\mathbf{u}) - \mathbf{v}^T \mathcal{T}(\mathbf{v})| \\
&\stackrel{(b)}{\leq} |\mathbf{u}^T \mathcal{T}(\mathbf{u}) - \mathbf{v}^T \mathcal{T}(\mathbf{u})| + |\mathbf{v}^T \mathcal{T}(\mathbf{u}) - \mathbf{v}^T \mathcal{T}(\mathbf{v})| \\
&\stackrel{(c)}{\leq} \|\mathcal{T}(\mathbf{u})\| \|\mathbf{u} - \mathbf{v}\| + \|\mathbf{v}\| \|\mathcal{T}(\mathbf{u}) - \mathcal{T}(\mathbf{v})\| \\
&\stackrel{(d)}{\leq} (\Lambda^{-1}nB^2 + (\Lambda^{-1} + 2)\sqrt{n_d}M) \|\mathbf{u} - \mathbf{v}\|,
\end{aligned}$$

where step (a) follows by  $\|\mathbf{g}\| \leq 1$  as established before, step (b) follows by triangle inequality, step (c) follows by Cauchy–Schwarz, and step (d) follows by

$$\begin{aligned}
\|\mathcal{T}(\mathbf{u})\| &= \|\Lambda^{-1}\mathbf{g}^T \mathbf{u}\mathbf{u} + \Lambda^{-1}\mathbf{u} - \text{diag}(\mathbf{g})\mathbf{u}\| \\
&\leq \Lambda^{-1}\|\mathbf{g}\| \|\mathbf{u}\|^2 + \Lambda^{-1}\|\mathbf{u}\| + \widehat{g}\|\mathbf{u}\| \\
&\leq \Lambda^{-1}n_b M^2 + \Lambda^{-1}\sqrt{n_b}M + \sqrt{n_b}M.
\end{aligned}$$

Similarly, we can show that

$$T_3 \leq (\Lambda^{-1}n_b M^2 + (\Lambda^{-1} + 2)\sqrt{n_b}M) \|\mathbf{u} - \mathbf{v}\|.$$

Proposition 3 is then obtained by combining the upper bounds for  $T_1$  through  $T_6$ .

## Chapter 6

# Conclusions

Given its sensitivity to the smallest quantities of light and its ability to tag detected photons with high temporal precision, single-photon lidar holds promise for depth mapping under challenging requirements of low-power illumination, long-range measurements, or real time acquisition. However, sifting through the streams of detection times to form useful depth and reflectivity images is a non-trivial task. Applying insufficiently rigorous models of detection time distributions yields output images that range from misleading to entirely erroneous.

After carefully outlining the various components and basic modeling approaches for SPL systems, this thesis focused on three conditions that caused previous estimation approaches to fail. In Chapter 3, we considered the problem of tolerating strong ambient light, which is important for the use of SPL systems in real-world environments. Key to solving this problem was deriving rules for deciding when an estimate was likely due to background light and should thus be improved or ignored. The resulting algorithm succeeded in reconstructing scenes for which each pixel detected only 2 signal photons and 50 background photons on average. Chapter 4 discussed the challenge of recovering precise depth estimates, even if the time stamps were captured with much coarser resolution. Mathematical modeling and simulation demonstrated that subtractive dither could be used to estimate Gaussian-distributed signals more efficiently and with greater precision than measurements with coarse quantization or nonsubtractive dither. Implementing an SPL system incorporating subtractive

dither and modifying estimators to account for non-Gaussian pulse shapes enabled up to 13-fold improvements in estimation accuracy. Finally, Chapter 5 addressed the dead time following each detection in an SPL system, which limits acquisition speeds or leads to erroneous depth estimates. The temporal dependence of one detection on the previous detection time led to identification of the detection time sequence as a Markov chain. The stationary distribution of the Markov chain matches the empirical distribution captured experimentally as a histogram and can thus be used to estimate depth directly or correct the histogram of dead time distortions.

The push for faster, safer, and cheaper lidar systems that can image the world with enough fidelity for autonomous navigation means there are countless problems to address that could not be contained in a single thesis. In terms of probabilistic modeling, there are more elaborate optical interactions with the scene—including multiple bounces, fog and other scattering or attenuating media, laser broadening as a function of distance and surface orientation, and unknown or varying background levels—that defy previous modeling assumptions but could extend functionality if included. Additionally, there are modifications to the acquisition procedure—e.g., temporally- or spatially-coded illumination, gating, the use of multiple or different types of detectors—that interact with existing considerations such as dead time in complex ways. Finally, there is the question of the next step in the processing pipeline: how those probabilistic models are integrated into scene reconstruction algorithms that use structural priors. Modern image processing and optimization approaches integrating Bayesian inference, plug-and-play priors, or deep neural networks may achieve significantly better results than initial methods using TV regularization. Ultimately, single-photon lidar implementations will be successful when their probabilistic models account for a wide variety of environmental conditions, are tailored to the specific hardware and acquisition method, and are still tractable enough for

useful computation within algorithmic procedures.

## References

- [1] 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 Transactions on Image Processing*, 5(6):1012–1025, June 1996. 36
- [2] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2274–2282, 2012. 49, 69
- [3] Y. Altmann, X. Ren, A. McCarthy, G. S. Buller, and S. McLaughlin. Lidar waveform-based analysis of depth images constructed using sparse single-photon data. *IEEE Transactions on Image Processing*, 25(5):1935–1946, May 2016. 7, 127
- [4] Y. Altmann, X. Ren, A. McCarthy, G. S. Buller, and S. McLaughlin. Robust Bayesian target detection algorithm for depth imaging from sparse single-photon data. *IEEE Transactions on Computational Imaging*, 2(4):456–467, Dec. 2016. 31
- [5] M.-C. Amann, T. Bosch, M. Lescure, R. Myllylä, and M. Rioux. Laser ranging: a critical review of usual techniques for distance measurement. *Optical Engineering*, 40(1):10–19, 2001. 5, 6
- [6] Anonymous. Dissertation sur la recherche du milieu le plus probable, entre les résultats de plusieurs observations ou expériences. *Annales de Mathématiques pures et appliquées*, 12:181–204, 1821. 85
- [7] I. M. Antolovic, S. Burri, C. Bruschini, R. Hoebe, and E. Charbon. Nonuniformity analysis of a 65-kpixel CMOS SPAD imager. *IEEE Transactions on Electron Devices*, 63(1):57–64, Jan. 2016. 126, 127
- [8] I. M. Antolovic, C. Bruschini, and E. Charbon. Dynamic range extension for photon counting arrays. *Optics Express*, 26(17):22234, Aug. 2018. 126, 127, 128
- [9] G. R. Arce. *Nonlinear Signal Processing: A Statistical Approach*. John Wiley & Sons, 2005. 84

- [10] J. AuYeung. Picosecond optical pulse generation at gigahertz rates by direct modulation of a semiconductor laser. *Applied Physics Letters*, 38(5):308–310, Mar. 1981. 14
- [11] I. Bar-David. Communication under the Poisson regime. *IEEE Transactions on Information Theory*, 15(1):31–37, Jan. 1969. 33, 34, 152
- [12] Z. W. Barber, J. R. Dahl, T. L. Sharpe, and B. I. Erkmen. Shot noise statistics and information theory of sensitivity limits in frequency-modulated continuous-wave lidar. *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, 30(7):1335, July 2013. 6
- [13] N. C. Beaulieu and Q. Guo. Novel estimator for the location parameter of the generalized Gaussian distribution. *IEEE Communications Letters*, 16(12):2064–2067, Dec. 2012. 86, 87, 89, 108
- [14] N. C. Beaulieu and Q. Guo. True ML estimator for the location parameter of the generalized Gaussian distribution with  $p = 4$ . *IEEE Communications Letters*, 17(1):155–157, Jan. 2013. 84
- [15] W. Becker. *Advanced Time-Correlated Single Photon Counting Techniques*. Springer, Berlin, 2005. 3, 17, 19, 109, 127, 128
- [16] W. Becker. Fluorescence lifetime imaging - techniques and applications. *Journal of Microscopy*, 247(2):119–136, Aug. 2012. 4
- [17] W. Becker, A. Bergmann, M. A. Hink, K. König, K. Benndorf, and C. Biskup. Fluorescence lifetime imaging by time-correlated single-photon counting. *Microscopy Research and Technique*, 63(1):58–66, Jan. 2004. 4
- [18] B. Behroozpour, P. A. Sandborn, M. C. Wu, and B. E. Boser. Lidar system architectures and circuits. *IEEE Communications Magazine*, 55(10):135–142, Oct. 2017. 5, 6
- [19] K. Bengler, K. Dietmayer, B. Farber, M. Maurer, C. Stiller, and H. Winner. Three decades of driver assistance systems: Review and future perspectives. *IEEE Intelligent Transportation Systems Magazine*, 6(4):6–22, 2014. 5
- [20] W. R. Bennett, Jr. Radiative lifetimes and collision transfer cross sections of excited atomic states. In J. R. Singer, editor, *Advances in Quantum Electronics*, pages 28–43. Columbia University Press, New York, 1961. 3
- [21] W. R. Bennett, Jr., P. J. Kindlmann, and G. N. Mercer. Measurement of excited state relaxation rates. *Applied Optics*, 4:34–57, Jan. 1965. 4
- [22] P. J. Bickel and E. L. Lehmann. Descriptive statistics for nonparametric models: II. location. *The Annals of Statistics*, 3(5):1045–1069, 1975. 112

- [23] G. Blom. On linear estimates with nearly minimum variance. *Arkiv för Matematik*, 3(4):365–369, Jan. 1957. 85
- [24] L. M. Bollinger and G. E. Thomas. Measurement of the time dependence of scintillation intensity by a delayed-coincidence method. *Review of Scientific Instruments*, 32(9):1044–1050, Sept. 1961. 3, 9, 19
- [25] A. Bovik, T. Huang, and D. Munson. Nonlinear filtering using linear combinations of order statistics. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 7, pages 2067–2070, May 1982. 85
- [26] A. Bovik, T. Huang, and D. Munson. A generalization of median filtering using linear combinations of order statistics. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 31(6):1342–1350, Dec 1983. 85
- [27] G. E. P. Box. Robustness in the strategy of scientific model building. In R. L. Launer and G. N. Wilkinson, editors, *Robustness in Statistics*, pages 201–236. Academic Press, Apr. 1978. 1
- [28] J. D. Broffitt. An example of the large sample behavior of the midrange. *The American Statistician*, 28(2):69–70, 1974. 77
- [29] G. S. Buller and R. J. Collins. Single-photon detectors for infrared wavelengths in the range 11.7  $\mu\text{m}$ . In P. Kapusta, M. Wahl, and R. Erdmann, editors, *Advanced Photon Counting*, Springer Series on Fluorescence (Methods and Applications), pages 43–69. Springer, Cham, 2014. 13
- [30] G. S. Buller and A. M. Wallace. Ranging and three-dimensional imaging using time-correlated single-photon counting and point-by-point acquisition. *IEEE Journal of Selected Topics in Quantum Electronics*, 13(4):1006–1015, 2007. 46
- [31] A. Bülter. Single-photon counting detectors for the visible range between 300 and 1,000 nm. In P. Kapusta, M. Wahl, and R. Erdmann, editors, *Advanced Photon Counting*, Springer Series on Fluorescence (Methods and Applications), pages 23–42. Springer, Cham, 2014. 13
- [32] J. Busck and H. Heiselberg. Gated viewing and high-accuracy three-dimensional laser radar. *Applied Optics*, 43(24):4705–4710, 2004. 106
- [33] M. Buttafava, G. Boso, A. Ruggeri, A. Dalla Mora, and A. Tosi. Time-gated single-photon detection module with 110 ps transition time and up to 80 MHz repetition rate. *Review of Scientific Instruments*, 85(8), 2014. 125, 126
- [34] A. B. Cambra, A. Muñoz, A. C. Murillo, J. J. Guerrero, and D. Gutierrez. Improving depth estimation using superpixels. In *Proceedings of the Spanish Computer Graphics Conference (CEIG)*, pages 1–9, 2014. 48, 49

- [35] M. A. Canuto, F. Estrada-Belli, T. G. Garrison, S. D. Houston, M. J. Acuña, M. Kováč, D. Marken, P. Nondédéo, L. Auld-Thomas, C. Castanet, D. Chate-lain, C. R. Chiriboga, T. Drápela, T. Lieskovský, A. Tokovinine, A. Velasquez, J. C. Fernández-Díaz, and R. Shrestha. Ancient lowland Maya complexity as revealed by airborne laser scanning of northern Guatemala. *Science*, 361(6409), 2018. 16
- [36] P. Carbone and D. Petri. Effect of additive dither on the resolution of ideal quantizers. *IEEE Transactions on Instrumentation and Measurement*, 43(3): 389–396, June 1994. 97
- [37] P. Carbone, C. Narduzzi, and D. Petri. Performance of stochastic quantizers employing nonlinear processing. *IEEE Transactions on Instrumentation and Measurement*, 45(2):435–439, Apr. 1996. 77
- [38] L. C. Chen, W. R. Lloyd, C. W. Chang, D. Sud, and M. A. Mycek. Fluorescence lifetime imaging microscopy for quantitative biological imaging. In *Methods in Cell Biology*, volume 114, pages 457–488. Academic Press, Jan. 2013. 3, 4
- [39] Z. Chen, R. Fan, X. Li, Z. Dong, Z. Zhou, G. Ye, and D. Chen. Accuracy improvement of imaging lidar based on time-correlated single-photon counting using three laser beams. *Optics Communications*, 429:175–179, Dec. 2018. 107
- [40] H. S. Cho, C. H. Kim, and S. G. Lee. A high-sensitivity and low-walk error LADAR receiver for military application. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 61(10):3007–3015, Oct. 2014. 129
- [41] F. Christnacher, S. Schertzer, N. Metzger, E. Bacher, M. Laurenzis, and R. Habermacher. Influence of gating and of the gate shape on the penetration capacity of range-gated active imaging in scattering environments. *Optics Express*, 23(26):32897, Dec. 2015. 106
- [42] C. J. Chunnillall, I. P. Degiovanni, S. Kück, I. Müller, and A. G. Sinclair. Metrology of single-photon sources and detectors: a review. *Optical Engineering*, 53(8):081910, 2014. 17, 18, 143
- [43] T. A. C. M. Claasen and A. Jongpier. Model for the power spectral density of quantization noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(4):914–917, Aug. 1981. 73
- [44] P. B. Coates. The correction for photon ‘pile-up’ in the measurement of radiative lifetimes. *Journal of Physics E: Scientific Instruments*, 1(8):878–879, Aug. 1968. 127

- [45] A. Colaço, A. Kirmani, G. A. Howland, J. C. Howell, and V. K. Goyal. Compressive depth map acquisition using a single photon-counting detector: Parametric signal processing meets sparsity. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 96–102, 2012. 22
- [46] A. Cominelli, G. Acconcia, P. Peronio, M. Ghioni, and I. Rech. High-speed and low-distortion solution for time-correlated single photon counting measurements: A theoretical analysis. *Review of Scientific Instruments*, 88(12):123701, Dec. 2017. 128, 133, 145
- [47] S. Cova, A. Lacaita, and G. Ripamonti. Trapping phenomena in avalanche photodiodes on nanosecond scale. *IEEE Electron Device Letters*, 12(12):685–687, Dec. 1991. 18
- [48] S. Cova, M. Ghioni, M. A. Itzler, J. C. Bienfang, and A. Restelli. Semiconductor-based detectors. In A. Migdall, S. V. Polyakov, J. Fan, and J. C. Bienfang, editors, *Single-Photon Generation and Detection*, chapter 4, pages 83–146. Academic Press, 2013. 109
- [49] O. Dabeer and A. Karnik. Signal parameter estimation using 1-bit dithered quantization. *IEEE Transactions on Information Theory*, 52(12):5389–5405, Dec. 2006. 74
- [50] P. Daniell. Observations weighted according to order. *American Journal of Mathematics*, 42(4):222–236, 1920. 85, 112
- [51] H. A. David and H. Nagaraja. *Order Statistics*. John Wiley & Sons, Hoboken, NJ, 3rd edition, 2003. 40
- [52] C. C. Davis and T. A. King. Photon pile-up corrections in the study of time-varying light sources. *Journal of Physics E: Scientific Instruments*, 5(11):1072–1074, 1972. 127
- [53] M. R. W. Dawson. Fitting the ex-Gaussian equation to reaction time distributions. *Behavior Research Methods, Instruments, & Computers*, 20(1):54–57, 1988. 110
- [54] J. O. Dickey, P. L. Bender, J. E. Faller, X. X. Newhall, R. L. Ricklefs, J. G. Ries, P. J. Shelus, C. Veillet, A. L. Whipple, J. R. Wiant, J. G. Williams, and C. F. Yoder. Lunar laser ranging: A continuing legacy of the Apollo program. *Science*, 265(5171):482–490, July 1994. 7
- [55] E. J. Dudewicz and S. N. Mishra. Moments. In *Modern Mathematical Statistics*, chapter 5, pages 216–228. Wiley, New York, 1988. 120

- [56] M. D. Eisaman, J. Fan, A. Migdall, and S. V. Polyakov. Invited review article: Single-photon sources and detectors. *Review of Scientific Instruments*, 82(7):071101, July 2011. 17
- [57] B. Erkmén and B. Moision. Maximum likelihood time-of-arrival estimation of optical pulses via photon-counting photodetectors. In *Proceedings of the IEEE International Symposium on Information Theory*, pages 1909–1913, 2009. 46
- [58] J. Fang and H. Li. Distributed adaptive quantization for wireless sensor networks: From delta modulation to maximum likelihood. *IEEE Transactions on Signal Processing*, 56(10):5246–5257, Oct. 2008. 74
- [59] R. C. Farias and J. M. Brossier. Scalar quantization for estimation: From an asymptotic design to a practical solution. *IEEE Transactions on Signal Processing*, 62(11):2860–2870, June 2014. 74
- [60] C. Favi and E. Charbon. A 17ps time-to-digital converter implemented in 65nm FPGA technology. In *Proceedings of the ACM/SIGDA International Symposium on Field-Programmable Gate Arrays*, page 113. ACM Press, 2009. 21
- [61] H. Fedder, S. Oesterwind, M. Wick, F. Olbrich, P. Michler, T. Veigel, M. Berroth, and M. Schlagmüller. Characterization of electro-optical devices with low jitter single photon detectors – towards an optical sampling oscilloscope beyond 100 GHz. In *Proceedings of the European Conference on Optical Communication (ECOC)*, pages 1–3, Sept. 2018. 21
- [62] M. Feigin, R. Whyte, A. Bhandari, A. Dorington, and R. Raskar. Modeling “wiggling” as a multi-path interference problem in AMCW ToF imaging. *Optics Express*, 23(15):19213–19225, July 2015. 5
- [63] W. Feller. On probability problems in the theory of counters. In R. L. Schilling, Z. Vondraček, and W. A. Woyczyński, editors, *Selected Papers I*, pages 751–759. Springer International Publishing, Cham, Switzerland, 2015. 125
- [64] G. G. Furman. *Improving performance of quantizer feedback systems by use of external dither*. M.S. thesis, Massachusetts Institute of Technology, 1959. 74
- [65] R. Garnett, T. Huegerich, C. Chui, and W. He. A universal noise removal algorithm with an impulse detector. *IEEE Transactions on Instrumentation and Measurement*, 14(11):1747–1754, 2005. 36
- [66] E. Gaviola. Die Abklingungszeiten der Fluoreszenz von Farbstofflösungen. *Annalen der Physik*, 386(23):681–710, Jan. 1926. 4

- [67] M. Ghioni, S. Cova, F. Zappa, and C. Samori. Compact active quenching circuit for fast photon counting with avalanche photodiodes. *Review of Scientific Instruments*, 67(10):3440–3448, 1996. 17, 18
- [68] C. Gianelli, L. Xu, J. Li, and P. Stoica. One-bit compressive sampling with time-varying thresholds: Maximum likelihood and the Cramér-Rao bound. In *Asilomar Conference on Signals, Systems, and Computers*, pages 399–403, Nov. 2016. 74
- [69] L. F. Gillespie. Apparent illuminance as a function of range in gated, laser night-viewing systems. *Journal of the Optical Society of America*, 56(7):883–887, July 1966. 106
- [70] A. Giudice, M. Ghioni, R. Biasi, F. Zappa, S. Cova, P. Maccagnani, and A. Gulinatti. High-rate photon counting and picosecond timing with silicon-SPAD based compact detector modules. *Journal of Modern Optics*, 54(2-3):225–237, 2007. 125
- [71] A. C. Giudice, M. Ghioni, S. Cova, and F. Zappa. A process and deep level evaluation tool: Afterpulsing in avalanche junctions. In *Proceedings of the European Solid-State Device Research Conference*, pages 347–350. IEEE, 2003. 18
- [72] M. Gladney, B. F. Dowden, J. D. Swalen, R. Lusebrink, and C. H. Sederholm. Computer-assisted gas-liquid chromatography. *Analytical Chemistry*, 41(7):883–888, 1969. 110
- [73] B. V. Gnedenko. On the theory of Geiger-Müller counters. *Zhurnal Èksperimental'noi i Teoreticheskoi Fiziki (Journal of Experimental and Theoretical Physics Letters)*, 11:101–106, 1941. 125
- [74] W. M. Goodall. Television by pulse code modulation. *Bell System Technical Journal*, 30(1):33–49, Jan. 1951. 74
- [75] R. M. Gray and T. G. Stockham. Dithered quantizers. *IEEE Transactions on Information Theory*, 39(3):805–812, 1993. 74, 75
- [76] E. Grushka. Characterisation of exponentially-modified Gaussian peaks in chromatography. *Analytical Chemistry*, 44(11):1733–1738, 1972. 110
- [77] A. Gupta, A. Ingle, and M. Gupta. Asynchronous single-photon 3D imaging. arXiv: 1908.06372v1, 2019. 124
- [78] A. Gupta, A. Ingle, A. Velten, and M. Gupta. Photon-flooded single-photon 3D cameras. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6770–6779, 2019. 68, 124

- [79] A. K. Gupta. Estimation of the mean and standard deviation of a normal population from a censored sample. *Biometrika*, 39(3):260–273, Dec. 1952. 85, 113
- [80] A. Halimi, A. Maccarone, A. McCarthy, S. McLaughlin, and G. Buller. Object depth profile and reflectivity restoration from sparse single-photon data acquired in underwater environments. *IEEE Transactions on Computational Imaging*, 3(3):472–484, Sept. 2017. 16
- [81] A. Halimi, R. Tobin, A. McCarthy, S. McLaughlin, and G. S. Buller. Restoration of multilayered single-photon 3D lidar images. In *Proceedings of the European Signal Processing Conference*, pages 708–712, Aug. 2017. 31
- [82] A. Halimi, R. Tobin, A. McCarthy, J. M. Bioucas-Dias, S. McLaughlin, and G. S. Buller. Robust restoration of sparse multidimensional single-photon LiDAR images. *IEEE Transactions on Computational Imaging*, pages 1–1, 2019. 31, 67
- [83] P. Hartzell, Z. Dang, Z. Pan, and C. Glennie. Radiometric evaluation of an airborne single photon LiDAR sensor. *IEEE Geoscience and Remote Sensing Letters*, 15(9):1466–1470, Sept. 2018. 7
- [84] W. He, B. Sima, H. Chen, Yunfei a Dai, Q. Chen, and G. Gu. A correction method for range walk error in photon counting 3D imaging LIDAR. *Optics Communications*, 308:211–217, Nov. 2013. 129, 154
- [85] J. Hecht. Lidar for self-driving cars. *Optics and Photonics News*, 29:26–33, Jan. 2018. 6
- [86] F. Heide, S. Diamond, D. B. Lindell, and G. Wetzstein. Sub-picosecond photon-efficient 3D imaging using single-photon sensors. *Scientific Reports*, 8(1):17726, Dec. 2018. 124, 128, 129, 138
- [87] M. Henriksson, H. Larsson, C. Grönwall, and G. Tolt. Continuously scanning time-correlated single-photon-counting single-pixel 3-D lidar. *Optical Engineering*, 56(3):031204, Sept. 2016. 23
- [88] Q. Hernandez, D. Gutierrez, and A. Jarabo. A computational model of a single-photon avalanche diode sensor for transient imaging. arXiv:1703.02635 [physics.ins-det], Feb. 2017. 109, 135
- [89] S. Hernandez-Marin, A. M. Wallace, and G. J. Gibson. Bayesian analysis of lidar signals with multiple returns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(12):2170–2180, 2007. 31

- [90] S. Hernandez-Marin, A. M. Wallace, and G. J. Gibson. Multilayered 3D Lidar image construction using spatial models in a Bayesian framework. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(6):1028–1040, June 2008. 31
- [91] P. A. Hiskett, C. S. Parry, A. McCarthy, and G. S. Buller. A photon-counting time-of-flight ranging technique developed for the avoidance of range ambiguity at gigahertz clock rates. *Optics Express*, 16(18):13685–13698, 2008. 24
- [92] G. A. Howland, D. J. Lum, M. R. Ware, and J. C. Howell. Photon counting compressive depth mapping. *Optics Express*, 21(20):23822–23837, Oct. 2013. 22
- [93] G. Humer, M. Peev, C. Schaeff, S. Ramelow, M. Stipčević, and R. Ursin. A simple and robust method for estimating afterpulsing in single photon detectors. *Journal of Lightwave Technology*, 33(14):3098–3107, July 2015. 18, 143
- [94] S. Isbaner, N. Karedla, D. Ruhlandt, S. C. Stein, A. Chizhik, I. Gregor, and J. Enderlein. Dead-time correction of fluorescence lifetime measurements and fluorescence lifetime imaging. *Optics Express*, 24(9):9429, 2016. 128, 139, 146, 155, 156
- [95] T. Ishikawa. Linearization of contactor control systems by external dither signals. Technical report, Electronics Laboratories, Stanford University, Stanford, CA, 1960. 74
- [96] M. A. Itzler, X. Jiang, and M. Entwistle. Power law temporal dependence of InGaAs/InP SPAD afterpulsing. *Journal of Modern Optics*, 59(17):1472–1480, 2012. 18
- [97] R. C. Jaffe. *Causal and statistical analyses of dithered systems containing three-level quantizers*. M.S. thesis, Massachusetts Institute of Technology, 1959. 74
- [98] N. S. Jayant and L. R. Rabiner. The application of dither to the quantization of speech signals. *Bell System Technical Journal*, 51(6):1293–1304, July 1972. 74
- [99] J. Kalisz. Review of methods for time interval measurements with picosecond resolution. *Metrologia*, 41(1):17–32, Feb. 2004. 19, 20, 21
- [100] S. M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, Upper Saddle River, NJ, 1993. 79
- [101] M. G. Kendall. The conditions under which Sheppard’s corrections are valid. *Journal of the Royal Statistical Society*, 101(3):592–605, 1938. 73

- [102] J. Kim, J. Seok, H. Lee, and M. Lee. Penalized maximum likelihood estimation of lifetime and amplitude images from multi-exponentially decaying fluorescence signals. *Optics Express*, 21(17):20240–53, Aug. 2013. 4
- [103] T. Kim, P. Bhargava, and V. Stojanovic. Optimal spectral estimation and system trade-off in long-distance frequency-modulated continuous-wave lidar. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 1583–1587, Apr. 2018. 6
- [104] A. Kirmani, D. Venkatraman, D. Shin, A. Colaco, F. N. C. Wong, J. H. Shapiro, and V. K. Goyal. First-photon imaging. *Science*, 343(6166):58–61, 2014. 2, 7, 23, 24, 27, 28, 32, 36, 37, 55, 127
- [105] A. Kirmani, D. Venkatraman, D. Shin, A. Colaco, F. N. C. Wong, J. H. Shapiro, and V. K. Goyal. Supporting materials for first-photon imaging. <http://www.sciencemag.org/content/343/6166/58/suppl/DC1>, 2014. 24, 55
- [106] B. Korzh, T. Lunghi, K. Kuzmenko, G. Boso, and H. Zbinden. Afterpulsing studies of low-noise InGaAs/InP single-photon negative-feedback avalanche diodes. *Journal of Modern Optics*, 62(14):1151–1157, 2015. 18
- [107] L. Kosten. On the frequency distribution of the number of discharges counted by a Geiger-Müller counter in a constant interval. *Physica*, 10(9):749–756, Nov. 1943. 125
- [108] N. J. Krichel, A. McCarthy, and G. S. Buller. Resolving range ambiguity in a photon counting depth imager operating at kilometer distances. *Optics Express*, 18(9):9192–9206, 2010. 24
- [109] J. D. Kurbatov and H. B. Mann. Correction of G-M counter data. *Physical Review*, 68(1-2):40–43, 1945. 125
- [110] S. Kurtti and J. Kostamovaara. Pulse width time walk compensation method for a pulsed time-of-flight laser rangefinder. In *Proceedings of the IEEE Instrumentation and Measurement Technology Conference (I2MTC)*, pages 1059–1062, May 2009. 129
- [111] Y. Lacouture and D. Cousineau. How to use MATLAB to fit the ex-Gaussian and other probability functions to a distribution of response times. *Tutorials in Quantitative Methods for Psychology*, 4(1):35–45, 2008. 111
- [112] J. R. Lakowicz. Fluorescence lifetime imaging. *Analytical Biochemistry*, 202(2):316–30, May 1992. 4
- [113] M. Lawton, R. C. Bolden, and M. J. Shaw. A 10 ns multichannel photon counter. *Journal of Physics E: Scientific Instruments*, 9(8):686–690, Aug. 1976. 4

- [114] C. Levert and W. Scheen. Probability fluctuations of discharges in a Geiger-Müller counter produced by cosmic radiation. *Physica*, 10(4):225–238, Apr. 1943. 125
- [115] H. Li and Z. Lin. Accelerated proximal gradient methods for nonconvex programming. In *Proceedings of the Neural Information Process. Syst.*, Montreal, Canada, December 7–12 2015. 149
- [116] Z.-P. Li, X. Huang, Y. Cao, B. Wang, Y.-H. Li, W. Jin, C. Yu, J. Zhang, Q. Zhang, C.-Z. Peng, F. Xu, and J.-W. Pan. Single-photon computational 3D imaging at 45 km. arXiv:1904.10341, Apr. 2019. 7, 67
- [117] C. Lin, P. Liu, T. Damen, D. Eilenberger, and R. Hartman. Simple picosecond pulse generation scheme for injection lasers. *Electronics Letters*, 16(15):16–18, 1980. 14
- [118] D. B. Lindell, M. O’Toole, and G. Wetzstein. Single-photon 3D imaging with deep sensor fusion. *ACM Transactions on Graphics*, 37(4):113, 2018. 66, 67
- [119] D. B. Lindell, M. O’Toole, and G. Wetzstein. Towards transient imaging at interactive rates with single-photon detectors. In *Proceedings of the IEEE International Conference on Computational Photography (ICCP)*, May 2018. 22
- [120] E. H. Lloyd. Least-squares estimation of location and scale parameters using order statistics. *Biometrika*, 39(1/2):88–95, Apr. 1952. 76, 77, 85, 112
- [121] C. Louchet and L. Moisan. Total variation denoising using posterior expectation. In *Proceedings of the European Signal Processing Conference*, pages 1–5, Aug. 2008. 47
- [122] R. Lussana, F. Villa, A. D. Mora, D. Contini, A. Tosi, and F. Zappa. Enhanced single-photon time-of-flight 3D ranging. *Optics Express*, 23(19):24962, Sept. 2015. 70, 106
- [123] L. R. A. MacColl. *Fundamental Theory of Servomechanisms*. D. Van Nostrand Company, inc., 1945. 74
- [124] H. Mahmoudabadi, T. Shoaf, and M. J. Olsen. Superpixel clustering and planar fit segmentation of 3D LIDAR point clouds. In *Proceedings of the International Conference on Computing for Geospatial Research and Application*, pages 1–7, July 2013. 48, 49
- [125] D. Marco and D. L. Neuhoff. The validity of the additive noise model for uniform scalar quantizers. *IEEE Transactions on Information Theory*, 51(5):1739–1755, May 2005. 73

- [126] J. S. Massa, A. M. Wallace, G. S. Buller, S. J. Fancey, and A. C. Walker. Laser depth measurement based on time-correlated single-photon counting. *Optics Letters*, 22(8):543–545, 1997. 2
- [127] J. S. Massa, G. S. Buller, A. C. Walker, S. Cova, M. Umasuthan, and A. M. Wallace. Time-of-flight optical ranging system based on time-correlated single-photon counting. *Applied Optics*, 37(31):7298–7304, 1998. 2
- [128] A. McCarthy, R. J. Collins, N. J. Krichel, V. Fernandez, A. M. Wallace, and G. S. Buller. Long-range time-of-flight scanning sensor based on high-speed time-correlated single-photon counting. *Applied Optics*, 48(32):6241–6251, Nov. 2009. 21, 23, 70
- [129] S. C. Medin, J. Murray-Bruce, D. Castañón, and V. K. Goyal. Beyond binomial and negative binomial: Adaptation in Bernoulli parameter estimation. *IEEE Transactions on Computational Imaging*, 5(4):570–584, Dec. 2019. 23, 159
- [130] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Springer Science & Business Media, 2012. 133
- [131] V. Molebny, P. McManamon, O. Steinvall, T. Kobayashi, and W. Chen. Laser radar: historical prospective—from the East to the West. *Optical Engineering*, 56(3):031220, Dec. 2016. 5
- [132] C. Moler. Matrix computation on distributed memory multiprocessors. In M. Heath, editor, *Hypercube Multiprocessors*, pages 181–195, Philadelphia, PA, 1986. SIAM. 69
- [133] A. M. Mood, F. A. Graybill, and D. C. Boes. *Introduction to the Theory of Statistics*. McGraw-Hill, 3 edition, 1974. 76, 81, 143
- [134] C. Moret-Tatay, D. Gamermann, E. Navarro-Pardo, and P. F. d. C. Castellá. ExGUtils: A python package for statistical analysis with the ex-Gaussian probability density. *Frontiers in Psychology*, 9:612, 2018. 111
- [135] A. Moschitta, J. Schoukens, and P. Carbone. Information and statistical efficiency when quantizing noisy DC values. *IEEE Transactions on Instrumentation and Measurement*, 64(2):308–317, Feb. 2015. 72, 77, 93
- [136] J. A. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, Jan. 1965. 92
- [137] D. V. O’Connor and D. D. Phillips. *Time-Correlated Single Photon Counting*. Academic Press, Orlando, FL, 1984. 4, 126
- [138] R. Öten and R. J. P. de Figueiredo. An efficient method for L-filter design. *IEEE Transactions on Signal Processing*, 51(1):193–203, Jan. 2003. 85, 86, 88

- [139] P. Palojärvi, T. Ruotsalainen, and J. Kostamovaara. A 250-MHz BiCMOS receiver channel with leading edge timing discriminator for a pulsed time-of-flight laser rangefinder. *IEEE Journal of Solid-State Circuits*, 40(6):1341–1349, June 2005. 129
- [140] H. C. Papadopoulos, G. W. Wornell, and A. V. Oppenheim. Sequential signal encoding from noisy measurements using quantizers with dynamic bias control. *IEEE Transactions on Information Theory*, 47(3):978–1002, 2001. 74, 87
- [141] M. Patting, P. Reisch, M. Sackrow, R. Dowler, M. Koenig, and M. Wahl. Fluorescence decay data analysis correcting for detector pulse pile-up at very high count rates. *Optical Engineering*, 57(03):1, Jan. 2018. 128
- [142] P. Paulus, R. Langenhorst, and D. Jäger. Generation and optimum control of picosecond optical pulses from gain-switched semiconductor lasers. *IEEE Journal of Quantum Electronics*, 24(8):1519–1523, 1988. 14
- [143] A. M. Pawlikowska, A. Halimi, R. A. Lamb, and G. S. Buller. Single-photon three-dimensional imaging at up to 10 kilometers range. *Optics Express*, 25(10):11919–11931, 2017. 7
- [144] A. K. Pediredla, A. C. Sankaranarayanan, M. Buttafava, A. Tosi, and A. Veeraghavan. Signal processing based pile-up compensation for gated single-photon avalanche diodes. arXiv:1806.07437 [physics.ins-det], June 2018. 124, 126, 127, 129
- [145] S. Pommé, R. Fitzgerald, and J. Keightley. Uncertainty of nuclear counting. *Metrologia*, 52(3):S3–S17, June 2015. 129
- [146] A. Raghuram, A. Pediredla, S. G. Narasimhan, I. Gkioulekas, and A. Veeraghavan. STORM: Super-resolving transients by oversampled measurements. In *Proceedings of the IEEE International Conference on Computational Photography (ICCP)*, pages 44–54, 2019. 107
- [147] J. Rapp and V. K. Goyal. A few photons among many: Unmixing signal and noise for photon-efficient active imaging. *IEEE Transactions on Computational Imaging*, 3(3):445–459, 2017. 8, 36
- [148] J. Rapp, R. M. A. Dawson, and V. K. Goyal. Dither-enhanced lidar. In *Proceedings of the OSA Imaging and Applied Optics Congress*, page JW4A.38, Orlando, FL, June 2018. 9, 70
- [149] J. Rapp, R. M. A. Dawson, and V. K. Goyal. Improving lidar depth resolution with dither. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, pages 1553–1557, Oct. 2018. 9, 70

- [150] J. Rapp, R. M. A. Dawson, and V. K. Goyal. Estimation from quantized Gaussian measurements: When and how to use dither. *IEEE Transactions on Signal Processing*, 67(13):3424–3438, July 2019. 9, 70
- [151] J. Rapp, Y. Ma, R. M. A. Dawson, and V. K. Goyal. Dead time compensation for high-flux depth imaging. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 7805–7809. IEEE, May 2019. 10, 123
- [152] J. Rapp, Y. Ma, R. M. A. Dawson, and V. K. Goyal. Dead time compensation for high-flux ranging. *IEEE Transactions on Signal Processing*, 67(13):3471–3486, July 2019. 10, 123
- [153] J. Rapp, Y. Ma, R. M. A. Dawson, and V. K. Goyal. Markov chain modeling for high-flux single-photon detection with dead times. In *Proceedings of the OSA Imaging and Applied Optics Congress*, page MM2D.3, Munich, Germany, June 2019. OSA. 10, 123
- [154] X. Ren and J. Malik. Learning a classification model for segmentation. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 10–17, 2003. 47
- [155] X. Ren, P. W. R. Connolly, A. Halimi, Y. Altmann, S. McLaughlin, I. Gyongy, R. K. Henderson, and G. S. Buller. High-resolution depth profiling using a range-gated CMOS SPAD quanta image sensor. *Optics Express*, 26(5):5541, 2018. 106
- [156] A. Restrepo and A. C. Bovik. Adaptive trimmed mean filters for image restoration. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(8):1326–1337, 1988. 86, 88
- [157] A. Ribeiro and G. B. Giannakis. Bandwidth-constrained distributed estimation for wireless sensor networks - part I: Gaussian case. *IEEE Transactions on Signal Processing*, 54(3):1131–1143, Mar. 2006. 74
- [158] G. W. Roberts and M. Ali-Bakhshian. A brief introduction to time-to-digital and digital-to-time converters. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 57(3):153–157, Mar. 2010. 20
- [159] L. Roberts. Picture coding using pseudo-random noise. *IEEE Transactions on Information Theory*, 8(2):145–154, Feb. 1962. 74
- [160] B. Rossi and N. Nereson. Experimental arrangement for the measurement of small time intervals between the discharges of Geiger-Müller counters. *Review of Scientific Instruments*, 17(2):65–71, Feb. 1946. 19

- [161] D. Rousseau and G. V. Anand. Nonlinear estimation from quantized signals: Quantizer optimization and stochastic resonance. In *Third International Symposium on Physics in Signal and Image Processing*, pages 89–92, 2003. 74
- [162] T. Sanchez, M. Zhang, D. Needleman, and E. Seli. Metabolic imaging via fluorescence lifetime imaging microscopy for egg and embryo assessment. *Fertility and Sterility*, 111(2):212–218, Feb. 2019. 4
- [163] G. Satat, M. Tancik, and R. Raskar. Towards photography through realistic fog. In *Proceedings of the IEEE International Conference on Computational Photography (ICCP)*, pages 1–10, May 2018. 16
- [164] D. Scharstein and C. Pal. Learning conditional random fields for stereo. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, June 2007. xvii, 56, 57, 58, 159
- [165] T. Schönau, S. Riecke, A. Bültner, and K. Lauritsen. Modern pulsed diode laser sources for time-correlated photon counting. In P. Kapusta, M. Wahl, and R. Erdmann, editors, *Advanced Photon Counting*, Springer Series on Fluorescence (Methods and Applications), pages 71–87. Springer, Cham, 2014. 14
- [166] L. Schuchman. Dither signals and their effect on quantization noise. *IEEE Transactions on Communications*, 12(4):162–165, Dec. 1964. 74, 75
- [167] Y. Sepulcre, T. Trigano, and Y. Ritov. Sparse regression algorithm for activity estimation in  $\gamma$  spectrometry. *IEEE Transactions on Signal Processing*, 61(17):4347–4359, Sept. 2013. 129
- [168] W. F. Sheppard. On the calculation of the average square, cube, of a large number of magnitudes. *Journal of the Royal Statistical Society*, 60(3):698–703, Sept. 1897. 73
- [169] W. F. Sheppard. On the calculation of the most probable values of frequency-constants, for data arranged according to equidistant divisions of a scale. *Proceedings of the London Mathematical Society*, 29:353–380, 1898. 73
- [170] D. Shin, A. Kirmani, V. K. Goyal, and J. H. Shapiro. Photon-efficient computational 3D and reflectivity imaging with single-photon detectors. *IEEE Transactions on Computational Imaging*, 1(2):112–125, June 2015. xvii, xviii, 7, 23, 28, 32, 33, 36, 37, 43, 45, 47, 49, 52, 54, 55, 57, 60, 62, 64, 127
- [171] D. Shin, F. Xu, D. Venkatraman, R. Lussana, F. Villa, F. Zappa, V. K. Goyal, F. N. Wong, and J. H. Shapiro. Photon-efficient imaging with a single-photon camera. *Nature Communications*, 7:12046, 2016. 7, 22, 28, 36, 37, 69, 70

- [172] D. Shin, F. Xu, F. N. C. Wong, J. H. Shapiro, and V. K. Goyal. Computational multi-depth single-photon imaging. *Optics Express*, 24(3):1873–1888, Feb. 2016. 31, 55
- [173] R. E. Slusher. Laser technology. *Reviews of Modern Physics*, 71(2):S471–S479, Mar. 1999. 13
- [174] D. L. Snyder and M. I. Miller. *Random Point Processes in Time and Space*. Springer Science & Business Media, 2012. 25, 27, 40, 130, 131
- [175] H. Soury and M. S. Alouini. New results on the sum of two generalized Gaussian random variables. In *Proceedings of the IEEE Global Conference on Signal and Information Processing*, pages 1017–1021, Dec. 2015. 81, 120
- [176] A. B. Sripad and D. L. Snyder. A necessary and sufficient condition for quantization errors to be uniform and white. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 25(5):442–448, Oct. 1977. 73, 74, 75
- [177] M. J. Stevens. Photon statistics, measurements, and measurements tools. In A. Migdall, S. V. Polyakov, J. Fan, and J. C. Bienfang, editors, *Single-Photon Generation and Detection*, chapter 2, pages 25–68. Academic Press, 2013. 17, 18
- [178] M. J. Stevens, R. H. Hadfield, R. E. Schwall, S. W. Nam, R. P. Mirin, and J. A. Gupta. Fast lifetime measurements of infrared emitters using a low-jitter superconducting single-photon detector. *Applied Physics Letters*, 89(3):031109, July 2006. 109
- [179] S. M. Stigler. Simon Newcomb, Percy Daniell, and the history of robust estimation 1885-1920. *Journal of the American Statistical Association*, 68(344):31, 1973. 85
- [180] S. M. Stigler. The anonymous Professor Gergonne. *Historia Mathematica*, 3(1):71–74, Feb. 1976. 85
- [181] K. Suhling, L. M. Hirvonen, J. A. Levitt, P. H. Chung, C. Tregidgo, A. Le Marois, D. A. Rusakov, K. Zheng, S. Ameer-Beg, S. Poland, S. Coelho, R. Henderson, and N. Krstajic. Fluorescence lifetime imaging (FLIM): Basic concepts and some recent developments. *Medical Photonics*, 27:3–40, 2015. 3, 4
- [182] M. J. Sun, M. P. Edgar, G. M. Gibson, B. Sun, N. Radwell, R. Lamb, and M. J. Padgett. Single-pixel three-dimensional imaging with time-based depth resolution. *Nature Communications*, 7:12010, July 2016. 22
- [183] A. Swatantran, H. Tang, T. Barrett, P. DeCola, and R. Dubayah. Rapid, high-resolution forest structure and terrain mapping over large areas using single photon lidar. *Scientific Reports*, 6(1):28277, Sept. 2016. 7

- [184] J. Tachella, Y. Altmann, S. McLaughlin, and J.-Y. Tourneret. Fast surface detection in single-photon lidar waveforms. In *Proceedings of the European Signal Processing Conference*, Sept. 2019. 31
- [185] J. Tachella, Y. Altmann, S. McLaughlin, and J.-Y. Tourneret. 3D reconstruction using single-photon lidar data exploiting the widths of the returns. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 7815–7819, May 2019. 16, 31
- [186] J. Tachella, Y. Altmann, X. Ren, A. McCarthy, G. S. Buller, S. McLaughlin, and J.-Y. Tourneret. Bayesian 3D reconstruction of complex scenes from single-photon lidar data. *SIAM Journal on Imaging Sciences*, 12(1):521–550, Jan. 2019. 31, 67
- [187] M. Tanveer, I. Nissinen, J. Nissinen, J. Kostamovaara, J. Borg, and J. Johansson. Time-to-digital converter based on analog time expansion for 3D time-of-flight cameras. In *Proceedings of SPIE The International Society for Optical Engineering*, volume 9022, page 90220A, San Francisco, CA, USA, Feb. 2014. 70
- [188] R. Tobin, A. Halimi, A. McCarthy, M. Laurenzis, F. Christnacher, and G. S. Buller. Three-dimensional single-photon imaging through obscurants. *Optics Express*, 27(4):4590–4611, Feb. 2019. 16, 67
- [189] I. Tasic and S. Drewes. Learning joint intensity-depth sparse representations. *IEEE Transactions on Image Processing*, 23(5):2122–2132, 2014. 69
- [190] T. Trigano, T. Dautremer, E. Barat, and A. Souloumiac. Pile-up correction algorithms for nuclear spectrometry. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 4, pages 441–444, 2005. 129
- [191] T. Trigano, Y. Sepulcre, and Y. Ritov. Sparse reconstruction algorithm for nonhomogeneous counting rate estimation. *IEEE Transactions on Signal Processing*, 65(2):372–385, Jan. 2017. 129
- [192] J. W. Tukey. The future of data analysis. *The Annals of Mathematical Statistics*, 33(1):1–67, 1962. 85
- [193] J. W. Tukey and D. H. McLaughlin. Less vulnerable confidence and significance procedures for location based on a single sample: Trimming/Winsorization 1. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, 25(3):331–352, 1963. 85

- [194] M. Van den Bergh, D. Carton, and L. Van Gool. Depth seeds: Recovering incomplete depth data using superpixels. In *Proceedings of the IEEE Workshop on Applications of Computer Vision (WACV)*, pages 363–368, 2013. 48, 49
- [195] H. L. Van Trees, K. L. Bell, and Z. Tian. *Detection, Estimation, and Modulation Theory: Detection, Estimation, and Filtering Theory*. John Wiley & Sons, Inc., Hoboken, NJ, 2 edition, 2013. 78, 119
- [196] M. K. Varanasi and B. Aazhang. Parametric generalized Gaussian density estimation. *Journal of the Acoustical Society of America*, 86(4):1404–1415, 1989. 80, 83
- [197] S. B. Vardeman. Sheppard’s correction for variances and the “quantization noise model”. *IEEE Transactions on Instrumentation and Measurement*, 54(5):2117–2119, Oct. 2005. 73
- [198] S. B. Vardeman and C.-S. Lee. Likelihood-based statistical estimation from quantized data. *IEEE Transactions on Instrumentation and Measurement*, 54(1):409–414, Feb. 2005. 72, 77, 93
- [199] D. Venkatraman. Data repository for photon-efficient pseudo-array imaging. <https://github.com/photon-efficient-imaging/sample-data>, 2015. 63
- [200] G. Verma and R. J. Drost. Inhomogeneous Poisson process rate function inference from dead-time limited observations. *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, 34(5):770, May 2017. 127
- [201] F. Villa, R. Lussana, D. Portaluppi, A. Tosi, and F. Zappa. Time-resolved CMOS SPAD arrays: architectures, applications and perspectives. In *Proceedings of SPIE The International Society for Optical Engineering*, volume 10212, page 102120J, Anaheim, CA, June 2017. 70
- [202] H. Viswanathan and R. Zamir. On the whiteness of high-resolution quantization errors. *IEEE Transactions on Information Theory*, 47(5):2029–2038, July 2001. 73
- [203] I. Vornicu, R. Carmona-Galan, and A. Rodriguez-Vazquez. Arrayable voltage-controlled ring-oscillator for direct time-of-flight image sensors. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 64(11):2821–2834, 2017. 70
- [204] M. F. Wagdy. Effect of various dither forms on quantization errors of ideal A/D converters. *IEEE Transactions on Instrumentation and Measurement*, 38(4):850–855, 1989. 103

- [205] M. Wahl. Modern TCSPC electronics: Principles and acquisition modes. In P. Kapusta, M. Wahl, and R. Erdmann, editors, *Advanced Photon Counting: Applications, Methods, Instrumentation*, pages 1–21. Springer International Publishing, Cham, 2015. 20
- [206] M. Wahl. Personal communication, July 2019. 20
- [207] M. Wahl, H. J. Rahn, I. Gregor, R. Erdmann, and J. Enderlein. Dead-time optimized time-correlated photon counting instrument with synchronized, independent timing channels. *Review of Scientific Instruments*, 78(3), 2007. 19, 128
- [208] M. Wahl, H. J. Rahn, T. Röhlicke, G. Kell, D. Nettels, F. Hillger, B. Schuler, and R. Erdmann. Scalable time-correlated photon counting system with multiple independent input channels. *Review of Scientific Instruments*, 79(12):123113, Dec. 2008. 20, 125
- [209] M. Wahl, T. Röhlicke, H.-J. Rahn, R. Erdmann, G. Kell, A. Ahlrichs, M. Kernbach, A. W. Schell, and O. Benson. Integrated multichannel photon timing instrument with very short dead time and high throughput. *Review of Scientific Instruments*, 84(4):043102, Apr. 2013. 20, 125
- [210] X. F. Wang, S. Kitajima, T. Uchida, D. M. Coleman, and S. Minami. Time-resolved fluorescence microscopy using multichannel photon counting. *Applied Spectroscopy*, 44(1):25–30, Jan. 1990. 4
- [211] R. A. Wannamaker, S. P. Lipshitz, J. Vanderkooy, and J. N. Wright. A theory of nonsubtractive dither. *IEEE Transactions on Signal Processing*, 48(2):499–516, 2000. 74
- [212] M. Ware, A. Migdall, J. C. Bienfang, and S. V. Polyakov. Calibrating photon-counting detectors to high accuracy: Background and deadtime issues. *Journal of Modern Optics*, 54(3-4):361–372, 2007. 18
- [213] R. Whyte, L. Streeter, M. J. Cree, and A. A. Dorrington. Application of lidar techniques to time-of-flight range imaging. *Applied Optics*, 54(33):9654, Nov. 2015. 5, 6
- [214] B. Widrow. Statistical analysis of amplitude-quantized sampled-data systems. *Transactions of the American Institute of Electrical Engineers Part 2: Applications and Industry*, 79(6):555–568, 1961. 73, 74
- [215] L. Xu, Y. Zhang, Y. Zhang, C. Yang, X. Yang, and Y. Zhao. Restraint of range walk error in a Geiger-mode avalanche photodiode lidar to acquire high-precision depth and intensity information. *Applied Optics*, 55(7):1683, Mar. 2016. 127

- [216] L. Ye, G. Gu, W. He, H. Dai, and Q. Chen. A real-time restraint method for range walk error in 3-D imaging lidar via dual detection. *IEEE Photonics Journal*, 10(2), 2018. 129
- [217] D. F. Yu and J. A. Fessler. Mean and variance of single photon counting with deadtime. *Physics in Medicine and Biology*, 45(7):2043–2056, July 2000. 126, 129
- [218] B. Zandbelt. exgauss: a MATLAB toolbox for fitting the ex-Gaussian distribution to response time data. June 2014. doi: 10.6084/m9.figshare.971318.v2. URL <https://figshare.com/articles/exgauss/971318>. 111
- [219] Q. Zhao, H.-W. Li, and Y.-T. Shen. On the sum of generalized Gaussian random signals. In *Proceedings of the International Conference on Signal Processing (ICSP)*, volume 1, pages 50–53, Aug. 2004. 81
- [220] J. Zheng. Coherence analysis of optical frequency-modulated continuous-wave interference. *Applied Optics*, 45(16):3681–3687, June 2006. 6
- [221] A. W. Ziarkash, S. K. Joshi, M. Stipčević, and R. Ursin. Comparative study of afterpulsing behavior and models in single photon counting avalanche photo diode detectors. *Scientific Reports*, 8(1):5076, Dec. 2018. 18
- [222] C. L. Zitnick and S. B. Kang. Stereo for image-based rendering using image over-segmentation. *International Journal of Computer Vision*, 75(1):49–65, 2007. 48, 49
- [223] A. Zymnis, S. Boyd, and E. Candès. Compressed sensing with quantized measurements. *IEEE Signal Processing Letters*, 17(2):149–152, Feb. 2010. 87

# CURRICULUM VITAE

