Optimal Threshold Design for Quanta Image Sensor

Omar A. Elgendy, *Student Member, IEEE* and Stanley H. Chan, *Member, IEEE*

Abstract—**Quanta Image Sensor (QIS) is a binary imaging device envisioned to be the next generation image sensor after CCD and CMOS. Equipped with a massive number of single photon detectors, the sensor has a threshold q above which the number of arriving photons will trigger a binary response "1", or "0" otherwise. Existing methods in the device literature typically assume that $q = 1$ uniformly. We argue that a spatially varying threshold can significantly improve the signal-to-noise ratio of the reconstructed image. In this paper, we present an optimal threshold design framework. We make two contributions. First, we derive a set of oracle results to theoretically inform the maximally achievable performance. We show that the oracle threshold should match exactly with the underlying pixel intensity. Second, we show that around the oracle threshold there exists a set of thresholds that give asymptotically unbiased reconstructions. The asymptotic unbiasedness has a phase transition behavior which allows us to develop a practical threshold update scheme using a bisection method. Experimentally, the new threshold design method achieves better rate of convergence than existing methods.**

Index Terms—**Quanta image sensor, single-photon imaging, high dynamic range, binary quantization, maximum likelihood.**

I. INTRODUCTION

A. Threshold Design for Quanta Image Sensor

Quanta Image Sensor (QIS) is a class of solid-state image sensors envisioned to be the next generation imaging device after CCD and CMOS. Originally proposed by Eric Fossum in 2005 [1], the sensor has gained significant momentum in the past decade, both in terms of hardware design [2]–[4] and image processing [5]–[9]. The advantage of QIS over the mainstream CCD and CMOS is attributed to its high spatial resolution (e.g., 10^9 pixels per sensor with 200nm pitch per pixel [10]) and high speed (e.g., 100k fps as reported in [11]). However, in order to simplify circuit, minimize power and reduce data transfer, QIS is operated in a binary mode: When the number of photons arriving at the sensor exceeds a threshold q, the sensor generates a binary bit "1". When the number of photons is less than q, the sensor generates a "0". The goal of this paper is to address the question of how to optimally choose q to maximize the signal-to-noise ratio of the reconstructed image.

Optimal threshold design for QIS is important as it directly affects the dynamic range of an image. Figure 1 illustrates an example where we simulate the raw binary data acquired by

(a) Observed, $q = 3$ (b) Observed, $q = q^*(c)$ (c) Observed, $q = 12$

(d) Reconstruction, $q = 3$ (e) Reconstruction, $q = 12$

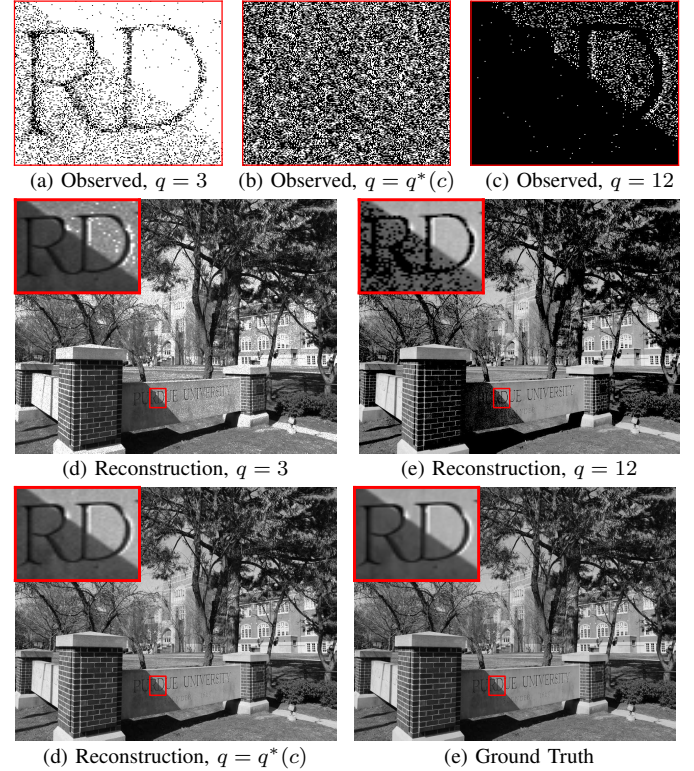(d) Reconstruction, $q = q^*(c)$ (e) Ground Truth

Fig. 1.   Simulated QIS data and the reconstructed gray-scale images using different thresholds. Top row: The binary measurements obtained using thresholds $q = 3$, $q = q^*(c)$, and $q = 12$. Bottom figures: The maximum likelihood estimates obtained from the binary measurements, with comparison to the ground truth. The results show that our spatially varying threshold $q^*(c)$ offers the best reconstruction. In this experiment, we spatially oversample each pixel by $K = 2 \times 2$ binary bits and we use $T = 25$ independent temporal measurements.

a QIS using a uniform threshold q. When q is low, most of the bits in the raw input are "1". The reconstructed image is therefore an over-exposed image. On the other hand, when q is high, most of the bits in the raw input are "0". The reconstructed image is then under-exposed. In both cases, it is evident from the simulation that a uniform threshold has limited performance. A better way is to allow q to vary spatially so that a pixel (or a group of pixels) has its own threshold value. The result in Figure 1(d) shows the reconstruction result using a spatially varying threshold obtained from our proposed technique, which is clearly better than the uniform thresholds.

B. Scope and Contributions

The goal of this paper is to present an optimal threshold design methodology and provide theoretical justifications. The two major contributions are summarized as follows.

First, we provide a rigorous theoretical analysis of the performance limit of the image reconstruction as a function of the threshold. These results form the basis of our subsequent discussions of the threshold update scheme. Some results are

TABLE I
LIST OF QIS PROTOTYPES AND PARAMETERS

Camera	Canon 5D CMOS	EMCCD [12]	GMAPD [13]	SPC SPAD [14]	SwissSPAD [11]	Fossum QIS [15]
Price	$5,000	$20,000	Prototype	Prototype	Prototype	Prototype
Resolution	4096×2160	1024×1024	256×256	320×240	512×128	1376×768
Pixel Pitch (μm)	2.3	13	25	8	24	3.6
Full-well capacity	69 ke- (@ISO100)	180 ke-	-	$56 - 125$ e-		$1 - 250$ e-
Frames per second (fps)	6	$26 - 92$	8×10^3	2×10^4	1.56×10^5	1×10^3
Sensor data rate	88.6 Mbps	0.48 Gbps	0.52 Gbps	1.54 Gbps	10.24 Gbps	1 Gbps

known, e.g., the signal-to-noise ratio is a function of the Fisher Information [16], [17], but a number of new results are shown. In particular, we show that (i) the maximum likelihood estimate has a closed-form expression in terms of the incomplete Gamma function (Section III.B); (ii) the oracle threshold can be derived in closed-form by maximizing the signal-to-noise ratio (Section III.C); (iii) the image reconstruction has a phase transition behavior (Section IV.A - Section IV.D).

Second, we propose an efficient threshold update scheme based on our theoretical results. The new scheme is a bisection method which iteratively updates the threshold *without* the need of reconstructing the image. By checking whether the proportion of one's and zero's approaches 0.5 in a spatial-temporal block, the threshold is guaranteed to be near optimal. Compared to other existing threshold update schemes such as [18] and [19]–[21], the new scheme offers significantly faster rate of convergence (Section IV.E). We also demonstrate how the dynamic range can be extended for high dynamic range (HDR) imaging (Section IV.F).

A preliminary version of this paper was presented in ICIP 2016 [8]. This journal version contains significantly more details including complete proofs of major results, more comprehensive comparisons with existing methods, and discussions of HDR imaging.

II. BACKGROUND

A. Current State of QIS

Quanta Image Sensor (QIS) belongs to the family of photon-counting devices. These photon-counting devices have been known for a long time. Some better-known examples are the electron-multiplying charge-coupled device (EMCCD) [22], [23], single-photon avalanche diode (SPAD) [11], [14], [24], Geiger-mode avalanche photodiode (GMAPD) [13], etc. The common feature of these devices is their single photon sensitivity, which makes them useful in medical imaging [25]–[27], astronomy [28], defense [29], nuclear engineering [30], depth and reflectivity reconstruction [31], ultra-fast low-light tracking [32], and recently in quantum random number generation used in cryptography [33], [34].

The concept of QIS was first proposed by Fossum in 2005 as a solution for sub-diffraction limit pixels. The sensor was called the digital film sensor, and later the quanta image sensor [15], [35], [36]. After the introduction of QIS, researchers in EPFL developed a similar concept called the Gigavision camera [6], [37], [38]. Recently, teams at the University of Edingburgh [14], [24], [39] and EPFL [33], [40] have made new progresses in QIS using binary single photon detectors.

In the industry, Rambus Inc. (Sunnyvale, CA) has developed binary image sensors for high dynamic range imaging [19]–[21]. Table I lists several recent QIS prototypes that are available or are currently being developed. As a comparison we also show a Canon 5D Mark III CMOS camera. Among many different features, the most noticeable is the frame rate. For example, SPS SPAD can be operated at 20k fps. SwissSPAD can even achieve 156k fps. Both are significantly faster than a standard CMOS camera.

B. Related Work on Threshold Design

Existing work on QIS threshold design study can be summarized into three classes of methods.

- Markov Chain [18]. The Markov Chain method developed by Hu and Lu [18] is a time-sequential update scheme. A Markov Chain probability is used to control how easy the threshold should be increased or decreased. While the method has provable convergence, the threshold of each single photon detector of the QIS has to be updated sequentially in time. In contrast, our proposed method allows a group of single photon detectors to share the same threshold. As a result, our proposed method has significantly faster rate of convergence.
- Conditional Reset [19]–[21]. The conditional reset method is a hardware solution proposed by Vogelsang and colleagues. The idea is to take a sequence of images with ascending (or descending) thresholds, and digitally integrate the sequence to form an image. The drawback of the method, besides the additional hardware cost of the per-pixel reset transistors, is the limited quality of the reconstructed image. For the same number of frames, our proposed method produces better images.
- Checkerboard Threshold [16]. This method constructs a checkerboard of thresholds by alternating two threshold values q_1 and q_2. The optimality criterion of q_1 and q_2 is based on minimizing the Cramér-Rao lower bound (CRLB) integrated over a range of light intensities, which is essentially an average case result. Our proposed method obtains the optimal threshold for each pixel. This per-pixel optimization has higher reconstruction performance compared to checkerboard threshold.

C. QIS Imaging Model

In this subsection we provide an overview of the QIS imaging model. The model has been previously discussed in several papers, e.g., [6]–[9]. Readers interested in details can refer to these papers for further explanations.
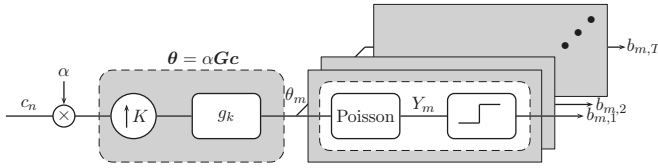
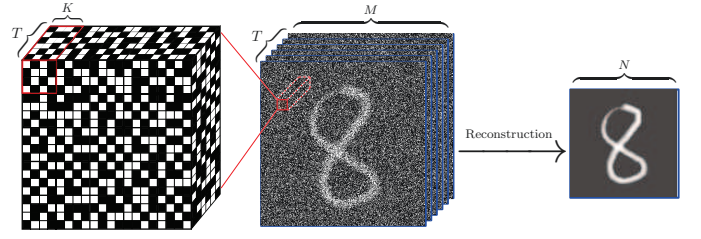Fig. 2. Block diagram illustrating the image formation process of QIS.

Fig. 3. Image reconstruction of QIS data. Given the binary bit planes, the reconstruction algorithm estimates the gray-scale image shown on the right.

1) Spatial Oversampling: We denote the discrete version of the light intensity as a vector $c = [c_0, \ldots, c_{N-1}]^T$, where $n = 0, \ldots, N-1$ specify the spatial coordinates. We assume that c_n is normalized to the range $[0, 1]$ for all n so that there is no scaling ambiguity. To model the actual light intensity, we multiply c_n by a constant α to yield αc_n, where $\alpha > 0$ is a fixed scalar constant.

Given the N-dimensional vector c, QIS uses $M \gg N$ tiny pixels called *jots* to sample c. The ratio $K \stackrel{\text{def}}{=} M/N$ is known as the spatial oversampling factor. The oversampling process is illustrated in Figure 2, where it first upsamples the vector c by a factor of K, and then filters the output by a lowpass filter $\{g_k\}$. Mathematically, the process can be expressed as

$$\theta = \alpha G c, \tag{1}$$

where $\theta = [\theta_0, \ldots, \theta_{M-1}]^T$ denotes the light intensity sampled at the M jots, and the matrix G is defined as

$$G = \frac{1}{K} I_{N \times N} \otimes 1_{K \times 1}, \tag{2}$$

where $1_{K \times 1}$ is a vector of all ones and $\otimes$ denotes the Kronecker product. Note that the choice of G in (2) is the result of simplifying the model by assuming that the lowpass filter is $g_k = 1/K$ for all k. This assumption is typically reasonable, because on each QIS jot there is a micro-lens to focus the incident light. Although previous papers, e.g., [6], [7], do not make such assumption, in this paper we decide to use a simplified G, for otherwise the theoretical analysis will become very complicated. Nevertheless, in the Supplementary Material we show comparison between a general G and the simplified G. The gap is usually insignificant.

2) Truncated Poisson Process: We assume that the operating speed of QIS is significantly faster than the scene motion. Therefore, for a given scene c (and also θ), we are able to acquire a set of T independent measurements. We illustrate this using the T channels in Figure 2.

The oversampled signal θ generates a sequence of Poisson random variables according to the distribution

$$\mathbb{P}(Y_{m,t} = y_{m,t}) = \frac{\theta_m^{y_{m,t}} e^{-\theta_m}}{y_{m,t}!}, \tag{3}$$

where $m = 0, 1, \ldots, M-1$ denotes the m-th jot of the QIS and $t = 0, 1, \ldots, T-1$ denotes the t-th independent measurement in time. Denoting $q \in \mathbb{N}$ as the quantization threshold, the final observed binary measurement $B_{m,t}$ is a truncation of $Y_{m,t}$:

$$B_{m,t} = \begin{cases} 0, & \text{if } Y_{m,t} < q. \\ 1, & \text{if } Y_{m,t} \geq q \end{cases}$$

The probability mass function of $B_{m,t}$ is given by

$$\mathbb{P}(B_{m,t} = b_{m,t}) = \begin{cases} \sum_{k=0}^{q-1} \frac{\theta_m^k e^{-\theta_m}}{k!}, & \text{if } b_{m,t} = 0, \\ \sum_{k=q}^{\infty} \frac{\theta_m^k e^{-\theta_m}}{k!}, & \text{if } b_{m,t} = 1. \end{cases} \tag{4}$$

The goal of image reconstruction is to recover the underlying image c from the binary measurements $\mathcal{B} = \{B_{m,t} \mid m = 0, \ldots, M-1, \text{and } t = 0, \ldots, T-1\}$. A pictorial illustration of the reconstruction is shown in Figure 3.

3) Properties of Truncated Poisson Processes: The probability mass function of $B_{m,t}$ in (4) is Bernoulli. However, the right hand side of (4) involves infinite sums which are difficult to interpret. To simplify the equations, we consider the upper incomplete Gamma function $\Psi_q : \mathbb{R}^+ \to [0, 1]$ defined in [41] as:

$$\Psi_q(\theta) \stackrel{\text{def}}{=} \frac{1}{\Gamma(q)} \int_\theta^\infty t^{q-1} e^{-t} dt, \quad \text{for } \theta > 0, \ q \in \mathbb{N}.$$

where $\Gamma(q) = (q-1)!$ is the standard Gamma function. The incomplete Gamma function allows us to rewrite the infinite sums in (4) using the following identity [41]:

$$\Psi_q(\theta) = \sum_{k=0}^{q-1} \frac{\theta^k}{k!} e^{-\theta}. \tag{5}$$

Consequently, the probabilities in (4) become

$$\mathbb{P}(B_{m,t} = 0) = \Psi_q(\theta_m),$$
$$\mathbb{P}(B_{m,t} = 1) = 1 - \Psi_q(\theta_m). \tag{6}$$

Example 1. *In the special case of $q = 1$, we obtain:*

$$\mathbb{P}(B_{m,t} = 0) = \frac{1}{\Gamma(1)} \int_{\theta_m}^\infty t^0 e^{-t} dt = e^{-\theta_m},$$

which coincides with the results shown in [6] and [7].

The incomplete Gamma function $\Psi_q(\theta)$ is a decreasing function of θ because the first order derivative of $\Psi_q(\theta)$ with respect to θ is negative:

$$\frac{d}{d\theta} \Psi_q(\theta) = \frac{-\theta^{q-1} e^{-\theta}}{\Gamma(q)} < 0, \quad \forall q \in \mathbb{N}, \text{ and } \theta > 0. \tag{7}$$

The limiting behavior of $\Psi_q(\theta)$ is important. For a fixed q, the function $\Psi_q(\theta) \to 1$ as $\theta \to 0$ and $\Psi_q(\theta) \to 0$ as $\theta \to \infty$. While Ψ_q^{-1} still exists in these situations because Ψ_q is monotonically decreasing, for a given z the value $\Psi_q^{-1}(z)$ could be numerically very difficult to evaluate. To characterize the sets of θ and q that Ψ_q is (numerically) invertible, we define the *θ-admissible set* and the *q-admissible set*.

Definition 1. *The θ-admissible set and q-admissible set of the incomplete Gamma function are*

$$\Theta_q \overset{def}{=} \{\theta \mid \varepsilon \le \Psi_q(\theta) \le 1 - \varepsilon\},$$
$$\mathcal{Q}_\theta \overset{def}{=} \{q \mid \varepsilon \le \Psi_q(\theta) \le 1 - \varepsilon\}, \qquad (8)$$

respectively, where $0 < \varepsilon < \frac{1}{2}$ is a constant.

More discussions of the incomplete Gamma function can be found in the Supplementary Material.

Remark 1. *In this paper, we assume that QIS is noise-free, i.e., the only source of randomness is the truncated Poisson random variable. In real sensors, there will be readout noise, photo-response non-uniformity caused by conversion gain variation, dark count rate (a.k.a. dark current), optical crosstalk and electronic crosstalk. See [42] for details.*

III. OPTIMAL THRESHOLD: THEORY

A. Image Reconstruction by MLE

We begin the optimal threshold design by discussing image reconstruction because the optimality of the threshold is measured with respect to the reconstructed image. However, since QIS is a new device, the number of reconstruction methods is limited. A few examples that can be found in the literature are the gradient descent [6], dynamic programming [43], ADMM [7], and Transform-Denoise method [9], and neural network [44]. In this paper, we shall focus on the maximum likelihood estimation (MLE) approach as it provides closed-form expressions.

Given $\mathcal{B}$, MLE solves the following optimization problem:

$$\widehat{c} \overset{(a)}{=} \underset{c}{\operatorname{argmax}} \prod_{t=0}^{T-1} \prod_{m=0}^{M-1} \mathbb{P}[B_{m,t} = 1\,;\,\theta_m]^{b_{m,t}}$$
$$\times \mathbb{P}[B_{m,t} = 0\,;\,\theta_m]^{1-b_{m,t}}$$
$$\overset{(b)}{=} \underset{c}{\operatorname{argmax}} \sum_{t=0}^{T-1} \sum_{m=0}^{M-1} \Big\{ b_{m,t} \log(1 - \Psi_q(\theta_m))$$
$$+ (1 - b_{m,t}) \log \Psi_q(\theta_m) \Big\}, \quad (9)$$

subject to the constraint that $\boldsymbol{\theta} = \alpha \boldsymbol{G} \boldsymbol{c}$. Here, the right hand side of (a) is the likelihood function of a Bernoulli random variable, and (b) follows from taking the logarithm. With the $\boldsymbol{G}$ defined in (2), we can partition $\mathcal{B}$ into N blocks $\{\mathcal{B}_1, \ldots, \mathcal{B}_N\}$ where each block is

$$\mathcal{B}_n \overset{def}{=} \{B_{Kn+k,t} \mid k = 0, \ldots, K-1, t = 0, \ldots, T-1\}.$$

Then, the pixel $\widehat{c}_n$ can be estimated as follows.

Proposition 1 (Closed-form ML Estimate). *The solution of the MLE in (9) is*

$$\widehat{c}_n = \frac{K}{\alpha} \Psi_q^{-1} \left(1 - \frac{S_n}{KT}\right), \qquad (10)$$

where $S_n \overset{def}{=} \sum_{t=0}^{T-1} \sum_{k=0}^{K-1} B_{Kn+k,t}$ is the sum of bits in the n-th block $\mathcal{B}_n$.

Proof: See [9]. ∎

B. Signal-to-Noise Ratio of ML Estimate

In order to determine the optimal threshold, we need to quantify the performance of the ML estimate. The performance metric we use is the signal-to-noise ratio of the ML estimate at every pixel $\widehat{c}_n$. Considering each $\widehat{c}_n$ individually is allowed here because they are independently determined according to (10). For notation simplicity we drop the subscript n in the subsequent discussions.

Definition 2. *The signal-to-noise ratio (SNR) of the ML estimate $\widehat{c}$ is defined as*

$$\operatorname{SNR}_q(c) \overset{def}{=} 10 \log_{10} \frac{c^2}{\mathbb{E}[(\widehat{c} - c)^2]}, \qquad (11)$$

where the expectation is taken over the probability mass function of the binary measurements in (6).

The difficulty of working with $\operatorname{SNR}_q(c)$ is that it does not have a simple closed-form expression. In view of this, Lu [17] showed that the SNR is asymptotically linear to the log of the Fisher Information.

Proposition 2. *As $KT \to \infty$,*

$$\operatorname{SNR}_q(c) \approx 10 \log_{10} \left(c^2 I_q(c)\right) + 10 \log_{10} KT, \qquad (12)$$

where $I_q(c) \overset{def}{=} \mathbb{E}_B \left[\frac{-\partial^2}{\partial c^2} \log \mathbb{P}(B = b; \theta) \right]$ is the Fisher Information measuring the amount of information that the random variable B carries about the unknown value c.

Proof: See [17]. ∎

While the asymptotic result shown in Proposition 2 has significantly simplified the SNR, we still need to determine the Fisher Information. The following proposition gives a new result of the Fisher Information with arbitrary q.

Proposition 3. *The Fisher Information $I_q(c)$ of the probability mass function in (6) under a threshold q is:*

$$I_q(c) = \left(\frac{\alpha}{K}\right)^2 \frac{e^{-2\left(\frac{\alpha c}{K}\right)} \left(\frac{\alpha c}{K}\right)^{2q-2}}{\Gamma^2(q) \Psi_q\left(\frac{\alpha c}{K}\right) \left(1 - \Psi_q\left(\frac{\alpha c}{K}\right)\right)}. \qquad (13)$$

Proof: See Appendix A-A. ∎

Substituting (13) into (12), we observe that the SNR can be approximated as

$$\operatorname{SNR}_q(c) \approx 10 \log_{10} \frac{KT e^{-2\left(\frac{\alpha c}{K}\right)} \left(\frac{\alpha c}{K}\right)^{2q}}{\Gamma(q)^2 \Psi_q\left(\frac{\alpha c}{K}\right) \left(1 - \Psi_q\left(\frac{\alpha c}{K}\right)\right)}, \qquad (14)$$

which is characterized by the unknown pixel value c, the threshold q, the spatial oversampling ratio K and the number of temporal measurements T. To understand the behavior of (14), we show in Figure 4 $\operatorname{SNR}_q(c)$ as a function of c for different thresholds $q \in \{1, \ldots, 16\}$. For a fixed q, $\operatorname{SNR}_q(c)$ is a convex function with a unique maximum. The goal of optimal threshold design is to determine a q which maximizes $\operatorname{SNR}_q(c)$ for a fixed c.

Remark 2. *The $\operatorname{SNR}_q(c)$ in (14) can also be derived from a concept in the device literature called the* exposure-referred SNR *[45]. See Supplementary Material for discussions.*
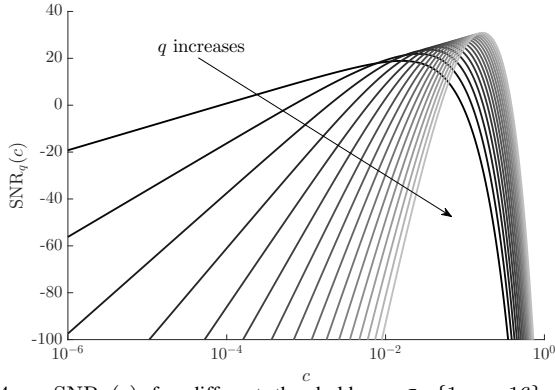
Fig. 4. $\mathrm{SNR}_q(c)$ for different thresholds $q \in \{1, \ldots, 16\}$. In this experiment, we set $\alpha = 400$, $K = 4$, and $T = 30$. For fixed q, $\mathrm{SNR}_q(c)$ is always a convex function.

C. Oracle Threshold

We now discuss the optimal threshold design in the oracle setting. We call the result oracle because the optimal threshold depends on the unknown pixel intensity c. The practical threshold design scheme will be discussed in Section IV.

Using the definition of the signal-to-noise ratio, the optimal threshold is determined by maximizing $\mathrm{SNR}_q(c)$ with respect to q:

$$q^* = \underset{q \in \mathbb{N}}{\mathrm{argmax}} \ \mathrm{SNR}_q(c) = \underset{q \in \mathbb{N}}{\mathrm{argmax}} \ \log(c^2 I_q(c)). \quad (15)$$

The second equality follows from Proposition 2. Substituting (13) yields an expression of the right hand side of (15). To further simplify the expression we derive the following lower bound.

Proposition 4. *The function $\log(c^2 I_q(c))$ is lower bounded as follows.*

$$\log(c^2 I_q(c)) \geq \underbrace{2 \left(\log 2 - \frac{\alpha c}{K} + q \log \frac{\alpha c}{K} - \log \Gamma(q) \right)}_{\overset{def}{=} L_q(c)}.$$

Proof: See Appendix A-B. ∎

Using this lower bound, we can derive the optimal threshold q as follows [1].

Proposition 5. *The optimal threshold $q^*(c)$ is*

$$q^*(c) = \underset{q \in \mathbb{N}}{\mathrm{argmax}} \ L_q(c) = \left\lfloor \frac{\alpha c}{K} \right\rfloor + 1, \quad (16)$$

where $\lfloor \cdot \rfloor$ denotes the flooring operator that returns the largest integer smaller than or equal to the argument.

Proof: See Appendix A-C. ∎

The result of Proposition 5 is important as it states that the oracle threshold is *exactly the same* as the light intensity $\alpha c/K$. The flooring operation and the addition of a constant 1 are not crucial here because they are only used to ensure that q is an integer. In [18], a special where $\alpha = 1$ was demonstrated experimentally. Proposition 5 now provides a theoretical justification.

[1] Straightly speaking, the result shown in Proposition 5 is a "near-optimal" result because we are minimizing the lower bound. From our experience, the gap between the near-optimality and the exact optimality is typically insignificant.

IV. OPTIMAL THRESHOLD: PRACTICE

The oracle threshold derived in the previous section provides a theoretical foundation but is practically infeasible as it requires knowledge of the ground truth c. In this section, we present an alternative solution by relaxing the optimality criteria. Our strategy is to consider a set of thresholds which are close to the oracle threshold $q^*(c)$, and show that they are asymptotically unbiased when the number of observed bits approaches infinity (Section IV.A). This result will allow us to characterize the estimate $\hat{c}$ (Section IV.B). We will then show that there exists a phase transition region where the asymptotic unbiasedness is maintained as q stays within a certain range around $q^*(c)$, and is lost rapidly as q falls outside this range (Section IV.C - IV.D). Based on these observations, we will present a practical threshold update scheme (Section IV.E).

A. Asymptotic Unbiasedness

In order to derive an alternative threshold that does not require the ground truth, we start by reconsidering the ML estimate $\hat{c}$ in Proposition 1. For a spatial-temporal block $\mathcal{B} = \{B_{k,t} \mid 0 \leq k \leq K-1, 0 \leq t < T-1\}$, the ML estimate $\hat{c}$ satisfies the condition

$$\Psi_q \left(\frac{\alpha \hat{c}}{K} \right) = 1 - \frac{S}{KT}, \quad (17)$$

where $S = \sum_{k,t} B_{k,t}$ is the sum of bits in $\mathcal{B}$. The right hand side of this equation is an important quantity. We denote it as

$$\gamma_q(c) \overset{\text{def}}{=} 1 - \frac{S}{KT}. \quad (18)$$

In the device literature (e.g., [45]), the term $1 - \gamma_q(c)$ is known as the *bit-density* as it is the proportion of ones in $\mathcal{B}$. Note that $\gamma_q(c)$ is a random variable because S is the sum of KT i.i.d. random binary bits. Therefore, if we want to understand (17), we must first derive the the mean and variance of $\gamma_q(c)$.

Proposition 6. *The mean and variance of $\gamma_q(c)$ are*

$$\mathbb{E}[\gamma_q(c)] = \Psi_q \left(\frac{\alpha c}{K} \right), \ \ and$$

$$\mathrm{Var}[\gamma_q(c)] = \frac{1}{KT} \Psi_q \left(\frac{\alpha c}{K} \right) \left[1 - \Psi_q \left(\frac{\alpha c}{K} \right) \right], \quad (19)$$

respectively.

Proof: See Appendix A-D. ∎

We can now look at the asymptotic behavior of $\gamma_q(c)$ to see if it offers any insight about the optimal threshold. Applying the strong law of large number to S/KT, we can show that as $KT \to \infty$,

$$\gamma_q(c) = 1 - S/KT \overset{a.s.}{\to} 1 - \mathbb{E}[B_{k,t}] = \Psi_q(\alpha c/K). \quad (20)$$

Going back to (17)-(18), the ML estimate $\hat{c}$ should have the expectation:

$$\mathbb{E}[\hat{c}] \overset{(a)}{=} \frac{K}{\alpha} \mathbb{E} \left[\Psi_q^{-1} \left(\gamma_q(c) \right) \right]$$
$$\overset{(b)}{\to} \frac{K}{\alpha} \Psi_q^{-1} \Psi_q \left(\frac{\alpha c}{K} \right) \overset{(c)}{=} c. \quad (21)$$

where (a) follows from the definition of $\hat{c}$, (b) follows from (20), and (c) holds because Ψ_q and Ψ_q^{-1} cancels each other.

What is the implication of (21)? It shows that the ML estimate $\widehat{c}$ is asymptotically unbiased. That is, as the number of independent measurements grows, the estimate $\widehat{c}$ approaches to the ground truth c. In other words, as long as KT is large enough, the random variable $\widehat{c}$ would be an accurate estimate of the ground truth. How can this be used to determine the threshold q? Let us look at $\mathcal{Q}_\theta$.

B. Set of Admissible Thresholds $\mathcal{Q}_\theta$

The result in (17)-(21) shows that for a given S (or equivalently $\gamma_q(c)$), the ML estimate can be found by

$$\widehat{c} = \frac{K}{\alpha} \Psi_q^{-1}\left(\gamma_q(c)\right). \tag{22}$$

When this happens, the $\widehat{c}$ given by (22) is asymptotically unbiased. However, the inversion Ψ_q^{-1} is not always allowed. There is a set of q's that can make Ψ_q invertible, which is defined as $\mathcal{Q}_\theta$ in Definition 1. The following proposition relates $\mathcal{Q}_\theta$ to $\gamma_q(c)$.

Proposition 7. *Let $0 < \delta < 1$ be a constant. Then, for any*

$$q \in \mathcal{Q}_\theta \stackrel{def}{=} \left\{ q \, \middle| \, 1 - \left(\frac{\delta}{2}\right)^{\frac{1}{KT}} \leq \Psi_q(\theta) \leq \left(\frac{\delta}{2}\right)^{\frac{1}{KT}} \right\}, \tag{23}$$

the random variable $\gamma_q(c)$ will not attain 0 or 1 with probability at least $1 - \delta$, i.e.,

$$\mathbb{P}[0 < \gamma_q(c) < 1] > 1 - \delta.$$

In this case, the ML estimate $\widehat{c}$ is uniquely defined by (22).

Proof: See Appendix A-E. ∎

Before we proceed, let us look at some rough magnitude of the parameters in the following example.

Example 2. *Let the ground truth pixel value be $c = 0.5$. The sensor parameters are set as $T = 50$, $K = 4$, $\alpha = 300$. For a constant $\delta = 2 \times 10^{-4}$, the tolerance level is $\varepsilon = 1 - (\delta/2)^{1/KT} = 0.045$. Therefore, as long as $q \in \{q \,|\, 0.045 \leq \Psi_q(\theta) \leq 1 - 0.045\}$, which is the set $\{q \,|\, 28 \leq q \leq 48\}$, the probability that $\gamma_q(c)$ equals to 0 or 1 is upper bounded by $\delta = 2 \times 10^{-4}$.*

C. Gap between $\mathcal{Q}_\theta$ and q^*

The result in the previous subsection shows that as long as $q \in \mathcal{Q}_\theta$, the ML estimate is asymptotic unbiased. However, how is a $q \in \mathcal{Q}_\theta$ compared to the oracle threshold q^*? We answer this question in three parts.

First, does an asymptotically unbiased estimate maximize the SNR? The answer is no, because Proposition 5 states that if q^* is the optimal threshold, then $\mathrm{SNR}_{q^*}(c) \geq \mathrm{SNR}_q(c)$ for any $q \neq q^*$. Therefore, moving from the exact optimal q^* to an asymptotically unbiased threshold is a relaxation of the optimality criteria.

If asymptotic unbiasedness is a relaxed optimality criteria, how much SNR drop will there be if we choose a $q \in \mathcal{Q}_\theta$ but not necessarily $q = q^*$? We show in Figure 5 the plot of a typical experiment with setup discussed in Example 2. As shown in the figure, the green zone is the set $\mathcal{Q}_\theta = \{q \,|\, 28 \leq$
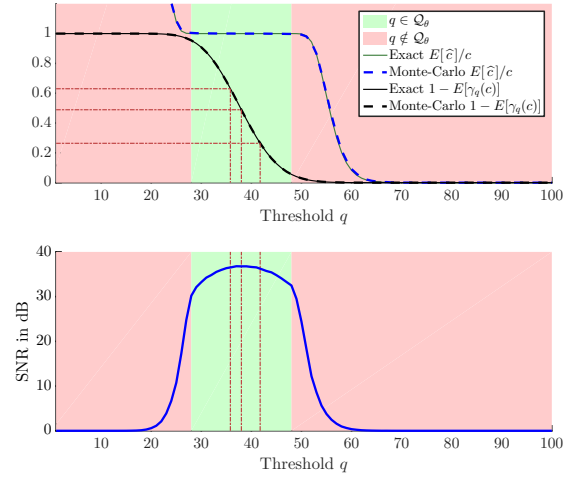

Fig. 5. Phase transition of the ML estimate and its relationship to the average bit density $1 - \mathbb{E}[\gamma_q(c)]$. The red region is where it is impossible to recover c, whereas the green region is where we can have perfect recovery.

$q \leq 48\}$, or equivalently $\mathcal{Q}_\theta = \{q \,|\, 0.045 \leq \Psi_q(\theta) \leq 0.9955\}$. For any q in this $\mathcal{Q}_\theta$, the reconstruction has a SNR at least 30dB. If we further tighten $\mathcal{Q}_\theta$ so that $\mathcal{Q}_\theta = \{q \,|\, 35 \leq q \leq 42\}$, or equivalently $\mathcal{Q}_\theta = \{q \,|\, 0.25 \leq \Psi_q(\theta) \leq 0.6\}$, the SNR stays in the range $36.15\mathrm{dB} \leq \mathrm{SNR}_q(c) \leq 36.65\mathrm{dB}$, which is reasonably narrow.

How tight should $\mathcal{Q}_\theta$ be? Ideally we want $\mathcal{Q}_\theta$ to be as tight as possible. But knowing the fact that the incomplete Gamma function has a rapid transition (See the black line in Figure 5), $\mathcal{Q}_\theta$ can be much wider. In fact, we can choose $\mathcal{Q}_\theta$ such that $1 - \gamma_q(c)$ stays close to 0.5, so that we are guaranteed to obtain a near optimal threshold. From an information theoretic point of view, $1 - \gamma_q(c) \approx 0.5$ is where the bit density attains the maximum information — if q is too high then most bits become 0 whereas if q is too low then most bits become 1. It is maximum when q leads to 50% zeros and 50% ones. [2]

D. Phase Transition Phenomenon

We can now point out a very interesting phenomenon in Figure 5. In the upper plot of Figure 5 we show two sets of curves: blue curves (solid and dotted), and black curves (solid and dotted). The black curves represent the ratio $\mathbb{E}[\widehat{c}]/c$, and the black curves represent the average bit density $1 - \mathbb{E}[\gamma_q(c)]$. For both sets of curves, we use dotted lines to illustrate the Monte-Carlo simulation using 10,000 random samples, where each sample refers to a spatial-temporal block $\mathcal{B}_n$ containing $KT = 200$ binary bits. Notice that these dotted lines overlap exactly with their expectations, and hence (17)-(21) are valid.

Let us take a closer look at the blue curve $\mathbb{E}[\widehat{c}]/c$. Let $\mathcal{Q}_\theta = \{q \,|\, q_L \leq q \leq q_H\}$, where q_L and q_H are the smallest and the largest integers in $\mathcal{Q}_\theta$ respectively. There are three distinct phases:

- When $q < q_L$, the threshold is low and so most bits become 1. Therefore, $\gamma_c(q) \to 0$ and hence $\widehat{c} \to \infty$. Thus, $\mathbb{E}[\widehat{c}]/c \to \infty$ as q decreases.

---

[2] The exact optimal value of $1 - \gamma_q(c)$ at q^* is slightly lower than 0.5 due to the nonlinearity of the Gamma function. See Supplementary Material for additional discussion.

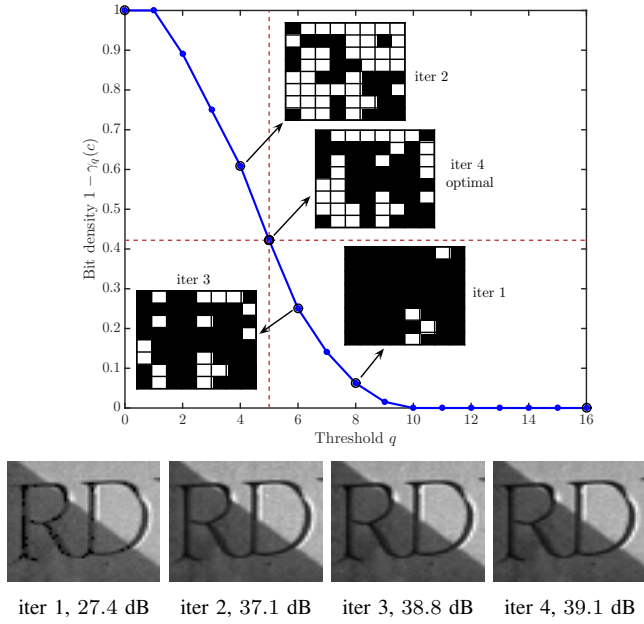iter 1, 27.4 dB   iter 2, 37.1 dB   iter 3, 38.8 dB   iter 4, 39.1 dB

Fig. 6. The proposed bisection update scheme adjusts the threshold q such that the bit density $1 - \gamma_q(c)$ approaches 0.5. The upper graph illustrates the bisection steps. Bottom row shows cropped patches from reconstructed images using threshold maps at different iterations and the PSNRs.

- When $q > q_H$, the threshold high and so most bits become 0. Therefore, $\gamma_c(q) \to 1$ and hence $\hat{c} \to 0$. Thus, $\mathbb{E}[\hat{c}]/c \to 0$ as q increases.
- When $q_L \leq q \leq q_H$, the ML estimate $\hat{c}$ is asymptotically unbiased. Therefore, $\mathbb{E}[\hat{c}]/c = 1$.

Essentially, Figure 5 demonstrates a phase transition behavior of the threshold. Such phase transition exists because Ψ_q is only invertible when $q \in \mathcal{Q}_\theta$.

E. Bisection Threshold Update Scheme

Now we present a practical threshold update scheme. As we discussed in Section IV.C, the oracle threshold q^* can be obtained when bit density $\gamma_q(c)$ is close to 0.5. Therefore, a practical procedure to determine q is to sweep through a range of q until the bit density reaches 0.5. To achieve this objective, we propose a bisection method illustrated in Figure 6 and Algorithm 1. Starting with initial thresholds q_A and q_B, we check whether the bit density satisfies $1 - \gamma_{q_A} > 0.5$ and $1 - \gamma_{q_B} < 0.5$. If this is the case, then we find a mid point $q_M = (q_A + q_B)/2$ and check whether $1 - \gamma_{q_M}$ is greater or less than 0.5. If $1 - \gamma_{q_M} > 0.5$, we replace q_A by q_M, otherwise we replace q_B by q_M. The process repeats until $1 - \gamma_{q_M}$ is sufficiently close to 0.5.

In our proposed threshold update scheme, we assume that the image has been partitioned into N blocks $\{\mathcal{B}_n \mid n = 0, \ldots, N - 1\}$. Each $\mathcal{B}_n$ contains KT binary bits and is used to estimate one pixel value c_n. This setting results in N different thresholds, one for every pixel. To generalize the setting, it is also possible to allow multiple pixels to share a common threshold. Figure 7 shows an example. The advantage of sharing a threshold for multiple pixels is that circuits associated with the sensor can be simplified. In terms of performance, since neighboring pixels are typically correlated, sharing the threshold causes little drop in the resulting SNR.

Algorithm 1 Bisection Threshold Update Scheme

Initial thresholds q_A and q_B such that $1 - \gamma_{q_A} > 0.5$ and $1 - \gamma_{q_B} < 0.5$.

Compute $q_M = \lceil (q_A + q_B)/2 \rceil$, where $\lceil \cdot \rceil$ denotes the ceiling operator.

while $|\gamma_{q_M} - 0.5| < $ `tol` **do**

    If $\gamma_{q_M} < 0.5$, then set $q_A = q_M$. Else, set $q_B = q_A$.

    Compute $q_M = \lceil (q_A + q_B)/2 \rceil$.
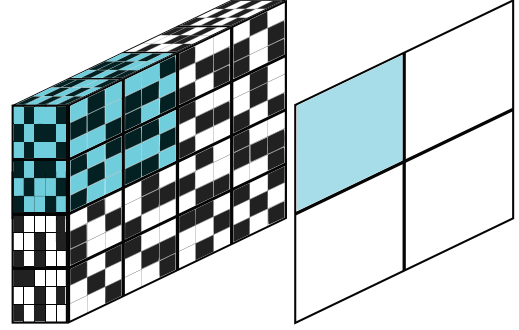
end while

return q_M

Fig. 7. Concept of shared thresholds. (Left) binary measurements, spatial oversampling $K = 3 \times 3$, Temporal oversampling $T = 5$. (Right) Threshold map, one threshold value is shared by 6×6 jots.

The price that the proposed bisection algorithm has to pay is the number of frames it requires to determine a good q. For every evaluation of γ_{q_M}, the sensor has to physically acquire one frame and compute the bit density in each of the N blocks. Therefore, the more bisection steps we need, the more frames that the sensor has to physically acquire. The rate of convergence of the proposed method and existing methods will be compared in Section V.

F. Extension to High Dynamic Range

While QIS is a photon counting device, it is designed to count a few photons to keep the full-well capacity small, e.g. 20 photoelectrons as reported in [46]. Therefore, for practical imaging tasks, we need to extend the dynamic range for QIS.

There are two ways to enable dynamic range extension:

- **Bright Scenes: Reduce Duty Cycle.** In the signal processing block diagram shown in Figure 2, we can replace the constant α by a fraction as $\alpha\tau$, where $0 \leq \tau \leq 1$ determines the ratio between the actual integration time and the readout scan time. It can also be referred to as the shutter duty cycle because the shutter is opened to collect photons during this proportion of time [47]. For very bright scenes, a low duty cycle will prevent QIS from saturating early.
- **Dark Scenes: Multiple Measurements.** For dark scenes, multiple measurements can be taken to ensure enough photons over the measurement period. This, however, is different from conventional HDR imaging. In conventional HDR imaging, the multiple shots are taken at different shutter speeds, e.g., 1/8192, 1/2048, 1/512, 1/128, 1/32, 1/8, 1/2 seconds [48], which is redundant. QIS's multiple shot functions more similar to burst photography
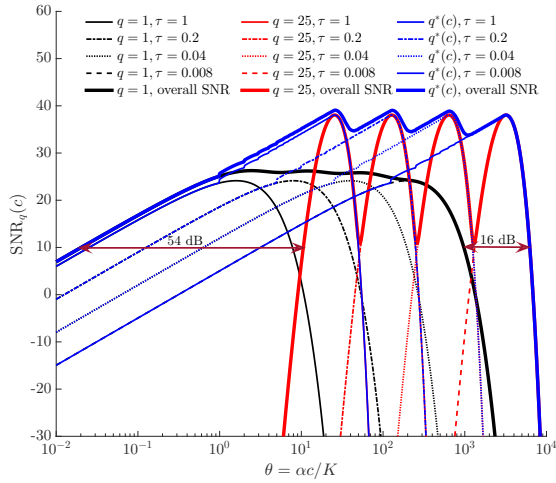
Fig. 8. SNR in dB vs. exposure θ for HDR imaging mode obtained by fusion of frames with shutter duty cycles $\tau \in \{1, 0.2, 0.04, 0.008\}$. Three scenarios are shown: constant threshold with $q = 1$ (black), $q = 25$ (red) and an optimal spatially varying threshold (blue).

[49]. The amount of acquisition time is significantly less than the conventional HDR imaging.

These two methods can be used for *any* threshold scheme, including ours and others. The benefit of using our proposed threshold scheme is that it supports a much wider dynamic range extension. In Figure 8, we illustrate the total dynamic range that can be covered using 4 multiple measurements at duty cycles $\tau = 1$, $\tau = 0.2$, $\tau = 0.04$, and $\tau = 0.008$. The maximum threshold level is $q_{\max} = 25$, and the minimum threshold level is $q_{\min} = 1$. It can be seen from the figure that with the optimal threshold q^*, the dynamic range is significantly more than the non-optimal ones. In particular, we observe a 16dB and a 54dB improvement compared to $q_{\min} = 1$ and $q_{\max} = 25$, respectively. Experimental results will be shown in Section V.C.

G. Hardware Consideration

Concerning the hardware implementation, we anticipate that future QIS will be equipped with per-pixel FPGAs to perform the proposed threshold update scheme. On-sensor FPGA is an actively developing technology. For example, MIT Lincoln Lab's digital focal plane array can achieve on-sensor image stabilization and edge detection [50] . For QIS threshold update, the complexity is low because we are only counting the number of ones in the bisection. More specifically, in order to perform the bisection, we only need K additions to compute $\sum_{k=0}^{K-1} b_{Kn+k,t}$; one comparison $\sum_{k=0}^{K-1} b_{Kn+k,t} \geq 0.5$; one addition and one multiplication (with a constant 0.5) to update the threshold $q_M = \lceil (q_A + q_B)/2 \rceil$. The dominating factor here is the K additions, which can be implemented efficiently by shifting bits in a buffer.

We should also point out that the proposed bisection method can be flexibly adjusted spatially and temporally for different hardware configurations. For example, we can use a spatial-temporal window $4 \times 4 \times 1$ for low-resolution high-speed imaging, or $1 \times 1 \times 16$ for high-resolution low-speed imaging. This flexibility offers additional advantages of QIS over conventional CCD and CMOS cameras.

V. Experimental Results

In this section we evaluate the proposed threshold update scheme by comparing it with existing methods. We consider two evaluation metrics: (1) convergence rate of the threshold update methods; (2) quality of the reconstructed images. For reconstruction evaluation, we create our own Purdue dataset comprising 77 images captured by a Canon EOS Rebel T6i camera. For HDR imaging, we use the HDR-Eye dataset by Nemoto et al. [51], [52]. In all experiments, we fix the spatial over-sampling factor as $K = 4 \times 4 = 16$, and number of temporal frames as $T = 13$. The maximum threshold level is set as $q_{\max} = 16$ to ensure that it is realistic for today's QIS.

A. Convergence

We compare the proposed threshold update scheme with the Markov Chain (MC) adaptation proposed by Hu and Lu [18]. The Markov Chain adaptation models the threshold as a variable with 2^L states. These 2^L states can be regarded as 2^L steps before reaching to the next threshold level. The probability of changing from one state to another is controlled by a parameter $1 - \beta$ with $0 < \beta < 1$. When a bit arrives, the state will be updated (increased or decreased) or will remain unchanged. Once the state is increased by 2^L times, the threshold will be increased by one.

When comparing Markov Chain adaptation with the proposed bisection algorithm, one should be aware of the difference between the two methods. Markov Chain adaptation is a *per-jot* update scheme whereas the proposed bisection algorithm is a *per-pixel* update scheme. For a pixel with $K \times K$ jots, Markov Chain adaptation needs K^2 iterations to update the threshold *sequentially*. In contrast, the proposed bisection algorithm updates a common threshold for all K^2 jots *simultaneously*. Thus in practice our bisection algorithm is significantly less complex to implement in hardware than the Markov Chain. In order to take the different forms of updates into account, we treat the K^2 iterations of Markov Chain adaptation as one "major iteration" and compare it with the one bisection step of the proposed algorithm.

The first comparison we make is to check the threshold at different jots. Figure 9 shows the results of three typical runs with underlying optimal thresholds $q^* = 1, 8, 16$. In this experiment, we generate 100 random binary blocks of size $K \times K$ and estimate the threshold at each major iteration. We report the average of these 100 estimates to minimize the randomness of the data. The results show that one iteration of the proposed bisection algorithm works as good as the K^2 iterations of the Markov Chain adaptation. In some cases, Markov Chain tends to oscillate whereas the bisection result is stable.

The second comparison we make is to check how close the estimated threshold is compared to the optimal threshold. The optimal threshold q^* is obtained using the oracle scheme. In Figure 10, we plot the mean squared error between the estimated threshold and the oracle threshold. For fairness, we show the results of the MSE averaged over the 77 images of our dataset, and 50 random samples per image. One threshold is shared by $K \times K$ jots, and each $K \times K$ jots correspond
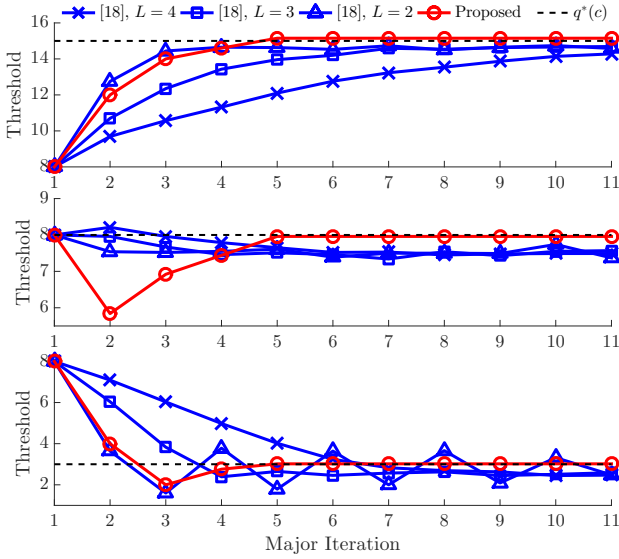
Fig. 9. Convergence of the threshold at 3 jots. Each curve is averaged over 100 random samples. The red curve indicates the proposed bisection method. The blue curves are the Markov chain adaptation [18] with $\beta = 0.25$. Note that one major iteration of Markov Chain adaptation corresponds to K^2 sequential updates, and one major iteration of the bisection method corresponds to a single update to K^2 jots simultaneously.
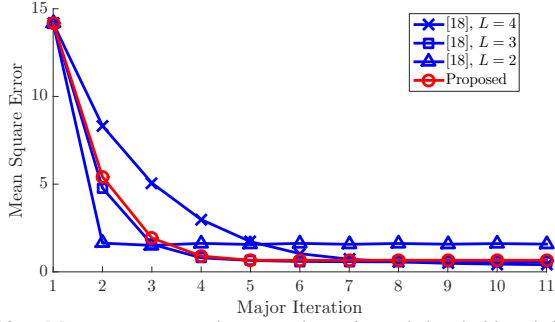

Fig. 10. Mean square error between the estimated threshold and the ideal oracle threshold. Each curve is averaged over 50 random samples and 77 images. The red curve indicates the proposed bisection method. The blue curves are the Markov chain adaptation [18] with $\beta = 0.25$.

to one pixel. The result is consistent with the ones shown in Figure 9.

B. Image Reconstruction Quality

The convergence comparison in the previous subsection is only useful to compare threshold update methods that actually return a threshold. In the QIS literature, there are methods that implicitly update the threshold, e.g., the conditional reset method [21]. For comparison with these methods, we have to compare the quality of the image reconstructed from the binary raw data. The image reconstruction is done using the closed-form ML estimate in Section III-A.

We consider three classes of methods:

- Uniform Threshold. Uniform threshold is commonly used in the device literature [5]–[7]. A uniform threshold is a single threshold applied to all pixels in the image. In this experiment, we consider the following choices of uniform thresholds: $q = 1$, $q = 5$, $q = 10$ and $q = 16$.
- Conditional Reset [21]. Conditional reset counts the number of photons and is reset when it is above the threshold.

TABLE II
AVERAGE PSNR AND STANDARD DEVIATION OF 77 RECOVERED IMAGES USING DIFFERENT Q-MAPS AND 50 RANDOM SAMPLES.

Configuration		Average PSNR	Std
Uniform Threshold	$q = 1$	10.30	0.01
	$q = 5$	28.80	0.04
	$q = 10$	23.22	0.02
	$q = 16$	12.95	0.01
Conditional Reset [21]	Ascending q sequence	23.77	0.52
	Descending q sequence	24.95	0.53
Proposed Method	$2K^2 \times 2K^2$	30.14	0.06
	$K^2 \times K^2$	31.18	0.06
	$K \times K$	32.78	0.02

The threshold in conditional reset is sequentially increasing or decreasing. The reconstructed image is obtained by digitally integrating the raw binary frames.
- Proposed Method. As we discussed in Section IV-E, the proposed method can be implemented to let multiple pixels share a common threshold. Thus, in this experiment we consider three sharing strategies: (1) Share a threshold between a neighborhood of $K \times K$ jots (i.e., one threshold for one pixel); (2) Share a threshold between a neighborhood of $K^2 \times K^2$ jots (i.e., one threshold for $K \times K$ pixel); (3) Share a threshold between a neighborhood of $2K^2 \times 2K^2$ jots (i.e., one threshold for $2K \times 2K$ pixels).

The result of the experiment is shown in Table II. The PSNR values reported are averaged over 77 images in our dataset. Each image generates 50 random realizations, and the PSNR of an image is averaged over these 50 random realizations to minimize the randomness. As shown in the table, while conditional reset generally performs better than a uniform threshold, it performs significantly worse than the proposed threshold update scheme.
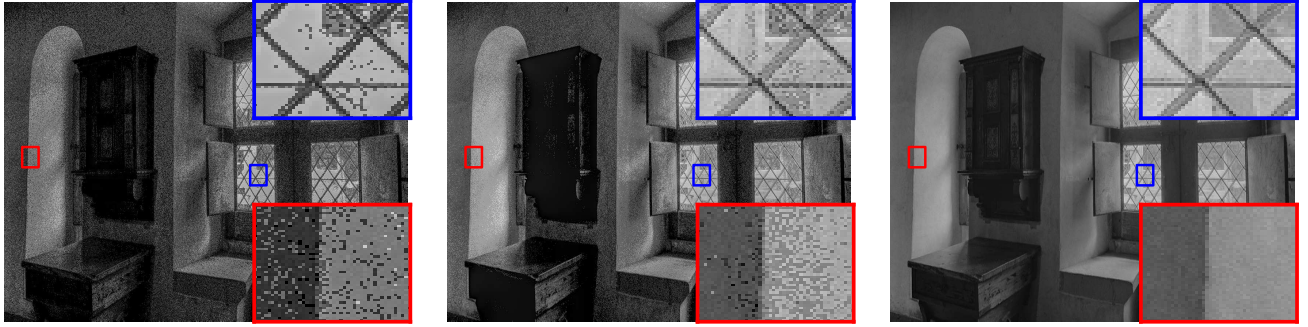
C. Influence of QIS Threshold on HDR Imaging

Since QIS does not have sufficient full well capacity to accumulate photons for HDR imaging, we apply the dynamic range extension method discussed in Section IV-F. When different threshold schemes are used, the reconstructed HDR images will be affected. The objective of this experiment is to evaluate the influence of the threshold in HDR imaging.

In this experiment, we consider the HDR-Eye image dataset [51], [52]. Each HDR image in this dataset contains 9 images acquired at different exposure settings ($-2.7, -2, -1.3, -0.7, 0, 0.7, 1.3, 2,$ and 2.7 EV). A snapshot of these images are shown in Figure 11. From each exposure, we simulate the photon counts resulting from the luminance channel. The sensor gain is set as $\alpha = K^2(q_{max} - 1)$ to ensure proper number of photons, where $K = 4 \times 4 = 16$ and $q_{max} = 16$. On the reconstruction side, we reconstruct the 9 images using the MLE discussed in Section III-A. Tone mapping and exposure fusion [53] are applied to the 9 imags to generate an HDR image. As a reference, we apply the same tone mapping and

Fig. 11.   Bracketed images with different exposure settings. From Left to Right: $-2.7, -2, -1.3, -0.7, 0, 0.7, 1.3, 2$, and 2.7 EV.

$q = 1$, PSNR $= 17.94$ dB $q = 16$, PSNR $= 20.77$ dB   Proposed, PSNR $= 31.46$ dB

Fig. 12.   The reconstructed HDR images using different thresholds. See supplementary material for additional results.

fusion algorithm to the 9 ground truth images. PSNR between the reference and the estimated is then recorded.

The result of this experiment is shown in Figure 12. With the proposed threshold update scheme, the reconstructed images achieve the highest PSNR value and visual quality. When $q = 1$, which is too low, the image appears under-exposed. When $q = 16$, which is too high, the image appears over-exposed. The spatially varying property of the proposed method mitigates the issue by allowing multiple thresholds.

In practice, one would typically add image denoisers to handle the randomness in the ML estimate and potentially other types of noise. This can be done using methods such as [9]. In HDR literature, there are also optical approaches that reduce the number of exposures, e.g., [54], [55]. These techniques are complementary to QIS, because QIS is a sensor of similar functionality of a CMOS. Thus optical techniques can always be added.

VI. CONCLUSION

Quanta Image Sensor is a new image sensor for high speed, high resolution and high dynamic range imaging. The sensor has a threshold which needs to be carefully adjusted so that the dynamic range can be maximized. We studied the threshold design problem by establishing several theoretical results. First, we showed that an oracle threshold can be obtained assuming that we know the underlying pixel value. Our result showed that the oracle threshold must match with the pixel value in order to maximize the signal-to-noise ratio. Second, we showed that around the oracle threshold, there exists a set of thresholds that can produce asymptotically unbiased estimates of the pixel value. Within this set of threshold, the signal-to-noise ratio stays very close to the oracle case. Third, we developed a bisection method to update the threshold scheme. We also discussed how QIS can be used in HDR imaging, and its advantages compared to standard sensors. Experimental results showed the effectiveness of our proposed approach compared to the standard approach that uses uniform threshold for all pixels.

APPENDIX A

A. Proof of Proposition 3

The Fisher Information metric is defined as:

$$I_q(c) \stackrel{\text{def}}{=} \mathbb{E}_B \left[\frac{-\partial^2}{\partial c^2} \log \mathbb{P}(B = b; \theta, q) \right], \qquad (24)$$

where $\theta = \alpha c / K$. Using the chain rule, we can derive the Fisher Information as follows

$$I_q(c) = \left(\frac{\alpha}{K} \right)^2 \mathbb{E}_B \left[\frac{-\partial^2}{\partial \theta^2} \log \mathbb{P}(B = b; \theta, q) \right]. \qquad (25)$$

The expectation can be calculated as follows

$$I_q(c) = \left(\frac{\alpha}{K} \right)^2 \left[\frac{-\partial^2}{\partial \theta^2} \log \mathbb{P}(B = 1; \theta, q) \right] \mathbb{P}(B = 1; \theta, q)$$
$$+ \left(\frac{\alpha}{K} \right)^2 \left[\frac{-\partial^2}{\partial \theta^2} \log \mathbb{P}(B = 0; \theta, q) \right] \mathbb{P}(B = 0; \theta, q) \qquad (26)$$

Using (7) to differentiate the 1st term, we get:

$$\frac{\partial^2}{\partial \theta^2} \log \mathbb{P}(B = 1; \theta, q) = \frac{\partial^2}{\partial \theta^2} \log \left(1 - \Psi_q(\theta) \right)$$
$$= \frac{R'(1 - \Psi_q(\theta)) - R^2 / \Gamma(q)}{\Gamma(q) \left(1 - \Psi_q(\theta) \right)^2}, \qquad (27)$$

where $R = e^{-\theta} \theta^{q-1}$ and $R' = \partial R / \partial \theta$. Similarly, the second term is

$$\frac{\partial^2}{\partial \theta^2} \log \mathbb{P}(B = 0; \theta, q) = \frac{\partial^2}{\partial \theta^2} \log \Psi_q(\theta)$$
$$= -\frac{R' \Psi_q(\theta) + R^2 / \Gamma(q)}{\Gamma(q) \left(\Psi_q(\theta) \right)^2}. \qquad (28)$$

Substitute (27) and (28) in (26) yields

$$I_q(\theta) = \left(\frac{\alpha}{K}\right)^2 \left[-\frac{R'\Gamma(q)(1-\Psi_q(\theta)) - R^2}{\Gamma^2(q)\left(1-\Psi_q(\theta)\right)} \right.$$
$$\left. + \frac{R'\Gamma(q)\Psi_q(\theta) + R^2}{\Gamma^2(q)\Psi_q(\theta)} \right]$$
$$= \left(\frac{\alpha}{K}\right)^2 \frac{e^{-2\theta}\theta^{2q-2}}{\Gamma^2(q)\Psi_q(\theta)\left(1-\Psi_q(\theta)\right)}.$$

B. Proof of Proposition 4

The lower bound is obtained by observing that the product $\Psi_q(\theta)\left(1-\Psi_q(\theta)\right)$ attains its maximum value when $\Psi_q(\theta) = 1/2$. Substituting with the upper bound $\Psi_q(\theta)\left(1-\Psi_q(\theta)\right) \leq 1/4$, we get:

$$\log(c^2 I_q(c)) = \log\left\{\left(\frac{\alpha c}{K}\right)^2 \frac{e^{-2\theta}\theta^{2q-2}}{\Gamma^2(q)\Psi_q(\theta)\left(1-\Psi_q(\theta)\right)}\right\}$$
$$= \log\frac{e^{-2\theta}\theta^{2q}}{\Gamma^2(q)\Psi_q(\theta)\left(1-\Psi_q(\theta)\right)}$$
$$\geq \log\frac{4e^{-2\theta}\theta^{2q}}{\Gamma^2(q)}$$
$$= 2\log 2 - 2\theta + 2q\log\theta - 2\log\Gamma(q)$$
$$= 2\left(\log 2 - \frac{\alpha c}{K} + q\log\frac{\alpha c}{K} - \log\Gamma(q)\right).$$

C. Proof of Proposition 5

Using the definition of Gamma function $\Gamma(q) = (q-1)!$ and $\theta = \frac{\alpha c}{K}$, we can rewrite the lower bound in Proposition 4 as follows.

$$L_q(c) = 2\left(\log 2 - \theta + q\log\theta - \log(q-1)!\right)$$
$$= 2\left(\log 2 - \theta + (q-1)\log\theta + \log\theta - \log\prod_{k=1}^{q-1} k\right)$$
$$= 2\left(\log 2 - \theta + \sum_{k=1}^{q-1}\log(\theta/k) + \log\theta\right)$$

The only dependence on q is in the second term, so we take a closer look at it. When $q-1 < \lfloor\theta\rfloor$, all summands $\log(\theta/k)$ are positive because $k < \lfloor\theta\rfloor$. Hence, the total sum increases by increasing q. On the other hand, when $q-1 > \lfloor\theta\rfloor$, we start to add negative summands $\log(\theta/k)$ because $k > \theta$. Therefore, the total sum decreases on increasing $q-1$ over $\lfloor\theta\rfloor$. Thus, maximum is obtained at $q = \lfloor\theta\rfloor + 1 = \lfloor\frac{\alpha c}{K}\rfloor + 1$.

D. Proof of Proposition 6

By definition, $S \stackrel{\text{def}}{=} \sum_{t=0}^{T-1}\sum_{k=0}^{K-1} B_{k,t}$ is the summation of KT independent i.i.d. Bernoulli random variables. Therefore, S is a binomial random variable with parameters $n \stackrel{\text{def}}{=} KT$ and $p \stackrel{\text{def}}{=} 1 - \Psi(\alpha c/K)$. The mean and variance of a binomial random variable is $\mathbb{E}[S] = np$, and $\text{Var}[S] = np(1-p)$. Therefore, we have

$$\mathbb{E}\left[\gamma_q(c)\right] = 1 - \frac{\mathbb{E}[S]}{KT} = \Psi_q\left(\frac{\alpha c}{K}\right), \quad \text{and}$$
$$\text{Var}\left[\gamma_q(c)\right] = \frac{\text{Var}[S]}{K^2 T^2} = \frac{1}{KT}\Psi_q\left(\frac{\alpha c}{K}\right)\left(1 - \Psi_q\left(\frac{\alpha c}{K}\right)\right).$$

E. Proof of Proposition 7

The probability $\mathbb{P}[0 < \gamma_q(c) < 1]$ can be evaluated by checking the complement when $\gamma_q(c) = 0$ or $\gamma_q(c) = 1$:

$$\mathbb{P}[0 < \gamma_q(c) < 1] = 1 - \mathbb{P}[\gamma_q(c) = 0] - \mathbb{P}[\gamma_q(c) = 1]$$
$$= 1 - \mathbb{P}[S = 0] - \mathbb{P}[S = KT]$$
$$\stackrel{(a)}{=} 1 - \Psi_q(\theta)^{KT} - [1 - \Psi_q(\theta)]^{KT},$$

where (a) follows from the fact that S, which is a sum of i.i.d. Bernoulli random variables, is a binomial random variable.

Let $0 < \delta < 1$. If

$$1 - \left(\frac{\delta}{2}\right)^{\frac{1}{KT}} \leq \Psi_q(\theta) \leq \left(\frac{\delta}{2}\right)^{\frac{1}{KT}},$$

then we have

$$\Psi_q(\theta)^{KT} < \frac{\delta}{2} \quad \text{and} \quad [1 - \Psi_q(\theta)]^{KT} < \frac{\delta}{2}.$$

Thus, it holds that

$$1 - \Psi_q(\theta)^{KT} - [1 - \Psi_q(\theta)]^{KT} > 1 - \delta.$$

REFERENCES

[1] E. R. Fossum, "What to do with sub-diffraction-limit (SDL) pixels?– A proposal for a gigapixel digital film sensor (DFS)," in *Proc. IEEE Workshop CCDs Adv. Image Sensors*, Sep. 2005, pp. 214–217.

[2] J. Ma, D. Hondongwa, and E. R. Fossum, "Jot devices and the quanta image sensor," in *Proc. IEEE Int. Electron Devices Meeting*, Dec. 2014, pp. 10.1.1–10.1.4.

[3] J. Ma and E. R. Fossum, "A pump-gate jot device with high conversion gain for a quanta image sensor," *IEEE J. Electron Devices Soc.*, vol. 3, no. 2, pp. 73–77, Mar. 2015.

[4] J. Ma, L. Anzagira, and E. R. Fossum, "A 1 μm-pitch quanta image sensor jot device with shared readout," *IEEE J. Electron Devices Soc.*, vol. 4, no. 2, pp. 83–89, Mar. 2016.

[5] F. Yang, Y. M. Lu, L. Sbaiz, and M. Vetterli, "An optimal algorithm for reconstructing images from binary measurements," in *Proc. IS&T/SPIE Electronic Imaging, Computational Imaging VIII*, Jan. 2010, vol. 7533, pp. 75330K–75330K–12.

[6] F. Yang, Y. M. Lu, L. Sbaiz, and M. Vetterli, "Bits from photons: Oversampled image acquisition using binary poisson statistics," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1421–1436, Apr. 2012.

[7] S. H. Chan and Y. M. Lu, "Efficient image reconstruction for gigapixel quantum image sensors," in *Proc. IEEE Global Conf. Signal and Information Processing (GlobalSIP'14)*, Dec. 2014, pp. 312–316.

[8] O. A. Elgendy and S. H. Chan, "Image reconstruction and threshold design for quanta image sensors," in *Proc. IEEE Int. Conf. Image Process. (ICIP'16)*, Sep. 2016, pp. 978–982.

[9] S. H. Chan, O. A. Elgendy, and X. Wang, "Images from bits: Non-iterative image reconstruction for quanta image sensors," *MDPI Sensors*, vol. 16, no. 11, pp. 1961, 2016.

[10] E. R. Fossum, J. Ma, and S. Masoodian, "Quanta image sensor: concepts and progress," in *Proc. SPIE 9858, Advanced Photon Counting Techniques X*, 2016, vol. 9858, pp. 985804–985804–14.

[11] I. M. Antolovic, S. Burri, C. Bruschini, R. Hoebe, and E. Charbon, "Nonuniformity analysis of a 65k pixel CMOS SPAD imager," *IEEE Trans. Electron Devices*, vol. 63, no. 1, pp. 57–64, Jan. 2016.

[12] "Andor ixon ultra 888 specifications," http://www.andor.com/pdfs/specifications/iXon_Ultra_888_EMCCD_Specifications.pdf, Accessed: 2016-08-03.

[13] B. F. Aull, D. R. Schuette, D. J. Young, D. M. Craig, B. J. Felton, and K. Warner, "A study of crosstalk in a 256×256 photon counting imager based on silicon Geiger-mode avalanche photodiodes," *IEEE Sens. J.*, vol. 15, no. 4, pp. 2123–2132, Apr. 2015.

[14] N. A. W. Dutton, I. Gyongy, L. Parmesan, and R. K. Henderson, "Single photon counting performance and noise analysis of CMOS SPAD-based image sensors," *MDPI Sensors*, vol. 16, no. 7, pp. 1122, Jul. 2016.

[15] S. Masoodian, A. Rao, J. Ma, K. Odame, and E. R. Fossum, "A 2.5 pj/b binary image sensor as a pathfinder for quanta image sensors," *IEEE Trans. Electron Devices*, vol. 63, no. 1, pp. 100–105, Jan. 2016.

[16] F. Yang, *Bits from Photons: Oversampled Binary Image Acquisition*, Ph.D. thesis, École Polytechnique Fédérale De Lausanne, 2012.

[17] Y. M. Lu, "Adaptive sensing and inference for single-photon imaging," in *Proc. 47th Annual Conf. Inf. Sci. and Sys. (CISS'13)*, Mar. 2013, pp. 1–6.

[18] C. Hu and Y. M. Lu, "Adaptive time-sequential binary sensing for high dynamic range imaging," in *Proc. SPIE Conf. Adv. Photon Counting*, 2012, vol. 8375, pp. 83750A–1.

[19] T. Vogelsang and D. G. Stork, "High-dynamic-range binary pixel processing using non-destructive reads and variable oversampling and thresholds," in *Proc. IEEE Sens*, Oct. 2012, pp. 1–4.

[20] T. Vogelsang, M. Guidash, and S. Xue, "Overcoming the full well capacity limit: high dynamic range imaging using multi-bit temporal oversampling and conditional reset," in *Proc. Int. Image Sens. Workshop (IISW)*, Snowbird, UT, 2013.

[21] T. Vogelsang, D. G. Stork, and M. Guidash, "Hardware validated unified model of multibit temporally and spatially oversampled image sensors with conditional reset," *J. Electronic Imaging*, vol. 23, no. 1, pp. 013021, 2014.

[22] J. Hynecek, "Impactron-a new solid state image intensifier," *IEEE Trans. Electron Devices*, vol. 48, no. 10, pp. 2238–2241, Oct. 2001.

[23] M. S. Robbins and B. J. Hadwen, "The noise performance of electron multiplying charge-coupled devices," *IEEE Trans. Electron Devices*, vol. 50, no. 5, pp. 1227–1232, May 2003.

[24] N. A. W. Dutton, I. Gyongy, L. Parmesan, S. Gnecchi, N. Calder, B. R. Rae, S. Pellegrini, L. A. Grant, and R. K. Henderson, "A SPAD-based QVGA image sensor for single-photon counting and quanta imaging," *IEEE Trans. Electron Devices*, vol. 63, no. 1, pp. 189–196, Jan. 2016.

[25] L. Liang, H. Shen, P. De Camilli, and J. S. Duncan, "A novel multiple hypothesis based particle tracking method for clathrin mediated endocytosis analysis using fluorescence microscopy," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1844–1857, Apr. 2014.

[26] L. H. C. Braga, L. Gasparini, L. Grant, R. K. Henderson, N. Massari, M. Perenzoni, D. Stoppa, and R. Walker, "A fully digital 8 × 16 SiPM array for PET applications with per-pixel TDCs and real-time energy output," *IEEE J. Solid-State Circuits*, vol. 49, no. 1, pp. 301–314, Jan. 2014.

[27] S. P. Poland, N. Krstajić, J. Monypenny, S. Coelho, D. Tyndall, R. J. Walker, V. Devauges, J. Richardson, N. Dutton, P. Barber, D. D. Li, K. Suhling, T. Ng, R. K. Henderson, and S. M. Ameer-Beg, "A high speed multifocal multiphoton fluorescence lifetime imaging microscope for live-cell fret imaging," *Biomed. Opt. Express*, vol. 6, no. 2, pp. 277–296, Feb. 2015.

[28] G. Grubbs, R. Michell, M. Samara, D. Hampton, and J.-M. Jahn, "A synthesis of star calibration techniques for ground-based narrowband electron-multiplying charge-coupled device imagers used in auroral photometry," *J. Geophysical Research: Space Physics*, vol. 121, no. 6, pp. 5991–6002, 2016, 2015JA022186.

[29] P. Seitz and A. J. Theuwissen, *Single-photon imaging*, vol. 160, Springer Science & Business Media, 2011.

[30] L. J. Meng, "An intensified EMCCD camera for low energy Gamma ray imaging applications," *IEEE Trans. Nuclear Science*, vol. 53, no. 4, pp. 2376–2384, Aug. 2006.

[31] D. Shin, F. Xu, D. Venkatraman, R. Lussana, F. Villa, F. Zappa, V. K. Goyal, F. N. C. Wong, and J. H. Shapiro, "Photon-efficient imaging with a single-photon camera," *Nature Communications*, vol. 7, pp. –, 2016.

[32] I. Gyongy, T. Al Abbas, N. A.W. Dutton, and R. K. Henderson, "Object tracking and reconstruction with a quanta image sensor," in *Proc. Int. Image Sens. Workshop (IISW)*, 2017.

[33] S. Burri, Y. Maruyama, X. Michalet, F. Regazzoni, C. Bruschini, and E. Charbon, "Architecture and applications of a high resolution gated spad image sensor," *Opt. Express*, vol. 22, no. 14, pp. 17573–17589, July 2014.

[34] E Amri, Y. Felk, D. Stucki, J. Ma, and E. R. Fossum, "Quantum random number generation using a quanta image sensor," *MDPI Sensors*, vol. 16, no. 7, pp. 1002, Jun. 2016.

[35] E. R. Fossum, "The Quanta Image Sensor (QIS): Concepts and Challenges," in *Proc. OSA Topical Mtg Computational Optical Sensing and Imaging*, Jul. 2011, paper JTuE1.

[36] J. Ma and E. R. Fossum, "Quanta image sensor jot with sub 0.3e-r.m.s. read noise and photon counting capability," *IEEE Electron Device Letters*, vol. 36, no. 9, pp. 926–928, Sep. 2015.

[37] L. Sbaiz, F. Yang, E. Charbon, S. Susstrunk, and M. Vetterli, "The gigavision camera," in *Proc. 2009 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP'09)*, Apr. 2009, pp. 1093–1096.

[38] F. Yang, L. Sbaiz, E. Charbon, S. Süsstrunk, and M. Vetterli, "On pixel detection threshold in the gigavision camera," in *Proc. IS&T/SPIE Electronic Imaging, Digital Photography VI*, 2010, vol. 7537, pp. 75370G–75370G–8.

[39] N. A. W. Dutton, L. Parmesan, A. J. Holmes, L. A. Grant, and R. K. Henderson, "320 × 240 oversampled digital single photon counting image sensor," in *Proc. Symp VLSI Circuits*, Jun. 2014, pp. 1–2.

[40] I. M. Antolovic, S. Burri, R. A. Hoebe, Y. Maruyama, C. Bruschini, and E. Charbon, "Photon-counting arrays for time-resolved imaging," *MDPI Sensors*, vol. 16, no. 7, pp. 1005, 2016.

[41] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*, Number 55. Courier Corporation, 1964.

[42] E. R. Fossum, "Photon counting error rates in single-bit and multi-bit quanta image sensors," *IEEE J. Electron Devices Soc.*, vol. 4, no. 3, pp. 136–143, May 2016.

[43] F. Yang, L. Sbaiz, E. Charbon, S. Ssstrunk, and M. Vetterli, "Image reconstruction in the gigavision camera," in *Proc. IEEE 12th Int. Conf. on Computer Vision Workshops (ICCV Workshops), 2009*, Sep. 2009, pp. 2212–2219.

[44] R. A. Rojas, W. Luo, V. Murray, and Y. M. Lu, "Learning optimal parameters for binary sensing image reconstruction algorithms," in *Proc. IEEE Int. Conf. Image Process. (ICIP'17)*, Sep. 2017, pp. 2791–2795.

[45] E. R. Fossum, "Modeling the performance of single-bit and multi-bit quanta image sensors," *IEEE J. Electron Devices Soc.*, vol. 1, no. 9, pp. 166–174, Sep. 2013.

[46] J. Ma, D. Starkey, A. Rao, K. Odame, and E. R. Fossum, "Characterization of quanta image sensor pump-gate jots with deep sub-electron read noise," *IEEE J. Electron Devices Soc.*, vol. 3, no. 6, pp. 472–480, Nov. 2015.

[47] E. R. Fossum, "Multi-bit quanta image sensors," in *Proc. Int. Image Sens. Workshop (IISW)*, Jun. 2015, pp. 292–295.

[48] I. Sprow, D. Kuepper, Z. Barańczuk, and P. Zolliker, "Image quality assessment using a high dynamic range display," in *Proc. 12th Congress of the International Colour Association, Newcastle upon Tyne, England*, Jul. 2013.

[49] S. Hasinoff, D. Sharlet, R. Geiss, A. Adams, J. T. Barron, F. Kainz, J. Chen, and M. Levoy, "Burst photography for high dynamic range and low-light imaging on mobile cameras," *ACM Transactions on Graphics (Proc. SIGGRAPH Asia)*, 2016.

[50] K. I. Schultz, M. W. Kelly, J. J. Baker, M. H. Blackwell, M. G. Brown, C. B. Colonero, C. L. David, B. M. Tyrrell, and J. R. Wey, "Digital-pixel focal plane array technology," *Lincoln Lab J*, vol. 20, no. 2, pp. 36–51, 2014.

[51] "HDR-eye dataset," http://mmspg.epfl.ch/hdr-eye, Accessed: 2016-11-30.

[52] H. Nemoto, P. Korshunov, P. Hanhart, and T. Ebrahimi, "Visual attention in LDR and HDR images," in *9th Int. Workshop on Video Process. and Quality Metrics for Consumer Electronics (VPQM)*, 2015.

[53] T. Mertens, J. Kautz, and F. V. Reeth, "Exposure fusion," in *Proc. 15th Pacific Conf. Comp. Graphics and App., 2007. PG '07.*, Oct 2007, pp. 382–390.

[54] S. K. Nayar and T. Mitsunaga, "High dynamic range imaging: spatially varying pixel exposures," in *Proc. IEEE Conf Computer Vis Pattern Recog (CVPR)*, 2000, vol. 1, pp. 472–479 vol.1.

[55] C. Aguerrebere, A. Almansa, J. Delon, Y. Gousseau, and P. Muse, "A Bayesian hyperprior approach for joint image denoising and interpolation, with an application to HDR imaging," *IEEE Trans Comput Imag*, vol. PP, no. 99, pp. 1–1, 2017.

Optimal Threshold Design for Quanta Image Sensor (Supplementary Material)

Omar A. Elgendy, *Student Member, IEEE* and Stanley H. Chan, *Member, IEEE*

Abstract

This supplementary report provides the following additional information of the main article.

- Derivation of $\text{SNR}_q(c)$ from exposure-referred SNR,
- Properties of the incomplete Gamma function,
- Comparison with the threshold design scheme by Yang [1],
- Phase transition under different configurations,
- Influence of Non-Boxcar Kernel $\mathbf{G}$, and
- Additional results for HDR image reconstruction.

I. DERIVATION OF $\text{SNR}_q(c)$ FROM EXPOSURE-REFERRED SNR

In the literature of QIS devices, one metric to quantify the image quality is the *exposure*-referred signal-to-noise [2]. In image processing, however, exposure-referred SNR is not commonly used. The goal of this section is to show that the SNR we showed in the main article is equivalent to the exposure-referred SNR.
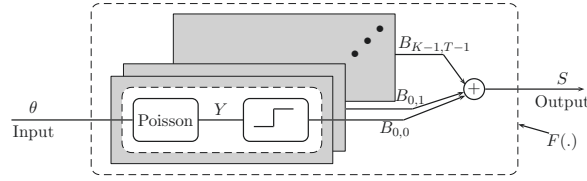

Fig. 1. Block diagram illustrating a QIS with input-output relation output $= F(\text{input})$

To understand the exposure-referred SNR, we have to first understand two common ways of defining a signal to noise ratio. Consider the truncated Poisson part of the QIS model shown in Figure 1. The input to this model is the over-sampled measurement θ. The truncated Poisson process can be considered as a black box function F which takes an input θ and generates an output S, defined as

$$S = \sum_{t=0}^{T-1} \sum_{k=0}^{K-1} B_{k,t},\tag{1}$$

where $\mathcal{B}_n = \{B_{k,t} \mid k = 0, 1, \ldots, K-1, \ t = 0, 1, \ldots, T-1\}$ is the spatial-temporal block containing all binary bits corresponding to θ. As shown in the main article, the mean and variance of S are

$$\mathbb{E}[S] = KT(1 - \Psi_q(\theta)), \quad \text{Var}[S] = KT\Psi_q(\theta)(1 - \Psi_q(\theta)),\tag{2}$$

respectively.

The first notion of signal-to-noise, which is the one used in CCD and CMOS, is called the output-referred SNR. SNR_{OR} is defined as the ratio between the output signal and the photon shot noise. Referring to Figure 1, this is

$$\text{SNR}_{\text{OR}} = \frac{\text{output signal}}{\text{noise}} = \frac{\mathbb{E}[S]}{\sqrt{\text{Var}[S]}} = \sqrt{KT\frac{1 - \Psi_q(\theta)}{\Psi_q(\theta)}}.\tag{3}$$

The authors are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907, USA. Email: { oelgendy, stanchan}@purdue.edu.
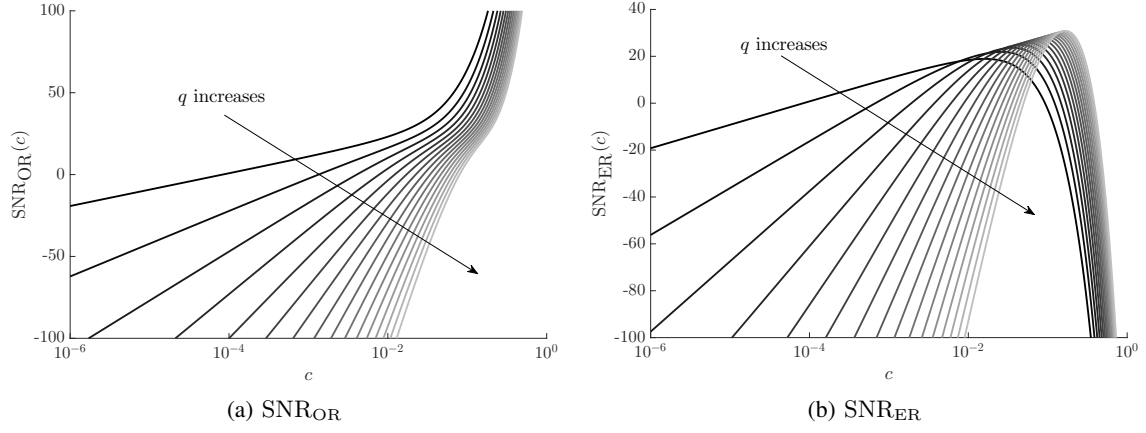
Fig. 2. Comparison of the SNRs for $q \in \{1, \ldots, 16\}$. In this experiment, we fix $\alpha = 400$, $K = 4$, and $T = 30$.

However, SNR_{OR} fails to work for QIS because the shot noise is arbitrarily small if all bits are 1 or 0. In [2], Fossum called it squeezing of the noise. If we plot SNR_{OR} as a function of θ, then we observe that SNR_{OR} approaches to infinity as θ grows.

The second notion of signal-to-noise, which is a modification of SNR_{OR}, is the exposure-referred SNR. SNR_{ER} is the ratio between the exposure signal θ and the exposure-referred noise. This noise is defined as [2]:

$$\text{Exposure-referred noise} = \frac{d\theta}{d\mathbb{E}[S]} \sqrt{\text{Var}[S]}$$

The factor $\frac{d\theta}{d\mathbb{E}[S]}$ can be considered as the "inverse" transfer function from the output to the input. $\frac{d\theta}{d\mathbb{E}[S]}$ can be determined by taking derivative of the expectation in (2) with respect to $\mathbb{E}[S]$

$$\frac{d\mathbb{E}[S]}{d\mathbb{E}[S]} = \frac{dKT\left(1 - \Psi_q(\theta)\right)}{d\mathbb{E}[S]}$$

Using chain rule, we observe that

$$1 = -KT \frac{d}{d\theta} \Psi_q(\theta) \frac{d\theta}{d\mathbb{E}[S]}$$

Since $\frac{d}{d\theta} \Psi_q(\theta) = \frac{-e^{-\theta}\theta^{q-1}}{\Gamma(q)}$, it holds that

$$1 = -KT \left(\frac{-e^{-\theta}\theta^{q-1}}{\Gamma(q)}\right) \frac{d\theta}{d\mathbb{E}[S]}$$

Hence,

$$\frac{d\theta}{d\mathbb{E}[S]} = \frac{\Gamma(q)}{KTe^{-\theta}\theta^{q-1}}$$

The exposure-referred SNR is defined as

$$\begin{aligned}
\text{SNR}_{\text{ER}} &= \frac{\text{exposure signal}}{\text{exposure-referred noise}} \\
&= \frac{\theta}{\sqrt{\text{Var}[S]} \frac{d\theta}{d\mathbb{E}[S]}} \\
&= \frac{e^{-\theta}\theta^q}{\Gamma(q)} \sqrt{\frac{KT}{\Psi_q(\theta)\left(1 - \Psi_q(\theta)\right)}}.
\end{aligned}$$

Taking logarithm shows that SNR_{ER} is identical to the SNR derived from the Fisher Information shown in the main article.
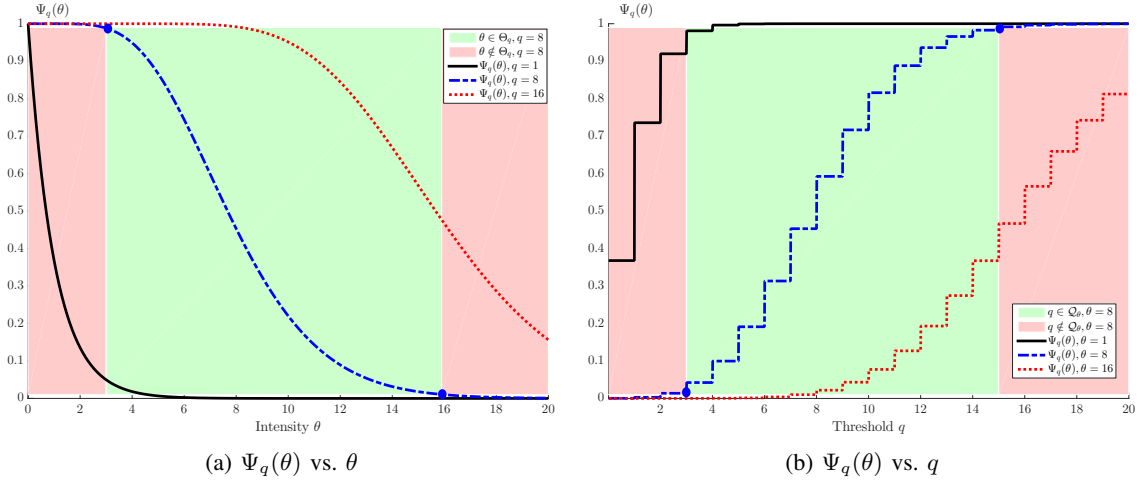
(a) $\Psi_q(\theta)$ vs. θ (b) $\Psi_q(\theta)$ vs. q

Fig. 3. $\Psi_q(\theta)$ as a function of θ and q. In defining, $\mathcal{Q}_\theta$ and Θ_q, we set $\epsilon = 0.01$.

II. PROPERTIES OF THE INCOMPLETE GAMMA FUNCTION

In the main article, we used the incomplete Gamma function for QIS analysis. In this section, we provide more details about the properties of the incomplete Gamma function.

First, we recall that the normalized upper incomplete Gamma function $\Psi_q : \mathbb{R}^+ \to [0, 1]$ is defined as

$$\Psi_q(\theta) \stackrel{\text{def}}{=} \frac{1}{\Gamma(q)} \int_\theta^\infty t^{q-1} e^{-t} dt, \quad \text{for } \theta > 0, \ q \in \mathbb{N}. \tag{4}$$

where $\Gamma(q) = (q-1)!$ is the standard Gamma function.

In this equation, we note that $\Psi_q(\theta)$ depends on two variables: q and θ.

- As a function of θ. As we showed in the main article, $\Psi_q(\theta)$ is a monotonically decreasing function of θ because the derivative is negative:

$$\frac{d}{d\theta} \Psi_q(\theta) = \frac{-\theta^{q-1} e^{-\theta}}{\Gamma(q)} < 0.$$

However, $\Psi_q(\theta)$ is very close to 1 when θ is small, and is very close to 0 when θ is large. Therefore, there exists a range of θ in which $\Psi_q(\theta)$ can attain a reasonably good inverse. We define this set as the θ-admissible set

$$\Theta_q \stackrel{\text{def}}{=} \{\theta \mid \varepsilon \le \Psi_q(\theta) \le 1 - \varepsilon\}, \tag{5}$$

for any fixed q and a tolerance level ε. An illustration of Θ_q is shown in Figure 3.

- As a function of q. The incomplete Gamma function $\Psi_q(\theta)$ can also be considered as a function of q. In this case, $\Psi_q(\theta)$ is only defined for integer values of q. We illustrate the behavior of $\Psi_q(\theta)$ as a function of q in Figure 3. The set of q in which $\Psi_q(\theta)$ is sufficiently away from 0 and 1 is defined as the q-admissible set.

$$\mathcal{Q}_\theta \stackrel{\text{def}}{=} \{q \mid \varepsilon \le \Psi_q(\theta) \le 1 - \varepsilon\}. \tag{6}$$

III. COMPARISON WITH THE THRESHOLD DESIGN SCHEME BY YANG [1]

In this section, we compare our threshold scheme with the one in [1].

First, we recall that the optimality of our method is based on a lower-bound $L_q(c)$ for the per-pixel SNR:

$$q^*(c) = \underset{q \in \mathbb{N}}{\operatorname{argmax}} \ \text{SNR}_q(c) \approx \underset{q \in \mathbb{N}}{\operatorname{argmax}} \ L_q(c) = \left\lfloor \frac{\alpha c}{K} \right\rfloor \tag{7}$$

Therefore, the optimal threshold is a function of c, which changes in space and in time.

In contrast, [1] uses a checkerboard pattern by alternating two thresholds (q_1^*, q_2^*). These two thresholds are obtained by maximizing the Cramér-Rao lower bound (CRLB) over a range of light intensity values $[c_{\min}, c_{\max}]$:

$$(q_1^*, q_2^*) = \underset{1 \le q_1, q_2 \le q_{\max}}{\operatorname{argmin}} \int_{c_{\min}}^{c_{\max}} CRLB(q_1, q_2, c) \, dc. \tag{8}$$

(a) $q = 1$, 15.97dB   (b) $q = 5$, 27.80dB   (c) $q = 8$, 25.84dB   (d) $q = 10$, 22.36dB

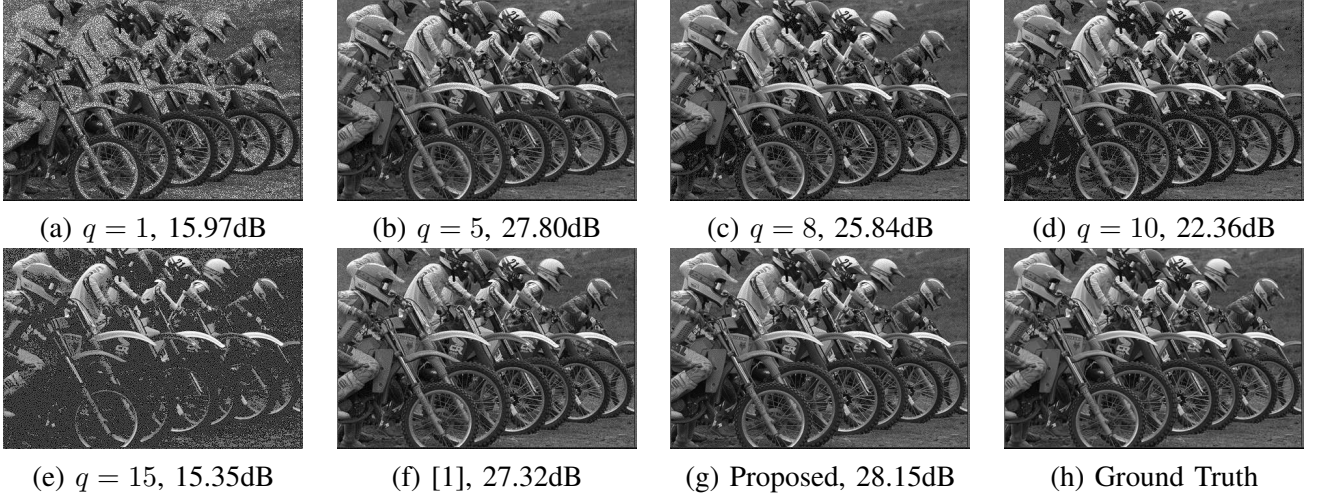(e) $q = 15$, 15.35dB   (f) [1], 27.32dB   (g) Proposed, 28.15dB   (h) Ground Truth

Fig. 4. Spatial oversampling $K = 4$. Temporal oversampling $T = 20$. Quadratic B-spline kernel is used in synthesis and reconstruction models. Gradient descent is used to obtain the ML estimate. For bisection threshold map, 8 frames are used for adapting the map, and 12 frames are used for reconstruction. For all other maps, the whole 20 frames are used for reconstruction.

As a result, the threshold is optimal in the *average sense*. To compare the two approaches, we followed the same steps in [1] to obtain $CRLB(q_1, q_2, c)$ for a checkerboard pattern in terms of $\Psi_q(c)$ as follows

$$CRLB(q_1, q_2, c) = \sum_{i=1}^{2} \frac{\alpha^2}{2K} \frac{e^{-2\theta}\theta^{(2(q_i-1))}}{\Gamma(q_i)^2 \Psi_{q_i}(\theta) \left[1 - \Psi_{q_i}(\theta)\right]} \tag{9}$$

where $\theta = \alpha c/K$. Using the parameters $\alpha = K(q_{\max-1})$, $q_{\max} = 16$, $K = 4$, and using trapezoidal technique for numerical integration over c, we obtain that $q_1^* = 4$ and $q_2^* = 12$. Figure 4 shows the reconstructed images using uniform threshold maps with thresholds $q \in \{1, 5, 8, 10, 15\}$, the checkerboard threshold map in [1] with $q_1^* = 4$ and $q_2^* = 12$, and the oracle threshold map obtained by (7). In this experiment, our proposed method achieves 28.15 dB, which is 0.83 dB higher than the checkerboard pattern.

IV. PHASE TRANSITION UNDER DIFFERENT CONFIGURATIONS

In the main article, we showed the phase transition behavior of the ML estimate using $K = 4$, $T = 50$, and $\delta = 2 \times 10^{-4}$. In this section, we study the effect of changing K, T, and δ on the phase transition region width.

As a function of T. Figure 5-Figure 6 illustrate the phase transition behavior when $T = 10, 25, 50$, and 100. As T increases, the width of the green region increases. However, if we fix the range of the bit density $1 - \mathbb{E}[\gamma_q(c)]$, we observe that the SNR does not vary significantly even as T changes.

As a function of K. The spatial oversampling K affects both the threshold $q^*(c) = \lfloor \alpha c/K \rfloor + 1$ and the phase transition width. Figure 7(a) illustrates the behavior of the threshold q^* as a function of K. As K increases, q^* decreases. However, the optimal q^* still stays within the set $\mathcal{Q}_\theta$.

As a function of δ. The constant δ is used to define the set $\mathcal{Q}_\theta$:

$$\mathcal{Q}_\theta \stackrel{\text{def}}{=} \left\{ q \ \Big| \ 1 - \left(\frac{\delta}{2}\right)^{\frac{1}{KT}} \le \Psi_q(\theta) \le \left(\frac{\delta}{2}\right)^{\frac{1}{KT}} \right\}. \tag{10}$$

The constant δ is the tolerance level. When δ increases, the size of the set $\mathcal{Q}_\theta$ should also increase. This result is shown in Figure 7(b).

Using the closed form expression of the average bit density $1 - \Psi_q(\theta)$, we can calculate the average bit density at the optimal threshold $q^* = \lfloor \theta \rfloor + 1$, which is shown in Figure 8. We notice that as long as $\theta \geq 1$, the average bit density is between 0.264 and 0.630. Within this range, we observe from Figure 5-Figure 6 that the SNR does not vary significantly if the estimated threshold is deviated from the optimal threshold. This observation relaxes the requirement of the bisection method from obtaining the exact optimal threshold to obtaining a threshold that make the bit density equal to 0.5. Since $0.5 \in [0.264, 0.630]$, we guarantee to achieve an SNR which is sufficiently close to the optimal SNR.

Controlling $\theta \geq 1$ can be achieved by tuning the constant α. Tuning α can be hardware-implemented by increasing the exposure period. Intuitively what $\theta \geq 1$ requires is that the average number of impinging photons per jot must be at least one. If θ is less than one, then most bits will become zeros. Increasing exposure period (i.e., increasing α) will ensure sufficient number of photons.

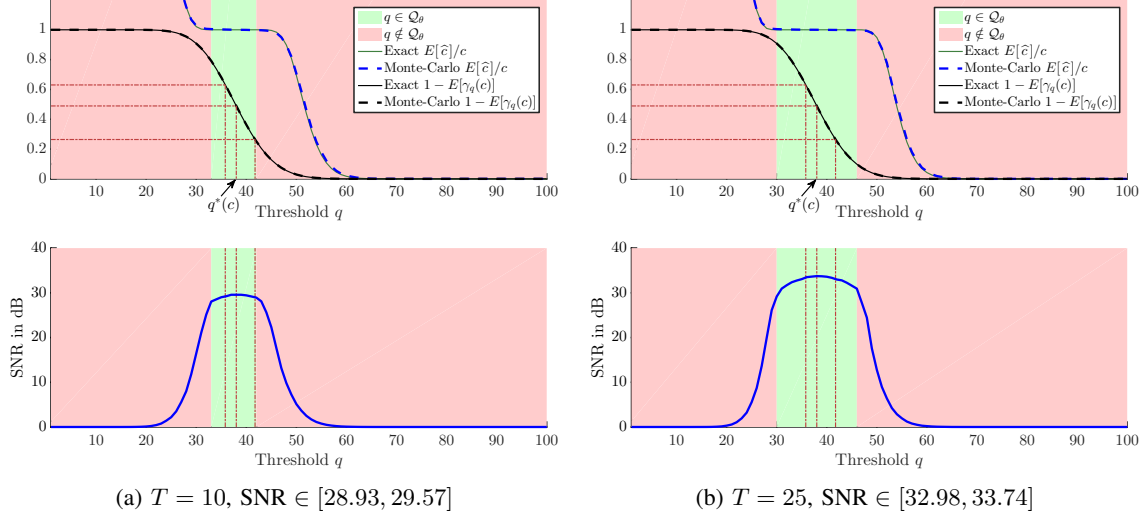(a) $T = 10$, SNR $\in [28.93, 29.57]$ (b) $T = 25$, SNR $\in [32.98, 33.74]$

Fig. 5. Phase transition for $T = 10$ and $T = 25$. SNR range is shown for average bit density $1 - \mathbb{E}[\gamma_q(c)]$ in the range $[0.264, 0.630]$. For all cases, we set $\delta = 2 \times 10^{-4}$, and $K = 4$.

(a) $T = 50$, SNR $\in [36.15, 36.80]$ (b) $T = 100$, SNR $\in [39.35, 39.82]$
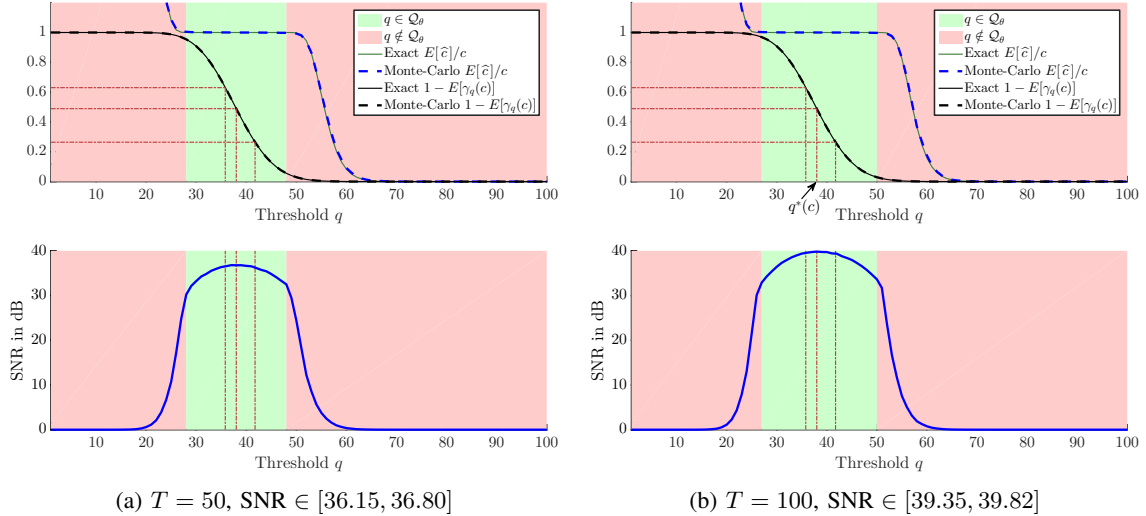
Fig. 6. Phase transition for $T = 50$ and $T = 100$. SNR range is shown for average bit density $1 - \mathbb{E}[\gamma_q(c)]$ in the range $[0.264, 0.630]$. For all cases, we set $\delta = 2 \times 10^{-4}$, and $K = 4$.

V. INFLUENCE OF NON-BOXCAR KERNEL $\mathbf{G}$

In this section, we discuss the boxcar kernel assumption in QIS model, i.e., $\boldsymbol{G} = \frac{1}{K}\boldsymbol{I}_{N \times N} \otimes \mathbf{1}_{K \times 1}$. We also study the effect of assuming a general kernel $\boldsymbol{G}$ on our results.

On QIS, we typically assume that there are micro-lenses on top of each jot or a group of jots. These micro-lenses ensure that the incident light converges onto the sensing site with no (or very minor) interference with adjacent jots or groups. As a result, we can model the incoming light using the boxcar kernel. This assumption is perhaps strong in some perspective, but it allows us to significantly simplify the theory and offer efficient implementations.

What if there is a mismatch between the physical model (e.g., using B-spline or Gaussian kernel $\boldsymbol{G}$) and the reconstruction (e.g., using boxcar)? To see the effect of this mismatch on the reconstruction quality, we conduct two sets of experiments.
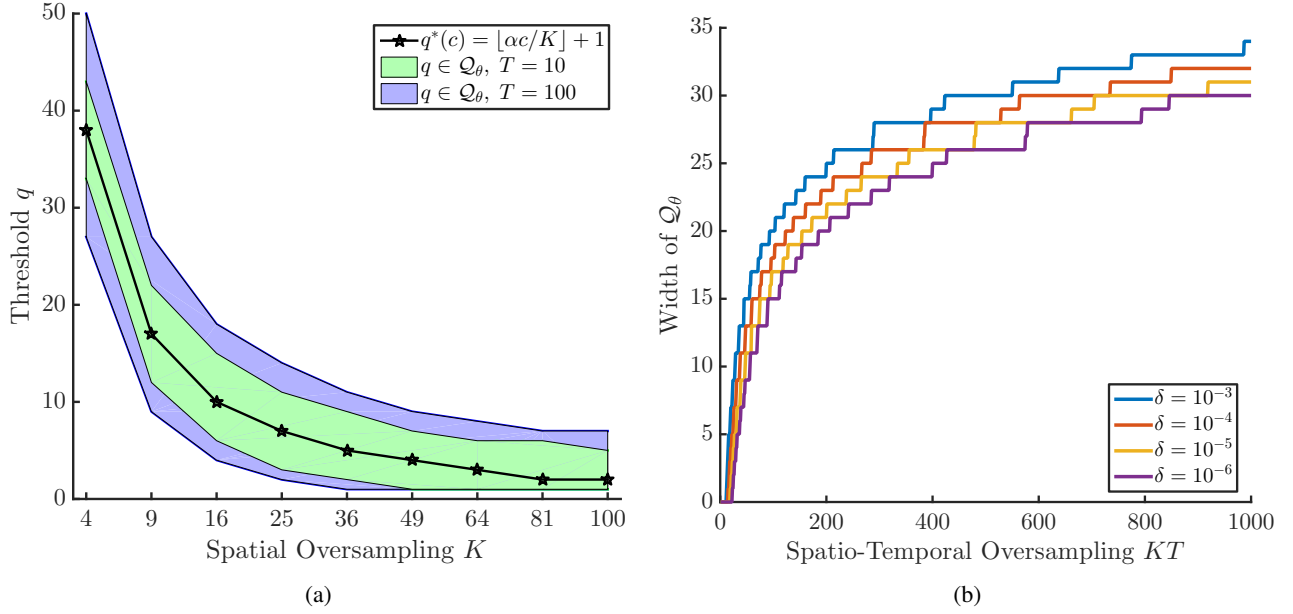
Fig. 7.   (a) The threshold q and $\mathcal{Q}_\theta$ as K increases. (b) The width of $\mathcal{Q}_\theta$ as KT and δ changes.
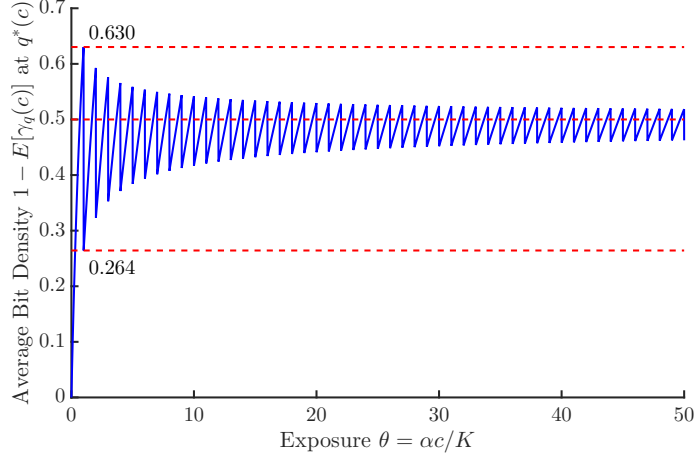

Fig. 8.   Average bit density $1 - \mathbb{E}[\gamma_q(c)]$ calculated at optimal threshold $q^* = \lfloor \theta \rfloor + 1$.

- **1D Signal**: We consider a 1D signal with 10 coefficients. These 10 coefficients are modulated with boxcar kernels and B-spline kernels to generate two sets of incident light. On the QIS simulator, we set the spatial and temporal oversampling factors as $K = 9$ and $T = 30$, respectively. Then we use the oracle threshold map for quantization. To reconstruct the images, we use boxcar kernel for both cases so that we have one matching case and one mismatching case. Figure 9 shows the reconstructed signals. As expected, when the forward model matches with the reconstruction model, the reconstructed image has the highest PSNR. However, the gap between the cases are not significant.
- **2D Signal**: Figure 10 shows a 2D example. Similar to the 1D case, boxcar kernel leads to the best reconstruction but its gap with the other cases are not significant.

The reader might think why we do not use B-spline on the reconstruction so that it will match with the forward model? In principle this is doable, but we need an iterative algorithm to compute the ML estimate such as gradient descent as reported in [1]. In contrast, the boxcar assumption allows us to use a closed-form ML estimate, which is practically much more affordable.

(a) Boxcar kernel, PSNR= 33.74 dB

(b) Linear B-spline, PSNR= 31.62 dB

(c) Quadratic B-spline, PSNR= 31.66 dB
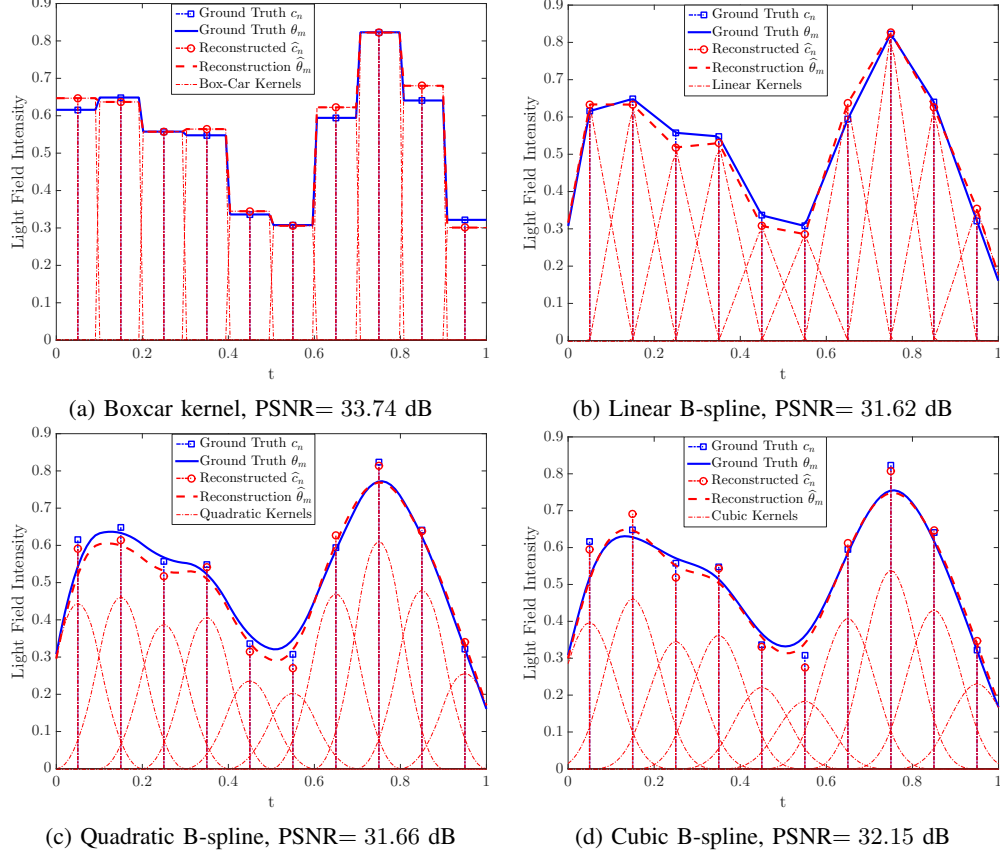
(d) Cubic B-spline, PSNR= 32.15 dB

Fig. 9. Spatial oversampling $K = 9$. Temporal oversampling $T = 30$. Oracle threshold map is used for quantization. Different kernels are used in synthesis and boxcar kernel is used in reconstruction. ML closed-form is used for reconstruction

(a) Boxcar Ground Truth

(b) Linear B-spline Ground Truth

(c) Quadratic B-spline Ground Truth

(d) Cubic B-spline Ground Truth

(e) Reconstruction PSNR= 33.33

(f) Reconstruction PSNR= 30.67

(g) Reconstruction PSNR= 32.54

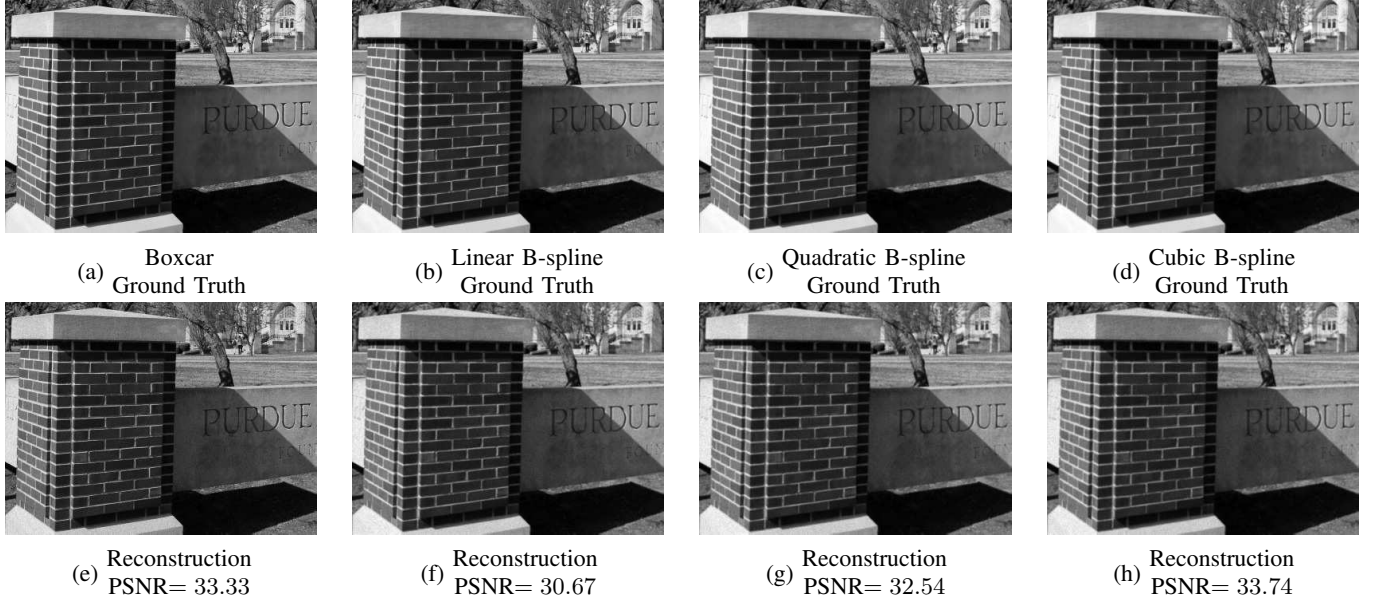(h) Reconstruction PSNR= 33.74

Fig. 10. Ground truth and reconstructed images using simulated binary measurements synthesized by (a)(e) Boxcar, (b)(f) Linear B-spline, (c)(g) Quadratic B-spline, and (d)(h) Cubic B-spline kernels. In this experiment, we spatially oversample each pixel by $K = 4 \times 4$ binary bits and we use $T = 15$ independent temporal measurements. We use 8 frames for learning the threshold map using bisection method, and the remaining 7 frames are used for image reconstruction using the ML closed-form by the boxcar kernel assumption.

VI. Supplementary HDR results

In this section, we show more results for HDR image reconstruction using our method compared to the fixed threshold approach. Figure 11 show reconstructed HDR images using adapted Q-map by the bisection algorithm, and fixed Q-maps with low threshold ($q = 1$) and high threshold ($q_{max} = 16$). The spatial and temporal oversampling factors are $K = 4$, and $T = 13$, respectively. Sensor gain is $\alpha = K^2/(q_{max} - 1)$.
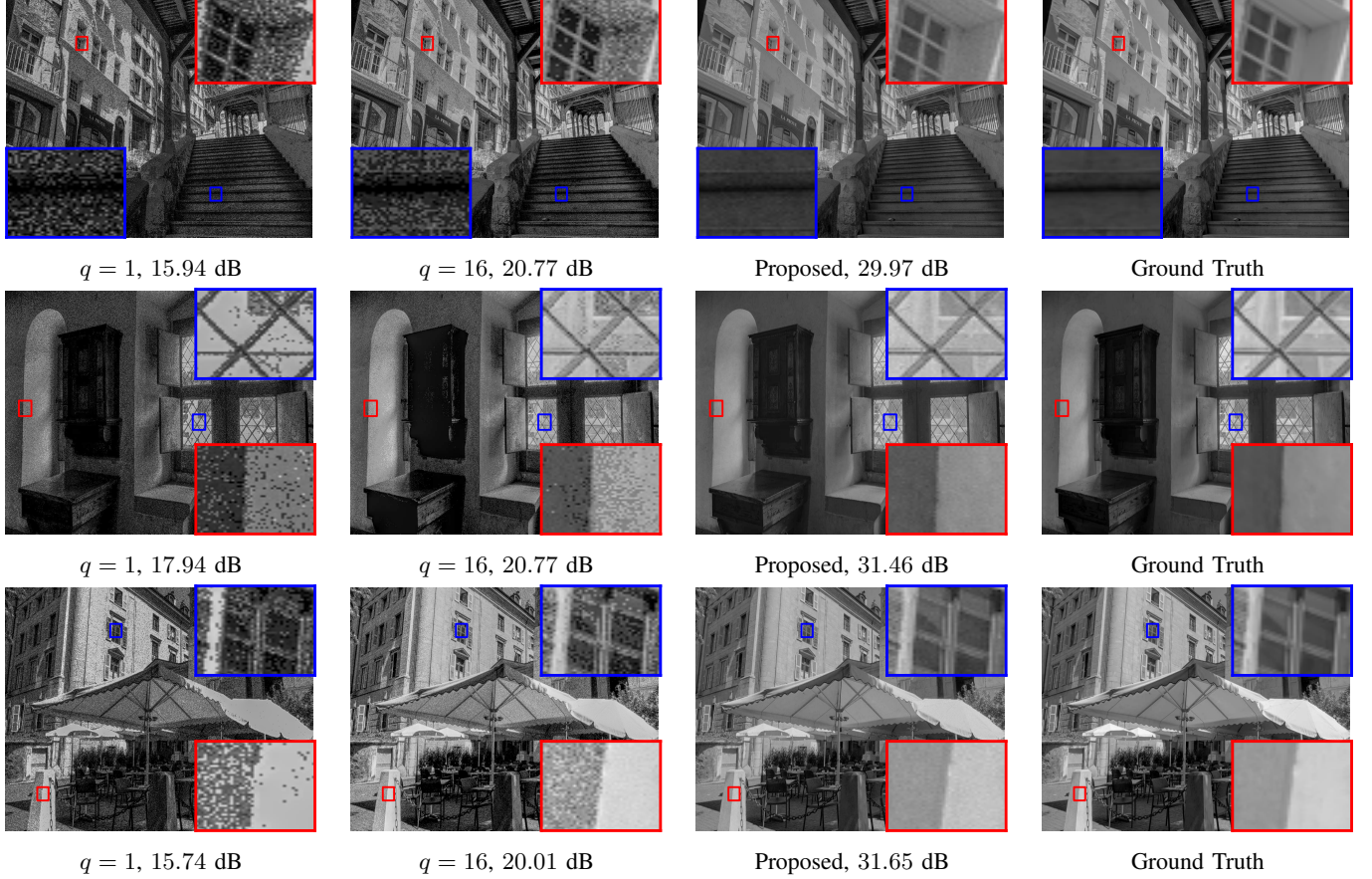

$q = 1$, 15.94 dB	$q = 16$, 20.77 dB	Proposed, 29.97 dB	Ground Truth
$q = 1$, 17.94 dB	$q = 16$, 20.77 dB	Proposed, 31.46 dB	Ground Truth
$q = 1$, 15.74 dB	$q = 16$, 20.01 dB	Proposed, 31.65 dB	Ground Truth

Fig. 11.   Reconstructed HDR images using different threshold maps

References

[1] F. Yang, *Bits from Photons: Oversampled Binary Image Acquisition*, Ph.D. thesis, École Polytechnique Fédérale De Lausanne, 2012.

[2] E. R. Fossum, "Modeling the performance of single-bit and multi-bit quanta image sensors," *IEEE J. Electron Devices Soc.*, vol. 1, no. 9, pp. 166–174, Sep. 2013.