ADAPTIVE SENSING AND INFERENCE FOR SINGLE-PHOTON IMAGING

Yue M. Lu

School of Engineering and Applied Sciences
Harvard University, Cambridge, MA 02138
Email: yuelu@seas.harvard.edu

ABSTRACT

In recent years, there have been increasing efforts to develop solid-state sensors with single-photon sensitivity, with applications ranging from bio-imaging to 3D computer vision. In this paper, we present adaptive sensing models, theory and algorithms for these single-photon sensors, aiming to improve their dynamic ranges. Mapping different sensor configurations onto a finite set of states, we represent adaptive sensing schemes as finite-state parametric Markov chains. After deriving an asymptotic expression for the Fisher information rate of these Markovian systems, we propose a design criterion for sensing policies based on minimax ratio regret. We also present a suboptimal yet effective sensing policy based on random walks. Numerical experiments demonstrate the strong performance of the proposed scheme, which expands the sensor dynamic ranges of existing nonadaptive approaches by several orders of magnitude.

Index Terms— Single-photon imaging, adaptive sensing, high dynamic range, photon counting

1. INTRODUCTION

Recent advances in materials, devices and fabrication technologies have led to an emerging class of solid-state sensors that can detect individual photons in space and time [1, 2]. Thanks to their single-photon sensitivity, sub-nanosecond time resolution, and rapidly increasing spatial resolutions, these new sensors have become a key enabling technology behind recent progresses in several domains. Examples include bio-imaging [3], time-of-flight 3D vision [4], LIDAR [5], quantum cryptography [6], optical communications [7], and astronomy [8].

Analogous to silver-halide grains on photographic film [9], each pixel of these single-photon sensors (SPS) has a binary response ("click" or no "click"), revealing only one-bit information of the photon flux at that pixel during short exposure periods. By using large pixel arrays and very high temporal sampling rates (*e.g.*, 10^6 frames/s), the SPS generates a massive spatiotemporal volume of bits that sample and encode the original visual information.

As a spatiotemporal sampling device of the light intensity field, the SPS poses several unique challenges: First, the quantum nature of light implies that measurements taken by the SPS are always stochastic, obeying Poisson statistics; Second, SPS pixels have single-photon sensitivity, but they do not have photon counting capabilities. That is, a "click" in the sensor output can be caused by a single photon or multiple photons. This ambiguity imposes a limit on the dynamic range of the SPS: When the light intensity is high enough, there is almost always one or more photons arriving during each sampling window (of length T). In this high intensity regime, the sensor becomes saturated, generating "clicks" constantly without being able to distinguish between different light intensities.

A typical solution to this problem is to increase the sampling rate (*i.e.*, to make T smaller,) thus lowering the probability of multiple photon arrivals during each window [10]. However, device metrics such as detector dead time and after pulsing [11] impose fundamental lower bounds on T; furthermore, the maximum sampling rate is also constrained by the finite bandwidth of the sensor's I/O system. In this paper, we present time-sequential adaptive sensing schemes for SPS, demonstrating that careful designs of sensing and inference algorithms allow one to improve the dynamic range of the SPS beyond the nominal limits imposed by current hardware.

The rest of the paper is organized as follows. The sensing model of the SPS is described in Section 2, where we also demonstrate the problem of limited dynamic ranges associated with existing (nonadaptive) approaches. We present in Section 3 a general framework for adaptive sensing, representing each sensing policy as a finite state parametric Markov chain. We analyze the performance of these Markovian systems in terms of their asymptotic Fisher information rates, and propose a minimax regret formulation for designing the optimal sensing policy. Finally, we present in Section 4 a suboptimal yet effective sensing scheme based on random walks, and demonstrate its strong performance through numerical experiments.

2. MODEL AND PROBLEM FORMULATION

2.1. Sensing Model

Let $p(t)$ denote the photon flux (units of photons/s) impinging upon one of the pixels of the SPS. Then, the arrival of *detectable* events (shown as vertical arrows in the first row of Figure 1) can be modeled by a non-homogeneous Poisson process [12] whose rate function is equal to

$$p_e(t) = \eta\, p(t) + p_{\text{dc}}. \tag{1}$$

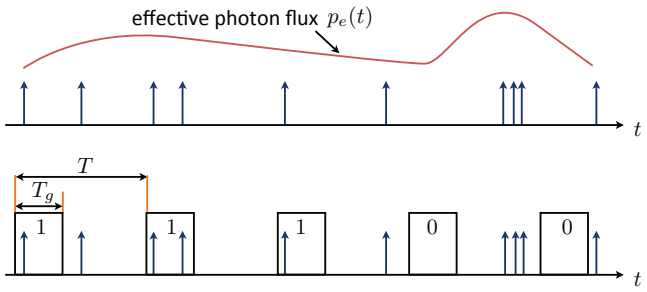

Fig. 1. The binary sensing model of the SPS over time.

Here, η is the *photon detecting efficiency* of the sensor, defined as the probability that a photon incident on the detector will actually trigger a "click" in the output signal; p_{dc} is the *dark count rate*, measuring the expected number of "false alarms" per second the sensor produces, even if it is put in total darkness (*i.e.*, when there is no incoming photon.) Incorporating the above sources of imperfections, $p_e(t)$ in (1) is thus the "effective" photon flux. Since $p_e(t)$ is the rate function of the Poisson process, the number of detectable events during any time window $[t_1, t_2]$ has a Poisson distribution

$$\mathbb{P}(Y = y; c) = \frac{c^y\, e^{-c}}{y!}, \quad \text{for } y \in \mathbb{Z}^+ \cup \{0\}\,, \tag{2}$$

where $c = \int_{t_1}^{t_2} p_e(t)\, dt$.

When operating in the gated mode, the SPS works as a uniform sampling device of $p_e(t)$. As illustrated in the second row of Figure 1, within each sampling period of length T, the sensor is set to "on" (for T_g seconds) and then "off" (for $T - T_g$ seconds). When the sensor is "off", no photon arrival will be recorded. During the "on" periods, the sensor outputs "1" if there is at least one photon (or dark count) detected; otherwise, the sensor outputs "0". Note that an output of "1" can be due to a single photon or multiple photons, as is the case in the second detection period shown in the figure.

2.2. Problem Formulation

In this work, we are concerned with the problem of estimating the photon flux $p(t)$ from the sensor measurements. For simplicity, we consider the case where $p(t)$ has constant values, *i.e.*, $p(t) \equiv p$, over some time period $\tau \gg T$. This is a reasonable assumption when the sampling rate $1/T$ (*e.g.*, 10^6 frames/s) is much higher than the temporal bandwidth of $p(t)$. The more general case of estimating time-varying photon flux will be addressed in a follow-up work.

Let $K \stackrel{\text{def}}{=} \tau/T$ be the *oversampling factor*, which, without loss of generality, is assumed to be an integer. Denote by

$$d \stackrel{\text{def}}{=} T_g/T$$

the *duty cycle* of the SPS, and by $B_1, B_2, \ldots, B_K$ the binary sensor measurements obtained over the K sampling windows. It follows from (2) that

$$\mathbb{P}(B_k|\theta) = \begin{cases} 1 - e^{-d\theta}, & \text{if } B_k = 1; \\ e^{-d\theta}, & \text{if } B_k = 0, \end{cases} \tag{3}$$

where $\theta \stackrel{\text{def}}{=} (\eta\, p + p_{\text{dc}})T$ is the expected number of detectable events within a time interval of length T.

We can then estimate the photon flux p from the binary sensor measurements by using maximum likelihood (ML) estimation.[1] The signal-to-noise (SNR) of the estimation is

$$\text{SNR}_{\text{P}}(p) \stackrel{\text{def}}{=} 10 \log_{10} \left(p^2/\mathbb{E}[(\widehat{p}_{\text{ML}} - p)^2]\right), \tag{4}$$

where $\widehat{p}_{\text{ML}}$ is the ML estimate of p. By the function invariance of the ML estimator, we can also approach the problem by estimating the parameter θ, and get $\widehat{\theta}_{\text{ML}} = (\eta\, \widehat{p}_{\text{ML}} + p_{dc})T$. Defining $\text{SNR}_{\Theta}(\theta)$ in a similar way as (4), we can then write

$$\text{SNR}_{\text{P}}(p) = \text{SNR}_{\Theta}\Big((\eta\, p + p_{\text{dc}})T\Big) - 20 \log_{10} \Big(1 + \frac{p_{\text{dc}}}{\eta\, p}\Big). \tag{5}$$

In what follows, we will focus on $\text{SNR}_{\Theta}(\theta)$, from which the target SNR function $\text{SNR}_{\text{P}}(p)$ can be derived according to (5).

When the duty cycle d is kept fixed for all K sampling intervals, the binary sensor measurements $B_1, B_2, \ldots, B_K$ are independent and identically distributed random variables with density function (3). We can verify that all the conditions for Cramér's theorem [13, p. 121] hold and that the ML estimate $\widehat{\theta}$ is asymptotically normal. In particular, we have $K\, \mathbb{E}[(\widehat{\theta}_{\text{ML}} - \theta)^2] \xrightarrow{K\to\infty} 1/I(\theta, d)$, where $I(\theta, d)$ is the Fisher information of the density function in (3). It follows that, for large K,

$$\text{SNR}_{\Theta}(\theta) \approx 10 \log_{10}(\theta^2 I(\theta, d)) + 10 \log_{10} K. \tag{6}$$

So, the expression in (6) can serve as an asymptotic surrogate for $\text{SNR}_{\Theta}(\theta)$.

The Fisher information can be computed as

$$\begin{aligned} I(\theta, d) &= \mathbb{E}_B \left[-\frac{\partial^2}{\partial \theta^2} \log \mathbb{P}(B|\theta)\right] \\ &= \mathbb{E}_B[B d^2 e^{-d\theta}/(1 - e^{-d\theta})^2] \\ &= d^2/(e^{d\theta} - 1). \end{aligned} \tag{7}$$

In Figure 2, we plot the function $10 \log_{10}(\theta^2 I(\theta, d))$ for three different choices of the duty cycles: $d = 0.95, d = 5 \times 10^{-3}$, and $d = 5 \times 10^{-5}$, respectively. It is clear from the figure that each fixed value of the duty cycle only corresponds to a fairly limited working range of θ, beyond which the performance drops rapidly. For example, for $d = 0.95$, the SNR drops significantly when $\theta > 10$, and this is due to the sensor becoming saturated at high light intensities; Choosing a smaller duty cycle $d = 5 \times 10^{-5}$ increases the saturation point, allowing the sensor to operate at higher values of θ; however, the performance under this setting becomes unacceptable for low values of θ. This happens because, in the low light regime, there is only a very small number of photon arrivals per sampling interval, and most of them will fall outside of the T_g and thus remain undetected due to the small duty cycle.

[1] It is easy to verify that the log-likelihood function $\log \mathbb{P}(B^K|\theta)$ is concave [10], and thus the ML estimate can be obtained by simple gradient ascent algorithms.

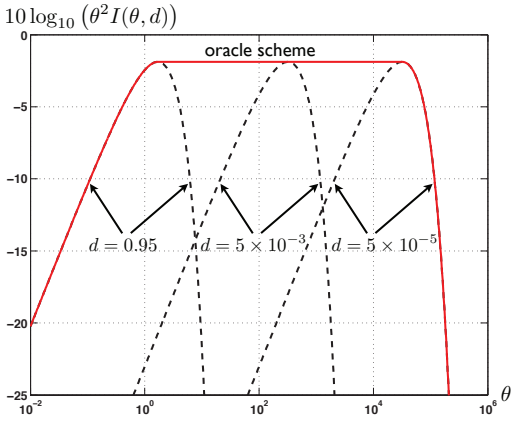

Fig. 2. The dashed lines show the SNRs corresponding to three different choices of the duty cycles. The red line shows the performance of the oracle scheme, achieved by choosing the optimal duty cycle $d^*(\theta)$ for each θ.

For each θ, there exists an optimal choice of the duty cycle, defined as

$$d^*(\theta) \stackrel{\text{def}}{=} \underset{d_{\min}\le d\le d_{\max}}{\arg\max}\ f(\theta,d),$$

where $0 < d_{\min}, d_{\max} < 1$ are, respectively, the minimum and maximum duty cycles the SPS can use. The corresponding Fisher information is denoted by

$$I^*(\theta) \stackrel{\text{def}}{=} I(\theta, d^*(\theta)). \tag{8}$$

In subsequent discussions, we refer to $d^*(\theta)$ as the *oracle choice* as it requires the knowledge of the unknown parameter θ.

From the specific form of the Fisher function in (7), it is easy to show the following lemma, whose proof is omitted.

Lemma 1 (Oracle choice)**.**

$$d^*(\theta) = \begin{cases} d_{\max}, & \text{if } \theta \le \gamma/d_{\max} \\ d_{\min}, & \text{if } \theta \ge \gamma/d_{\min} \\ \gamma/\theta & \text{otherwise}, \end{cases} \tag{9}$$

where $\gamma \approx 1.5936$ *is the unique positive solution to* $\gamma = 2 - 2e^{-\gamma}$.

The red line in Figure 2 shows the performance of the *oracle* scheme, obtained by always choosing $d^*(\theta)$ for each θ. Equivalently, this curve can be obtained as the upper envelope of all the SNR curves corresponding to different choices of the duty cycles between $d_{\min}$ and $d_{\max}$. Compared with nonadaptive schemes, the oracle scheme has a much wider dynamic range. The goal of this work is to develop models, theory and algorithms for time-sequential adaptive sensing schemes that can emulate the oracle performance.

3. TIME-SEQUENTIAL ADAPTIVE SENSING

3.1. Adaptive Sensing with Finite Memory

Consider a discrete setup where, at each sampling interval, the SPS can choose to use one of M different duty cycles. We denote by $\mathcal{D} \stackrel{\text{def}}{=} \{d_1, d_2, \ldots, d_M\}$ the collection of these duty cycles, with $d_{\min} \le d_m \le d_{\max}$ for $1 \le m \le M$. The proposed adaptive sensing scheme has a finite state space $\mathcal{S} \stackrel{\text{def}}{=} \{1, 2, \ldots, N\}$, where $N \ge M$. Let $g : \mathcal{S} \to \mathcal{D}$ be the mapping between the state space and the collection of possible duty cycles. Thus, each state $i \in \mathcal{S}$ is associated with one of the duty cycles $g(i) \in \mathcal{D}$.

The dynamics of the proposed adaptive sensing scheme can be represented by a parametric Markov chain on $\mathcal{S}$. Let $X_1, X_2, \ldots X_K \in \mathcal{S}$ represent the sequence of states, with X_k corresponding to the kth sampling window. The first state, X_1, is drawn from an initial distribution $\pi^{(0)}$ defined on $\mathcal{S}$. At the kth sampling window (for $k \ge 1$), if the state of the sensor is X_k, then the sensor will set the duty cycle to $g(X_k)$. Let $B_k \in \{0, 1\}$ be the binary sensor measurement obtained at that window. Then, depending on the value of B_k, the probability that the system moves from state i to state j is

$$\mathbb{P}(X_{k+1} = j | X_k = i, B_k) = P_{i,j}^{B_k} Q_{i,j}^{1-B_k}, \tag{10}$$

where $\boldsymbol{P}, \boldsymbol{Q} \in \mathbb{R}^{N\times N}$ are two right stochastic matrices. Using (3), we can rewrite the above transition probability as

$$\begin{aligned}\mathbb{P}(X_{k+1} = j | X_k = i, \theta) &= (1 - e^{-g(i)\theta})\boldsymbol{P}_{i,j} + e^{-g(i)\theta}\boldsymbol{Q}_{i,j} \\ &\stackrel{\text{def}}{=} T_{i,j}(\theta).\end{aligned} \tag{11}$$

Thus, for any given θ, the dynamics of the system is completely describe by a homogeneous Markov chain with a transition matrix $\boldsymbol{T}(\theta) \in \mathbb{R}^{N\times N}$. Since θ is unknown, the proposed adaptive sensing scheme is an infinite family of Markov chains indexed by θ. The transition matrix $\boldsymbol{T}(\theta)$ can be partially "influenced" by using different $\boldsymbol{P}$ and $\boldsymbol{Q}$, as in (11). Thus, designing the adaptive sensing policy boils down to choosing the two stochastic matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$.

3.2. Asymptotic Analysis and Optimal Policy Design

In what follows, we analyze the asymptotic performance of the proposed finite-state adaptive sensing scheme, and present a minimax formulation for designing the optimal sensing policy.

We start by deriving the log-likelihood function

$$\ell(X_1, B_1, \ldots, X_K, B_K|\theta) \stackrel{\text{def}}{=} \log \mathbb{P}(X_1, B_1, \ldots, X_K, B_K|\theta),$$

where X_k and B_k are the system state and binary sensor measurement at the kth sampling window, respectively.

Proposition 1 (Log-likelihood)**.**

$$\begin{aligned}&\ell(X_1, B_1, \ldots, X_K, B_K|\theta) \\ &= \log \pi^{(0)}(X_1) + \sum_{k=1}^{K-1} \log P_{X_k,X_{k+1}}^{B_k} Q_{X_k,X_{k+1}}^{1-B_k} \\ &+ \sum_{k=1}^{K} \left(B_k \log\left(1 - e^{-g(X_k)\theta}\right) - (1 - B_k) g(X_k)\theta \right).\end{aligned} \tag{12}$$

Proof. For notational simplicity, we prove this proposition for $K = 2$. The proof under the general case is similar.

Using the multiplication rule for joint probabilities, we have

$$\begin{aligned}&\mathbb{P}(X_1, B_1, X_2, B_2|\theta)\\ &= \mathbb{P}(X_1|\theta)\mathbb{P}(B_1|X_1,\theta)\mathbb{P}(X_2|X_1,B_1,\theta)\mathbb{P}(B_2|X_2,X_1,B_1,\theta)\\ &= \pi^{(0)}(X_1)\mathbb{P}(B_1|X_1,\theta)\mathbb{P}(X_2|X_1,B_1)\mathbb{P}(B_2|X_2,\theta),\end{aligned}$$

where $\pi^{(0)}(\cdot)$ is the initial distribution from which we generate the first state X_1. Substituting (3) and (10) into the above equation and taking the logarithm, we get (12). □

Denote by $I_K(\theta; \boldsymbol{P}, \boldsymbol{Q})$ the Fisher information associated with the random variables $X_1, B_1, \ldots, X_K, B_K$. The following proposition provides an asymptotic formula for this quantity. A similar result was stated, without proof, in our earlier paper on adaptive binary sensing [14].

Proposition 2. *Let $\mathcal{X}$ be the range of possible values for the parameter θ. If the Markov chain associated with the transition matrix $\boldsymbol{T}(\theta)$ is irreducible and aperiodic for every $\theta \in \mathcal{X}$, then*

$$\lim_{K\to\infty} \frac{I_K(\theta; \boldsymbol{P}, \boldsymbol{Q})}{K} = \sum_{m=1}^{M} I(\theta, d_m) \sum_{i\in\mathcal{S}_m} \pi_i(\theta; \boldsymbol{P}, \boldsymbol{Q}) \le I^*(\theta), \tag{13}$$

where $\pi_i(\theta; \boldsymbol{P}, \boldsymbol{Q}), i \in \mathcal{S}$ is the stationary distribution of $\boldsymbol{T}(\theta)$, $\mathcal{S}_m \stackrel{def}{=} \{i \in \mathcal{S} : g(i) = d_m\}$ is the subset of states corresponding to the mth duty cycle d_m, and $I^*(\theta)$ is the oracle Fisher information defined in (8)*.*

Proof. Using the expression for the log-likelihood function in (12), we compute the Fisher information as

$$\begin{aligned}&I_K(\theta; \boldsymbol{P}, \boldsymbol{Q})\\ &= \mathbb{E}\left[-\frac{\partial^2}{\partial\theta^2}\ell(X_1, B_1, \ldots, X_K, B_K|\theta)\right]\\ &= \sum_{k=1}^{K} \mathbb{E}_{X_k}\left[\mathbb{E}_{B_K}\left[B_k g^2(X_k) e^{-g(X_k)\theta}/(1 - e^{-g(X_k)\theta})^2 | X_k\right]\right]\\ &= \sum_{k=1}^{K} \mathbb{E}_{X_k}[I(\theta, g(X_k)]\\ &= \sum_{m=1}^{M} I(\theta, d_m) \sum_{i\in\mathcal{S}_m} \sum_{k=1}^{K} \mathbb{P}(X_k = i).\end{aligned} \tag{14}$$

By assumption, $\boldsymbol{T}(\theta)$ is irreducible and aperiodic, and thus the corresponding Markov chain has a unique stationary distribution $\pi_i(\theta; \boldsymbol{P}, \boldsymbol{Q})$, *i.e.*, for any $i \in \mathcal{S}$, $\lim_{K\to\infty} \mathbb{P}(X_k = i) = \pi_i(\theta; \boldsymbol{P}, \boldsymbol{Q})$. It follows that

$$\lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \mathbb{P}(X_k = i) = \pi_i(\theta; \boldsymbol{P}, \boldsymbol{Q}). \tag{15}$$

Dividing both sides of (14) by K and using the limit expression in (15), we obtain the equality in (13). Finally, the inequality in (13) follows from the definition of $I^*(\theta)$ and from the fact that $0 \le \sum_{i\in\mathcal{S}_m} \pi_i(\theta; \boldsymbol{P}, \boldsymbol{Q}) \le 1$. □

Remark 1. *The result in Proposition 2 indicates that the asymptotic Fisher information rate associated with the adaptive sensing scheme is determined by the stationary distributions of the parametric Markov chain. Thus, a qualitative criterion in designing the adaptive sensing scheme is that the stationary distributions $\pi(\theta; \boldsymbol{P}, \boldsymbol{Q})$ should be as tightly concentrated around the oracle choice $d^*(\theta)$ as possible.*

Quantitatively, we propose the following minimax formulation for designing optimal adaptive sensing policies:

$$\arg\min_{\boldsymbol{P},\boldsymbol{Q}} \max_{\theta\in\Theta} I^*(\theta) / \Big(\sum_{m=1}^{M} I(\theta, d_m) \sum_{i\in\mathcal{S}_m} \pi_i(\theta; \boldsymbol{P}, \boldsymbol{Q}) \Big), \tag{16}$$

In decision theory [15], the cost function in (16) is called a *ratio regret*, computing the ratio between the best possible outcome that could have been achieved and the outcome obtained by the adaptive sensing scheme. The best sensing policy $(\boldsymbol{P}, \boldsymbol{Q})$ is then the one that can minimize the worst-case regret.

The cost function in (16) is also related to performances measured in SNRs. To see this, we use (6) to write the SNR of the oracle scheme as

$$\text{SNR}^*_\Theta(\theta) \approx 10\log_{10}(\theta^2 I^*(\theta)) + 10\log_{10} K,$$

which is a good approximation when the oversampling factor K is large. Similarly, the SNR of the adaptive sensing scheme can be asymptotically approximated by

$$\text{SNR}_\Theta(\theta; \boldsymbol{P}, \boldsymbol{Q}) \approx 10\log_{10}(\theta^2 I_K(\theta; \boldsymbol{P}, \boldsymbol{Q})).$$

Their difference is thus equal to

$$10\log_{10}(I^*(\theta)) - 10\log_{10}(I_K(\theta; \boldsymbol{P}, \boldsymbol{Q})/K),$$

which is directly linked to the cost function after we replace the Fisher information rate by its asymptotic expression in (13).

4. ADAPTIVE SENSING BY RANDOM WALKS

We leave a rigorous study of the minimax problem in (16) to a follow-up work. In this section, we present a suboptimal yet intuitively sound adaptive sensing policy based on random walks.

4.1. Adaptive Sensing Policy

Let the available choices of the duty cycles be ordered in decreasing order so that $d_1 > d_2 > \ldots > d_M$. We consider a system with $N = LM$ states, where L is a positive integer. Imagine that these N states are placed on a line. Then, from left to right, the first L states are mapped to duty cycle d_1, the next L states to d_2, and so on (see Figure 3 for an example.) In general, the mapping between the state space $\mathcal{S}$ and the collection of duty cycles $\mathcal{D}$ can be specified as

$$g(i) = d_{\lfloor (i-1)/L \rfloor + 1}, \quad \text{for } 1 \le i \le N,$$

where $\lfloor x \rfloor$ is the largest integer smaller than or equal to x.

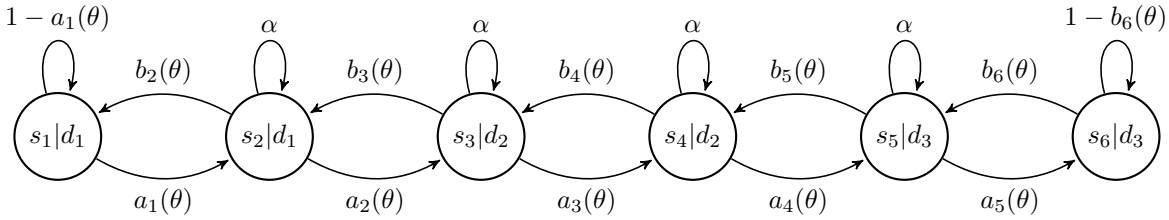

Fig. 3. The transition probability graph of the proposed adaptive sensing policy.

The intuition behind the proposed adaptive sensing policy can be described as follows: Suppose we get several consecutive "1"s from the sensor measurements. This suggests that the current observation window T_g (see Figure 1) might be too long for the incoming light intensity and that the sensor might be saturating. In this case, we reduce the duty cycle d (and thus $T_g = dT$), lowering the chance that the sensor becomes saturated. Due to the way we order the states and duty cycles, the Markov chain of the adaptive sensing scheme should move to the right; Similarly, when we see several consecutive "0"s, the Markov chain should move to the left.

The transition matrices for such random walks are

$$\boldsymbol{P} = \alpha \boldsymbol{I}_N + (1-\alpha) \begin{bmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 1 & 0 \\ \vdots & & & \ddots & 0 & 1 \\ 0 & \dots & \dots & \dots & 0 & 1 \end{bmatrix}, \quad (17)$$

and $\boldsymbol{Q}$ is similarly defined. In (17), $\boldsymbol{I}_N$ is the $N \times N$ identity matrix, and $\alpha \in (0,1)$ is the probability that the random walk stays at the current state.[2] Given this choice of $\boldsymbol{P}$ and $\boldsymbol{Q}$, the transition probability matrix $\boldsymbol{T}(\theta)$ as defined in (10) can be represented by a transition probability graph.

Figure 3 shows one such graph, for the special case of $M = 3$ and $N = 6$. Using (17) and (10), we can specify the transition probabilities in the figure as

$$a_i(\theta) = (1-\alpha)\left(1 - e^{-g(s_i)\theta}\right), \quad \text{for } 1 \le i \le 5,$$

and

$$b_i(\theta) = (1-\alpha)e^{-g(s_i)\theta}, \quad \text{for } 2 \le i \le 6.$$

It is easy to verify that, for all $\theta \in \mathcal{X}$, the Markov chain is irreducible and aperiodic, and thus it admits a unique stationary distribution. The local balance equation [16] yields

$$a_1(\theta)\pi_1(\theta) = b_2(\theta)\pi_2(\theta), a_2(\theta)\pi_2(\theta) = b_3(\theta)\pi_3(\theta), \dots$$

It follows that the stationary distribution can be obtained as

$$\pi_i(\theta) = \lambda_i(\theta) / \sum_{1 \le j \le 6} \lambda_j(\theta),$$

[2]The introduction of the "staying probability" α does not change the stationary distribution of the Markov chain, but suitable choices of α can speed up its convergence towards the steady state.

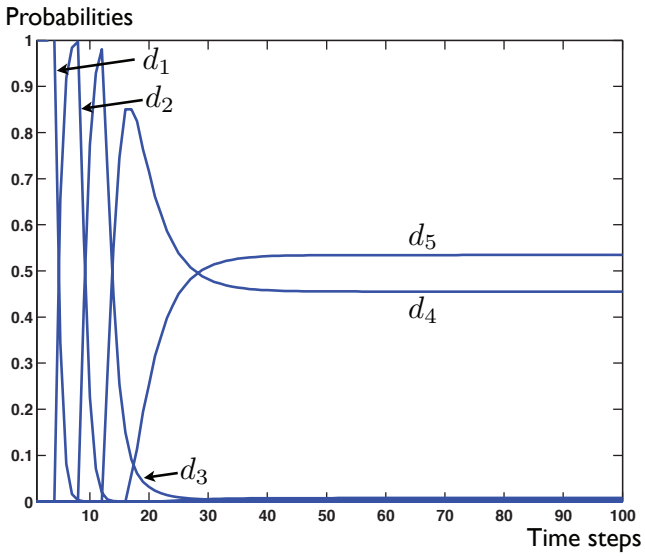

Fig. 4. The evolutions of the probabilities distributions of the proposed adaptive sensing scheme based on random walks.

where $\lambda_1(\theta) = 1$ and $\lambda_i(\theta) = \prod_{j=1}^{i-1} a_j(\theta) / \prod_{j=2}^{i} b_j(\theta)$ for $i \ge 2$.

4.2. Numerical Results

We demonstrate the performance of the proposed adaptive sensing scheme in numerical experiments with the following choices of the parameters: We assume that the incoming photon flux p remains (approximately) constant over time intervals of length $\tau = 5$ ms. The sampling period of the SPS is $T = 1.22$ μs, leading to a temporal oversampling factor of $K = \tau/T = 4096$. The photon detection efficiency of the sensor is set to be $\eta = 0.5$ and the dark count rate $p_{\text{dc}} = 100$ photons/s. In the adaptive sensing scheme, the maximum and minimum duty cycle is $d_{\max} = 0.95$ and $d_{\min} = 5 \times 10^{-5}$, respectively, and the sensor can choose $M = 8$ different duty cycle values, uniformly spaced (in log scale) between $d_{\max}$ and $d_{\min}$. The Markov chain of the system uses $N = 32$ states, with a staying probability $\alpha = 0.1$.

Figure 4 shows the evolutions of the probability distributions of the Markov chain, for $\theta = 100$. Each curve in the figure corresponds to one duty cycle choice and the mth curve plots $\mathbb{P}(g(X_k) = d_m)$ as a function of the time step k. The results of Figure 4 confirms the rapid convergence of the Markov

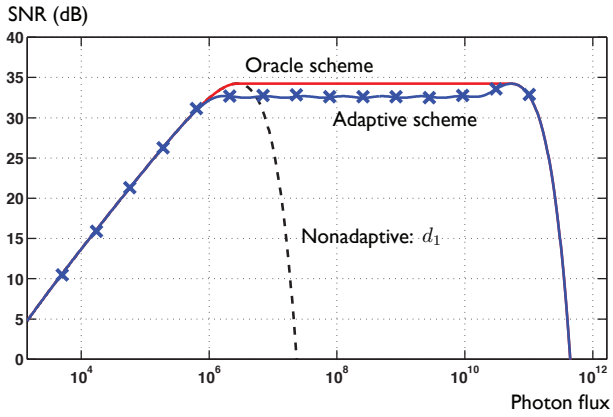

Fig. 5. Comparisons between the proposed adaptive sensing scheme, the oracle scheme, and a nonadaptive scheme.

chain, with the probability distributions stay almost constant after only 50 steps. Moreover, the stationary distribution is tightly concentrated on $d_4 = 1.393 \times 10^{-2}$ and $d_5 = 3.410 \times 10^{-3}$, which are the two duty cycle choices in $\mathcal{D}$ that are closest to the oracle choice $d^*(\theta)$ (obtained according to (9) as $\gamma/100 = 1.594 \times 10^{-2}$.)

In Figure 5, we show the performance of the adaptive sensing scheme in terms of its SNR values (the blue line.) For comparisons, we also show the oracle SNR (the red line) and the SNR values of a nonadaptive scheme with a fixed duty cycle (the dashed line.) All the SNR curves are computed analytically by using the Fisher information as a surrogate for the mean squared errors. These analytical approximation matches very well with Monte-Carlo estimations of the true SNRs (shown as "x"-shaped markers), obtained by simulating the adaptive sensing process and by using ML estimates for reconstructions. Each point is averaged over 2000 independent simulations. The proposed adaptive sensing scheme based on random walks significantly outperforms the nonadaptive scheme, expanding the dynamic range of the latter by 4 orders of magnitude.

5. CONCLUSION

We presented adaptive sensing models, theory and algorithms for single-photon imaging. We considered a discrete setting, where the adaptive sensing policy can be described by a parametric Markov chain on a finite number of states corresponding to different configurations of the SPS. We presented an analytical formula for the Fisher information rate of the Markovian system, showing that it is equal to the weighted average of individual Fisher information associated with different sensor configurations. Furthermore, the weights are simply the stationary probabilities of different configurations. Based on this result, we proposed a minimax regret formulation for designing the optimal sensing policy. We also presented a suboptimal yet effective adaptive sensing scheme based on random walks. Numerical experiments verify the strong performance of the proposed scheme, showing that it can significantly expand the dynamic range of the SPS over nonadaptive approaches.

6. REFERENCES

[1] M. A. Itzler, S. Cova, M. Wahl, and A. Tomita, "Introduction to the issue on single photon counting: Detectors and applications," *IEEE J. Sel. Topics Quantum Electron.*, vol. 13, no. 4, pp. 849 –851, Aug. 2007.

[2] W. H. P. Pernice, C. Schuck, O. Minaeva, M. Li, G. N. Goltsman, A. V. Sergienko, and H. X. Tang, "High-speed and high-efficiency travelling wave single-photon detectors embedded in nanophotonic circuits," *Nature Communications*, vol. 3, pp. 1325, Dec. 2012.

[3] D. E. Schwartz, E. Charbon, and K. L. Shepard, "A single-photon avalanche diode array for fluorescence lifetime imaging microscopy," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 11, pp. 2546–2557, 2008.

[4] C. Niclass, A. Rochas, P. A. Besse, and E. Charbon, "Design and characterization of a CMOS 3-D image sensor based on single photon avalanche diodes," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 9, pp. 1847–1854, 2005.

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

[6] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, 1984, vol. 175.

[7] B. S. Robinson, A. J. Kerman, E. A. Dauler, R. J. Barron, D. O. Caplan, M. L. Stevens, J. J. Carney, S. A. Hamilton, J. K. Yang, and K. K. Berggren, "781 mbit/s photon-counting optical communications using a superconducting nanowire detector," *Optics Letters*, vol. 31, no. 4, pp. 444–446, 2006.

[8] D. F. Figer, J. Leea, B. J. Hanolda, B. F. Aullb, J. A. Gregoryb, and D. R. Schuetteb, "A photon-counting detector for exoplanet missions," in *Proc. SPIE Techniques and Instrumentation for Detection of Exoplanets V*, 2011.

[9] E. R. Fossum, "What to do with sub-diffraction-limit (SDL) pixels? – A proposal for a gigapixel digital film sensor (DFS)," in *IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors*, Nagano, June 2005, pp. 214–217.

[10] 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, 2012.

[11] C. M. Natarajan, M. G. Tanner, and R. H. Hadfield, "Superconducting nanowire single-photon detectors: Physics and applications," *Superconductor Science and Technology*, vol. 25, 2012.

[12] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*, Wiley-Interscience, 2nd edition, Mar. 2007.

[13] T. S. Ferguson, *A Course in Large Sample Theory*, vol. 38, Chapman & Hall/CRC, 1996.

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

[15] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, Springer, 2nd edition, Aug. 1985.

[16] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, Oxford University Press, 3rd edition, 2001.