Photon Inhibition for Energy-Efficient Single-Photon Imaging: Supplementary Information

Lucas J. Koerner
University of St. Thomas
Department of Electrical and Computer Engineering, St. Paul MN, 55105, USA
koerner.lucas@stthomas.edu

Shantanu Gupta
University of Wisconsin Madison
Department of Computer Sciences, Madison, WI, 53706, USA
sgupta@cs.wisc.edu

Atul Ingle
Portland State University
Department of Computer Science
Portland, OR, 97201, USA
ingle2@pdx.edu

Mohit Gupta
University of Wisconsin Madison
Department of Computer Sciences, Madison, WI, 53706, USA
mohitg@cs.wisc.edu

Contents

S 1. Photon Flux Values in Photometric Units

Fig. 1(a) in the main text shows the rate of growth in camera power consumption as a function of photon flux in photons per second. Unfortunately, due to the complications involved in direct conversion of radiometric photon flux into photometric quantities, it is not possible to provide exact numbers (in lux) for the photon flux. For some intuition on the real-world lighting conditions that these photon flux values correspond to, we provide "back of the envelope" estimates in terms of lux levels using the following relationship:

$$\text{illuminance (lux)} \approx \frac{hc}{\lambda}\frac{K\phi}{A}$$

where $h = 6.626 \times 10^{-34}$ is the Planck's constant, $c = 3 \times 10^8$ is the speed of light in vacuum, λ is the nominal visible wavelength of light which we assume is $555\,\mathrm{nm}$ at which the luminous efficacy of an ideal monochromatic light source is $K = 683\,\mathrm{lm/W}$, ϕ is the incident photon flux (adjusted for the SPAD pixel's non-ideal quantum efficiency of 20%), and A is the effective pixel area (assuming a pixel pitch $\sim 4\,\mu\mathrm{m}$ and fill factor of 10%). Plugging in the range of photon fluxes $10^3 - 10^6$ gives a range of lux levels from $< 1\,\mathrm{lx}$ to over $1000\,\mathrm{lx}$. These lux levels are approximately denoted by icons along the x-axis in Fig. 1(a) as $< 1\,\mathrm{lx}$ for a moonless night, $\sim 1 - 10\,\mathrm{lx}$ for twilight, $\sim 100\,\mathrm{lx}$ at sunrise or sunset, and $>\sim 1000\,\mathrm{lx}$ on a clear sunny day outdoors.

S 2. Power consumption estimates

Since we do not have a hardware prototype of our own, we make a very rough estimate here based on previously published works. From Fig. 4.15 on pg. 97 of Andrei Ardelean's thesis [2], computations in the UltraPhase imager effectively use $\sim$ 0.21 mW to constantly perform "MAC operations with data in registers" in a tight loop (= 1.19 mW total $-$ 0.98 mW "standby" power use). Since UltraPhase has 12x24 = 288 pixels, this comes out to about 729 nW per pixel. Energy consumption by avalanches is estimated in [19] as 11.6 pJ/avalanche, for a different SPAD sensor (not the SwissSPAD2). This suggests that we should come out ahead if we can inhibit at least $\frac{729nW}{11.6pJ/detection}$ = 62,845 detections/sec. Under the parameters of Sec. S1 above, an example of this would be going from 90,000 det./sec. ($\sim$230 lux, mapping to daylight or office light) to 25,000 det./sec. (effectively $\sim$30 lux, around dusk) — still enough for a reasonable image with a SPAD sensor. Therefore there exists a very plausible application setting where inhibition can make an impact.

The numbers above are clearly not specific to our sensor and computations. The UltraPhase processing is reconfigurable and has a 32-bit wide arithmetic logic unit; whereas, inhibition processing would use fixed logic with smaller bit widths. As such we expect the above to be an over-estimate of the computation power. We would also need to measure the avalanche energy expenditure of the SwissSPAD2 sensor instead of re-using the estimates from [19].

S 3. Spatio-temporal Policies for Static Imaging

S 3.1. Simulation Implementation Details

In this section, the methods for the simulations of Section 6 of the main manuscript are described. Images from the BSDS500 dataset [1] were used to simulate binary-frames (specifically 20 images were randomly selected from the official test set). This dataset was chosen due to the availability of ground truth edge maps. Images were gamma-decompressed using the sRGB to CIE XYZ transformation ($\gamma \approx 2.2$) and converted to grayscale using the OpenCV color space conversion function (*cvtColor* with COLOR_BGR2YUV) to create a reference image. For each reference image, 1,000 binary frames were simulated using Monte Carlo methods for each exposure time of interest and saved to disk. Exposure times are reported in units of the average number of photon arrivals per pixel (ppp), since absolute radiometric quantities are not available.

For static imaging, inhibition policy simulations were run for each exposure time separately. Once inhibition patterns are found for each frame index t, the cumulative detections and measurements are calculated for each frame index. This approach allows for extraction of performance metrics and images at a continuous range of average detections per pixel by selecting the number of accumulated binary frames. For exposure bracketing simulations an HDR reconstruction was generated at each frame index using SNR^2 weighting [9]. Metrics of SSIM [23] and mean squared error (MSE) were calculated on binary rate images for the accumulated binary frames with and without inhibition at each frame index using the original image as the reference.

S 3.2. Assessing Inhibition

Pixels that are inhibited are known at the beginning of a frame. An inhibited pixel is insensitive to photon arrivals and, as such, does not consume (avalanche) power when a photon converts. Inhibition is expected to be implemented by lowering or keeping the pixel SPAD bias below the threshold voltage for an avalanche. A pixel that is not inhibited measures either a '0' (if no photons arrived) or a '1'. The avalanche energy is assumed the same for one and more than one photon in a single frame, which has been demonstrated in hardware [15]. To assess the energy efficiency, measurement efficiency, and energy reduction enabled by inhibition we track the number of measurements (W) at each pixel and total measurements (W_T), and similarly the number of detections (D) for each pixel and total detections (D_T). These quantities are defined as follows:

$$W(i,j) = \sum_{t=0}^{t=N-1} (1 - M(i,j,t)) \tag{S1}$$

$$W_T = \sum_{i,j} W(i,j) \tag{S2}$$

$$D(i,j) = \sum_{t=0}^{t=N-1} F(i,j,t) = \sum_{t=0}^{t=N-1} (1 - M(i,j,t))Y(i,j,t) \tag{S3}$$

$$D_T = \sum_{i,j} D(i,j), \tag{S4}$$

where N denotes the number of binary frames. Measurements $W(i,j)$ are the total number of frames during which pixel (i,j) was not inhibited (i.e., the inhibition pattern $M(i,j) = 0$). The number of measurements may correlate with the readout energy if an unconventional readout architecture, such as token passing [7] or asynchronous event readout [3], is combined with the inhibition pattern. Detections $D(i,j)$ are the total number of frames during which a photon was detected by pixel (i,j) when enabled. The number of detections tracks the total avalanche energy consumed by that pixel. Suppl. Table 1 summarizes relevant parameters used for evaluating different inhibition policies.

Description	Variable	Values / Units
Photon flux	$\phi(i,j,t)$	photons/s
Binary frame exposure time	T	s
Exposure	$H(i,j,t) = \phi(i,j,t)T$	photons
Inhibition pattern (disabled = 1)	$M(i,j,t)$	0/1
Incident binary frame	$Y(i,j,t) \sim \text{Bernoulli}(1 - e^{-H(i,j,t)})$	0/1
Binary frame (after inhibition)	$F(i,j,t) = Y(i,j,t) \cdot (1 - M(i,j,t))$	0/1
Binary rate estimate	$\widehat{Y}(i,j) = \sum_t \frac{F(i,j,t)}{M(i,j,t)}$	[0,1]
Explicitly inhibited photons	$I(i,j,t) = Y(i,j,t) \cdot M(i,j,t)$	0/1
Total photon detections	$D_T = \sum_t D(i,j) = \sum_t F(i,j,t)$	photons
Fraction of (possible) photons inhibited	$I_F = \frac{\sum I(i,j,t)}{\sum Y(i,j,t)}$	[0,1]

Suppl. Table 1. Relevant quantities for assessing inhibition policies. Pixels are indexed by i and j while $t = 0, 1, ..., N-1$ is the discrete frame number. Explicitly inhibited photons are due to the inhibition pattern itself and not clocked recharge policy.

S 3.3. Details of Imaging Policies

Sec. 5 and Figures 5, 6 in the main text describe static inhibition policies that use spatio-temporal information to compute inhibition patterns. The aggressiveness of these policies is controlled through two parameters η and τ_H: lower values of η and higher values of τ_H can be used to reduce the number of measurements, and hence reduce the total avalanche energy consumption. The policies shown below are the best performing combinations of η and τ_H on exposure bracket captures shown in Fig. 7 (main text) for each of the four spatial policies presented. Policies are designed so that multiplications can be implemented using bit shifts (powers of two) for ease of future in-pixel hardware implementation.

$$\mathbf{P_{cr}}:\ K_s = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 8 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad K_T = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix},\ \eta = 12,\ \tau_H = 32;\ \text{“Center + ring”}$$

$$\mathbf{P_{L}}:\ K_s = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix},\ K_T = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix},\ \eta = 24,\ \tau_H = 4;\ \text{“Laplacian”}$$

$$\mathbf{P_{avg}}: K_s = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad K_T = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix},\ \eta = 6,\ \tau_H = 32,\ \text{“Average”}$$

$$\mathbf{P_{s}}:\ K_s = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad K_T = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix},\ \eta = 2,\ \tau_H = 32; \text{“Single pixel”}$$

The policies described below (and annotated with a $'$) are the best performing combinations of η and τ_H on single-exposure time captures in Fig. 7 (sub-figures f and g) of the main manuscript for each of the four spatial policies presented.

$$\mathbf{P'_{cr}}:\ K_s = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 8 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad K_T = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix},\ \eta = 12,\ \tau_H = 4;\ \text{“Center + ring”}$$

$$\mathbf{P'_{L}}: K_s = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix},\ K_T = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix},\ \eta = 24,\ \tau_H = 4;\ \text{“Laplacian”}$$

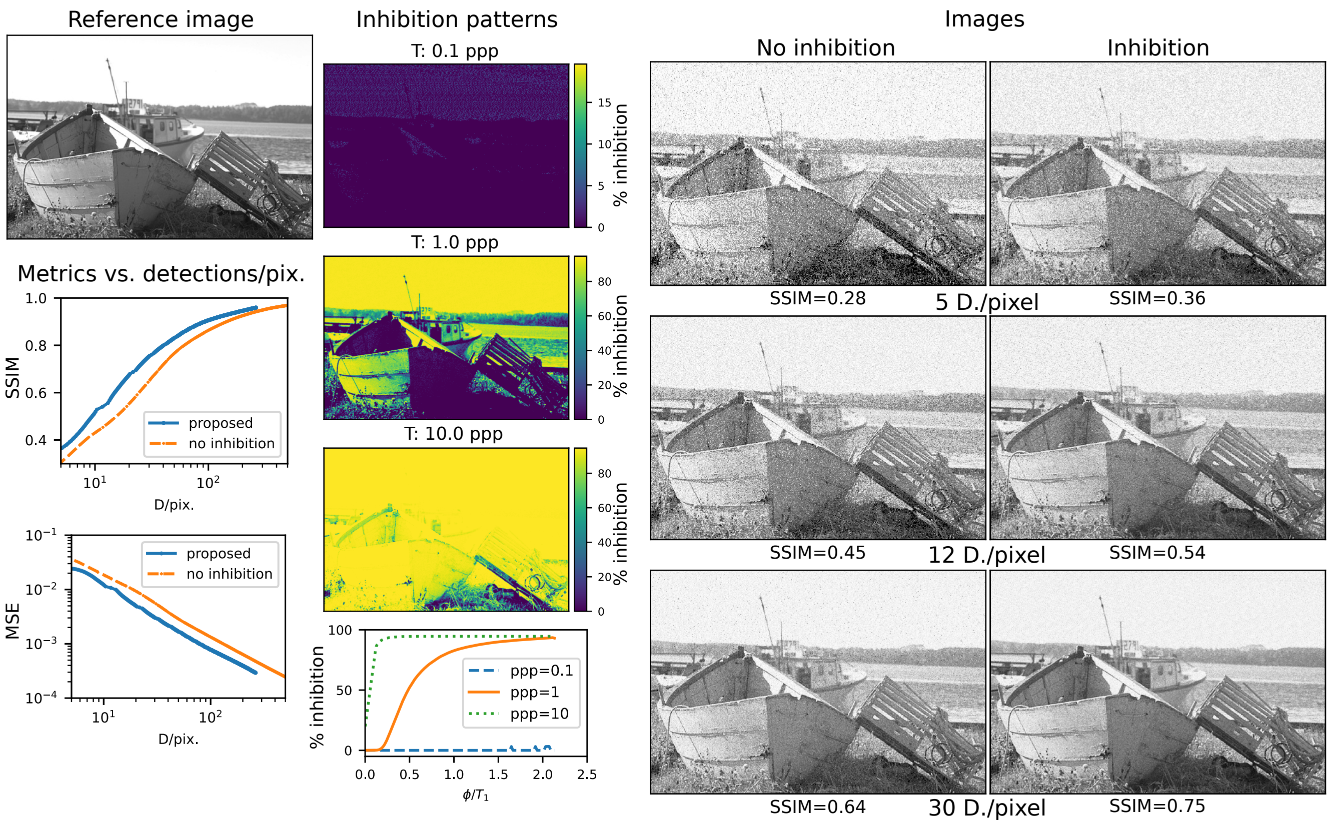
$$\mathbf{P'_{avg}}:\ K_s = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad K_T = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix},\ \eta = 12,\ \tau_H = 4;\ \text{“Average”}$$

$$\mathbf{P'_{s}}:,\ K_s = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad K_T = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix},\ \eta = 2,\ \tau_H = 8;\ \text{“Single pixel”}$$

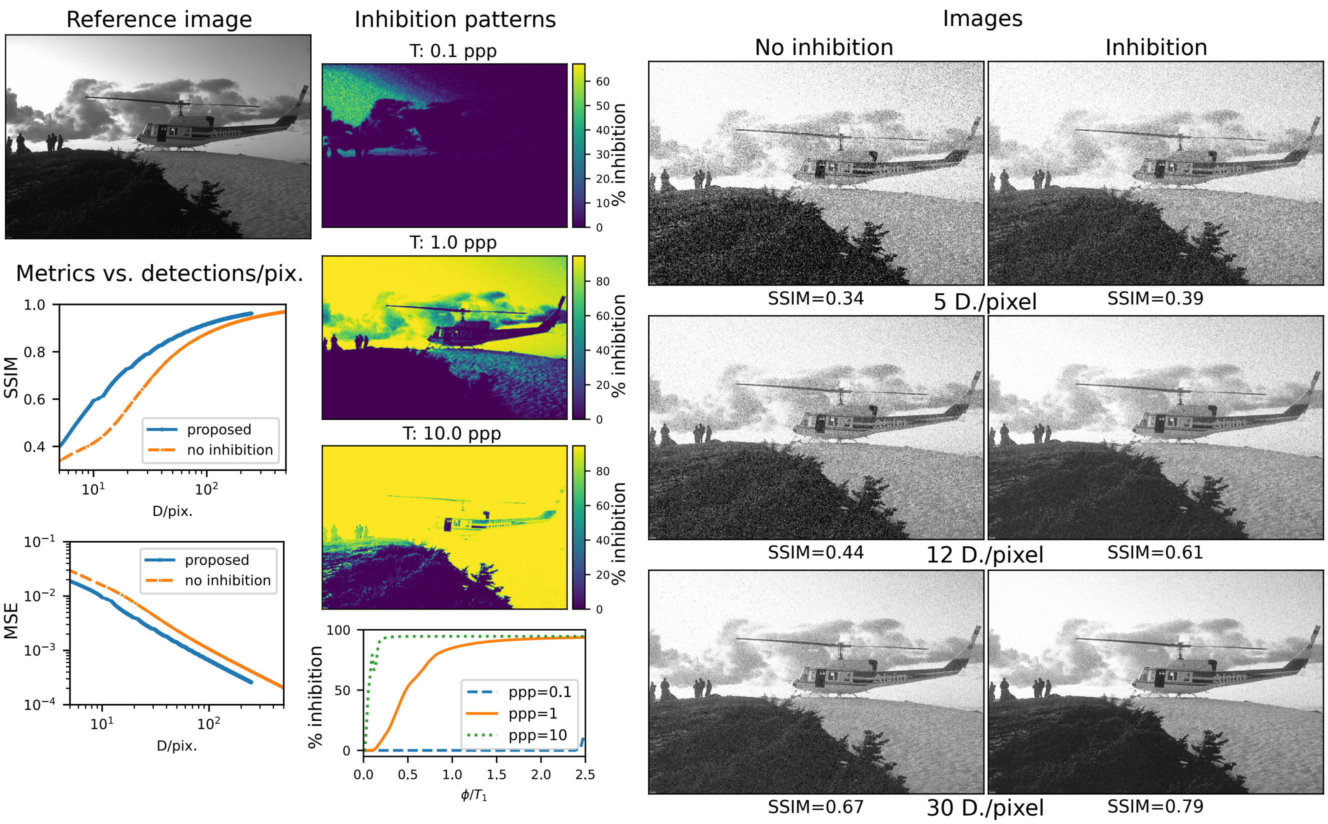
Inhibition policies for static imaging were studied by simulations at parameters values of $\eta = [2, 6, 12, 24]$ and $\tau_H = [4, 8, 16, 32]$. A more extensive search was not attempted due to computation time and disk usage.

S 3.4. Additional Static Image Simulation Results

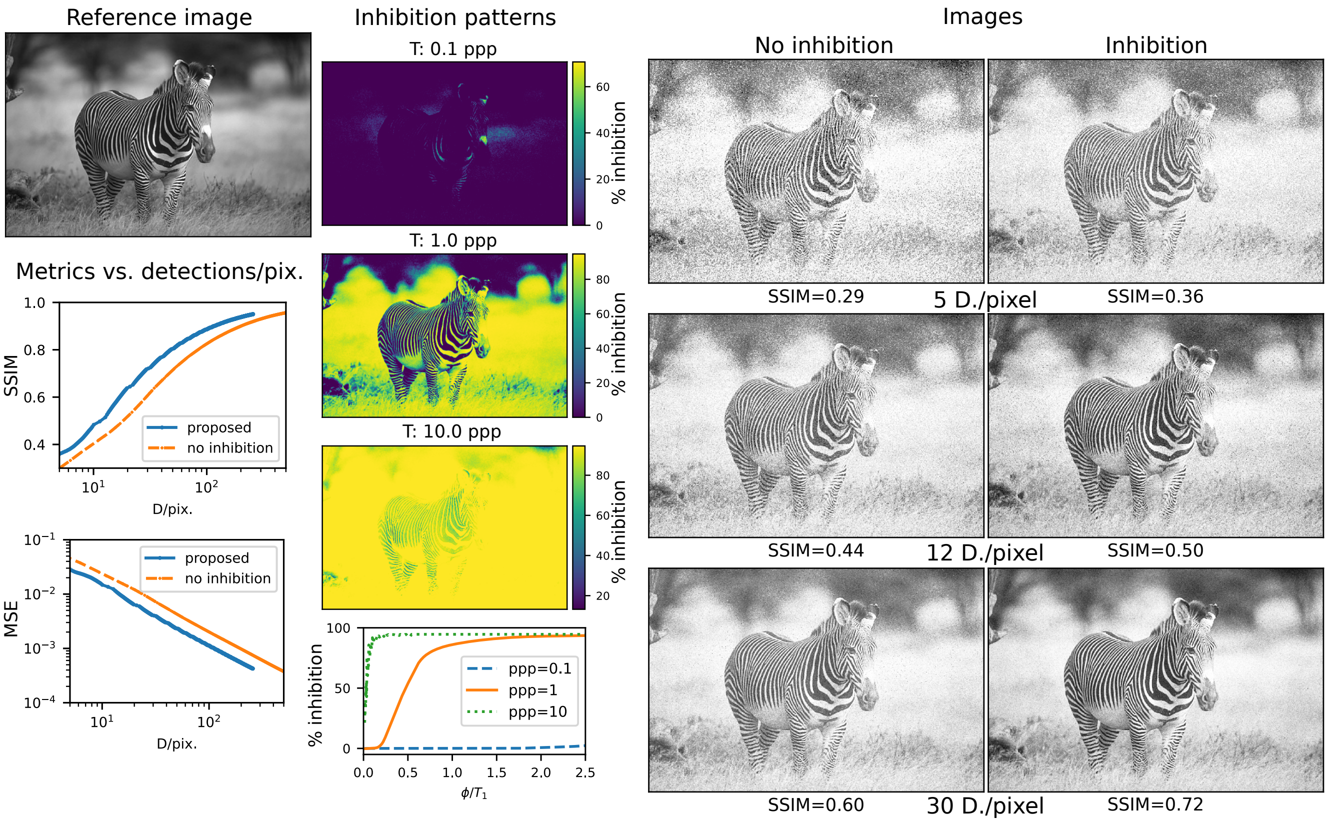
Suppl. Figs. 1, 2, and 3 expand upon the results of Fig. 5 of the main text to show inhibition patterns and resulting images at three levels of average detections per pixel using the P_{cr} policy described above with $\eta = 12$ and $\tau_H = 32$.



Suppl. Fig. 1. **Power-efficient static single-photon imaging via inhibition.** A reference image (BSDS500: 393035) displayed in the top left is captured using a bracket of three exposure times with 1,000 binary frames for each exposure time. The second column displays the resulting average inhibition patterns for each exposure time. The top most pattern from the shortest exposure time modestly inhibits and does so at the brightest pixels only. The inhibition pattern of the longest exposure time allocates most measurements to the darkest areas of the scene (in the shadows to the right of boat in the the foreground). The bottom chart summarizes the inhibition patterns using smoothed curves of the inhibition percent versus the flux of each pixel for each of the three exposure times (Lowess filter with a fraction of 1/5). The right-most columns show binary rate images using gamma compression ($\gamma = 0.4$) without (left) and with (right) inhibition at equal average detections per pixel. Detections increase moving down with averages of 5, 12, and 30 detections per pixel shown. Image quality metrics versus detections per pixel are summarized in the center and bottom of the left most column.

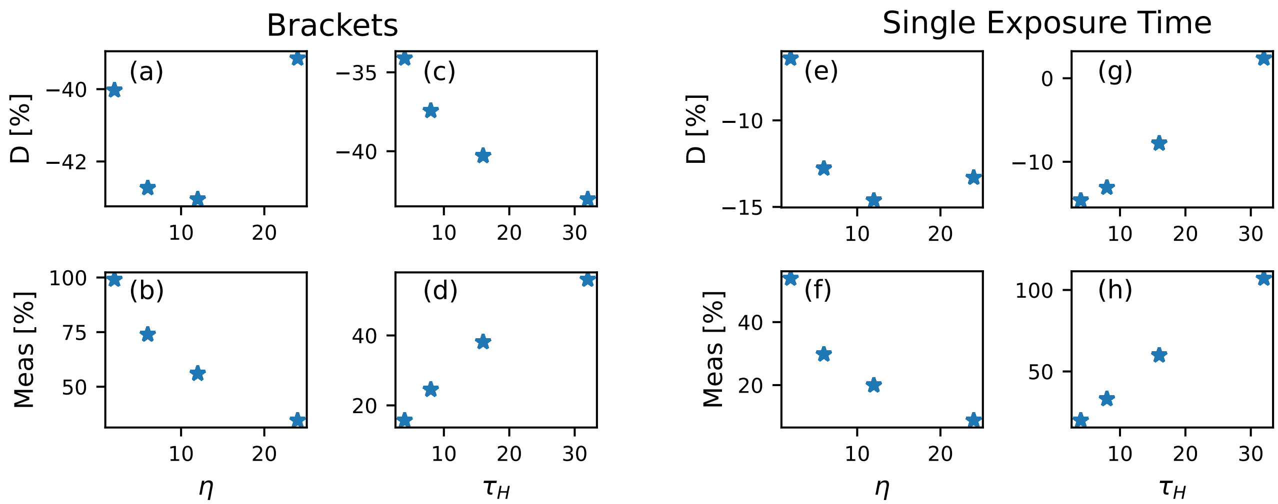


Suppl. Fig. 2. **Power-efficient static single-photon imaging via inhibition.** A reference image (BSDS500: 179084) displayed in the top left is captured using a bracket of three exposure times with 1,000 binary frames for each exposure time. The second column displays the resulting average inhibition patterns for each exposure time. The top most pattern from the shortest exposure time only modestly inhibits and does so at the brightest pixels only (maximum of ∼60% inhibition, primarily in the sky). The inhibition pattern of the longest exposure time allocates most measurements to the darkest areas of the scene (the hilltop and the dark areas of the helicopter). The bottom chart summarizes the inhibition patterns using smoothed curves of the inhibition percent versus the flux of each pixel for each of the three exposure times (Lowess filter with a fraction of 1/5). The right-most columns show binary rate images using gamma compression ($\gamma = 0.4$) without (left) and with (right) inhibition at equal average detections per pixel. Detections increase moving down with averages of 5, 12, and 30 detections per pixel shown. Image quality metrics versus detections per pixel are summarized in the center and bottom of the left most column.



Suppl. Fig. 3. **Power-efficient static single-photon imaging via inhibition.** A reference image (BSDS500: 130066) displayed in the top left is captured using a bracket of three exposure times with 1,000 binary frames for each exposure time. The second column displays the resulting average inhibition patterns for each exposure time. The top most pattern from the shortest exposure time only modestly inhibits and does so at the brightest pixels only. The inhibition pattern of the longest exposure time allocates most measurements to the darkest areas of the scene (the dark stripes of the zebra). The bottom chart summarizes the inhibition patterns using smoothed curves of the inhibition percent versus the flux of each pixel for each of the three exposure times. (Lowess filter with a fraction of 1/5). The right-most columns show binary rate images using gamma compression ($\gamma = 0.4$) without (left) and with (right) inhibition at equal average detections per pixel. Detections increase moving down with averages of 5, 12, and 30 detections per pixel shown. Note the improved contrast of the image captured using an inhibition policy in the bottom row (30 D./pix). Image quality metrics versus detections per pixel are summarized in the center and bottom of the left most column.

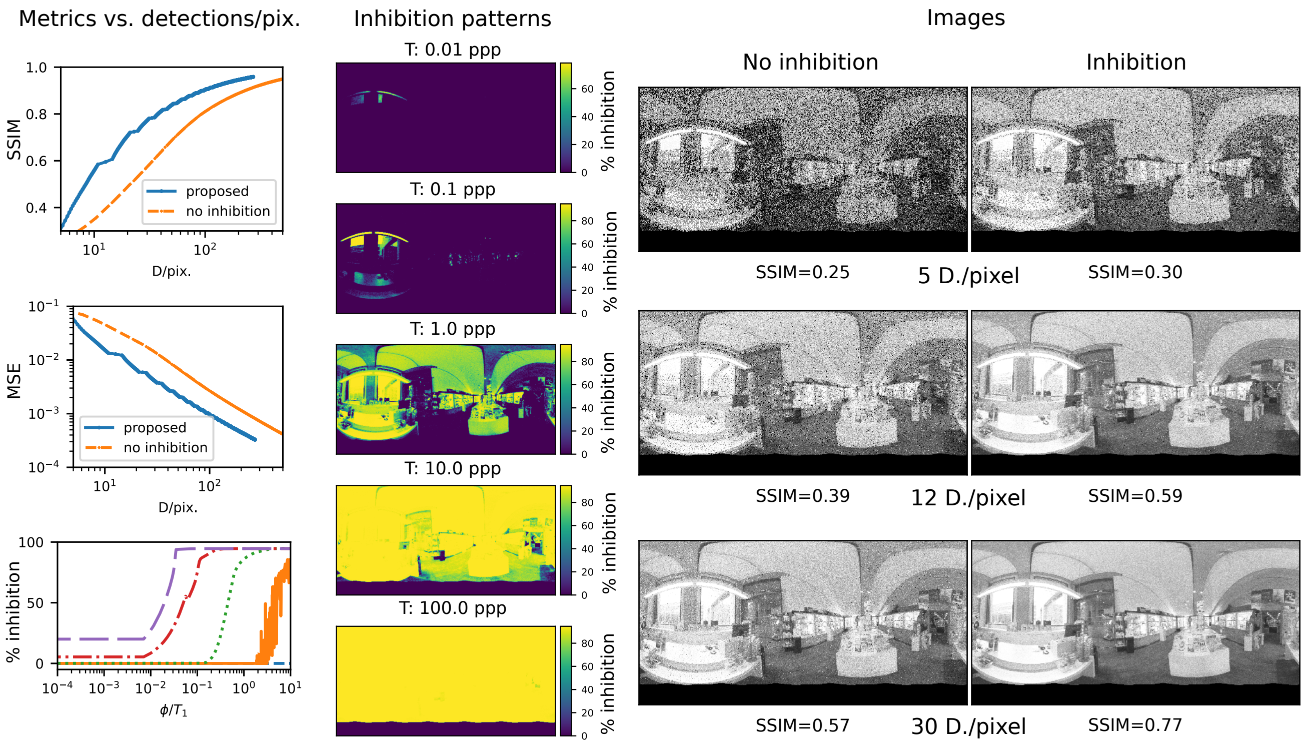
Suppl. Fig. 4 shows the average percent change from no inhibition to inhibition in detections (top row) and measurements (bottom row) at equal image quality versus two parameters of the P_{cr} proposed inhibition policy. These charts demonstrate the balance of detection efficiency and measurement efficiency. Efficiency improvements via inhibition correspond to negative values. For the exposure bracket scenario, the improvement in detection efficiency by more aggressive inhibition ($\tau_H \uparrow$) shown in (c) increases the measurements, and hence degrades measurement efficiency (d). As seen in (e, f, g, h) a single exposure time policy performs best with less aggressive inhibition ($\tau_H = 4$) since frames with a shorter exposure time are not available to fill in missing information for the brightest and aggressively inhibited pixels. Yet, single exposure time policies still reduce detections by nearly 15%.



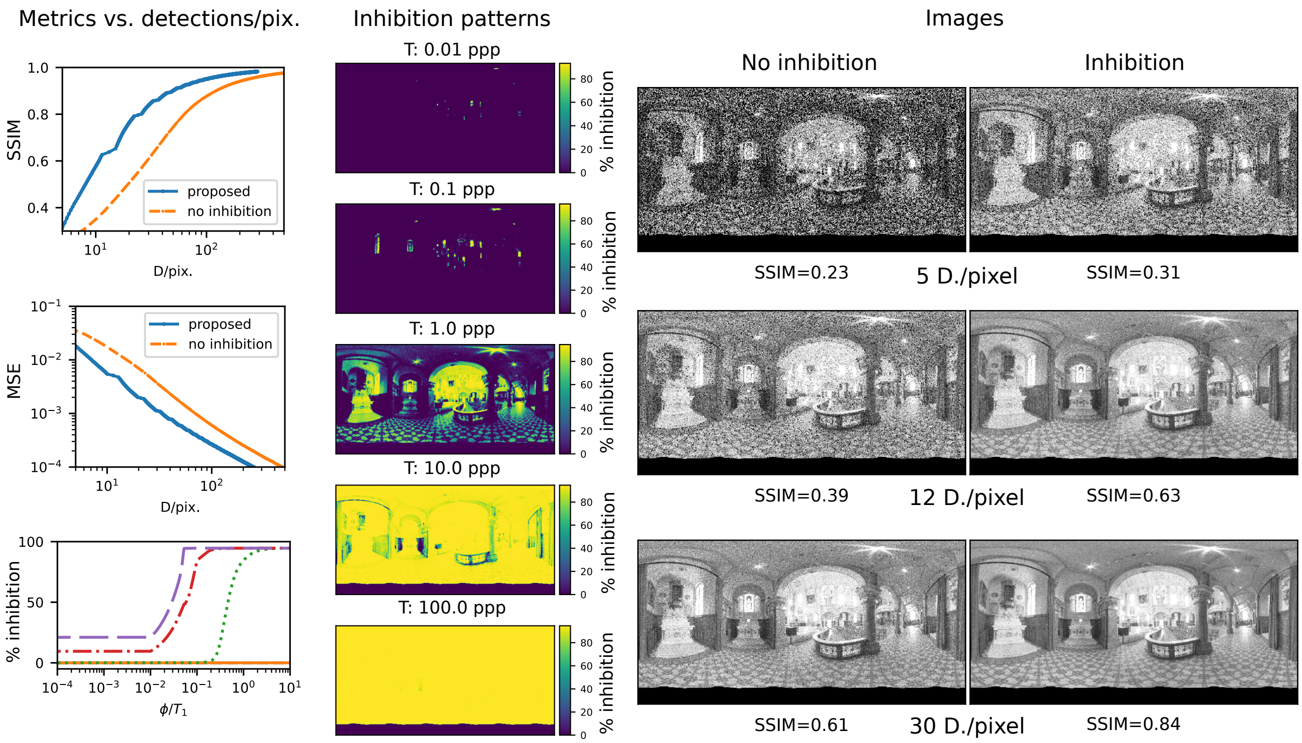
Suppl. Fig. 4. **Inhibition tuning parameters tradeoff detections and measurements.** For the P_{cr} (exposure brackets, left) and P'_{cr} (single exposure, right) policies percent change in detections (D) and measurements at SSIM=0.7 with one parameter varied. (a) D% (as compared to without inhibition) versus the inhibition threshold η at a constant holdoff time ($\tau_H = 32$). A negative value, as in (a), indicates that the inhibition policy required fewer detections for equal SSIM. Notice in (b) how measurements (as $\%$ of total possible) increase (measurement efficiency degrades) at more aggressive inhibition thresholds (smaller η). (c) Shows the impact of the holdoff time (τ_H) at a constant threshold of $\eta = 12$. (e, f, g, h) show the same for a single exposure time capture. For (e, f) $\tau_H = 4$ and for (g, h) $\eta = 12$.

S 3.5. High Dynamic Range Simulation Results

Score-based inhibition: We also assessed the impact of scored-based inhibition to high dynamic range images from the Laval indoor HDR dataset [8]. Suppl. Figs. 5 and 6 summarizes the results from the P_{cr} policy described above with $\eta = 12$ and $\tau_H = 32$. To accommodate the wide range of illumination in these images the simulations used five logarithmically spaced exposure times (in steps of ×10). Before simulation, the images were reduced in size by ×4 along both dimensions using openCV resize with the default bilinear interpolation method to decrease the time required for simulation. When capturing high dynamic range scenes, inhibition allows for a wide range of exposure times to efficiently measure bright and dim pixels with a reduced increase in avalanches at the brightest pixels. For these experiments the exposure time sequence was not carefully explored. Future work could optimize the sequence of exposure times in concert with the inhibition policy while using a more holistic energy cost model.



Suppl. Fig. 5. **Power-efficient static single-photon imaging via inhibition.** A reference image (9C4A0599; 143.5 dB dynamic range) is captured using a bracket of five exposure times with 1,000 binary frames for each exposure time. The second column displays the resulting average inhibition patterns for each exposure time. The top most pattern from the shortest exposure time modestly inhibits and does so at the brightest pixels only. The inhibition pattern of the longest exposure time allocates most measurements to the darkest areas of the scene. The bottom chart in the leftmost column summarizes the allocations of measurements using smoothed curves of the inhibition percent versus the flux of each pixel for each of the three exposure times (Lowess filter with a fraction of 1/5). The right-most columns show binary rate images using gamma compression ($\gamma = 0.4$) without (left) and with (right) inhibition at equal average detections per pixel. Detections increase moving down with averages of 5, 12, and 30 detections per pixel shown. Image quality metrics versus detections per pixel are summarized in the leftmost column.



Suppl. Fig. 6. **Power-efficient static single-photon imaging via inhibition.** A reference image (AG8A7597; $173.6\,\mathrm{dB}$ dynamic range) is captured using a bracket of five exposure times with 1,000 binary frames for each exposure time. The second column displays the resulting average inhibition patterns for each exposure time. The top most pattern from the shortest exposure time modestly inhibits and does so at the brightest pixels only. The inhibition pattern of the longest exposure time allocates most measurements to the darkest areas of the scene. The bottom chart summarizes the allocation of measurements using smoothed curves of the inhibition percent versus the flux of each pixel for each of the three exposure times (Lowess filter with a fraction of 1/5). The right-most columns show binary rate images using gamma compression ($\gamma = 0.4$) without (left) and with (right) inhibition at equal average detections per pixel. Detections increase moving down with averages of 5, 12, and 30 detections per pixel shown. Image quality metrics versus detections per pixel are summarized in the leftmost column.

Saturation look-ahead inhibition: We estimated the benefits of saturation look-ahead inhibition analytically using the 10 images with the widest dynamic range in the Laval indoor HDR dataset (up to $178\,\mathrm{dB}$ DR) [8]. The specific policy uses an exponential bracketing scheme with 5 exposure times scaled as $T_{n+1} = 5T_n$, and 10 measurements are taken with each one. The thresholds to inhibit subsequent exposures under the saturation look-ahead policy were set as $D = \{6, 6, 6, 6\}$, to ensure that on average, the flux range crossing them has a probability of photon detection of 0.99 or greater in the next (longer) exposure, representing (near) saturation. That is to say, taking $T_1 < T_2$, the detection rate Y_1 at the exposure time of T_1 that corresponds to a near-saturating detection rate of Y_2 at exposure time T_2 is computed as:

$$Y_1 = 1 - e^{\frac{T_1}{T_2}\log(1-Y_2)} = 1 - e^{0.2\times\log(1-0.99)} \approxeq 0.60 = 6/10. \tag{S5}$$

The analytical results calculated using a subset of pixels from each of the images show that exposure brackets inhibit 90.6% of detections as compared to a minimum exposure time of equal observation length. The saturation look-ahead policy further reduces detections by an average of -38.4% as compared to bracketing alone. In total, averaged over the 10 images, look-ahead inhibition with exposure bracketing inhibits 94.0% of the detections. The images studied were 9C4A6135, AG8A3343, AG8A2979, AG8A5920, AG8A7597, AG8A6813, 9C4A3821, 9C4A3335, 9C4A1696, and 9C4A0599.

S 3.6. Details of Edge Detection Policies

A high performing edge detection policy is presented in Section 6.2 and Fig. 6 of the main manuscript. This policy calculates a score from the 3×3 Laplacian (S_1) and a 3×3 average filter (S_2). The final inhibition decision is the Boolean operation of these scores as $((\eta_1{<}S_1{<}\eta_2) \wedge (S_2 > \eta_3)) \vee (S_2 > \eta_4)$. The thresholds for the Laplacian score S_1 are $\eta_1 = -12$, $\eta_2 = 12$ which detects regions of minimal spatial variation. To inhibit based on minimal spatial contrast, we additionally require that the average score exceeds a modest threshold, $\eta_3 = 4$, so that dim neighborhoods are not inhibited. Finally, independent of the Laplacian calculation, the pixel is inhibited if the average score is excessive, $\eta_4 = 16$. For this policy, $\tau_H = 16$. As in the policies for static imaging, the temporal kernel was $K_T = \begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix}$.

S 4. Photon Efficiency and Metrics

S 4.1. Oracle Measurement Allocation

Avg. D/pix.	Allocation method	MSE↓	SSIM↑	equal MSE: D %	equal SSIM: D %
5	Uniform	$5.23\times10^{-3} \pm 2.89\times10^{-3}$	0.696 ± 0.117		
	MSE optimal	$4.47\times10^{-3} \pm 2.64\times10^{-3}$	0.724 ± 0.116	-10.0 ± 7.8	-13.7 ± 6.87
12	Uniform	$2.18\times10^{-3} \pm 1.21\times10^{-3}$	0.815 ± 0.084		
	MSE optimal	$1.86\times10^{-3} \pm 1.10\times10^{-3}$	0.837 ± 0.081	-10.0 ± 7.8	-13.4 ± 7.09
30	Uniform	$8.71\times10^{-4} \pm 4.82\times10^{-4}$	0.905 ± 0.048		
	MSE optimal	$7.45\times10^{-4} \pm 4.40\times10^{-4}$	0.918 ± 0.046	-10.0 ± 7.8	-12.8 ± 7.44
100	Uniform	$2.61\times10^{-4} \pm 1.32\times10^{-4}$	0.967 ± 0.018		
	MSE optimal	$2.23\times10^{-4} \pm 1.45\times10^{-4}$	0.971 ± 0.017	-10.0 ± 7.8	-12.2 ± 7.43

Suppl. Table 2. Average results from simulations of 20 images using an oracle allocation policy that is optimized for MSE as compared to a uniform distribution of pixel measurements. In alignment with the main manuscript, the last two columns show the percent change in detections for equal MSE and SSIM enabled by the oracle allocation policy as compared to a uniform allocation of measurements. $\pm$ indicates the standard deviation over the 20 images simulated.

This supplemental section motivates the assertion that an optimal measurement allocation that minimizes image MSE allocates measurements in proportion to $\sqrt{1-Y}$. Photon inhibition enables an unequal distribution of measurements among the pixels of the sensor. To guide the design of inhibition policies presented in the main manuscript we considered two questions. How should a fixed number of detections be allocated among pixels of a single-photon sensor to optimize the image mean squared error (MSE)? By what amount can image metrics be improved by measurement allocation policies when total detections are constrained?

The variance of the binary rate estimate of pixel i with binary rate Y_i and allocated W_i measurements is

$$\sigma^2_{Y_i} = \frac{1}{W_i} Y_i(1-Y_i). \tag{S6}$$

The MSE is the sum of the variances over all pixels P. A minimum of the MSE is found by differentiating with respect to W_i and constraining the total detections to $D_T = \sum_i^P Y_i W_i$. The optimal measurements allocated to pixel i for this minimum MSE is then

$$W_i^* = \frac{D_T}{\sqrt{Y_i}} \left(\frac{\sqrt{Y_i(1-Y_i)}}{\sum_{j=1}^{j=P} Y_j\sqrt{1-Y_j}} \right). \tag{S7}$$

Equation (S7) shows that when constrained by total detections, total image MSE is minimized when measurements are allocated in proportion to $\sqrt{1-Y}$. This analytical approach requires perfect knowledge of the photon flux at each pixel, a non-causal "oracle", and is thus an upper bound on the improvements enabled by photon inhibition. Similar optimal allocation policies may be developed for a scenario with a constraint on measurements rather than detections. These allocations were derived separately and then later very similar allocations were found in the work of Medin et al. which derives optimal stopping rules for active imaging systems designed to estimate Bernoulli parameters [14].

To validate Eq. (S7) we simulated the same 20 BSDS500 images as used in the main manuscript at a single effective exposure time of an average of 1 photon per pixel. To prevent unbounded allocations, pixels were forced to a minimum rate of $Y = 0.01$ and a maximum rate of $Y = 0.99$. The variance of each pixel was calculated as Eq. (S6) and the image MSE was evaluated as the average of the variances of all pixels. Two measurement allocations were evaluated. The first approach allocated measurements equally between all pixels ("Uniform"); the second approach allocated measurements as defined in Eq. (S7) ("MSE optimal"). With the image variance calculated from the ground truth binary rate and the measurements allocated to each pixel, a noise corrupted image was generated by adding noise from Gaussian distributed samples at each pixel to the ground truth image. The noise image was then clipped to between [0,1]. This noise image was used to assess image quality metrics, such as SSIM, by comparison to the reference image. Supp. Table 2 demonstrates these simulated results using the oracle "MSE optimal" measurement allocations. The simulations suggest that the inhibition policy results presented in the main manuscript ($\sim$15% detection reduction) approach the limits established by an oracle allocation for these specific images. Potential sources of discrepancy include that an analytical allocation was only available for image MSE and because the oracle simulations added Gaussian distributed noise.

A Generalized Formulation

A binary quanta sensor artificially "squeezes" noise in the high-flux regime [6], and therefore one may say that exposure-referred noise is a more appropriate objective than the binomial mean-squared error (MSE) considered so far.

To accommodate a more general analysis we define the following image-space loss function

$$\mathcal{L}_{im.} \coloneqq \sum_i \frac{1}{W_i} E_i, \tag{S8}$$

where E_i is an arbitrary per-pixel normalized loss which is subsequently driven down by averaging W_i measurements. E_i is therefore assumed to be a function of H_i (or equivalently, Y_i). We maintain the same constraint on the total number of detections:

$$\sum_i W_i Y_i = D_T. \tag{S9}$$

Applying the method of Lagrange multipliers to the loss function $\mathcal{L} \coloneqq \mathcal{L}_{im.} + \lambda \left(\sum_i Y_i W_i - D_T\right)$, yields the optimal allocation

$$W_i^{opt.} = D_T \left(\frac{\sqrt{E_i/Y_i}}{\sum_j Y_j \sqrt{E_j/Y_j}} \right) \propto \sqrt{\frac{E_i}{Y_i}}. \tag{S10}$$

It can be verified that on setting $E_i = Y_i(1 - Y_i)$, the $\sqrt{1-Y}$ weighting of Eq. (S7) is recovered.

Suppl. Table 3. Optimal measurement allocations and expected number of detections for four loss functions. $\mathsf{SNR}_{H/D}$ is defined as $\sqrt{\mathsf{SNR}^2_{H/D}} = \frac{H\sqrt{1-Y}}{Y}$, and similarly $\mathsf{SNR}_{H/W} \coloneqq \sqrt{\mathsf{SNR}^2_{H/W}} = \frac{H\sqrt{1-Y}}{\sqrt{Y}}$.

Metric name	E_i	$W_i^{opt.}$	$\mathbb{E}[D_i]^{opt.} = Y_i W_i^{opt.}$
Binomial MSE	$Y_i(1-Y_i)$	$\propto \sqrt{1-Y_i}$	$\propto Y_i\sqrt{1-Y_i}$
Exposure-referred MSE	$\left(\frac{\mathrm{d}H_i}{\mathrm{d}Y_i}\right)^2 \cdot Y_i(1-Y_i)$ $= \frac{Y_i}{1-Y_i}$	$\propto \frac{1}{\sqrt{1-Y_i}}$	$\propto \frac{Y_i}{\sqrt{1-Y_i}}$
Relative exposure-referred MSE ($\equiv \frac{1}{\mathsf{SNR}^2_H}$)	$\frac{1}{H_i^2} \cdot \left(\frac{\mathrm{d}H_i}{\mathrm{d}Y_i}\right)^2 \cdot Y_i(1-Y_i)$ $= \frac{1}{H_i^2}\frac{Y_i}{1-Y_i}$	$\propto \frac{1}{H_i\sqrt{1-Y_i}} = \frac{\mathsf{SNR}_{H_i/D_i}}{\mathsf{SNR}^2_{H_i/W_i}}$ $= \frac{1}{Y_i\,\mathsf{SNR}_{H_i/D_i}}$	$\propto \frac{Y_i}{H_i\sqrt{1-Y_i}} = \frac{1}{\mathsf{SNR}_{H_i/D_i}}$
$\mathsf{SNR}_{H/W}$-*tracker*	$\frac{1}{Y_i}\mathsf{SNR}^{2k}_{H_i/W_i}$ $k \in \{1, 2, \dots\}$	$\propto \frac{1}{Y_i}\mathsf{SNR}^{k}_{H_i/W_i}$	$\propto \mathsf{SNR}^{k}_{H_i/W_i}$
For $k=2$	$\mathsf{SNR}^2_{H_i/D_i} \cdot \mathsf{SNR}^2_{H_i/W_i}$	$\propto \frac{1}{Y_i}\mathsf{SNR}^2_{H_i/W_i} = \mathsf{SNR}^2_{H_i/D_i}$	$\propto \mathsf{SNR}^2_{H_i/W_i}$

Suppl. Table 3 shows three other possible loss functions with the optimal allocations obtained using the process above. The binomial MSE has already been discussed. "Exposure-referred MSE" transfers the error to the linear radiance domain, for which the optimal allocation is proportional to $\frac{1}{\sqrt{1-Y}}$. This loss function has a potential problem: the optimal allocation *diverges* for very high flux ($Y_i \to 1$). Further, this metric generally encourages measurements to be spread more densely over *bright* pixels – the opposite of the binomial MSE discussed previously. Images acquired under this allocation do have slightly more detail (less noise) in highlights compared to both a uniform spread as well as weighting by $\sqrt{1-Y}$, but at the cost of almost complete loss of detail in dark regions (Suppl. Fig. 7).

The problem with exposure-referred MSE is partially compensated by defining a *relative* version, which normalizes it by the squared radiance – minimizing this is equivalent to maximizing SNR_H directly. The optimal allocation is now proportional to $\frac{1}{H\sqrt{1-Y}}$. Images acquired under this allocation do preserve much more detail in dim regions (Suppl. Fig. 7) –

perhaps excessively so, in fact, since the allocation now diverges for $H \to 0$ as well! This is particularly a problem under high dynamic range: while we can reasonably assume a finite upper bound on flux, there is no obvious lower limit of flux.

The issues with exposure-referred error measures stem from the under-specified nature of the problem so far. A simple approach to addressing the issue of diverging allocations is to regularize the problem: clamping the detection rate Y to the range $[0.01, 0.99]$ can be considered one form of this. Another principled alternative would be to formally include the several forces limiting the total number of measurements allowed: the amount of camera and scene motion, readout bandwidth and power consumption, and the latency-sensitivity of the vision task. Placing explicit inequality constraints on total measurements results in a more complex optimization problem which does not have a closed-form solution. This is an important question which we do not explore further here and leave for future work.

The fourth row of Suppl. Table 3 places the saturation look-ahead policy of the main paper in the current context. In Fig. 4 of the main paper, the expected number of detections $\mathbb{E}[D_i]^{opt.}$ tracks the $\mathsf{SNR}_{H/W}$ plot in log-log space, which suggests a power-law relation (at least approximately). Working backwards from that result to the loss function for which that allocation is optimal, provides insight into the behavior of the policy (which was definitely *not* designed with any formal optimization problem in mind, only to yield adequate performance and be practically feasible – see Suppl. Sec. S 7.2 for related discussion). Focusing on the $k = 2$ case from the table, we see that the so-called "loss" function that the saturation look-ahead policy appears to minimize is a sum of (powers of) SNRs, which does seem extremely counter-intuitive. A way to make sense of this observation is to realize that if treated abstractly, the function $\mathsf{SNR}^2_{H/D} \cdot \mathsf{SNR}^2_{H/W}$ is larger for dim pixels, and therefore an allocation that favors dim pixels drives down the "loss" further. This is the essence of the behavior we intuitively seek from an inhibition policy: $\sqrt{1-Y}$ weighting does the same. However, this is by no means an ideal metric, and tracking the $\mathsf{SNR}_{H/W}$ curve precisely is not an absolute necessity. Even confining ourselves to the space of the four choices considered in Suppl. Table 3, the allocation under binomial MSE appears to yield visually better images than the $\mathsf{SNR}^k_{H/W}$-tracker, and so an inhibition policy that can practically realize it is likely to perform even better than saturation look-ahead.

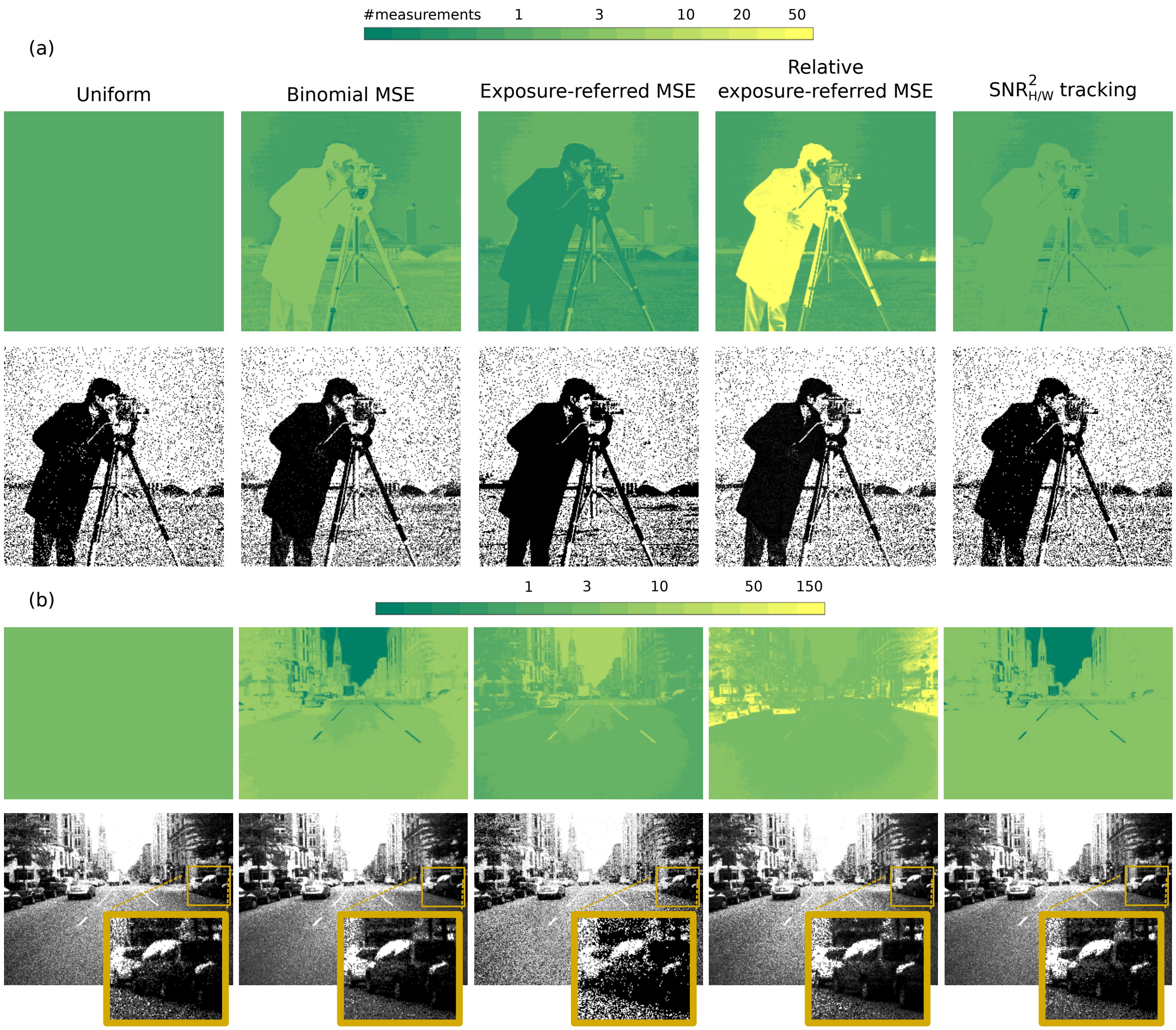
Is there a point to oracle-type analysis? These optimal allocations are derived assuming the pixel intensities are known ahead of time. These oracle policies would not appear to be realistic, yet recent work on optimal spatially varying exposures proposes a two-step capture sequence with the first step being a pilot image that guides configuration [18]. A similar pilot image approach may be useful for practical implementation of the optimal measurement allocations discussed here.

S 4.2. Binary Rate Efficiency Metrics

The efficiency metrics developed in Sec. 4.1 of the main manuscript evaluate the SNR of the measurement of photon exposure (H). An alternative representation of the scene is the binary rate (Y). Exposure and binary rate are related by a nonlinear transform of $H = -\ln(1 - Y)$. Entropy is a possible alternative to SNR for single-photon image sensors [10] if the application processes binary rate images. With a binary rate of Y the entropy is

$$S(Y) = -Y \log_2(Y) - (1 - Y) \log_2(1 - Y) \tag{S11}$$

with a maximum of 1 bit when $Y = 0.5$ [10]. Similar to the main manuscript, an entropy detection efficiency may be defined as $S(Y)^2/Y$. Like detection efficiency in the main manuscript, this metric also demonstrates that detections of nearly saturated pixels are less informative and should be inhibited for energy-efficient single-photon imaging.



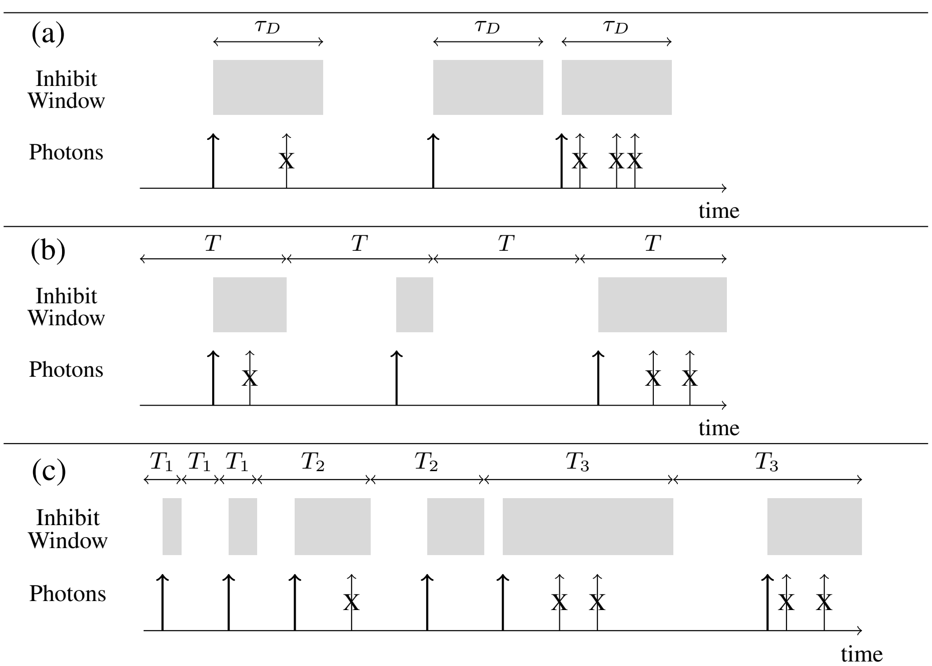
Suppl. Fig. 7. Two example images simulated under optimal allocations for the metrics of Table 3, in bottom rows of (a) and (b). The pseudo-color images in the top rows display the number of binary exposures allocated to each pixel by the expressions in the table (the colormap is in log-scale) – the total expected number of detections is held constant for all allocations. The actual images are simulated by rounding those expressions, and assuming a mean flux of 1.59 photons/pixel over the complete image for a single binary exposure (the peak of the SNR_H curve [4]). Pixels with zero measurements are replaced with either zero or the maximum flux of the true image, as appropriate for each individual metric (the choice is independent of the image). The binomial MSE, the relative exposure-referred MSE, and the $\mathsf{SNR}_{H/W}$-tracking loss function result in denser allocation towards dim pixels (and generally improved image quality), while the plain exposure-referred MSE does the opposite in both aspects. The peaky nature of the allocation with relative exposure-referred MSE can be seen through its strong highlights.
(The reader is requested to zoom in to observe finer details.)

S 5. Background on Static Single-Pixel Inhibition Policies

Current single-photon sensor designs implicitly inhibit photons by setting a maximum count [16, 17] or lengthening the exposure time [21] to limit detections and reduce avalanche energy. These architectures comprise a family of inhibition policies that operate at the individual pixel level and do not adapt as a function of the history of photon detections. Below is background on these already existing policies because our proposed inhibition policies build on top of these.

Suppl. Fig. 8 shows a subset of static single-pixel inhibition policies. SPADs require recharge after an avalanche-inducing photon detection during which recording a subsequent photon is not possible (dead time, τ_D). This detector response of *asynchronous recharge with dead time* inhibits photons at high exposure [11, 12], yet, power consumption is excessive when the average inter-photon arrival interval is shorter than the SPAD dead time [15, 17]. *Clocked recharge* is an alternative that establishes time windows, similar to a conventional exposure time, during which 0 or 1 photon may be detected [17, 21]. After the first photon, any subsequent arrivals during the same predefined exposure window are inhibited. The average number of inhibited photons is equal to $\sum_{k=2}^{\infty}(k-1)\,\mathsf{P}(K=k;H)$ where P denotes the Poisson probability mass function. At the measurement-limited SNR-optimal exposure $H=1.6$ there is a $Y=0.80$ chance of detecting a photon with an average of 0.83 photons inhibited per measurement. With $H\gg 1$ the average number of inhibited photons approaches $H-1$ yet the signal-to-noise ratio degrades because the pixel is nearly saturated. Clocked recharge considerably reduces power in bright light as compared to asynchronous recharge [15, 21]. Because of this, our proposed policies typically maintain and extend clocked recharge.

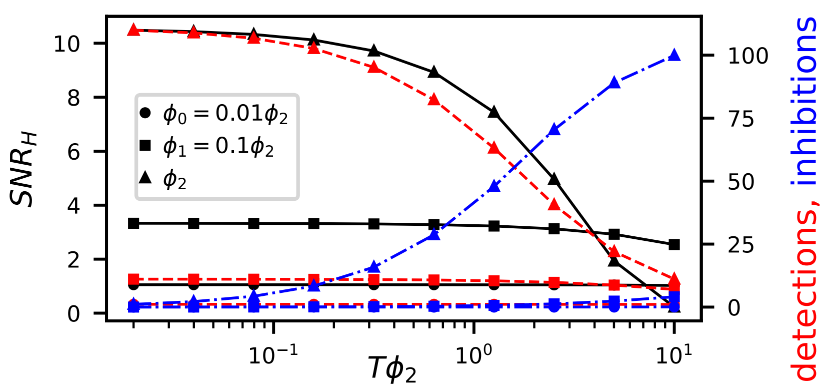
Clocked recharge with exposure brackets [5, 21], shown in Suppl. Fig. 8(c), is a static inhibition policy that uses multiple exposure times to balance constraints on detections and measurements while maintaining SNR over a range of illumination levels. Longer exposure times measure dim pixels with good SNR and limit the detections of bright pixels; short exposure times measure bright pixels with good SNR. Suppl. Fig. 9 guides the tradeoffs between detections, inhibitions, and SNR when selecting a single exposure time of a bracketing sequence. Suppl. Table 4 selects three specific exposure times of an exposure bracket inhibition policy and tracks detections, inhibitions, and the contributions of each exposure time to the HDR reconstruction. Due to near saturation, the detections by the brightest pixel at the longest exposure time(s) have a low weighting for SNR-based HDR reconstruction [9] but still represent $10/25.7 = 38.9\%$ of the total detections, suggesting a clear opportunity for more advanced inhibition policies to reduce avalanche power.



Suppl. Fig. 8. **Single-pixel static inhibition policies without computations.** Arrows represent incoming photons with an 'X' for inhibition. (a) *Asynchronous recharge with dead time (τ_D)*. After a photon detection, the bias voltage of a SPAD must be recharged. During this dead time (τ_D) photons are inhibited. (b) *Clocked recharge*. The recharge period of T sets a window in which 0 or 1 photons can be detected. Subsequent photons within the window are inhibited. (c) *Clocked recharge with exposure brackets*. An extension of clocked recharge with a sequence of different periods.

$\phi[\phi_2]$	$T = 0.1/\phi_2$ wt.	D	I	$T = 1.0/\phi_2$ wt.	D	I	$T = 10.0/\phi_2$ wt.	D	I	HDR SNR
0.01	0.10	0.01	0.00	0.09	0.10	0.00	0.90	0.95	0.05	1.03
0.10	0.32	0.10	0.00	0.14	0.95	0.05	0.85	6.32	3.68	2.62
1.00	0.98	0.95	0.05	0.85	6.32	3.68	0.01	**10.00**	88.11	2.61
Total		1.06	0.05		7.37	3.73		17.27	91.83	

Suppl. Table 4. Clocked recharge with exposure bracket results for three pixel fluxes with each exposure time using $W = 10$ of measurements. Flux values are in units of the maximum flux ϕ_2. *wt.* is the weighting for HDR reconstruction [9], *D* is detections, and *I* is inhibitions. A bold value indicates an opportunity to improve detection efficiency by using a more advanced inhibition policy.



Suppl. Fig. 9. (left) The SNR_H (black, $-$) versus the exposure time with the total sensing latency, T_L, maintained by varying the number of measurements ($W = T_L/T$) at three different flux levels. (right) The number of detections (red, $--$) and the number of inhibitions (blue, $-\cdot-$). A vertical slice represents one exposure time of an exposure bracket policy.

S 6. Details of Experiments on Dynamic Scenes

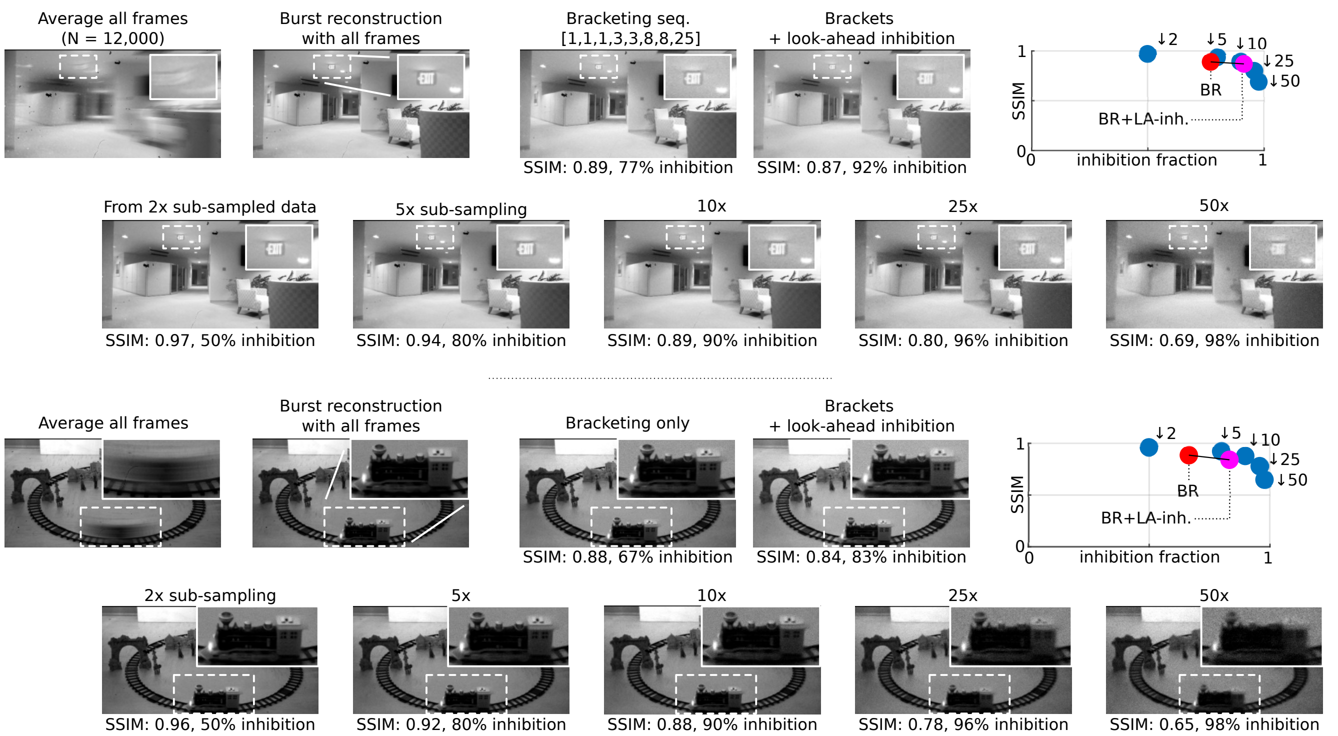
This section details the methods for Section 7 of the main text and motivates the sub-sampling ratio chosen for comparison in Figure 7.

S 6.1. Sub-sampling Factor Tradeoff

Sub-sampling a binary frame sequence is equivalent to setting a longer period for clocked recharge, but with the distinction that the actual exposure duration for which the pixel is photo-sensitive is kept constant (the length of the original binary frame exposures). Suppl. Fig. 10 shows example results for two real images with varying sub-sampling factors, evaluated on image quality using SSIM [23]. For sub-sampling factors greater than $10\times$ a large drop in SSIM can be seen, and similar behavior was obtained with the images of Fig. 7 in the main paper.

S 6.2. Additional Results

A video of the entire sequence of Figure 7 of the main text is included as `long_HDR_seq_burst_reconstructions.mp4`. Individual frames from the full video sequence result are provided in the folder `individual_output_frames`.



Suppl. Fig. 10. **Quality versus inhibition for a single image**. For two separate sequences of 12,000 binary frames each, burst reconstruction [13] is performed first directly (with all photons), and then with various inhibition policies: (top) exposure bracketing, with and without saturation look-ahead inhibition, and (bottom) sub-sampling by dropping frames. The binary frames were captured using the SwissSPAD2 sensor [22], similar to the main paper. The plots on the top-right measure the image quality relative to the reference (no-inhibition) result using SSIM [23], versus the fraction of photons inhibited/dropped by the policy. Sub-sampling factors larger than $10\times$ incur substantial image quality loss for these two images: analogous results are expected for other scenes, possibly for different sub-sampling factors depending on light levels. Separately, the saturation look-ahead policy provides significant inhibition on top of bracketing, with minimal loss in image quality.

S 7. Implementation of Proposed Policies

In Sec. S 7.1 we *estimate* the circuitry required to implement the proposed calculation-based inhibition policy (of Fig. 2). A design assumption is that the implementation will be more constrained by area than computation latency — at $400\,\mathrm{MHz}$ logic clock frequency [2] 4,000 clock cycles are available for computation during a $10\,\mu\mathrm{s}$ recharge period. Therefore, the approach prioritizes minimization of the required in-pixel memory. We emphasize that these are "back of the envelope" estimates; we have not fabricated a chip or created synthesizable digital logic yet. Sec. S 7.2 describes the bracketing-based saturation look-ahead policy in terms of the computation and memory required, as well as the likelihood-maximization process used to convert the bracketed measurements to an estimate of incident flux.

S 7.1. Calculation-based Inhibition

As a reminder from the main text, the inhibition score at each pixel is calculated as

$$S(i,j,t) = K * [(2F(i,j,t) - \mathbf{1}) \cdot M(i,j,t)] \tag{S12}$$

which applies a spatio-temporal filtering kernel, K, of dimensions L, H, T to a ternary representation of the pixel result (1, 0, or -1 for a detection, a disabled pixel, or a measurement that does not detect a photon, respectively). The kernel K can typically be separated into spatial and temporal components as $K = K_s \otimes K_t$ with dimensions $L \times H \times 1$ and $1 \times 1 \times T$, respectively. After each binary frame, the score is compared to a threshold η and the pixel is disabled for the subsequent τ_H frames: $M(i,j,t') = 0$ for $\{t' | t+1 \leq t' \leq t+1+\tau_H\}$ if $S(i,j,t) > \eta$. Suppl. Sec. S 3.3 shows the spatial and temporal kernels (K_s and K_T) used in the simulations.

Suppl. Table 5 describes possible on-chip and in-pixel circuitry for the calculation-based inhibition policy. A subset of the circuit elements must be independent for each pixel as indicated by an entry of "no" in the column titled "Share?". However, other computation circuitry could be shared among a local neighborhood of pixels and housed in a macropixel computation unit [2, 11]. The circuitry is separated into subcircuits of:

1) *SPAD control*: directly controls the bias voltage of the SPAD and allows for enable/disable of the pixel.
2) *Inhibition score*: short-length memory for detection results and arithmetic circuits (adder and shift left) for spatio-temporal computations.
3) *Inhibition control*: evaluates if the score, S, exceeds the inhibition threshold, η. If $S > \eta$, disables the pixel for a count of τ_H clocked-recharge exposure periods.
4) *Measurement results*: in-pixel counters to record the number of detections and inhibitions. These results must be readout to reconstruct the image.

(text continues after next page)

Circuit element	Description	Subcircuit	Notes	Share?
PMOS transistor	SPAD bias control	SPAD control		no
OR gate	Logic to determine SPAD enable, and drive PMOS gate	SPAD control		no
Logic gate	Converts SPAD detection and SPAD enable state to two bit ternary of 1, 0, -1.	Inhibition score		no
Register[1:0] $\times 4$	2-bit register to store signed detection results. T copies arranged as a shift register to create the temporal filter.	Inhibition score	Implements K_T. Proposed has $T = 4 \Rightarrow$ 8 registers as $K_T = [1, 1, 1, 1]$.	no
Adder	Adds each of T detection results.	Inhibition score	Implements K_T. Proposed allows results from [-4, 4].	yes
Shift left	K_s multiplications that are constrained to powers to 2.	Inhibition score	Implements K_s. Must allow for $\times(-1)$. Proposed needs max shift of $\ll 3$ (for $\times 8$).	yes
Adder	Adds pixel and nearest neighbor pixels for a spatial convolution.	Inhibition score	Implements K_s. Proposed results from [-64, 64]. This result does not need to be stored.	yes
Comparison	$S > \eta$	Inhibition control	Proposed implementation has common η among all pixels.	yes
Counter	Count for holdoff period, τ_H, triggered by $S > \eta$. Counts recharge periods and releases SPAD disable.	Inhibition control	Output combines with global clock recharge to enable/disable SPAD. The proposed parameters in simulations require a maximum counter depth of 5.	no
Counter	Count of detections. Depth set by maximum number of frames.	Measurement results	Readout to create the final pixel intensity value. Simulations used a maximum of 1,000 frames which would require a 10 bit counter.	no
Counter	Count of inhibition starts. Depth less than detection counter by $\times \log_2(\tau_H)$. For the best performing imaging policy $\tau_H = 32$ such that a 5 bit counter would be required to count inhibition starts ($2^{10}/2^5 = 2^5$).	Measurement results	Only counting inhibition *starts* reduces the required depth of the counter. Readout to recreate final pixel intensity value. End effects will be observed if the composite frame ends during a holdoff time.	no

Suppl. Table 5. Circuitry to implement the calculation-based inhibition policy. The SPAD control and detection counter is part of the circuitry of a conventional SPAD pixel (for example, see Ota et al. [20]). The last column indicates whether computations may be shared (time-multiplexed) among multiple pixels at a macropixel arithmetic unit. The order of implementation to store the single-pixel results of the temporal filter (K_T) and then re-calculate K_s at the end of each exposure period is designed to to minimize the required in-pixel memory (since the results of K_s span a wider range than K_T). The measurement results subcircuit is needed to recreate the total photon rate after readout. This requires a count of the detections and a count of the number of exposures during which the pixel was active or disabled. An approach that reduces the required circuitry counts only the inhibitions starts with the total number of disabled frame periods as $starts \times \tau_H$.

S 7.2. Saturation Look-ahead Inhibition

This policy is described with an example in Fig. 4a of the main paper.

Pseudo-code A MATLAB-like code listing describing the complete implementation is provided below.

```
function [B_LA, M] = bracket_LA(B, seq, policy)
% Look-ahead inhibition with exposure brackets:
% Input arguments:
%  B:   Nx1 binary vector (for single SPAD pixel)
%  seq: Tx1 bracketing sequence (integer) lengths
%       assumed sorted in non-decreasing order
%  policy: (length(unique(seq))-1) x 1, integer
% Outputs:
%  B_LA: Tx1 binary vector
%  M:    Nx1 "inhibition pattern" of line 299
%        M[n] = 1 means the pixel is _enabled_.
    B_LA = false([T 1]);
    M = true([N 1]);
    n = 1; sp = 0; Bsum_sp = 0; sat = false;
    for s = 1:length(seq)
        % Bracketing from binary SPAD frames
        for ns = 1:seq(s)
            B_LA(s) = B_LA(s) || (M(n) && B(n));
            if ns < seq(s)  % inh. within bracket
                M(n+1) = M(n) && ~B_LA(s);
            n = n + 1;
        % Look-ahead inhibition implementation
        if (s == 1) || (seq(s) ~= seq(s-1))
            % new unique sequence length
            sp = sp + 1; Bsum_sp = 0;
        Bsum_sp = Bsum_sp + B_LA(s);
        if sat || (Bsum_sp > policy(sp))
            M(n) = 0; sat = true;    % disable
```

Complexity The look-ahead inhibition policy is single-pixel and thus is expected to allow for a lighter-weight in-pixel implementation than the calculation-based policies. For a bracketing exposure time sequence of $T := \{T_i\}_{i=1}^K$, the measurement results are represented by the binary sequence of detections $B_T := \{b_i\}_{i=1}^K$ and the inhibition pattern $M_T := \{m_i\}_{i=1}^K$, where $m_i = 1$ denotes the pixel being enabled during that exposure. The memory footprint of both B_T and M_T is already relatively small, but further efficiency is realized by recognizing that for a given unique exposure time, the order of detections with that setting is not important; instead just the sum of detections contributes to the flux estimate.

Furthermore, certain detection sequences are precluded by the inhibition policy. As an example, for the Fibonacci bracketing sequence $[1, 1, 2, 3, 5, 8, 13, 21]$ used in the main text with the inhibition policy of $[2, 1, 1, 1, 1, 1]$, only 15 unique combinations of (B_T, M_T) are possible. Thus, inhibition may even result in *greater* efficiency than the standard exposures, not less (at least in terms of memory and bandwidth use) — the original Fibonacci brackets have 192 possible unique measurements. When multiple bracketing cycles are aggregated on-chip, it may be possible to map the binary detection sequence to an index in a histogram via an encoding implemented on-chip, with only the histogram read out later.

Control signal flow. At the conclusion of each exposure time sequence within a bracketing sequence, the number of detections must be compared to the threshold count in the inhibition policy. If the comparison triggers inhibition, this status is stored and used to disable the SPAD, as in the implementation of S 7.1. At the end of a bracketing sequence a global signal is required to reset the inhibition status of a pixel.

Maximum Likelihood Estimation (MLE) Computation for Exposure Bracketing

As above, the bracketing exposure time sequence is denoted by $T := \{T_i\}_{i=1}^K$, and the inhibition pattern by $M_T := \{m_i\}_{i=1}^K$. After bracketing, every sequence of $\mathrm{sum}(T) = \sum_i T_i$ binary measurements at the original rate is replaced with a binary sequence $B_T := \{b_i\}_{i=1}^K$.

The likelihood L is given as a function of incident flux ϕ:

$$L(\phi) = \prod_{i=1}^{K} \left[(1 - m_i) + m_i \cdot \left(\exp(-\phi T_i)^{1-b_i} \cdot (1 - \exp(-\phi T_i))^{b_i} \right) \right]. \tag{S13}$$

This expression does not have a closed-form expression for its maximum in ϕ. Therefore, we optimize ϕ numerically given a particular combination of M_T and B_T, searching exhaustively over $2{,}000$ uniformly spaced points in the range $[0, 10]$. For a fixed bracket cycle T, the MLE may be found offline for all possible combinations of B_T and M_T and stored in a look-up table (LUT) of maximal possible size $2^{2\times \mathrm{count}(T)}$. But as stated above, only 15 sequences of detections are possible for the Fibonacci bracketing sequence used in the main text when combined with saturation look-ahead inhibition. Therefore the corresponding LUT is also extremely small in practice.

References

[1] Pablo Arbeláez, Michael Maire, Charless Fowlkes, and Jitendra Malik. Contour Detection and Hierarchical Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(5):898–916, 2011. 2

[2] Andrei Ardelean. *Computational Imaging SPAD Cameras*. PhD thesis, EPFL, 2023. 2, 19

[3] Andrew Berkovich, Timir Datta-Chaudhuri, and Pamela Abshire. A scalable 20 ×20 fully asynchronous SPAD-based imaging sensor with AER readout. In *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 1110–1113, 2015. 3

[4] Stanley H. Chan. What Does a One-Bit Quanta Image Sensor Offer? *IEEE Transactions on Computational Imaging*, 8:770–783, 2022. 15

[5] Neale A. W. Dutton, Tarek Al Abbas, Istvan Gyongy, Francescopaolo Mattioli Della Rocca, and Robert K. Henderson. High Dynamic Range Imaging at the Quantum Limit with Single Photon Avalanche Diode-Based Image Sensors. *Sensors*, 18(4):1166, 2018. 16

[6] Eric R. Fossum. Modeling the Performance of Single-Bit and Multi-Bit Quanta Image Sensors. *IEEE Journal of the Electron Devices Society*, 1(9):166–174, 2013. 13

[7] A. Gabrielli. Fast readout for pixel devices. *Measurement Science and Technology*, 19(8):085101, 2008. 3

[8] Marc-André Gardner, Kalyan Sunkavalli, Ersin Yumer, Xiaohui Shen, Emiliano Gambaretto, Christian Gagné, and Jean-François Lalonde. Learning to predict indoor illumination from a single image. *ACM Trans. Graph.*, 36(6):176:1–176:14, 2017. 9, 11

[9] Abhiram Gnanasambandam and Stanley H. Chan. HDR Imaging With Quanta Image Sensors: Theoretical Limits and Optimal Reconstruction. *IEEE Transactions on Computational Imaging*, 6:1571–1585, 2020. 2, 16, 17

[10] Abhiram Gnanasambandam and Stanley H. Chan. Exposure-Referred Signal-to-Noise Ratio for Digital Image Sensors. *IEEE Transactions on Computational Imaging*, 8:561–575, 2022. 14

[11] Robert K Henderson, Nick Johnston, Sam W Hutchings, Istvan Gyongy, Tarek Al Abbas, Neale Dutton, Max Tyler, Susan Chan, and Jonathan Leach. 256 × 256 40nm/90nm CMOS 3D-stacked 120 dB dynamic-range reconfigurable time-resolved SPAD imager. In *2019 IEEE International Solid-State Circuits Conference-(ISSCC)*, pages 106–108. IEEE, 2019. 16, 19

[12] Atul Ingle, Andreas Velten, and Mohit Gupta. High Flux Passive Imaging With Single-Photon Sensors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6760–6769, 2019. 16

[13] Sizhuo Ma, Shantanu Gupta, Arin C. Ulku, Claudio Bruschini, Edoardo Charbon, and Mohit Gupta. Quanta burst photography. *ACM Transactions on Graphics*, 39(4):79:1–79:16, 2020. 18

[14] Safa C. Medin, John Murray-Bruce, David Castañón, and Vivek K Goyal. Beyond Binomial and Negative Binomial: Adaptation in Bernoulli Parameter Estimation. *IEEE Transactions on Computational Imaging*, 5(4):570–584, 2019. 12

[15] Kazuhiro Morimoto, Andrei Ardelean, Ming-Lo Wu, Arin Can Ulku, Ivan Michel Antolovic, Claudio Bruschini, and Edoardo Charbon. Megapixel time-gated SPAD image sensor for 2D and 3D imaging applications. *Optica*, 7(4):346–354, 2020. 3, 16

[16] Jun Ogi, Takafumi Takatsuka, Kazuki Hizu, Yutaka Inaoka, Hongbo Zhu, Yasuhisa Tochigi, Yoshiaki Tashiro, Fumiaki Sano, Yusuke Murakawa, Makoto Nakamura, and Yusuke Oike. A 250fps 124dB Dynamic-Range SPAD Image Sensor Stacked with Pixel-Parallel Photon Counter Employing Sub-Frame Extrapolating Architecture for Motion Artifact Suppression. In *2021 IEEE International Solid- State Circuits Conference (ISSCC)*, pages 113–115, 2021. 16

[17] Yasuharu Ota, Kazuhiro Morimoto, Tomoya Sasago, Mahito Shinohara, Yukihiro Kuroda, Wataru Endo, Yu Maehashi, Shintaro Maekawa, Hiroyuki Tsuchiya, Aymantarek Abdelahafar, Shingo Hikosaka, Masanao Motoyama, Kenzo Tojima, Kosei Uehira, Junji Iwata, Fumihiro Inui, Yasushi Matsuno, Katsuhito Sakurai, and Takeshi Ichikawa. A 0.37W 143dB-Dynamic-Range 1Mpixel Backside-Illuminated Charge-Focusing SPAD Image Sensor with Pixel-Wise Exposure Control and Adaptive Clocked Recharging. In *2022 IEEE International Solid- State Circuits Conference (ISSCC)*, pages 94–96, 2022. 16

[18] Xiangyu Qu, Yiheng Chi, and Stanley H. Chan. Spatially Varying Exposure With 2-by-2 Multiplexing: Optimality and Universality. *IEEE Transactions on Computational Imaging*, 10:261–276, 2024. 14

[19] Fabio Severini, Iris Cusini, Davide Berretta, Klaus Pasquinelli, Alfonso Incoronato, and Federica Villa. SPAD pixel with sub-ns dead-time for high-count rate applications. *IEEE Journal of Selected Topics in Quantum Electronics*, 28(2: Optical Detectors):1–8, 2021. 2

[20] T. Takatsuka, J. Ogi, Y. Ikeda, K. Hizu, Y. Inaoka, S. Sakama, I. Watanabe, T. Ishikawa, S. Shimada, J. Suzuki, H. Maeda, K. Toshima, Y. Nonaka, A. Yamamura, H. Ozawa, F. Koga, and Y. Oike. A 3.36 μm-pitch SPAD photon-counting image sensor using clustered multi-cycle clocked recharging technique with intermediate most-significant-bit readout. In *2023 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)*, pages 1–2, 2023. 20

[21] T. Takatsuka, J. Ogi, Y. Ikeda, K. Hizu, Y. Inaoka, S. Sakama, I. Watanabe, T. Ishikawa, S. Shimada, J. Suzuki, H. Maeda, K. Toshima, Y. Nonaka, A. Yamamura, H. Ozawa, F. Koga, and Y. Oike. A 3.36 μm-pitch SPAD photon-counting image sensor using clustered multi-cycle clocked recharging technique with intermediate most-significant-bit readout. In *2023 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)*, pages 1–2, 2023. 16

[22] Arin Can Ulku, Claudio Bruschini, Ivan Michel Antolovic, Yung Kuo, Rinat Ankri, Shimon Weiss, Xavier Michalet, and Edoardo Charbon. A 512 × 512 SPAD Image Sensor With Integrated Gating for Widefield FLIM. *IEEE Journal of Selected Topics in Quantum Electronics*, 25(1):1–12, 2019. 18

[23] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004. 2, 18