Photon-Flooded Single-Photon 3D Cameras
Anant Gupta  Atul Ingle  Andreas Velten  Mohit Gupta
{agupta225,ingle,velten,mgupta37}@wisc.edu
University of Wisconsin-Madison

Abstract

Single-photon avalanche diodes (SPADs) are starting to play a pivotal role in the development of photon-efficient, long-range LiDAR systems. However, due to non-linearities in their image formation model, a high photon flux (e.g., due to strong sunlight) leads to distortion of the incident temporal waveform, and potentially, large depth errors. Operating SPADs in low flux regimes can mitigate these distortions, but, often requires attenuating the signal and thus, results in low signal-to-noise ratio. In this paper, we address the following basic question: what is the optimal photon flux that a SPAD-based LiDAR should be operated in? We derive a closed form expression for the optimal flux, which is quasi-depth-invariant, and depends on the ambient light strength. The optimal flux is lower than what a SPAD typically measures in real world scenarios, but surprisingly, considerably higher than what is conventionally suggested for avoiding distortions. We propose a simple, adaptive approach for achieving the optimal flux by attenuating incident flux based on an estimate of ambient light strength. Using extensive simulations and a hardware prototype, we show that the optimal flux criterion holds for several depth estimators, under a wide range of illumination conditions.

1. Introduction

Single-photon avalanche diodes (SPADs) are increasingly being used in active vision applications such as fluorescence lifetime-imaging microscopy (FLIM) [34], non-line-of-sight (NLOS) imaging [25], and transient imaging [24]. Due to their extreme sensitivity and timing resolution, these sensors can play an enabling role in demanding imaging scenarios, for instance, long-range LiDAR [7] for automotive applications [21], with only limited power budgets [26].

A SPAD-based LiDAR (Fig. 1) typically consists of a laser which sends out periodic light pulses. The SPAD detects the first incident photon in each laser period, after which it enters a dead time, during which it cannot detect any further photons. The first photon detections in each period are then used to create a histogram (over several periods) of the time-of-arrival of the photons. If the incident flux level is sufficiently low, the histogram is approximately a scaled version of the received temporal waveform, and thus, can be used to estimate scene depths and reflectivity.

Although SPAD-based LiDARs hold considerable promise due to their single-photon sensitivity and extremely high timing (hence, depth) resolution, the peculiar histogram formation procedure causes severe non-linear distortions due to ambient light [13]. This is because of intriguing characteristics of SPADs under high incident flux: the detection of a photon depends on the time of arrival of previous photons. This leads to non-linearities in the image formation model; the measured histogram gets skewed towards earlier time bins, as illustrated in Figs. 1 and 2. This distortion, also called “pile-up” [13], becomes increasingly severe as the amount of ambient light increases, and can lead to large depth errors. This can severely limit the performance of SPAD-based LiDAR in outdoor conditions, for example, imagine a power-constrained automotive LiDAR operating on a bright sunny day [21].

One way to mitigate these distortions is to attenuate the incident flux sufficiently so that the image formation model becomes approximately linear [27, 14]. However,

†This research was supported in part by ONR grants N00014-15-1-2652 and N00014-16-1-2995 and DARPA grant HR0011-16-C-0025.
in a LiDAR application, most of the incident flux may be due to ambient light. In this case, lowering the flux (e.g., by reducing aperture size), requires attenuating both the ambient and the signal light\(^2\). While this mitigates distortions, it also leads to signal loss. This fundamental tradeoff between distortion (at high flux) and low signal (at low flux) raises a natural question: Is there an optimal incident flux for SPAD-based active 3D imaging systems?

**Optimal incident flux for SPAD-based LiDAR:** We address this question by analyzing the non-linear imaging model of SPAD LiDAR. Given a fixed ratio of source-to-ambient light strengths, we derive a closed-form expression for the optimal incident flux. Under certain assumptions, the optimal flux is quasi-invariant to source strength and scene depths, and surprisingly, depends only on the ambient strength and the unambiguous depth range of the system. Furthermore, the optimal flux is lower than that encountered by LiDARs in typical outdoor conditions. This suggests that, somewhat counter-intuitively, reducing the total flux improves performance, even if that means attenuating the signal. On the other hand, the optimal flux is considerably higher than that needed for the image formation to be in the linear regime\(^2, 16\). As a result, while the optimal flux still results in some degree of distortion, with appropriate computational depth estimators, it achieves high performance across a wide range of imaging scenarios.

Based on this theoretical result, we develop a simple adaptive scheme for SPAD LiDAR where the incident flux is adapted based on an estimate of the ambient light strength. We perform extensive simulation and hardware experiments to demonstrate that the proposed approach achieves up to an order of magnitude higher depth precision compared to existing rule-of-thumb approaches\(^2, 16\) that require lowering flux levels to linear regimes.

**Implications:** The theoretical results derived in this paper can lead to a better understanding of this novel and exciting sensing technology. Although our analysis is performed for an analytical pixel-wise depth estimator\(^8\), we show that in practice, the improvements in depth estimation are achieved for several reconstruction approaches, including pixel-wise statistical approaches such as MAP, as well as estimators that account for spatial correlations and scene priors (e.g., neural network estimators\(^18\)). These results may motivate the design of practical, low-power LiDAR systems that can work in a wide range of illumination conditions, ranging from dark to extreme sunlight.

### 2. Related Work

**SPAD-based active vision systems:** Most SPAD-based LiDAR, FLIM and NLOS imaging systems\(^6, 17, 35, 30, 18, 3\) rely on the incident flux being sufficiently low so that pile-up distortions can be ignored. Recent work\(^14\) has addressed the problem of source light pile-up for SPAD-based LiDAR using a realistic model of the laser pulse shape and statistical priors on scene structure to achieve sub-pulse-width depth precision. Our goal is different—we provide theoretical analysis and design of SPAD LiDAR that can perform robustly even in strong ambient light.

**Theoretical analysis and computational methods for pile-up correction:** Pile-up distortion can be removed in post-processing by computationally inverting the non-linear image formation model\(^8, 36\). While these approaches can mitigate relatively low amount of pile-up, they have limited success in high flux levels, where a computational approach alone results in strong amplification of noise. Previous work has performed theoretical analysis similar to ours in a range-gating scenario where scene depths are known\(^11, 37, 10\). In contrast, we derive an optimal flux criterion that minimizes pile-up errors at capture time, is applicable for a broad range of, including extremely high, lighting levels, and does not require prior knowledge of scene depths.

**Alternative sensor architectures:** Pile-up can be suppressed by modifying the detector hardware, e.g., by using multiple SPADs per pixel connected to a single time-correlated single-photon counting (TCSPC) circuit to distribute the high incident flux over multiple SPADs\(^3\). Multi-SPAD schemes with parallel timing units and multi-photon thresholds can be used to detect correlated signal photons\(^29\) and reject ambient light photons that are temporally randomly distributed. The theoretical criteria derived here can be used in conjunction with these hardware architectures for optimal LiDAR design.

**Active 3D imaging in sunlight:** Prior work in the structured light and time-of-flight literature proposes various coding and illumination schemes to address the problem of low signal-to-noise ratios (SNR) due to strong ambient light\(^19, 12, 23, 1\). The present work deals with a different problem of optimal photon detection for SPAD-based pulsed time-of-flight. These previous strategies can potentially be applied in combination with our method to further improve depth estimation performance.

### 3. Background: SPAD LiDAR Imaging Model

This section provides mathematical background on the image formation model for SPAD-based pulsed LiDAR. Such a system typically consists of a laser source which transmits periodic short pulses of light at a scene point, and a co-located SPAD detector\(^22, 32, 9\) which observes the reflected light, as shown in Fig. 1. We model an ideal laser pulse as a Dirac delta function \(\delta(t)\). Let \(d\) be the distance of the scene point from the sensor, and \(\tilde{r} = 2d/c\) be the round trip time-of-flight for the light pulse. The photon flux incident on the SPAD is given by:

\[
\Phi(t) = \tilde{\Phi}_{\text{sig}} \delta(t - \tilde{r}) + \tilde{\Phi}_{\text{bkg}},
\]

where \(\tilde{\Phi}_{\text{sig}}\) is the signal component of the received waveform; it encapsulates the laser source power, distance-

\(^2\)Ambient light can be reduced to a limited extent via spectral filtering.
Figure 2. Effect of ambient light on SPAD LiDAR. A SPAD-based pulsed LiDAR builds a histogram of the time-of-arrival of the incident photons, over multiple laser pulse cycles. In each cycle, at most one photon is recorded, whose timestamp is used to increment the counts in the corresponding histogram bin. (Left) When there is no ambient light, the histogram is simply a discretized, scaled version of the incident light waveform. (Right) Ambient light photons arriving before the laser pulse skew the shape of the histogram, causing a non-linear distortion, called pile-up. This results in large depth errors, especially as ambient light increases.

squared fall-off, scene brightness and BRDF. $\tilde{\Phi}_{bkg}$ denotes the background component, assumed to be a constant due to ambient light. Since SPADs have a finite time resolution (few tens of picoseconds), we consider a discretized version of the continuous waveform in Eq. (1), using uniformly spaced time bins of size $\Delta$. Let $M_i$ be the number of photons incident on the SPAD in the $i$th time bin. Due to arrival statistics of photons, $M_i$ follows a Poisson distribution. The mean of the Poisson distribution, $E[M_i]$, i.e., the average number $r_i$ of photons incident in the $i$th bin, is given as:

$$r_i = \Phi_{sig} \delta_{i,\tau} + \Phi_{bkg}.$$  \hfill (2)

Here, $\delta_{i,j}$ is the Kronecker delta, $\Phi_{sig}$ is the mean number of signal photons received per bin, and $\Phi_{bkg}$ is the (undesirable) background and dark count photon flux per bin. Let $B$ be the total number of time bins. Then, we define the vector of values $(r_1, r_2, \ldots, r_B)$ as the ideal incident waveform.

SPAD histogram formation: SPAD-based LiDAR systems operate on the TCSPC principle \[16\]. A scene point is illuminated by a periodic train of laser pulses. Each period starting with a laser pulse is referred to as a cycle. The SPAD detects only the first incident photon in each cycle, after which it enters a dead time (\(\sim 100\) ns), during which it cannot detect any further photons. The time of arrival of the first photon is recorded with respect to the start of the most recent cycle. A histogram of first photon arrival times is constructed over many laser cycles, as shown in Fig. 2.

If the histogram consists of $B$ time bins, the laser repetition period is $B\Delta$, corresponding to an unambiguous depth range of $d_{\text{max}} = cB\Delta/2$. Since the SPAD only records the first photon in each cycle, a photon is detected in the $i$th bin only if at least one photon is incident on the SPAD during the $i$th bin, and, no photons are incident in the preceding bins. The probability $q_i$ that at least one photon is incident during the $i$th bin can be computed using the Poisson distribution with mean $r_i$ \[8\]:

$$q_i = P(M_i \geq 1) = 1 - e^{-r_i}.$$  \hfill (3)

Thus, the probability $p_i$ of detecting a photon in the $i$th bin, in any laser cycle, is given by \[28\]:

$$p_i = q_i \prod_{k=1}^{i-1} (1 - q_k) = (1 - e^{-r_i}) e^{-\sum_{k=1}^{i-1} r_k}.$$  \hfill (4)

Let $N$ be the total number of laser cycles used for forming a histogram and $N_i$ be the number of photons detected in the $i$th histogram bin. The vector $(N_1, N_2, \ldots, N_B)$ of the histogram counts follows a multinomial distribution:

$$(N_1, N_2, \ldots, N_B) \sim \text{Mult}(N, (p_1, p_2, \ldots, p_B)).$$  \hfill (5)

where, for convenience, we have introduced an additional $(B+1)$th index in the histogram to record the number of cycles with no detected photons. Note that $p_{B+1} =$
\[ 1 - \sum_{i=1}^{B} p_i \text{ and } N = \sum_{i=1}^{B+1} N_i. \] Eq. (4) describes a general probabilistic model for the histogram of photon counts acquired by a SPAD-based pulsed LiDAR.

Fig. 2 (a) shows the histogram formation in the case of negligible ambient light. In this case, all the photon arrival times line up with the location of the peak of the incident waveform. As a result, \( r_i = 0 \) for all the bins except that corresponding to the laser impulse peak. In this case, the measured histogram vector \( (N_1, N_2, \ldots, N_B) \), on average, is simply a scaled version of the incident waveform \( (r_1, r_2, \ldots, r_B) \). The time-of-flight can be estimated by locating the bin index with the highest photon counts:

\[ \hat{r} = \arg \max_{1 \leq i \leq B} N_i, \tag{5} \]

and the scene depth can be estimated as \( \hat{d} = \frac{c \hat{r} \Delta}{2} \).

For ease of theoretical analysis, we assume the laser pulse is a perfect Dirac-impulse with a duration of a single time bin. We also ignore other SPAD non-idealities such as jitter and afterpulsing. We show in the supplement that the results presented here can potentially be improved by combining our optimal photon flux criterion with recent work \[14\] that explicitly models the laser pulse shape and SPAD timing jitter.

4. Effect of Ambient Light on SPAD LiDAR

If there is ambient light, the waveform incident on the SPAD can be modeled as an impulse with a constant vertical shift, as shown in the top of Fig. 2 (b). The measured histogram, however, does not reliably reproduce this “DC shift” due to the peculiar histogram formation procedure that only captures the first photon for each laser cycle. When the ambient flux is high, the SPAD detects an ambient photon in the earlier histogram bins with high probability, resulting in a distortion with an exponentially decaying shape. This is illustrated in the bottom of Fig. 2 (b), where the peak due to laser source appears only as a small blip in the exponentially decaying tail of the measured histogram. The problem is exacerbated for scene points that are farther from the imaging system. This distortion, called pile-up, significantly lowers the accuracy of depth estimates because the bin corresponding to the true depth no longer receives the maximum number of photons. In the extreme case, the later histogram bins receive no photons, making depth reconstruction at those bins impossible.

**Computational Pile-up Correction:** In theory, it is possible to “undo” the distortion by inverting the exponential nonlinearity of Eq. (3), and finding an estimate of the incident waveform \( r_i \) in terms of the measured histogram \( N_i \):

\[ \hat{r}_i = \ln \left( \frac{N - \sum_{k=1}^{i-1} N_k}{N - \sum_{k=1}^{i-1} N_k - N_i} \right). \tag{6} \]

This method is called the Coates’s correction \[8\], and it can be shown to be equivalent to the maximum-likelihood estimate of \( r_i \) \[28\]. See supplementary document for a self-contained proof. The depth can then be estimated as:

\[ \hat{r} = \arg \max_{1 \leq i \leq B} \hat{r}_i. \tag{7} \]

Although this computational approach removes distortion, the non-linear mapping from measurements \( N_i \) to the estimate \( \hat{r}_i \) significantly amplifies measurement noise at later time bins, as shown in Fig. 3.

**Pile-up vs. Low Signal Tradeoff:** One way to mitigate pile up is to reduce the total incident photon flux (e.g., by reducing the aperture or SPAD size). Various rules-of-thumb \[2, 16\] advocate maintaining a low enough photon flux so that only 1-5% of the laser cycles result in a photon being detected by the SPAD. In this case, \( r_i \ll 1 \forall i \) and Eq. (3) simplifies to \( p_i \approx r_i \). Therefore, the mean photon counts \( N_i \) become proportional to the incident waveform \( r_i \), i.e., \( \mathbb{E}[N_i] = NP_i \approx NR_i \). This is called the linear operation regime because the measured histogram \( (N_i)_{i=1}^B \) is, on average, simply a scaled version of the true incident waveform \( (r_i)_{i=1}^B \). This is similar to the case of no ambient light as discussed above, where depths can be estimated by locating the histogram bin with the highest photon counts.

Although lowering the overall photon flux to operate in the linear regime reduces ambient light and prevents pile-up distortion, unfortunately, it also reduces the source signal considerably. On the other hand, if the incident photon flux is allowed to remain high, the histogram suffers from pile-up, undoing which leads to amplification of noise. This fundamental tradeoff between pile-up distortion and low signal raises a natural question: What is the optimal incident flux level for the problem of depth estimation using SPADs?
is already at the minimum level that is achievable by spectral filtering. We assume that the ambient level $\Phi_{bkg}$ is limited extent, via spectral filtering. The optimal attenuation level achieves a BRC with both low distortion, and high signal. (c) The optimal attenuation factor is given by the maxima location (unique) of the minimum value of BRC.

5. Bin Receptivity and Optimal Flux Criterion

In this section, we formalize the notion of optimal incident photon flux for a SPAD-based LiDAR. We model the original incident waveform as a constant ambient light level $\Phi_{bkg}$, with a single source light pulse of height $\Phi_{sig}$. We assume that we can modify the incident waveform only by attenuating it with a scale factor $\Upsilon \leq 1$. This attenuates both the ambient $\Phi_{bkg}$ and source $\Phi_{sig}$ components proportionally. Then, given a $\Phi_{bkg}$ and $\Phi_{sig}$, the total photon flux incident on the SPAD is determined by the factor $\Upsilon$. Therefore, the problem of finding the optimal total incident flux can be posed as determining the optimal attenuation $\Upsilon$. To aid further analysis, we define the following term.

**Definition 1. [Bin Receptivity Coefficient]** The bin receptivity coefficient $C_i$ of the $i$th histogram bin is defined as:

$$C_i = \frac{p_i}{r_i},$$

where $p_i$ is the probability of detecting a photon (Eq. (3)), and $r_i$ is the average number of incident photons (Eq. (2)) in the $i$th bin. $r$ is the total incident flux $r = \sum_{i=1}^{B} r_i$. The bin receptivity curve (BRC) is defined as the plot of the bin receptivity coefficients $C_i$ as a function of the bin index $i$.

The BRC can be considered an intuitive indicator of the performance of a SPAD LiDAR system, since it captures the pile-up vs. shot noise tradeoff. The first term $\frac{p_i}{r_i}$ quantifies the distortion in the shape of the measured histogram with respect to the ideal incident waveform, while the second term $r_i$ quantifies the strength of the signal. Figs. 4 (a-b) show the BRCs for high and low incident flux, achieved by using a high and low attenuation $\Upsilon$, respectively. For small $\Upsilon$ (low flux), the BRC is uniform (negligible pile-up), but skewed towards earlier bins (strong pile-up, as $\frac{p_i}{r_i}$ varies considerably from $\approx 1$ for earlier bins to $\ll 1$ for later bins). Higher the flux, larger the variation in $\frac{p_i}{r_i}$ over $i$.

**BRC as a function of attenuation factor $\Upsilon$:** Assuming total background flux $B\Phi_{bkg}$ over the entire laser period to be considerably stronger than the total source flux, i.e., $\Phi_{sig} \ll B\Phi_{bkg}$, the flux incident in the $i$th time bin can be approximated as $r_i \approx r/B$. Then, using Eqs. (8) and (3), the BRC can be expressed as:

$$C_i(\Upsilon) = B(1 - e^{-\Upsilon}) e^{-(1-\Upsilon)}.$$

(9)

Since total incident flux $r = \Upsilon(\Phi_{sig} + B\Phi_{bkg})$, and we assume $\Phi_{sig} \ll B\Phi_{bkg}$, $r$ can be approximated as $r \approx \Upsilon B\Phi_{bkg}$. Substituting in Eq. (9), we get an expression for BRC as a function only of the attenuation $\Upsilon$, for a given number of bins $B$ and a background flux $\Phi_{bkg}$:

$$C_i(\Upsilon) = B(1 - e^{-\Upsilon\Phi_{bkg}}) e^{-(1-\Upsilon)\Phi_{bkg}}.$$  

(10)

Eq. (10) allows us to navigate the space of BRCs, and hence, the shot noise vs. pile-up tradeoff, by varying a single parameter: the attenuation factor $\Upsilon$. Based on Eq. (10), we are now ready to define the optimal $\Upsilon$.

**Result 1 (Attenuation and Probability of Depth Error).** Let $\tau$ be the true depth bin and $\tilde{\tau}$ the estimate obtained using the Coates’s estimator (Eq. (7)). An upper bound on the average probability of depth error $\sum_{\tau=1}^{B} P(\tilde{\tau} \neq \tau)$ is minimized when the attenuation fraction is given by:

$$\Upsilon_{opt} = \arg \max_{\Upsilon} \min_{i} C_i(\Upsilon).$$

(11)

See the supplementary technical report for a proof. This result states that, given a signal and background flux, the optimal depth estimation performance is achieved when the minimum bin receptivity coefficient is maximized.

From Eq. (10) we note that for a fixed $\Upsilon$, the smallest receptivity value is attained at the last bin $i = B$, i.e., $\min_{i} C_i(\Upsilon) = C_B(\Upsilon)$. Substituting in Eq. (11), we get:

$$\Upsilon_{opt} = \arg \max_{\Upsilon} C_B(\Upsilon).$$

Using $C_B(\Upsilon)$ from Eq. (10) and solving for $\Upsilon$, we get:

$$\Upsilon_{opt} = \frac{1}{\Phi_{bkg}} \log \left( \frac{B}{B - 1} \right).$$

Finally, assuming that $B \gg 1$, we get $\log \left( \frac{B}{B - 1} \right) \approx \frac{1}{B}$. Since $B = 2^{d_{max}/c\Delta}$, where $d_{max}$ is the unambiguous depth range, the final optimality condition can be written as:

$$\Upsilon_{opt} = \frac{c \Delta}{2 d_{max} \Phi_{bkg}}.$$  

(12)

Optimal Flux Attenuation Factor

**Geometric interpretation of the optimality criterion:** Result 1 can be intuitively understood in terms of the space of shapes of the BRC. Figs. 4 (a-b) shows the effect of...
three different attenuation levels on the BRC of a SPAD exposed to high ambient light. When no attenuation is used, the BRC decays rapidly due to strong pile-up. Current approaches [2, 16] that use extreme attenuation make the BRC approximately uniform across all histogram bins, but very low on average, resulting in extremely low signal. With optimal attenuation, the curve displays some degree of pile-up, albeit much lower distortion than the case of no attenuation, but considerably higher values, on average, compared to extreme attenuation. Fig. 4 (c) shows that the optimal attenuation factor is given by the unique maxima location of the minimum value of BRC.

**Choice of optimality criterion:** Ideally, we should minimize the root-mean-squared depth error (RMSE or \( L^2 \)) in the design of optimal attenuation. However, this leads to an intractable optimization problem. Instead, we choose an upper bound on mean probability of depth error (\( L^0 \)) as a surrogate metric, which leads to a closed form minimizer. Our simulations and experimental results show that even though \( T_{\text{opt}} \) is derived using a surrogate metric, it also approximately minimizes \( L^2 \) error, and provides nearly an order of magnitude improvement in \( L^2 \) error.

**Estimating \( \Phi_{\text{bkg}} \):** In practice, \( \Phi_{\text{bkg}} \) is unknown and may vary for each scene point due to distance and albedo. We propose a simple adaptive algorithm (see supplement) that first estimates \( \Phi_{\text{bkg}} \) by capturing data over a few initial cycles with the laser source turned off, and then adapts the attenuation at each point by using the estimated \( \Phi_{\text{bkg}} \) in Eq. (11) on a per-pixel basis.

**Implications of the optimality criterion:** Note that \( T_{\text{opt}} \) is quasi-invariant to scene depths, number of cycles, as well as the signal strength \( \Phi_{\text{sig}} \) (assuming \( \Phi_{\text{sig}} \ll B \Phi_{\text{bkg}} \)). Depth-invariance is by design—the optimization objective in Result 1 assumes a uniform prior on the true depth. As seen from Eq. (11), this results in an \( T_{\text{opt}} \) that doesn’t depend on any prior knowledge of scene depths, and can be easily computed using quantities that are either known (\( \Delta \) and \( d_{\text{max}} \)) or can be easily estimated in real-time (\( \Phi_{\text{bkg}} \)). The optimal attenuation fraction can be achieved in practice using a variety of methods including aperture stops, varying the SPAD quantum efficiency, or with ND-filters.

### 6. Empirical Validation using Simulations

**Simulated single-pixel mean depth errors:** We performed Monte Carlo simulations to demonstrate the effect of varying attenuation on the mean depth error. We assumed a uniform depth distribution over a range of 1000 time bins.

---

5For example, consider a depth range of 100 m and a bin resolution of \( \Delta = 100 \text{ ps} \). Then, the 1% rule of thumb recommends extreme attenuation so that each bin receives \( \approx 1.5 \times 10^{-6} \) photons. In contrast, the proposed optimality condition requires that, on average, one background photon should be incident on the SPAD, per laser cycle. This translates to \( \approx 1.5 \times 10^{-4} \) photons per bin, which is orders of magnitude higher than extreme attenuation, and, results in considerably larger signal and SNR.

---

**Figure 5. Simulation based validation.** (Top row) The values of no, extreme, and optimal attenuation are indicated by dotted vertical lines. In each of the three plots, the value of optimal attenuation is approximately invariant to source power level. The optimal attenuation factor depends only on the fixed ambient light level. (Bottom row) For fixed values of source power, the optimal attenuation factor increases as ambient light decreases. The locations of theoretically predicted optimal attenuation (dotted vertical lines) line up with the valleys of the depth error curves.

**Figure 6. Neural network based reconstruction for simulations.** Depth and error maps for neural network-based depth estimation, under different levels of ambient light and attenuation. Extreme attenuation denotes average \( TB \Phi_{\text{bkg}} = 0.05 \). Optimal attenuation denotes \( TB \Phi_{\text{bkg}} = 1 \). % inliers denotes the percentage of pixels with absolute error \( < 30 \text{ cm} \). \( \Phi_{\text{sig}} = 2 \) for all cases.

Eq. (6) was used to estimate depths. Fig. 5 shows plots of the relative RMSE as a function of attenuation factor \( T \), for a wide range of \( \Phi_{\text{bkg}} \) and \( \Phi_{\text{sig}} \) values.

Each plot in the top row corresponds to a fixed ambient flux \( \Phi_{\text{bkg}} \). Different lines in a plot correspond to different signal flux levels \( \Phi_{\text{sig}} \). There are two main observations...
Validation of optimal attenuation using hardware experiments. These plots have the same layout as the simulations of Fig. 5. As in simulations, the theoretically predicted locations of the optimal attenuation match the valleys of the depth error curves. First, the optimal attenuation predicted by Eq. (12) (dotted vertical line) agrees with the locations of the minimum depth error valleys in these error plots. Second, the optimal attenuation is quasi-independent of the signal flux $\Phi_{\text{sig}}$, as predicted by Eq. (12). Each plot in the second row corresponds to a fixed source flux $\Phi_{\text{sig}}$; different lines represent different ambient flux levels. The predicted optimal attenuation align well with the valleys of respective lines, and as expected, are different for different lines.

Improvements in depth estimation performance: As seen from all the plots, the proposed optimal attenuation criterion can achieve up to 1 order of magnitude improvement in depth estimation error as compared to extreme or no attenuation. Since most valleys are relatively flat, in general, the proposed approach is robust to uncertainties in the estimated background flux, and thus, can achieve high depth precision across a wide range of illumination conditions.

Validation on neural networks-based depth estimation: Although the optimality condition is derived using an analytic pixel-wise depth estimator [8], in practice, it is valid for state-of-the-art deep neural network (DNN) based methods that exploit spatio-temporal correlations in natural scenes. We trained a convolutional DNN [18] using simulated pile-up corrupted histograms, generated using ground truth depth maps from the NYU depth dataset V2 [20], and tested on the Middlebury dataset [33]. For each combination of ambient flux, source flux and attenuation factor, a separate instance of the DNN was trained on corresponding training data, and tested on corresponding test data to ensure a fair comparison across the different attenuation methods.

Fig. 6 shows depth map reconstructions at different levels of ambient light. If no attenuation is used with high ambient light, the acquired data is severely distorted by pile-up, resulting in large depth errors. With extreme attenuation, the DNN is able to smooth out the effects of shot noise, but results in blocky edges. With optimal attenuation, the DNN successfully recovers the depth map with considerably higher accuracy, at all ambient light levels.

7. Hardware Prototype and Experiments

Our hardware prototype is similar to the schematic shown in Fig. 1. We used a 405 nm wavelength, pulsed, picosecond laser (PicoQuant LDH P-C-405B) and a co-located fast-gated single-pixel SPAD detector [5] with a 200 ns dead time. The laser repetition rate was set to 5 MHz corresponding to $d_{\text{max}} = 30 \text{ m}$. Photon timestamps were acquired using a TCSPC module (PicoQuant HydraHarp 400). Due to practical space constraints, various depths covering the full 30 m of unambiguous depth range in Fig. 7 were emulated using a programmable delay module (Micro Photon Devices PSD). Similarly, all scenes in Figs. 8, 9 and 10 were provided with a depth offset of 15 m using the PSD, to mimic long range LiDAR.

Single-pixel Depth Reconstruction Errors: Fig. 7 shows the relative depth errors that were experimentally acquired.
Figure 9. **Depth estimation with varying attenuation.** The average ambient illuminance of the scene was 15 000 lx. With no attenuation, most parts are affected by strong pile-up, resulting in several outliers. For extreme attenuation, large parts of the scene have very low SNR. In contrast, optimal attenuation achieves high depth estimation performance for nearly the entire object. (15 m depth offset removed.)

Figure 10. **Ambient-adaptive Υopt.** This scene has large ambient brightness variations, with both brightly lit regions (right) and shadows (left). Pixel-wise ambient flux estimates were used to adapt the optimal attenuation, as shown in the attenuation map. The resulting reconstruction achieves accurate estimates, both in shadows and brightly lit regions. (15 m depth offset removed.)

over a wide range of ambient and source flux levels and different attenuation factors. These experimental curves follow the same trends observed in the simulated plots of Fig. 5 and provide experimental validation for the optimal flux criterion in the presence of non-idealities like jitter and afterpulsing effects, and for a non-delta waveform.

**3D Reconstructions with Point Scanning:** Figs. 8 and 9 show 3D reconstruction results of objects under varying attenuation levels, acquired by raster-scanning the laser spot with a two-axis galvo-mirror system (Thorlabs GVS-012). It can be seen from the histograms in Fig. 8 (b) that extreme attenuation almost completely removes pile-up, but also reduces the signal to very low levels. In contrast, optimal attenuation has some residual pile-up, and yet, achieves approximately an order of magnitude higher depth precision as compared to extreme and no attenuation. Due to relatively uniform albedos and illumination, a single attenuation factor for the whole scene was sufficient.

Fig. 10 shows depth maps for a complex scene containing a wider range of illumination levels, albedo variations and multiple objects over a wider depth range. The optimal scheme for the “Blocks” scene adaptively chooses different attenuation factors for the parts of the scene in direct and indirect ambient light.7 Adaptive attenuation enables depth reconstruction over a wide range of ambient flux levels.

---

7 In this proof-of-concept, we acquired multiple scans at different attenuations, and stitched together the final depth map in post-processing.

---

**8. Limitations and Future Outlook**

**Achieving uniform depth precision across depths:** The optimal attenuation derived in this paper results in a high and relatively less skewed BRC (as shown in Fig. 4), resulting in high depth precision across the entire depth range. However, since the optimal curve has some degree of pile-up and is monotonically decreasing, later bins corresponding to larger depths still incur larger errors. It may be possible to design a time-varying attenuation scheme that gives uniform depth estimation performance.

**Handling non-impulse waveforms:** Our analysis assumes ideal delta waveform, as well as low source power, which allows ignoring the effect of pile-up due to the source itself. For applications where source power is comparable to ambient flux, a next step is to optimize over non-delta waveforms and derive the optimal flux accordingly.

**Multi-photon SPAD LiDAR:** With recent improvements in detector technology, SPADs with lower dead times (tens of ns) can be realized, which enable capturing more than one photon per laser cycle. This includes multi-stop TC-SPC electronics and SPADs that can be operated in the free-running mode, for which imaging models and estimators have been proposed recently [31, 15]. An interesting future direction is to derive optimal flux criterion for such multi-photon SPAD-based LiDARs.
References


S. 1. Computational Pile-up Correction via Analytic Inversion (Coates’s Method)

Theoretically, it is possible to “undo” the pile-up distortion in the measured histogram by analytically inverting the SPAD image formation model. This method, also called as Coates’s correction in the paper [8], provides a closed form expression for the true incident waveform $r_i$ as a function of the measured (distorted) histogram $N_i$ (Section 4.1 of the main paper).

In this section, we provide theoretical justification for using this method, and show that it is equivalent to computing the maximum likelihood estimate (MLE) of the true incident waveform, and therefore, under certain settings, provably optimal. This result was also proved in [28], and is provided here for completeness. This method has an additional desirable property of providing unbiased estimates of the incident waveform. Furthermore, this method assumes no prior knowledge about the shape of the incident waveform, and thus, can be used to estimate arbitrary incident waveforms, including those with a single dominant peak (e.g., typically received by a LIDAR sensor) for estimating scene depths.

S. 1.1. Derivation of MLE

In any given laser cycle, the detection of a photon in the $i$th bin is a Bernoulli trial with probability $q_i = 1 - e^{-r_i}$, conditioned on no photon being detected in the preceding bins. Therefore, in $N$ cycles, the number of photons $N_i$ detected in the $i$th bin is a binomial random variable when conditioned on the number of cycles with no photons detected in the preceding bins.

$$N_i \mid D_i \sim \text{Binomial} \left( D_i, q_i \right), \quad (S1)$$

where $D_i$ is the number of cycles with no photons detected in bins 1 to $i - 1$ and can be expressed in terms of the histogram counts as:

$$D_i = N - \sum_{j=1}^{i-1} N_j.$$

Therefore, the likelihood function of the probabilities $(q_1, q_2, ..., q_B)$ is given by:

$$L(q_1, q_2, ..., q_B) = P(N_1, N_2, ..., N_B | q_1, q_2, ..., q_B)$$

$$= P(N_1 | q_1) \prod_{i=2}^{B} P(N_i | q_i, N_1, N_2, ..., N_{i-1})$$

$$= P(N_1 | q_1, D_1) \prod_{i=2}^{B} P(N_i | q_i, D_i).$$

by the chain rule of probability, and using the fact that $N_i$ only depends on its probability $q_i$ and preceding histogram counts. Each term of the product is given by the binomial probability from Eq. (S1). Since each $q_i$ only affects a single term, we can calculate its MLE separately as:

$$\hat{q}_i = \arg \max_{q_i} P(N_i | q_i, D_i)$$

$$= \arg \max_{q_i} \left( \frac{D_i}{N_i} \right) q_i^{N_i} (1 - q_i)^{D_i - N_i}$$

$$= \frac{N_i}{D_i} = \frac{N_i}{N - \sum_{j=1}^{i-1} N_j}. \quad (S2)$$
S. 1.2. Calculating the bias of Coates’s corrected estimates

From Eq. (S2) for the MLE, we have for each $1 \leq i \leq B$:

$$E[\hat{q}_i] = E\left[\frac{N_i}{D_i}\right]$$

By the law of iterated expectations:

$$E[\hat{q}_i] = E\left[E\left[\frac{N_i}{D_i} \mid N_1, N_2, \ldots, N_{i-1}\right]\right]$$

$$= E\left[\frac{q_i D_i}{D_i}\right] = q_i$$  \hspace{1cm} (S3)

where the last step uses the mean of the binomial distribution.

Therefore, $\hat{q}_i$ is an unbiased estimate of $q_i$. By combining the expression for $\hat{q}_i$ with $\hat{r}_i = \ln \left(\frac{1}{1 - q_i}\right)$, we get the Coates’s formula mentioned in Section 4.1 of the main text.

S. 2. Derivation of the Optimal Attenuation Factor $\Upsilon^{opt}$

In this section, we derive the expression for optimal attenuation factor $\Upsilon^{opt}$ in terms of the bin receptivities $C_i$. We first compute some properties of the Coates’s estimator which are needed for the derivation. Then we derive an upper bound on the probability that Coates’s estimator produces the incorrect depth. This upper bound is a function of $\Upsilon$. The optimal $\Upsilon$ then follows by minimizing the upper bound.

We assume that the incident waveform is the sum of a constant ambient light level $\Phi_{bkg}$ and a single laser source pulse of height $\Phi_{sig}$. Following the notation used in the main text, we have:

$$r_i = \Phi_{sig} \delta_{i,r} + \Phi_{bkg}.$$  \hspace{1cm} (S7)

Furthermore, we assume that $r_i$ is small enough so that $q_i = 1 - e^{-r_i} \approx r_i$.  \hspace{1cm} (8)

S. 2.1. Variance of Coates’s estimates

From the previous section, the Coates’s estimator is given by:

$$\hat{q}_i = \frac{N_i}{D_i}$$

and the Coates’s time-of-flight estimator is given by:

$$\hat{r} = \arg \max_i \hat{q}_i$$

Note that locating the peak in the waveform is equivalent to locating the maximum $q_i$. From the previous section, we know that $E[\hat{q}_i] = q_i$. Intuitively, this means that the estimates of $q_i$ are correct on average, and we can pick the maximum $\hat{q}_i$ to get the correct depth, on average. However, in order to bound the probability of error, we need information about variance of the estimates. Let $\sigma^2_i$ denote the diagonal terms and $\sigma^2_{i,j}$ denote the off-diagonal terms of the covariance matrix of $(\hat{q}_1, \hat{q}_2, \ldots, \hat{q}_B)$. We have:

$$\sigma^2_i = E[(\hat{q}_i - q_i)^2]$$

$$= E\left[\left(\frac{N_i}{D_i} - q_i\right)^2\right]$$

$$= E\left[\frac{1}{D_i} \left(\frac{N_i}{D_i} - q_i\right)^2\right]$$

$$= E\left[\frac{q_i (1 - q_i)}{D_i}\right]$$  \hspace{1cm} (S5)

Note that this assumption is different from low flux assumption used in the linear operation regime which requires even lower flux levels satisfying $r_i \ll 1/B$.  \hspace{1cm} (S6)
where Eq. (S5) uses the law of iterated expectations and Eq. (S6) uses the variance of the binomial distribution. Note that $D_i$ is also a binomial random variable, therefore,

$$\sigma_i^2 = E \left[ \frac{q_i(1-q_i)}{D_i} \right] \approx \frac{q_i(1-q_i)}{E[D_i]}$$  \hspace{1cm} (S7)

where in the last step, we have interchanged expectation and reciprocal. This can be seen to be true when $D_i$ is large enough so that $D_i \approx D_i + 1$, by writing out $E[1/D_i + 1]$ explicitly. Recalling the definition of $D_i$ and using the mean of the multinomial distribution, we have:

$$E[D_i] = E \left[ N - \sum_{j=1}^{i-1} N_j \right] = N \left( 1 - \sum_{j=1}^{i-1} p_j \right) = \frac{Np_i}{q_i}$$

where the last step follows after some algebraic manipulation involving the definition of $p_i$. Substituting this into Eq. (S7) and using the definition of bin receptivity, we get:

$$\sigma_i^2 = \frac{q_i^2(1-q_i)}{Np_i} = \frac{q_i^2(1-q_i)r}{NC_ir_i} \approx \frac{r_i r}{NC_i}$$

since $r_i \approx q_i \ll 1$ by assumption.

Next we compute $\sigma_{i,j}, i \neq j$. Without loss of generality, assuming $i < j$, we have:

$$\sigma_{i,j}^2 = E \left[ (\tilde{q}_i - q_i)(\tilde{q}_j - q_j) \right]$$

$$= E_{N_1,N_2,...,N_i} \left[ (\tilde{q}_i - q_i)E_{N_{i+1},...,N_N | N_1,...,N_i} (\tilde{q}_j - q_j) \right]$$

$$= E_{N_1,N_2,...,N_i} \left[ (\tilde{q}_i - q_i)E_{N_j,D_j | N_1,...,N_i} (\tilde{q}_j - q_j) \right]$$

$$= E_{N_1,N_2,...,N_i} \left[ (\tilde{q}_i - q_i)E_{(N_j,D_j)} (N_j/D_j - q_j) \right] = 0$$  \hspace{1cm} (S8)

where Eq. (S8) uses the fact that $\tilde{q}_j = N_j/D_j$ only depends on $N_j$ and $D_j$, and Eq. (S9) uses the fact that the innermost expectation is zero. Therefore, $\sigma_{i,j}^2 = 0$ and $\tilde{q}_i$ and $\tilde{q}_j$ are uncorrelated for $i \neq j$.

### S. 2.2. Upper bound on depth error probability

To ensure that the estimated depth is correct, the bin corresponding to the actual depth should have the highest Coates-corrected count. Therefore, for a given true depth $\tau$, we want to minimize the probability of error $P(\tilde{\tau}_{\text{Coates}} \neq \tau)$.

$$P(\tilde{\tau}_{\text{Coates}} \neq \tau) = P \left( \bigcup_{i \neq \tau} \left( \tilde{q}_i > \tilde{q}_\tau \right) \right)$$

$$\leq \sum_{i \neq \tau} P \left( \tilde{q}_i > \tilde{q}_\tau \right)$$

$$= \sum_{i \neq \tau} P \left( \tilde{q}_i - \tilde{q}_\tau > 0 \right).$$

Note that $\tilde{q}_i - \tilde{q}_\tau$ has a mean $q_i - q_\tau$ and variance $\sigma_i^2 + \sigma_\tau^2$, since they are uncorrelated. For large $N$, by the central limit theorem, we have:

$$\tilde{q}_i - \tilde{q}_\tau \sim \mathcal{N}(q_i - q_\tau, \sigma_i^2 + \sigma_\tau^2).$$

Using the Chernoff bound for Gaussian random variables, we get:

$$P(\tilde{q}_i > \tilde{q}_\tau) \leq \frac{1}{2} \exp \left( -\frac{(q_i - q_\tau)^2}{2(\sigma_i^2 + \sigma_\tau^2)} \right)$$

$$\approx \frac{1}{2} \exp \left( -\frac{N(r_i - r_\tau)^2}{2(C_i + C_\tau)} \right)$$

$$= \frac{1}{2} \exp \left( -\frac{N(r_i - r_\tau)^2}{2(C_i + C_\tau)} \right)$$

$$= \frac{1}{2} \exp \left( -\frac{N(\Phi_i + \Phi_\tau)^2}{2(C_i + C_\tau)} \right)$$

where $\Phi_i = \Phi_i$, and $\Phi_\tau = \Phi_\tau$, so that $\Phi_i + \Phi_\tau = \Phi_i + \Phi_\tau$, and $\Phi_i + \Phi_\tau = \Phi_i + \Phi_\tau$, and $\Phi_i + \Phi_\tau = \Phi_i + \Phi_\tau$. Therefore, $P(\tilde{q}_i > \tilde{q}_\tau) \leq \frac{1}{2} \exp \left( -\frac{N(\Phi_i + \Phi_\tau)^2}{2(C_i + C_\tau)} \right)$. 


where the last step uses the definition of $r_i$. Since we are interested in the case of high ambient light and low source power, we assume $\Phi_{\text{sig}} \ll B\Phi_{\text{bkg}}$. The above expression then simplifies to:

$$
P(\hat{q}_i > \hat{q}_\tau) \leq \frac{1}{2} \exp \left( -\frac{N_B \theta^2}{2 \left( \frac{1}{C_i} + \frac{1+\theta}{C_\tau} \right)} \right)
$$

where $\theta = \Phi_{\text{sig}} / \Phi_{\text{bkg}}$ denotes the SBR. Assuming a uniform prior on $\tau$ over the whole depth range, we get the following upper bound on the average probability of error:

$$
\frac{1}{B} \sum_{\tau=1}^{B} P(\hat{q}_{\text{Coates}} \neq \tau) \leq \frac{1}{B} \sum_{\tau=1}^{B} \sum_{i \neq \tau} \frac{1}{2} \exp \left( -\frac{N_B \theta^2}{2 \left( \frac{1}{C_i} + \frac{1+\theta}{C_\tau} \right)} \right)
$$

We can minimize the probability of error indirectly by minimizing this upper bound. The upper bound involves exponential quantities which will be dominated by the least negative exponent. Therefore, the optimal attenuation is given by:

$$
\Upsilon_{\text{opt}} = \arg \min \limits_{\Upsilon} \frac{1}{B} \sum_{i,\tau=1}^{B} \frac{1}{2} \exp \left( -\frac{N_B \theta^2}{2 \left( \frac{1}{C_i} + \frac{1+\theta}{C_\tau} \right)} \right)
$$

The last step is true since the term inside the exponent is maximized for $i = \tau = \arg \min \limits_{i} C_i(r)$. Furthermore, the expression depends inversely on $C_i$ and $C_\tau$, and all other quantities ($N, B, \theta$) are independent of $\Upsilon$. Therefore, minimizing the expression is equivalent to maximizing the minimum bin receptivity.

### S. 2.3. Interpretation of the optimality criterion as a geometric tradeoff

We now provide a justification of our intuition that the optimal flux should make the BRC both uniform and high on average. The optimization objective $\min_i C_i(\Upsilon)$ (Eq. (11)) of Section 4.2 can be decomposed as:

$$
\min_i C_i(\Upsilon) = C_B(\Upsilon) = B \left( 1 - e^{-\Upsilon \Phi_{\text{bkg}}} \right) e^{-(B-1)\Upsilon \Phi_{\text{bkg}}}
$$

$$
= \left[ 1 - e^{-\Upsilon \Phi_{\text{bkg}}} \right] \frac{1}{B(1-e^{-\Upsilon \Phi_{\text{bkg}}}e^{-(B-1)\Upsilon \Phi_{\text{bkg}}})} \frac{1}{B(1-e^{-\Upsilon \Phi_{\text{bkg}}})}
$$

$$
= \frac{1}{B} \sum_{i=1}^{B} C_i(\Upsilon) \left( \frac{1}{C_B(\Upsilon)} - \frac{1}{C_1(\Upsilon)} \right) .
$$

The first term is the mean receptivity (area under the BRC). The second term is a measure of the non-uniformity (skew) of the BRC. Since the optimal $\Upsilon$ maximizes the objective $\min_i C_i(\Upsilon)$, which is the ratio of mean receptivity and skew, it simultaneously achieves low distortion and large mean values.

**Summary:** We derived the optimal flux criterion of Section 5 in the main text, using an argument about bounding the mean probability of error. The expression for optimal attenuation depends on a geometric quantity, the bin receptivity curve, which also has an intuitive interpretation.
S. 3. Alternative computational methods for pile-up correction

In this section we present depth estimation methods that can be used as alternatives to the Coates’s estimator in situations where additional information about the scene is available.

Suboptimality of Coates’s method for restricted waveform types: In our analysis of depth estimation in SPADs, we used the Coates’s estimator for convenience and ease of exposition. The Coates’s method estimates depth indirectly by first estimating the flux for each histogram bin. Although this is optimal for depth estimation with arbitrary waveforms, it is suboptimal in our setting where we assume some structure on the waveform. First, it does not utilize the shared parameter space of the incident waveform, which can be described using just three parameters: background flux $\Phi_{\text{bkg}}$, source flux $\Phi_{\text{sig}}$ and depth $d$. Instead, the Coates method allows an arbitrary waveform shape described by $B$ independent parameters for the flux values at each time bin. Moreover, it does not assume any prior knowledge of $\Phi_{\text{bkg}}$ and $\Phi_{\text{sig}}$.

MAP and Bayes estimators: In the extreme case, if we assume $\Phi_{\text{bkg}}$ and $\Phi_{\text{sig}}$ are known, the only parameter to be estimated is $d$. We can then explicitly calculate the posterior distribution of the depth using Bayes’s rule:

$$P(d|N_1, N_2, ..., N_B) = \frac{P(d)P(N_1, N_2, ..., N_B|d)}{P(N_1, N_2, ..., N_B)}.$$  

Assuming a uniform prior on depth, this can be simplified further:

$$P(d|N_1, ..., N_B) = \frac{P(N_1, ..., N_B|d)}{\sum_{i=1}^{N_B} P(N_1, ..., N_B|i)} \propto P(N_1, ..., N_B|d)$$

$$= \prod_{i=1}^{B} (q_{i|d})^{N_i} (1 - q_{i|d})^{N - \sum_{j=i+1}^{B} N_j}$$

$$= \exp \left\{ \sum_{i=1}^{B} N_i \ln (q_{i|d}) + \sum_{i=1}^{B} D_i \ln (1 - q_{i|d}) \right\}$$  \hspace{1cm} (S10)

where $q_{i|d}$ denotes the incident photon probability at the $i^{\text{th}}$ bin when the true depth is $d$. Note that $q_{i|d}$ for different depths are related through a rotation of the indices $q_{i|d} = q_{(i-d) \mod B|0}$. Therefore, the expression in the exponent of Eq. (S10) can be computed efficiently by a sum of two correlations. The Bayes and MAP estimators are then given by the mean and mode of the posterior distribution respectively.

Advantages of MAP Estimation: It can be shown that Bayes and MAP estimators are optimal in terms of mean squared loss and 0-1 loss respectively [4]. Unlike the Coates method, these methods are affected by the high variance of the later histogram bins only if the true depth corresponds to a later bin. Moreover, it can be seen from Supplementary Fig. 2, that using optimal attenuation improves performance when used in conjunction with a MAP estimator.

Disadvantages of MAP Estimation: The downside of these estimators is that they require knowledge $\Phi_{\text{bkg}}$ and $\Phi_{\text{sig}}$. While $\Phi_{\text{bkg}}$ can be estimated easily from data, estimating $\Phi_{\text{sig}}$ is difficult to estimate in real-time when the SPAD is already exposed to strong ambient light. In comparison, the Coates’s estimator is general and can be applied to any arbitrary flux scenario.

S. 4. Simulation details and results

In this section, we provide details of the Monte Carlo simulations that were used for the results in the main text. We then provide additional simulation results illustrating the effect of attenuation.

Details of Monte Carlo Simulation: We simulate the first photon measurements using a multinomial distribution as described earlier, for various background and source conditions. The true depth is selected uniformly at random from 1 to $B$, and the simulations are repeated on an average of 200 times. The root-mean-squared depth error (RMSE) is estimated using:

$$\text{RMSE} = \left( \frac{1}{200} \sum_{i=1}^{200} \left( \frac{\hat{\tau}_i - \tau_i^{\text{true}} + B}{2} \right) \mod B - \frac{B}{2} \right)^2$$
and the relative depth error is calculated as the ratio of the RMSE to the total depth range:

\[
\text{relative depth error} = \frac{\text{RMSE}}{B} \times 100.
\]

Here \( \tau_{i}^{\text{true}} \) denotes the true depth on the \( i \)th simulation run. It is chosen uniformly randomly from one of the \( B \) bins. Since the unambiguous depth range wraps around every \( B \) bins, we compute the errors modulo \( B \). The addition and subtraction of \( B/2 \) ensures that the errors lie in \((-B/2, B/2)\).

S. 4.1. Relative depth error under various signal and background flux conditions

Supplementary Fig. 1(a) shows the effect of attenuation on relative depth error, as a 2D function of \( \Phi_{\text{sig}} \) and \( \Phi_{\text{bkg}} \) for a wide range of flux conditions. It can be seen that with no attenuation, the operable flux range is limited to extremely low flux conditions. Extreme attenuation extends this range to intermediate ambient flux levels, but only when a strong enough source flux level is used. Using optimal attenuation not only provides lower reconstruction errors at high ambient flux levels but it also extends the range of SBR values over which SPAD-based LiDARs can be operated. For some \((\Phi_{\text{bkg}}, \Phi_{\text{sig}})\) combinations, optimal attenuation achieves zero depth errors, while extreme attenuation has the maximum possible error of 30%.

Supplementary Fig. 1(b) shows the same surface plot from a different viewing angle. It reveals various intersections between the three surface plots. Optimal attenuation provides lower errors than the other two methods for all flux combinations. The error surface when no attenuation is used intersects the extreme attenuation surface around the optimal flux level of \( \Phi_{\text{bkg}} \) of 0.001. For higher \( \Phi_{\text{bkg}} \), using no attenuation is worse than using extreme attenuation, and the trend is reversed for lower \( \Phi_{\text{bkg}} \) values. This is because when \( \Phi_{\text{bkg}} \leq 0.001 \), the optimal strategy is to use no attenuation at all. On the other hand, extreme attenuation reduces the flux even more to \( \Phi_{\text{bkg}} = 5e^{-4} \), with a proportional decrease in signal flux. Therefore, extreme attenuation incurs a higher error.

Also note that while the optimal attenuation and extreme attenuation curves are monotonic in \( \Phi_{\text{bkg}} \) and \( \Phi_{\text{sig}} \), the error surface with no attenuation has a ridge near the high \( \Phi_{\text{sig}} \) values. This is an artifact of the Coates’s estimator which we discuss in the next section.

\[9\text{Note that the maximum error of the Coates’s estimator is equal to that of a random estimator, which will have an error of 30% using the error metric defined earlier.}\]
S. 4.2. Explanation of anomalous second dip in error curves

Here we provide an explanation for an anomaly in the single-pixel error curves that is visible in both simulations and experimental results. When $\Phi_{\text{sig}}$ is high, increasing $\Upsilon$ beyond optimal has two effects: the Coates’s estimate of the true depth bin becomes higher (due to increasing effective $\Phi_{\text{sig}}$), and the Coates’s estimates of the later bins become noisier (due to pile-up). As $\Upsilon$ increases, the pile-up due to both $\Phi_{\text{bkg}}$ and $\Phi_{\text{sig}}$ increases up to a point where all photons are recorded at or before the true depth bins. Beyond this high flux level, the Coates’s estimates for all later bins become indeterminate ($N_i = D_i = 0$), and the Coates’s depth estimate corresponding to the location of the highest ratio is undefined.

This shortcoming of the Coates’s estimator can be fixed ignoring these bins when computing the depth estimate. However, since these later bins do not correspond to the true depth bins, the error goes down. As $\Upsilon$ is increased further, the pile-up due to ambient flux increases and starts affecting the estimates of earlier bins too, including the true depth bin. The number of bins with non-zero estimates keeps decreasing and the error approaches that of a random estimator.

Note that other estimators like MAP and Bayes do not suffer from the degeneracy of Coates’s estimator since they do not rely on intensity estimates, and should have U-shaped error curves.

S. 4.3. Visualization of depth estimation results using 3D mesh reconstructions

Supplementary Figure 2. 3D mesh reconstructions for a castle scene. (Top row) The raw point clouds obtained by pixel-wise depth estimation using the MAP estimator. The haze indicates points with noisy depth estimates. (Bottom row and inset) The reconstructed surfaces obtained after outlier removal, using ground truth triangulation. With insufficient attenuation, only the points that are nearby are estimated correctly, and far away points are totally corrupted. With extreme attenuation, points at all depths are corrupted uniformly. With optimal attenuation, most points are estimated correctly, with large depths incurring slightly more noise due to residual pile-up.

Supplementary Fig. 2 shows 3D mesh reconstructions for a “castle” scene. For each vertex in the mesh, the true depth was used to simulate a single SPAD measurement (500 cycles), which was then used to compute the MAP depth estimate. These
formed the raw point cloud. The mesh triangulation was done after an outlier removal step. These reconstructions show that nature reconstruction errors is like salt-and-pepper noise, unlike the Gaussian errors typically seen in other depth imaging methods such as continuous-wave time-of-flight. Also, it can be seen that as depth increases, so does the noise (number of outliers). This is because the pile-up effect increases along depth exponentially. All this suggests that ordinary denoising methods won’t be effective here, and more sophisticated procedures are needed.

### S. 4.4. Improvements from modeling laser pulse shape and SPAD jitter

Supplementary Figure 3. **Effect of modeling laser pulse shape and SPAD jitter, with and without optimal attenuation.** This figure compares Coates’s estimator and Heide *et al.*’s method for the baseline extreme attenuation and the proposed optimal attenuation, under three levels of ambient light. When the depth errors using Coates’s estimator are already low (red pixels in the error maps), Heide *et al.*’s method further reduces error to achieve sub centimeter accuracy (dark red or black pixels). However, for pixels with large errors (white pixels with error > 10 cm), Heide *et al.*’s method provides no improvement. The overall RMSE, being dominated by large errors, remains the same. On the other hand, going from extreme attenuation to optimal attenuation reduces depth errors (both visually and in terms of RMSE) for both estimators.

The depth estimate obtained using the Coates’s method (Eq. (6)) makes the simplifying assumption that the laser pulse is a perfect Dirac impulse that spans only one histogram bin, even though our simulation model and experiments use a non-impulse laser pulse shape. In recent work, Heide *et al.* [14] propose a computational method for pile-up mitigation which includes explicitly modeling laser pulse shape non-idealities to improve depth precision. Suppl. Fig. 3 shows simulated depth map reconstructions using the Coates’s estimator and compares them with results obtained using the point-wise depth estimator of Heide *et al.* for a range of ambient illumination levels. Observe that at low ambient light levels, pixels with low error values with the Coates’s estimator appear to be slightly improved in the depth error maps using the algorithm of [14]. The method, however, does not improve the overall RMSE value which is dominated by pixels with very high errors that stay unchanged. At high ambient flux levels, pile-up distortion becomes the main source of depth error and optimal attenuation becomes necessary to obtain good depth error performance with any depth estimation algorithm. The results using the total-variation based spatial regularization reconstruction of [14] did not provide further improvements and are not
shown here. In the next section, we show the effect of using DNN based methods that use spatial information on the depth estimation performance for the same simulation scenarios as Suppl. Fig. 3.

**S. 4.5. Combining attenuation with neural networks-based depth estimation methods**

![Effect of attenuation on neural-network based estimator.](image)

Supplementary Figure 4. This figure is an extension of Fig. 6 from the main text, with three levels of ambient light. Even when ambient light is low, optimal attenuation leads to an improvement in RMSE compared to extreme and no attenuation.

In this section, we provide additional simulation results validating the improvements obtained from optimal attenuation when used in conjunction with other state-of-the-art depth reconstruction algorithms. In addition to neural network based methods, we implemented the method from a paper by Rapp et al. [30] which exploits spatio-temporal correlations to censor background photons. Supplementary Fig. 4 is an extended version of the Fig. 6 in the main text and shows reconstruction results for three different ambient light levels.

Suppl. Fig. 5 shows the estimated depth maps and errors obtained using the method from [30] on simulated SPAD measurement data, for different attenuation and ambient flux levels. These results are similar to the neural network reconstructions. For high to moderate ambient flux levels, the depth estimates appear too noisy to be useful if no attenuation is used. With extreme attenuation the errors are lower but degrade when the ambient flux is high. Optimal attenuation provides the lowest RMSE at all ambient flux levels.

For the optimal attenuation results shown here, a single attenuation level was used for the entire scene. The average ambient flux for the whole scene was used to estimate $Y_{\text{opt}}$. This shows that as long as there are not too many flux variations in the scene, using a single attenuation level is sufficient to get good performance. For challenging scenes with large albedo or lighting variations, a single level may not be sufficient and it may become necessary to use a patch-based or pixel-based adaptive attenuation. This strategy is discussed in the next section.

**S. 5. Ambient-adaptive attenuation**

This section describes an algorithm for implementing the idea of optimal attenuation in practice. The only variable in the expression for optimal $Y$ is the background flux $\Phi_{bkg}$, which can be estimated separately, prior to beginning the depth measurements. For estimating $\Phi_{bkg}$, the laser is turned off, and $N'$ SPAD cycles are acquired. Since the background flux $\Phi_{bkg}$ is assumed to be constant, there is only one unknown parameter, and it can be estimated from the acquired histogram $(N_1', N_2', ..., N_B')$ using the MLE (Step 3 of Algorithm 1). Moreover, as mentioned in the main text, our method is quite
Supplementary Figure 5. **Effect of attenuation on Rapp and Goyal's method** [30]. The results follow the same trend as for the neural network.

robust to the choice of $\Upsilon$, which means that our estimate of $\Phi_{bkg}$ does not need to be very accurate. Therefore, we can set $N'$ to be as low as 20–30 cycles, which causes negligible increase in acquisition time.

**Algorithm 1** Adaptive ND-attenuation

1. Focus the laser source and SPAD detector at a given scene point.
2. With the laser power set to zero, acquire a histogram of photon counts $(N'_1, N'_2, \ldots, N'_{B+1})$ over $N'$ laser cycles.
3. Estimate the background flux level using:

   \[
   \tilde{\Phi}_{bkg} = \ln \left( \frac{\sum_{i=1}^B iN'_i + BN'_{B+1}}{\sum_{i=1}^{B+1} iN'_i - N'} \right).
   \]

4. Set the ND-attenuation fraction to $1/\tilde{\Phi}_{bkg}$.
5. Set the laser power to the maximum available level and acquire a histogram of photon counts $(N_1, N_2, \ldots, N_{B+1})$ over $N$ laser cycles.
6. Estimate the photon flux waveform using the Coates’s correction Equation (6), and scene depth using Equation (5).
7. Repeat for all scene points.

S. 6. Dependence of reconstruction errors on true depth value

In this section, we study the effect of true depth on depth estimation errors. Due to the non-linear nature of the image formation model, as well as the non-linear estimators used to rectify pile-up, the estimation error shows some non-linear variations as a function of the true depth.

Suppl. Fig. 6 compares depth error curves across various attenuation levels, for a several signal values. The first observation is that optimal attenuation has a lower error, on average, than extreme and no attenuation. The error curve for optimal attenuation lies below the other curves for most values of true depth (except for very low signal). Therefore, not only does optimal attenuation minimize the average error, it makes the error curve more uniform across all values of the true depth.
Supplementary Figure 6. **Effect of attenuation on error vs true depth curve.** For extreme, insufficient and no attenuation, the error curves are not only high on average, but also highly non-uniform (either decreasing or increasing with depth). In contrast, the optimal attenuation curve is both low on average, and relatively uniform across depth.
S. 7. Additional Experimental Results

Supplementary Figure 7. **Depth estimation with different attenuation factors.** (Top row) Depth maps for a staircase scene, with a brightly lit right half, and shadow on the left half. With no attenuation, the right half is completely corrupted with noise due to strong pile-up. (Bottom row) A challenging tabletop scene with large albedo and depth variations. The optimal attenuation method still gives a reasonably good reconstruction, and is significantly better than either no attenuation or extreme attenuation.
Supplementary Figure 8. **Reconstructing extremely dark objects.** Our method works for scenes with a large dynamic range of flux conditions, like this scene with an extremely dark black vase placed next to a white vase. $\Phi_{\text{sig}}$ was 10 times higher for the white vase. This scan was done with negligible ambient light ($< 10$ lux).