Supplementary Document for "Photon-Flooded Single-Photon 3D Cameras"

Anant Gupta, Atul Ingle, Andreas Velten, Mohit Gupta.

Correspondence to: agupta225@wisc.edu

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.

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

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:

$$\begin{aligned} L(q_1, q_2, ..., q_B) &= \mathbb{P}(N_1, N_2, ..., N_B | q_1, q_2, ..., q_B) \\ &= \mathbb{P}(N_1|q_1) \prod_{i=2}^{B} \mathbb{P}(N_i | q_i, N_1, N_2, ..., N_{i-1}) \\ &= \mathbb{P}(N_1|q_1, D_1) \prod_{i=2}^{B} \mathbb{P}(N_i | q_i, D_i). \end{aligned}$$

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:

$$\begin{aligned} \widehat{q_i} &= \arg\max_{q_i} \ \mathbb{P}(N_i | q_i, D_i) \\ &= \arg\max_{q_i} \ \binom{D_i}{N_i} q_i^{N_i} (1-q_i)^{D_i - N_i} \\ &= \frac{N_i}{D_i} = \frac{N_i}{N - \sum_{j=1}^{i-1} N_j}. \end{aligned} \tag{S2}$$

S. 1.2. Calculating the bias of Coates's corrected estimates

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

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

By the law of iterated expectations:

$$\mathbb{E}[q_i] = \mathbb{E}\left[\mathbb{E}\left[\frac{N_i}{D_i}\middle| N_1, N_2, ..., N_{i-1}\right]\right] \tag{S3}$$

$$= \mathbb{E}\left[\frac{q_i D_i}{D_i}\right] = q_i \tag{S4}$$

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

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

S. 2. Derivation of the Optimal Attenuation Factor Υ^{opt}

In this section, we derive the expression for optimal attenuation factor Υ^{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 Υ. The optimal Υ then follows by minimizing the upper bound.

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

$$r_i = \Phi_{\text{sig}}\delta_{i,r} + \Phi_{\text{bkg}}.$$

Furthermore, we assume that r_i is small enough so that $q_i = 1 - e^{-r_i} \approx r_i$.[8]

S. 2.1. Variance of Coates's estimates

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

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

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

$$\widehat{\tau} = \arg\max_i \widehat{q}_i$$

Note that locating the peak in the waveform is equivalent to locating the maximum q_i. From the previous section, we know that $\mathbb{E}[\widehat{q}_i] = q_i$. Intuitively, this means that the estimates of q_i are correct on average, and we can pick the maximum $\widehat{q}_i$ 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 σ_i^2 denote the diagonal terms and $\sigma_{i,j}^2$ denote the off-diagonal terms of the covariance matrix of $(\widehat{q}_1, \widehat{q}_2, \ldots, \widehat{q}_B)$. We have:

$$\sigma_i^2 = \mathbb{E}[(\widehat{q}_i - q_i)^2]$$

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

$$= \mathbb{E}\left[\mathbb{E}\left[\left(\frac{N_i}{D_i} - q_i\right)^2\middle| D_i\right]\right] \tag{S5}$$

$$= \mathbb{E}\left[\frac{q_i(1-q_i)}{D_i}\right] \tag{S6}$$

[8] Note that this assumption is different from low flux assumption used in the linear operation regime which requires even lower flux levels satisfying $r_i \ll 1/B$.

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,

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

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 $\mathbb{E}[1/D_i+1]$ explicitly. Recalling the definition of D_i and using the mean of the multinomial distribution, we have:

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

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:

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

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:

$$\begin{aligned}
\sigma_{i,j}^2 &= \mathbb{E}\left[(\widehat{q}_i - q_i)(\widehat{q}_j - q_j)\right] \\
&= \mathbb{E}_{N_1,N_2,\ldots,N_i}\left[(\widehat{q}_i - q_i)\mathbb{E}_{N_{i+1},\ldots,N_B|N_1,\ldots N_i}(\widehat{q}_j - q_j)\right] \\
&= \mathbb{E}_{N_1,N_2,\ldots,N_i}\left[(\widehat{q}_i - q_i)\mathbb{E}_{N_j,D_j|N_1,\ldots N_i}(\widehat{q}_j - q_j)\right] &\text{(S8)} \\
&= \mathbb{E}_{N_1,N_2,\ldots,N_i}\left[(\widehat{q}_i - q_i)\mathbb{E}_{D_j|N_1,\ldots N_i}\mathbb{E}_{N_j|D_j}(N_j/D_j - q_j)\right] = 0 &\text{(S9)}
\end{aligned}$$

where Eq. (S8) uses the fact that $\widehat{q}_j = N_j/D_j$ 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 $\widehat{q}_i$ and $\widehat{q}_j$ 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 τ, we want to minimize the probability of error $\mathbb{P}(\widehat{\tau}_{\text{Coates}} \neq \tau)$.

$$\begin{aligned}
\mathbb{P}(\widehat{\tau}_{\text{Coates}} \neq \tau) &= \mathbb{P}\left(\bigcup_{i\neq\tau}(\widehat{q}_i > \widehat{q}_\tau)\right) \\
&\leq \sum_{i\neq\tau}\mathbb{P}\left(\widehat{q}_i > \widehat{q}_\tau\right) \\
&= \sum_{i\neq\tau}\mathbb{P}\left(\widehat{q}_i - \widehat{q}_\tau > 0\right).
\end{aligned}$$

Note that $\widehat{q}_i - \widehat{q}_\tau$ 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:

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

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

$$\begin{aligned}
\mathbb{P}(\widehat{q}_i > \widehat{q}_\tau) &\leq \frac{1}{2}\exp\left(-\frac{(q_i - q_\tau)^2}{2(\sigma_i^2 + \sigma_\tau^2)}\right) \\
&\approx \frac{1}{2}\exp\left(-\frac{N(r_i - r_\tau)^2}{2(\frac{r_i r}{C_i} + \frac{r_\tau r}{C_\tau})}\right) \\
&= \frac{1}{2}\exp\left(-\frac{N(\frac{r_i}{r} - \frac{r_\tau}{r})^2}{2(\frac{r_i}{rC_i} + \frac{r_\tau}{rC_\tau})}\right) \\
&= \frac{1}{2}\exp\left(-\frac{N\left(\frac{\Phi_{\text{sig}}}{B\Phi_{\text{bkg}}+\Phi_{\text{sig}}}\right)^2}{2\left(\frac{\Phi_{\text{bkg}}}{(B\Phi_{\text{bkg}}+\Phi_{\text{sig}})C_i} + \frac{\Phi_{\text{bkg}}+\Phi_{\text{sig}}}{(B\Phi_{\text{bkg}}+\Phi_{\text{sig}})C_\tau}\right)}\right)
\end{aligned}$$

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:

$$\mathbb{P}(\widehat{q}_i > \widehat{q}_\tau) \leq \frac{1}{2} \exp\left(-\frac{\frac{N}{B}\theta^2}{2\left(\frac{1}{C_i} + \frac{1+\theta}{C_j}\right)} \right)$$

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

$$\begin{aligned}\frac{1}{B}\sum_{\tau=1}^{B} \mathbb{P}(\widehat{\tau}_{\text{Coates}} \neq \tau) &\leq \frac{1}{B}\sum_{\tau=1}^{B}\sum_{i\neq\tau} \frac{1}{2} \exp\left(-\frac{\frac{N}{B}\theta^2}{2\left(\frac{1}{C_i} + \frac{1+\theta}{C_\tau}\right)} \right) \\ &\approx \frac{1}{B}\sum_{\tau=1}^{B}\sum_{i=1}^{B} \frac{1}{2} \exp\left(-\frac{\frac{N}{B}\theta^2}{2\left(\frac{1}{C_i} + \frac{1+\theta}{C_\tau}\right)} \right)\end{aligned}$$

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:

$$\begin{aligned}\Upsilon^{\text{opt}} &= \arg\min_{\Upsilon} \frac{1}{B}\sum_{i,\tau=1}^{B} \frac{1}{2} \exp\left(-\frac{\frac{N}{B}\theta^2}{2\left(\frac{1}{C_i} + \frac{1+\theta}{C_\tau}\right)} \right) \\ &\approx \arg\min_{\Upsilon} \max_{i,\tau} \left[\frac{1}{2} \exp\left(-\frac{\frac{N}{B}\theta^2}{2\left(\frac{1}{C_i} + \frac{1+\theta}{C_\tau}\right)} \right) \right] \\ &= \arg\max_{\Upsilon} \min_{i} C_i\end{aligned}$$

The last step is true since the term inside the exponent is maximized for $i = \tau = \arg\min_i C_i(r)$. Furthermore, the expression depends inversely on C_i and C_τ, and all other quantities (N, B, θ) are independent of Υ. 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:

$$\begin{aligned}&\min_i C_i(\Upsilon) = C_B(\Upsilon) = B\left(1 - e^{-\Upsilon\,\Phi_{\text{bkg}}}\right) e^{-(B-1)\Upsilon\,\Phi_{\text{bkg}}} \\ &= \left[1 - e^{-\Upsilon B\Phi_{\text{bkg}}}\right] \frac{1}{\frac{1}{B(1-e^{-\Upsilon\Phi_{\text{bkg}}})e^{(-\Upsilon B\Phi_{\text{bkg}})}} - \frac{1}{B(1-e^{-\Upsilon\Phi_{\text{bkg}}})}} \\ &= \underbrace{\frac{1}{B}\sum_{i=1}^{B} C_i(\Upsilon)}_{\text{Mean receptivity}} \left(\underbrace{\frac{1}{C_B(\Upsilon)} - \frac{1}{C_1(\Upsilon)}}_{\text{Skew}} \right)^{-1}.\end{aligned}$$

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 Υ 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 Φ_{bkg}, source flux Φ_{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 Φ_{bkg} and Φ_{sig}.

MAP and Bayes estimators: In the extreme case, if we assume Φ_{bkg} and Φ_{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:

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

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

$$\begin{aligned}
\mathbb{P}(d|N_1, \ldots, N_B) &= \frac{\mathbb{P}(N_1, \ldots, N_B|d)}{\sum_{i=1}^{n} \mathbb{P}(N_1, \ldots, N_B|i)} \\
&\propto \mathbb{P}(N_1, \ldots, N_B|d) \\
&= \prod_{i=1}^{B} (q_{i|d})^{N_i} (1 - q_{i|d})^{N - \sum_{j=0}^{i-1} N_j} \\
&= \exp\left\{ \sum_{i=1}^{B} N_i \ln\left(q_{i|d}\right) + \sum_{i=1}^{B} D_i \ln\left(1 - q_{i|d}\right) \right\} \qquad \text{(S10)}
\end{aligned}$$

where $q_{i|d}$ denotes the incident photon probability at the i^{th} bin when the true depth is d. Note that $q_{i|d}$ for different depths are related through a rotation of the indices $q_{i|d} = q_{(i-d) \bmod B|0}$. 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 Φ_{bkg} and Φ_{sig}. While Φ_{bkg} can be estimated easily from data, estimating Φ_{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.

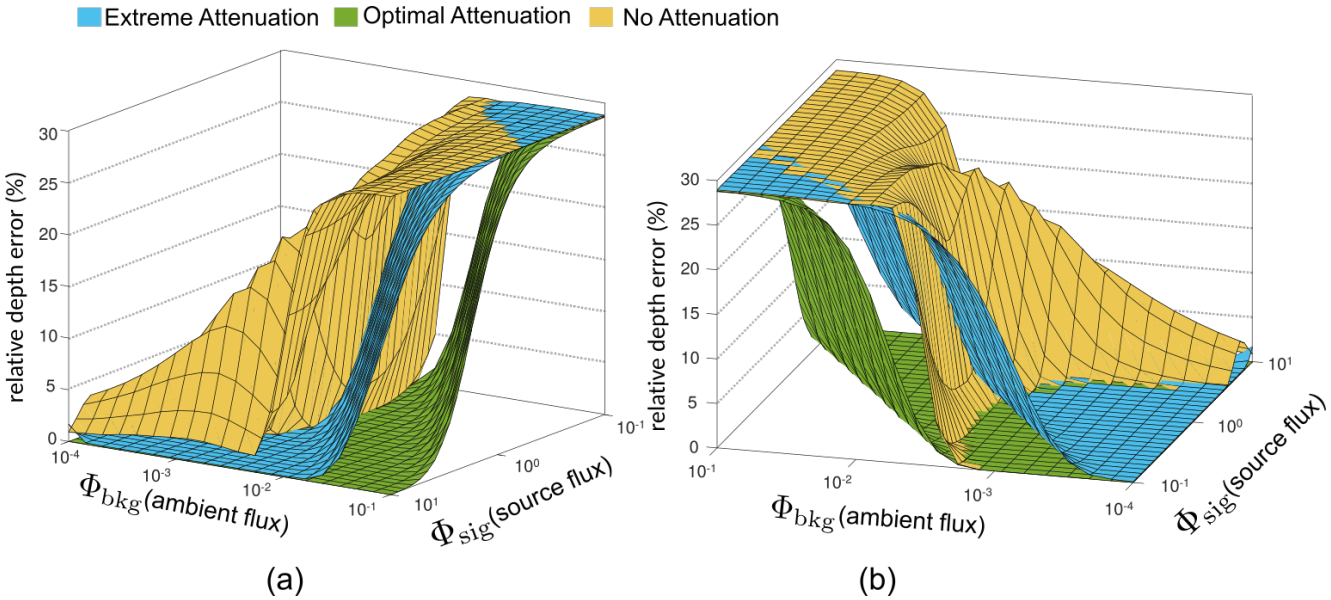
Details of Monte Carlo Simulation: We simulate the first photon measurements using a multinomial distribution as described earlier, for various background and source conditions. The true depth is selected uniformly at random from 1 to B, and the simulations are repeated on an average of 200 times. The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

$$\text{RMSE} = \sqrt{\frac{1}{200} \sum_{i=1}^{200} \left(\left(\widehat{\tau}_i - \tau_i^{\text{true}} + \frac{B}{2} \right) \bmod B - \frac{B}{2} \right)^2}$$

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 τ_i^{true} denotes the true depth on the i^{th} simulation run. It is chosen uniformly randomly from 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 $(-B/2, B/2)$.



Supplementary Figure 1. **Surface plots of relative depth reconstruction errors as a function of ambient and source light levels at different attenuation levels.** These figures show two different views of surface plots of depth reconstruction errors for three different attenuation levels. The optimal attenuation level chosen using the $B\Phi_{\text{bkg}} = 1$ performs better than the state-of-the-art methods that use extreme attenuation.

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 Φ_{sig} and Φ_{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%.[9]

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 Φ_{bkg} of 0.001. For higher Φ_{bkg}, using no attenuation is worse than using extreme attenuation, and the trend is reversed for lower Φ_{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 Φ_{bkg} and Φ_{sig}, the error surface with no attenuation has a ridge near the high Φ_{sig} values. This is an artifact of the Coates's estimator which we discuss in the next section.

[9] 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.

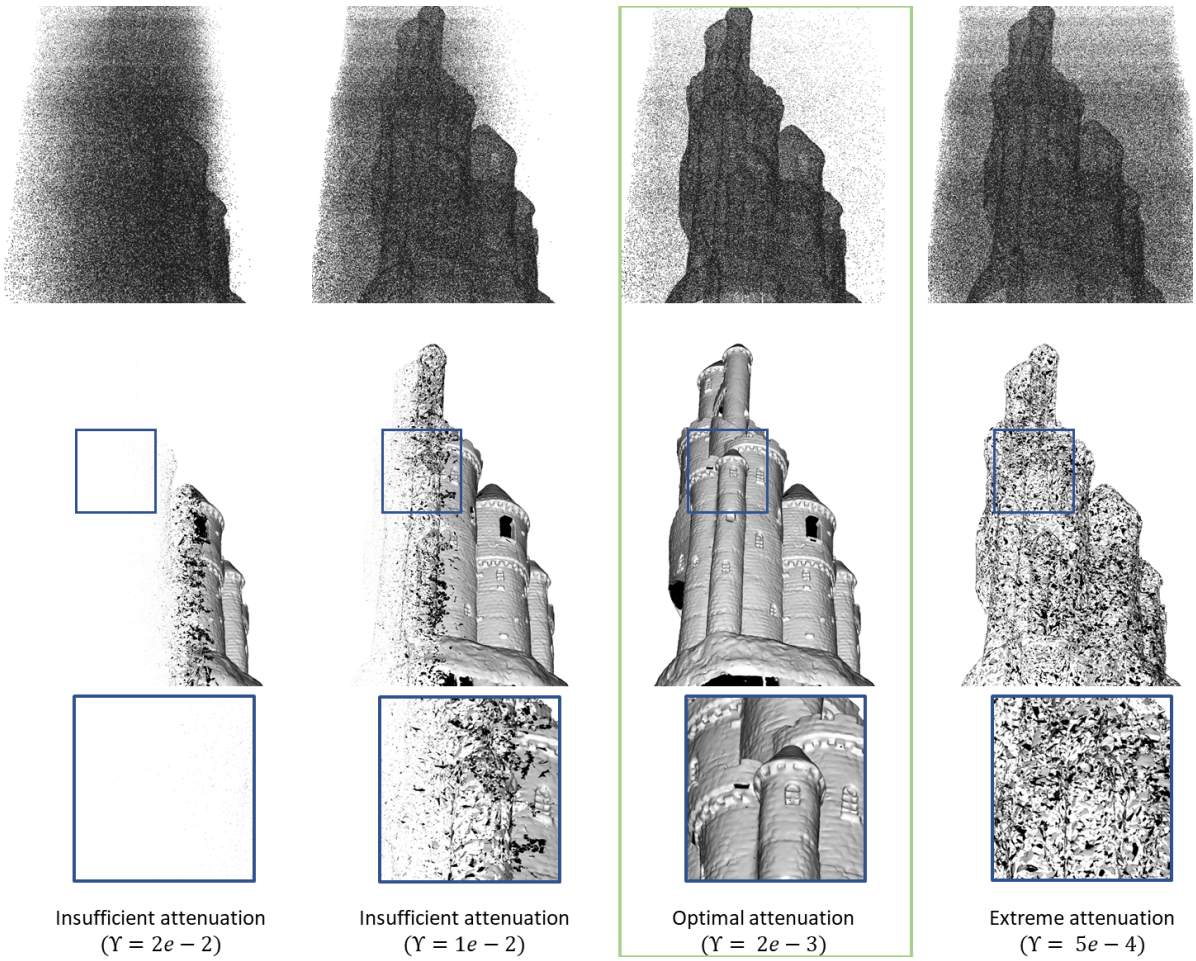
S. 4.2. Explanation of anomalous second dip in error curves

Here we provide an explanation for an anomaly in the single-pixel error curves that is visible in both simulations and experimental results. When Φ_{sig} is high, increasing Υ beyond optimal has two effects: the Coates's estimate of the true depth bin becomes higher (due to increasing effective Φ_{sig}), and the Coates's estimates of the later bins become noisier (due to pile-up). As Υ increases, the pile-up due to both Φ_{bkg} and Φ_{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 Υ 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

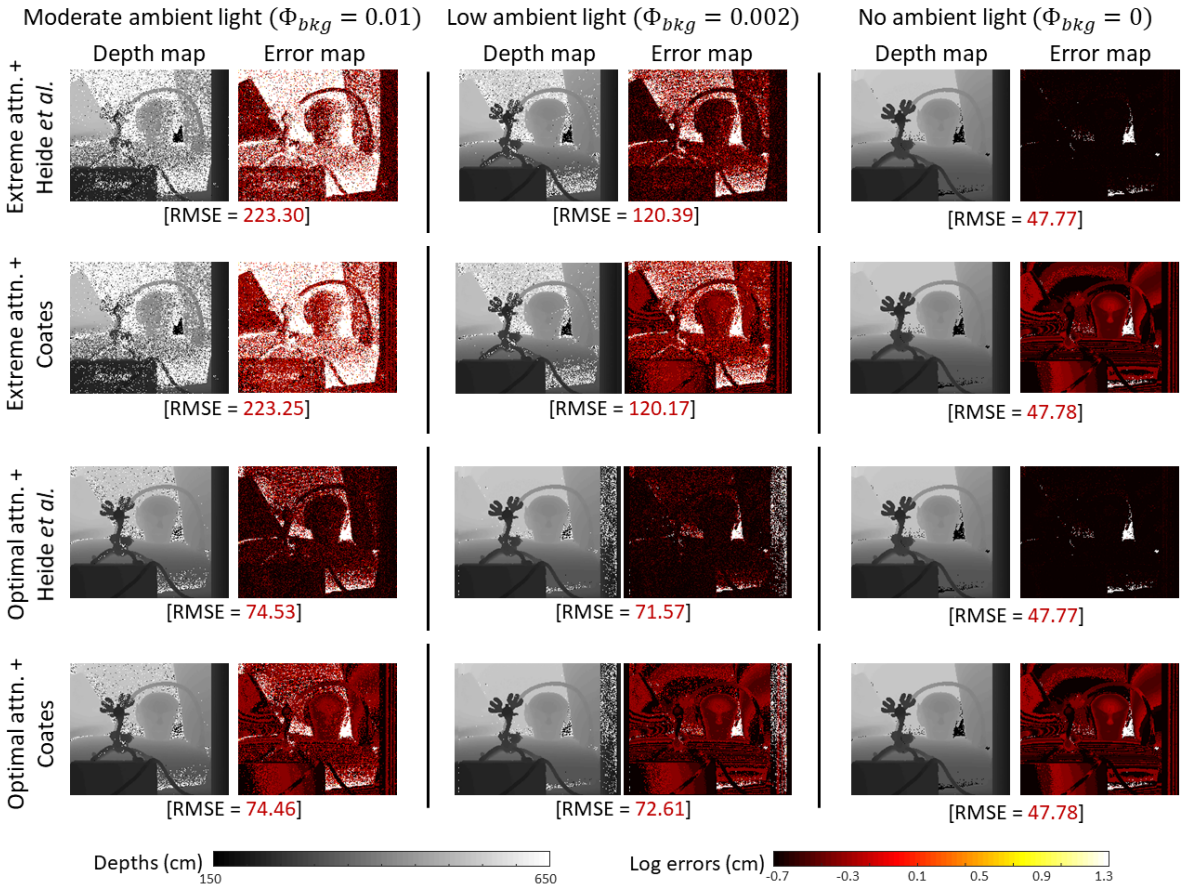


Supplementary Figure 2. **3D mesh reconstructions for a castle scene.** (Top row) The raw point clouds obtained by pixel-wise depth estimation using the MAP estimator. The haze indicates points with noisy depth estimates. (Bottom row and inset) The reconstructed surfaces obtained after outlier removal, 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

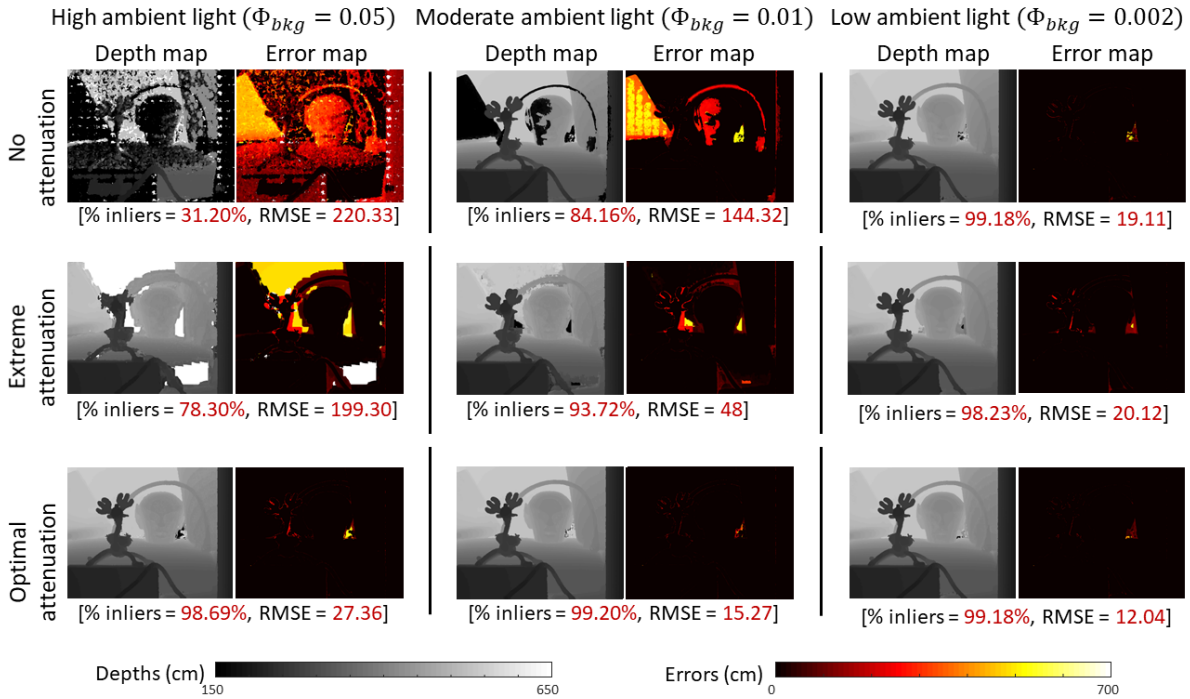


Supplementary Figure 3. **Effect of modeling laser pulse shape and SPAD jitter, with and without optimal attenuation.** This figure compares Coates's estimator and Heide *et al.*'s method for the baseline extreme attenuation and the proposed optimal attenuation, under three levels of ambient light. When 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



Supplementary Figure 4. **Effect of attenuation on neural-network based estimator.** 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.

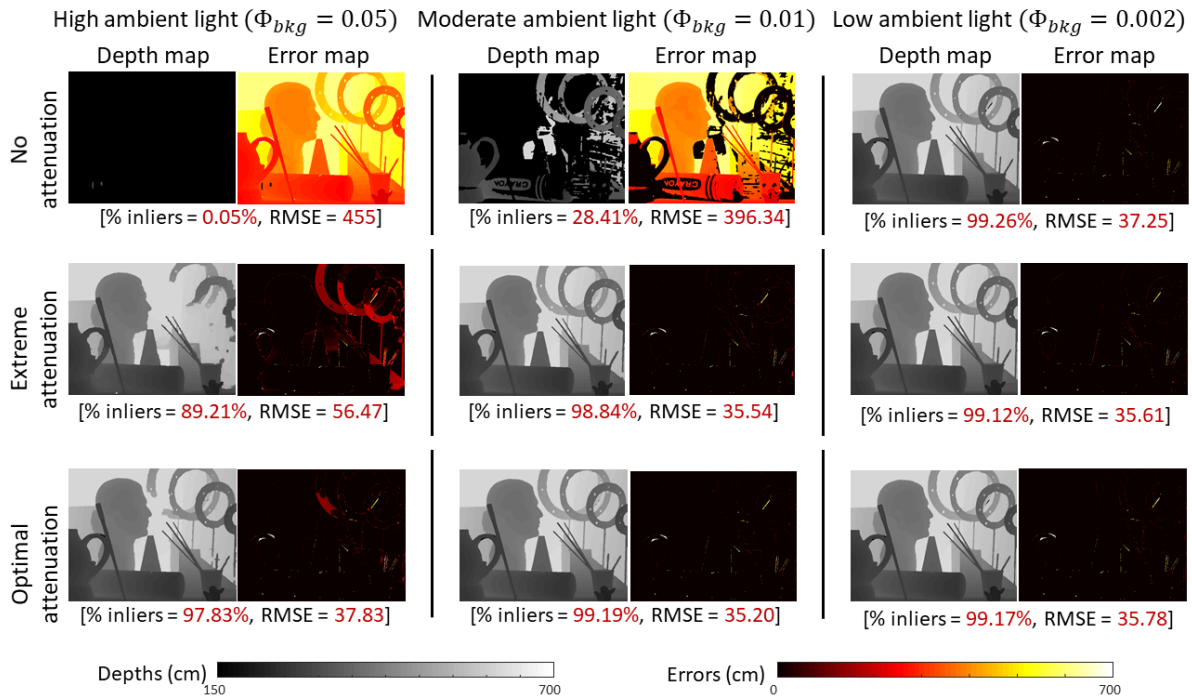
In this section, we provide additional simulation results validating the improvements obtained from optimal attenuation when used in conjunction with other state-of-the-art depth reconstruction algorithms. In addition to neural network based methods, we implemented the method from a paper by Rapp *et 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 Υ^{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 Υ is the background flux Φ_{bkg}, which can be estimated separately, prior to beginning the depth measurements. For estimating Φ_{bkg}, the laser is turned off, and N' SPAD cycles are acquired. Since the background flux Φ_{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 Υ, which means that our estimate of Φ_{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:

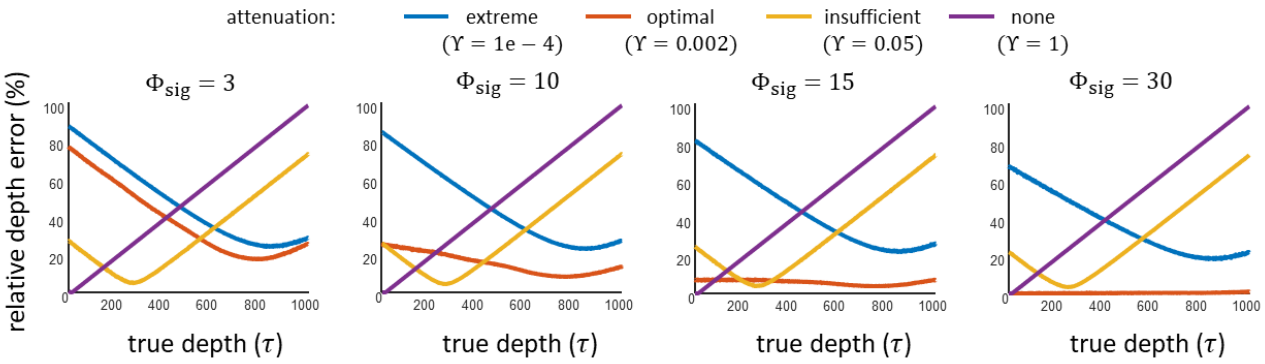
$$\widehat{\Phi}_{\text{bkg}} = \ln\left(\frac{\sum_{i=1}^{B} iN'_i + BN'_{B+1}}{\sum_{i=1}^{B+1} iN'_i - N'}\right).$$

4. Set the ND-attenuation fraction to $1/B\widehat{\Phi}_{\text{bkg}}$.
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

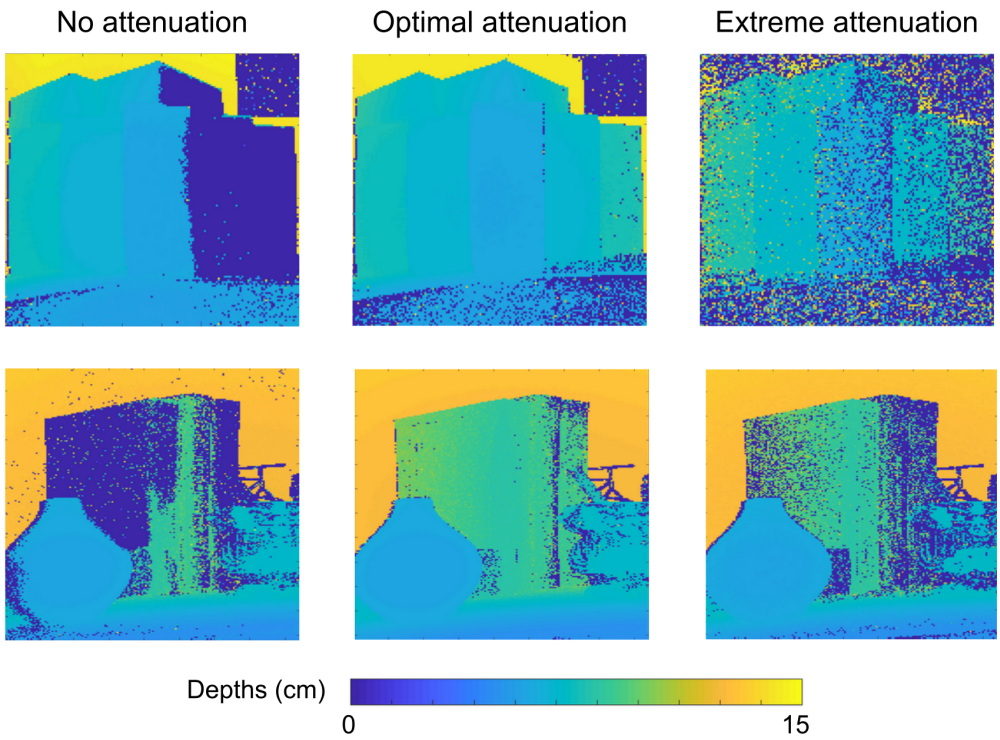
In this section, we study the effect of true depth on depth estimation errors. Due to the non-linear nature of the image formation model, as well as the non-linear estimators used to rectify pile-up, the estimation error shows some non-linear variations as a function of the true depth.

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.

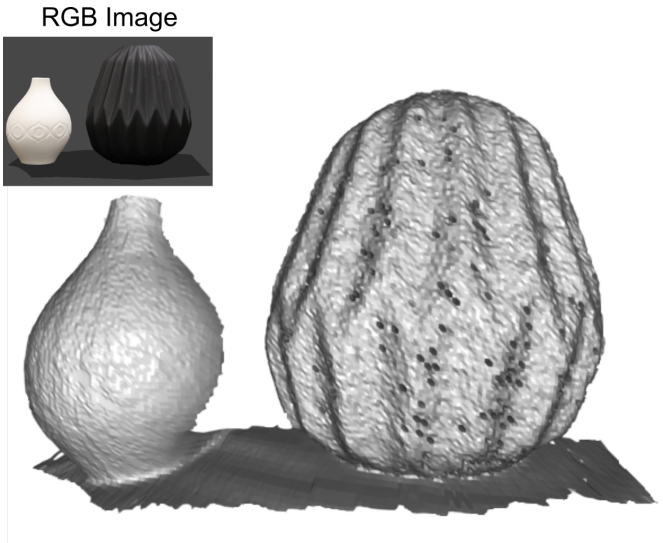


Supplementary Figure 6. **Effect of attenuation on error vs true depth curve.** For extreme, insufficient and no attenuation, the error curves are not only high on average, but also highly non-uniform (either decreasing or increasing with depth). In contrast, the optimal attenuation curve is both low 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. Φ_{sig} was 10 times higher for the white vase. This scan was done with negligible ambient light (< 10 lux).