

Supplementary Materials

CONTENTS

VIII Proofs of Additional Results	1
VIII-A Joint MLE in Eq. (3)	1
VIII-B Proof of Corollary 1	1
VIII-C Reflectivity estimator w/o depth in Eq. (6)	2
VIII-D Proof of Corollary 2	2
VIII-E Reflectivity estimator w/ depth in Eq. (8)	2
VIII-F Proof of Corollary 3	2
VIII-G Lemma 1 and its proof	3
IX Numerical Verification of Information Sharing under MLE	3
IX-A Simulation and experimental setups	3
IX-B Depth estimation and a search algorithm	4
IX-C Reflectivity estimation and a robust algorithm	4
X Simulation Pipeline	5
XI Visualization of Common Features	7
References	8

VIII. PROOFS OF ADDITIONAL RESULTS

We provide detailed derivations of the equations and corollaries in Sec. III.

A. Joint MLE in Eq. (3)

We derive the joint constrained Maximum Likelihood Estimation (MLE) problem.

$$\begin{aligned}
 (\hat{\tau}, \hat{\alpha}) &= \underset{0 < \tau < t_r, \alpha \geq 0}{\operatorname{argmax}} p[\mathbf{t}_M, M = m] \\
 &= \underset{0 < \tau < t_r, \alpha \geq 0}{\operatorname{argmax}} \log p[\mathbf{t}_M, M = m] \\
 &\stackrel{(i)}{=} \underset{0 < \tau < t_r, \alpha \geq 0}{\operatorname{argmax}} \left\{ -N_r(\eta S\alpha + \mathcal{B}) - \underline{\log(m!)} \right. \\
 &\quad \left. + \underline{m \log(N_r)} + \sum_{k=1}^m \log \lambda(t_k) \right\} \\
 &\stackrel{(ii)}{=} \underset{0 < \tau < t_r, \alpha \geq 0}{\operatorname{argmax}} \left\{ -N_r\eta S\alpha \right. \\
 &\quad \left. + \sum_{k=1}^m \log(\eta \alpha s(t_k - \tau) + b_\lambda) \right\},
 \end{aligned}$$

where (i) is based on the substitution of Eq. (2) and (ii) is based on the elimination of independent terms regarding the optimization variables.

B. Proof of Corollary 1

Derivation of Eq. (4):

Assume $b_\lambda = 0$ and τ is a known constant. Then, Eq. (3) becomes

$$\begin{aligned}
 \hat{\alpha} &= \underset{\alpha \geq 0}{\operatorname{argmax}} \left\{ -N_r\eta S\alpha + \sum_{k=1}^m \log(\eta \alpha s(t_k - \tau)) \right\} \\
 &\stackrel{(i)}{=} \underset{\alpha \geq 0}{\operatorname{argmax}} \left\{ m \log \alpha + \sum_{k=1}^m \log(\eta s(t_k - \tau)) - N_r\eta S\alpha \right\} \\
 &\stackrel{(ii)}{=} \underset{\alpha \geq 0}{\operatorname{argmax}} \underbrace{\left\{ m \log \alpha - N_r\eta S\alpha \right\}}_{L(\alpha)},
 \end{aligned}$$

where (i) is based on the logarithmic product rule and (ii) is based on the elimination of terms independent of α .

To solve $\hat{\alpha}$, we take the derivative of $L(\alpha)$ with respect to α and equate it with zero. This will give us

$$\frac{dL(\alpha)}{d\alpha} = \frac{m}{\alpha} - N_r\eta S = 0,$$

and thus $\hat{\alpha} = m/N_r\eta S$. We remark that the constraint $\alpha \geq 0$ is satisfied because all components are nonnegative.

Derivation of Eq. (5):

Assume $b_\lambda = 0$ and α is a known constant. Then, Eq. (3) becomes

$$\begin{aligned}
 \hat{\tau} &= \underset{0 < \tau < t_r}{\operatorname{argmax}} \left\{ -N_r\eta S\alpha + \sum_{k=1}^m \log[\eta \alpha s(t_k - \tau)] \right\} \\
 &\stackrel{(i)}{=} \underset{0 < \tau < t_r}{\operatorname{argmax}} \left\{ \sum_{k=1}^m \log[s(t_k - \tau)] - \underline{N_r\eta S\alpha} + \underline{m \log(\eta \alpha)} \right\} \\
 &\stackrel{(ii)}{=} \underset{0 < \tau < t_r}{\operatorname{argmax}} \underbrace{\left\{ \sum_{k=1}^m \log[s(t_k - \tau)] \right\}}_{L(\tau)},
 \end{aligned}$$

where (i) is based on the logarithmic product rule, (ii) is based on the removal of parameters independent of τ , and $s(\cdot)$ is a Gaussian-shaped function.

To solve $\hat{\tau}$, we take the derivative of $L(\tau)$ with respect to τ and set it to zero. This will give us

$$\frac{dL(\tau)}{d\tau} = \sum_{k=1}^m \frac{s(t_k - \tau) \cdot [(t_k - \tau)/\sigma_t^2]}{s(t_k - \tau)} = 0,$$

which is equivalent to

$$\frac{dL(\tau)}{d\tau} = \sum_{k=1}^m (t_k - \tau) = \sum_{k=1}^m t_k - m\tau = 0.$$

Therefore, $\hat{\tau} = \frac{1}{m} \sum_{k=1}^m t_k$. We observe that the constraint $0 < \tau < t_r$ is automatically satisfied.

C. Reflectivity estimator w/o depth in Eq. (6)

First, we derive an **unconstrained** MLE from Eq. (20). The log-likelihood is

$$\log p_M(m; \alpha) = -N_r \Lambda(\alpha) + m \log(N_r \Lambda(\alpha)) - \log(m!).$$

To maximize the log-likelihood function, the first order optimality condition implies that

$$\frac{\partial \log p_M(m; \alpha)}{\partial \alpha} = -N_r \eta S + \frac{m \eta S}{\eta S \alpha + B} = 0. \quad (22)$$

Solving the equation above, the **unconstrained** MLE becomes

$$\hat{\alpha}_c^* = \frac{1}{\eta S} \left(\frac{m}{N_r} - B \right).$$

Accordingly, the **constrained** MLE for the reflectivity without the help of depth is

$$\hat{\alpha}_c = \max \{ \hat{\alpha}_c^*, 0 \}.$$

D. Proof of Corollary 2

Based on Eq. (22), the second order derivative is

$$\begin{aligned} \frac{\partial^2 \log p_M(m; \alpha)}{\partial \alpha^2} &= \frac{\partial}{\partial \alpha} \left(-N_r \eta S + \frac{m \eta S}{\eta S \alpha + B} \right) \\ &= -\frac{m \eta^2 S^2}{(\eta S \alpha + B)^2}. \end{aligned}$$

The Fisher information is

$$\begin{aligned} I(\alpha) &\stackrel{\text{def}}{=} -\mathbb{E} \left[\frac{\partial^2 \log p_M(m; \alpha)}{\partial \alpha^2} \right] \\ &\stackrel{(i)}{=} \frac{\eta^2 S^2}{(\eta S \alpha + B)^2} \mathbb{E}[m] \\ &\stackrel{(ii)}{=} \frac{\eta^2 S^2}{(\eta S \alpha + B)^2} N_r (\eta S \alpha + B) \\ &= \frac{N_r \eta^2 S^2}{\eta S \alpha + B}, \end{aligned}$$

where (i) is based on the fact that deterministic parameters can be moved out of the expectation, and (ii) follows from the mean of the Poisson random variable m being the total energy over N_r cycles.

Therefore, the Cramer-Rao Lower Bound (CRLB) is

$$\text{Var}[\hat{\alpha}_c^*] \geq \frac{1}{I(\alpha)} = \frac{\eta S \alpha + B}{N_r \eta^2 S^2} = \frac{1 + 1/\text{SBR}}{N_r (\eta S / \alpha)},$$

where signal-to-background ratio (SBR) is defined as $\text{SBR} = S/B$.

To the best of our knowledge, although this result matches the CRLB of a Binomial photon count estimator reported in [1], the CRLB derivation for this Poisson photon count reflectivity estimator is new.

E. Reflectivity estimator w/ depth in Eq. (8)

Assume that τ is a known constant. Then, Eq. (3) becomes

$$\hat{\alpha} = \underset{\alpha \geq 0}{\operatorname{argmax}} \underbrace{\left\{ -N_r \eta S \alpha + \sum_{k=1}^m \log(\eta \alpha s(t_k - \tau) + b_\lambda) \right\}}_{L_t(\alpha)}.$$

To solve the **unconstrained** MLE $\hat{\alpha}_t^*$, the first order optimality condition implies that

$$\frac{d L_t(\alpha)}{d \alpha} = -N_r \eta S + \sum_{k=1}^m \frac{\eta s(t_k - \tau)}{\eta \alpha s(t_k - \tau) + b_\lambda} = 0. \quad (23)$$

Rearranging the terms, we obtain Eq. (8) as

$$\sum_{k=1}^m \frac{\eta s(t_k - \tau)}{\eta \alpha s(t_k - \tau) + B/t_r} = N_r \eta S.$$

The result is consistent with [2]. We notice that the left-hand side of Eq. (8) is monotonically decreasing for $0 \leq \alpha < \infty$. A robust estimation algorithm for α is proposed in Sec. IX-C.

F. Proof of Corollary 3

Similar to Sec. VIII-D, we derive the CRLB for the **unconstrained** MLE $\hat{\alpha}_t^*$.

The second order derivative of the log-likelihood of the joint density is

$$\begin{aligned} &\frac{\partial^2}{\partial \alpha^2} (\log p[\mathbf{t}_M, M = m; \alpha]) \\ &= \frac{d}{d \alpha} \left(\frac{d L_t(\alpha)}{d \alpha} \right) \\ &\stackrel{(i)}{=} \frac{d}{d \alpha} \left(-N_r \eta S + \sum_{k=1}^m \frac{\eta s(t_k - \tau)}{\eta \alpha s(t_k - \tau) + b_\lambda} \right) \\ &= -\sum_{k=1}^m \frac{\eta^2 s^2(t_k - \tau)}{(\eta \alpha s(t_k - \tau) + b_\lambda)^2}, \end{aligned}$$

where (i) is based on the substitution of Eq. (23).

The Fisher information is

$$\begin{aligned} I(\alpha) &= -\mathbb{E}_{M, \tilde{\mathbf{t}}_M} \left[\frac{\partial^2}{\partial \alpha^2} (\log p[\mathbf{t}_M, M = m; \alpha]) \right] \\ &= \mathbb{E}_{M, \tilde{\mathbf{t}}_M} \left[\sum_{k=1}^m \frac{\eta^2 s^2(t_k - \tau)}{(\eta \alpha s(t_k - \tau) + b_\lambda)^2} \right] \\ &\stackrel{(i)}{=} N_r \Lambda \mathbb{E}_{\tilde{t}|M=m} \left[\frac{\eta^2 s^2(t - \tau)}{(\eta \alpha s(t - \tau) + b_\lambda)^2} \middle| M = m \right] \\ &\stackrel{(ii)}{=} N_r \Lambda \int_{-\infty}^{\infty} \frac{\eta^2 s^2(t - \tau)}{(\eta \alpha s(t - \tau) + b_\lambda)^2} f_{\tilde{t}|M=m}(t|m) dt \\ &= N_r \Lambda \int_{-\infty}^{\infty} \frac{\eta^2 s^2(t - \tau)}{(\eta \alpha s(t - \tau) + b_\lambda)^2} \frac{\eta \alpha s(t - \tau) + b_\lambda}{\Lambda} dt \\ &= N_r \eta^2 \int_0^{t_r} \frac{s^2(t - \tau)}{\eta \alpha s(t - \tau) + b_\lambda} dt, \end{aligned}$$

where (i) is due to Lemma 1 in Sec. VIII-G and (ii) is based on the definition of the expectation of a function of the random variable t .

Thus, the CRLB is

$$\text{Var}[\hat{\alpha}_t^*] \geq \frac{1}{I(\alpha)} = \left[N_r \eta^2 \int_0^{t_r} \frac{s^2(t-\tau)}{\eta\alpha s(t-\tau) + b_\lambda} dt \right]^{-1}.$$

Compared to [3], our result provides an explicit formula.

G. Lemma 1 and its proof

Lemma 1. Define $g(t) = \frac{\eta^2 s^2(t-\tau)}{(\eta\alpha s(t-\tau) + b_\lambda)^2}$. Then,

$$\mathbb{E}_{M, \tilde{\mathbf{t}}_M} \left[\sum_{k=1}^m g(t_k) \right] = N_r \Lambda \mathbb{E}_{\tilde{\mathbf{t}}_M | M=m} [g(t) | M=m],$$

where the expectation on the left is taken jointly over M and $\tilde{\mathbf{t}}_M$ while the one on the right is conditional on $M = m$ over one single timestamp random variable $t \in [0, t_r]$.

Proof of Lemma 1.

$$\begin{aligned} & \mathbb{E}_{M, \tilde{\mathbf{t}}_M} \left[\sum_{k=1}^m \frac{\eta^2 s^2(t_k - \tau)}{(\eta\alpha s(t_k - \tau) + b_\lambda)^2} \right] \\ &= \mathbb{E}_{M, \tilde{\mathbf{t}}_M} \left[\sum_{k=1}^m g(t_k) \right] \\ &\stackrel{(i)}{=} \mathbb{E}_M \left[\mathbb{E}_{\tilde{\mathbf{t}}_M | M=m} \left[\sum_{k=1}^m g(t_k) \middle| M=m \right] \right] \\ &\stackrel{(ii)}{=} \mathbb{E}_M \left[\underbrace{\int_0^{t_r} \dots \int_0^{t_r}}_m \sum_{k=1}^m g(t_k) \right. \\ &\quad \left. \prod_{k=1}^m f_{\tilde{t}_k | M=m}(t_k | m) dt_1 \dots dt_m \right] \\ &= \mathbb{E}_M \left[\underbrace{\int_0^{t_r} \dots \int_0^{t_r}}_m \left[g(t_1) + \sum_{k=2}^m g(t_k) \right] \right. \\ &\quad \left. f_{\tilde{t}_1 | M=m}(t_1 | m) dt_1 \dots f_{\tilde{t}_m | M=m}(t_m | m) dt_m \right] \\ &\stackrel{(iii)}{=} \mathbb{E}_M \left[\underbrace{\int_0^{t_r} \dots \int_0^{t_r}}_{m-1} \left[\mathbb{E}[g(t_1)] | M=m + \sum_{k=2}^m g(t_k) \right] \right. \\ &\quad \left. f_{\tilde{t}_2 | M=m}(t_2 | m) dt_2 \dots f_{\tilde{t}_m | M=m}(t_m | m) dt_m \right] \\ &\stackrel{(iv)}{=} \mathbb{E}_M \left[\sum_{k=1}^m \mathbb{E}[g(t_k)] | M=m \right] \\ &\stackrel{(v)}{=} \sum_{m=0}^{\infty} p_M(m) \sum_{k=1}^m \mathbb{E}_{\tilde{\mathbf{t}}_M | M=m} [g(t) | M=m] \\ &\stackrel{(vi)}{=} \mathbb{E}_{\tilde{\mathbf{t}}_M | M=m} [g(t) | M=m] \mathbb{E}_M [M] \\ &= N_r \Lambda \mathbb{E}_{\tilde{\mathbf{t}}_M | M=m} [g(t) | M=m] \\ &= N_r \Lambda \mathbb{E}_{\tilde{\mathbf{t}}_M | M=m} \left[\frac{\eta^2 s^2(t-\tau)}{(\eta\alpha s(t-\tau) + b_\lambda)^2} \middle| M=m \right]. \end{aligned}$$

Here, (i) is based on the law of iterated expectation, and the outer expectation is over M . (ii) is due to the definition of conditional expectation, and we write the joint density as the product of m marginal densities as in Eq. (21) due to conditional independence. In (iii), we calculate the integral over t_1 . In (iv), we calculate the integrals over the other variables and group the results. In (v), we expand the expectation over m and replace the dummy variables t_k ($k = 1, \dots, m$) by t . In (vi), we move the conditional expectation outside the summation. The reason is that the conditional probability density function does not depend on the condition $M = m$ as in Eq. (21).

IX. NUMERICAL VERIFICATION OF INFORMATION SHARING UNDER MLE

In this section, we conduct additional numerical simulations to verify the mutual benefit between depth and reflectivity.

A. Simulation and experimental setups

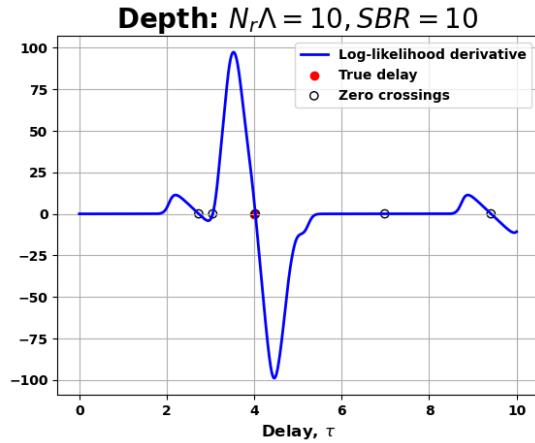
We simulate timestamps for one pixel using the Poisson photon arrival statistics under a low photon level and different SBRs. Due to the randomness, we repeat the experiment multiple times. The parameters used in the simulation are summarized in Tab. V.

TABLE V: Simulation specifications. $\{\cdot\}$ means a list of values used.

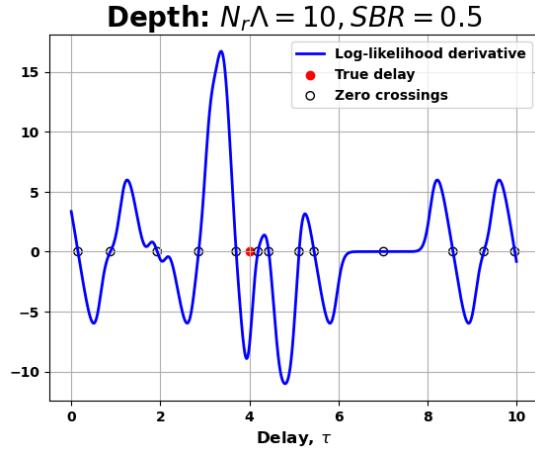
Symbols	Meaning	Values
t_r	repetition period	10
N_r	# repetition	1000
τ	ground truth time delay	4
α	ground truth reflectivity	0.5
t	temporal duration	$0 \leq t < 10$
dt	temporal resolution	$1/1000$
σ_t	pulse width	0.2
$N_r \Lambda$	average photon level	10
SBR	signal-to-background ratio	$\{0.5, 1, 2, 5, 10\}$
iter_num	# repetitive experiments	$\{1000, 50\}$

To verify the information sharing under the Maximum-Likelihood Estimation (MLE) framework, we conduct two parallel experiments to estimate the depth and reflectivity for one pixel from simulated timestamps. One is based on mutual help between depth and reflectivity while the other is not. We choose $\text{iter_num} = 1000$. The MSE results validate the benefit of information sharing, as can be seen in Fig. 2.

In the case where we use a Gaussian distribution without any background noise to estimate depth and a Poisson distribution to estimate reflectivity, both estimation problems have closed-form solutions, as given by Eqs. (5) and (6). However, when $\lambda_b \neq 0$, the two estimators help each other as in Eq. (3), and no closed-form solution is available. In this case, a search-based algorithm is required. We demonstrate the challenge of local minima in the two optimization problems and propose corresponding solutions in the next two subsections.



(a) The number of local minima is sparsely located when the SBR is high



(b) Low SBR incurs much more local minima near the ground truth.

Fig. 15: Visualization of $\frac{d}{d\tau} L_d(\tau)$ for two different SBR values: one high and one low.

B. Depth estimation and a search algorithm

Assume α is a known constant. Then, Eq. (3) becomes

$$\hat{\tau} = \operatorname{argmax}_{0 < \tau < t_r} \underbrace{\left\{ \sum_{k=1}^m \log (\eta \alpha s(t_k - \tau) + b_\lambda) \right\}}_{L_d(\tau)}. \quad (24)$$

To solve the unconstrained MLE $\hat{\tau}_t^*$ embedded in Eq. (24), it is necessary that

$$\frac{dL_d(\tau)}{d\tau} = \sum_{k=1}^m \frac{\eta \alpha \dot{s}(t_k - \tau)}{\eta \alpha s(t_k - \tau) + b_\lambda} = 0.$$

We note that $\frac{d}{d\tau} L_d(\tau)$ is random because it is a function of random variables t_k ($k = 1, \dots, m$). We plot two typical realizations of $\frac{d}{d\tau} L_d(\tau)$ in Fig. 15. Within the constraint interval $0 < \tau < t_r$, the zero crossings are fewer when the SBR is higher as shown in Fig. 15a while they are more when the SBR is lower in Fig. 15b. More severely for the latter case, there are a few zero crossings clustered near the ground truth.

Therefore, an ordinary numerical solver tends to be trapped in local minima.

Our goal is to find the zero crossing closest to the ground truth efficiently and robustly. In the proposed Algorithm 1, we start from a small interval centered at the ground truth and examine the signs of $\frac{d}{d\tau} L_d(\tau)$ at the two boundaries. If they are the same, we gradually increase the interval until the two signs are different. After applying a simple root-finding solver, we find the zero crossing on this interval, and this is the estimated depth. Although the starting point is the ground truth, it is almost impossible for the estimated depth to be the same due to the randomness of $\frac{d}{d\tau} L_d(\tau)$.

Algorithm 1 Robust Depth Estimation Algorithm

Require: Initial guess τ_0 , step size step , maximum iterations N_{\max} , iteration number iter

Ensure: Estimated depth $\hat{\tau}$

- 1: Initialize $\hat{\tau} \leftarrow \tau_0$, $a \leftarrow \tau_0$, $b \leftarrow \tau_0$, $\text{iter} \leftarrow 1$
 - 2: Define $L'_d(\tau)$ \triangleright Derivative of likelihood function
 - 3: **while** $\text{sign}[L'_d(a)] = \text{sign}[L'_d(b)]$ and $\text{iter} < N_{\max}$ **do**
 - 4: $a \leftarrow a - \text{step}$
 - 5: $b \leftarrow b + \text{step}$
 - 6: **end while**
 - 7: **if** $\text{sign}[L'_d(a)] \neq \text{sign}[L'_d(b)]$ **then**
 - 8: Solve $\hat{\tau}$ such that $L'_d(\hat{\tau}) = 0$ using a root-finding method on $[a, b]$
 - 9: **else**
 - 10: **Return Error:** "Bracket not found"
 - 11: **end if**
 - 12: **Return** $\hat{\tau}$
-

Depth estimation results. In addition to Fig. 2 from the main text, we provide the scatter plot of 50 realizations to visualize the estimation performance with and without knowledge of reflectivity, in Fig. 16. According to the figure, with the reflectivity information, the variance of the depth estimation is much smaller, leading to a more accurate estimate.

C. Reflectivity estimation and a robust algorithm

Eq. (23) is required to solve the reflectivity estimation with the help of depth. Likewise, this optimization problem also suffers from local minima, as illustrated in Fig. 17.

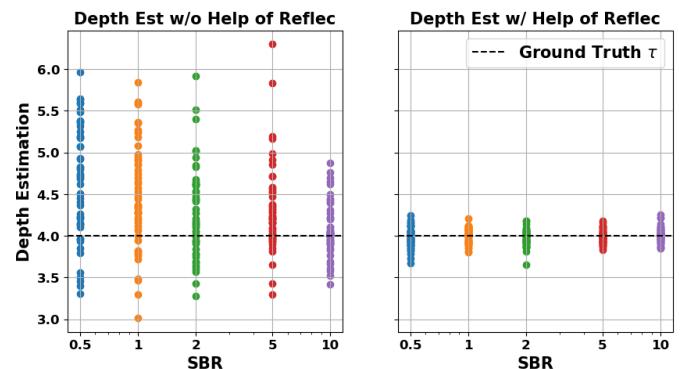
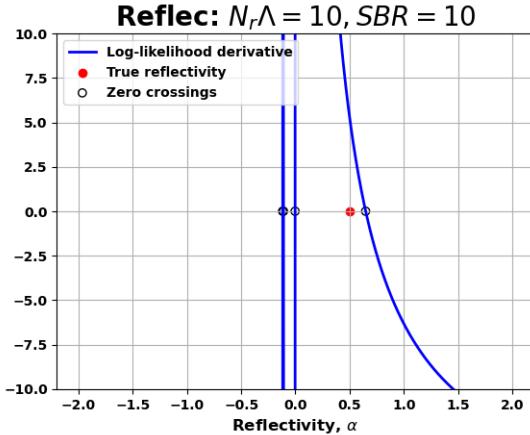
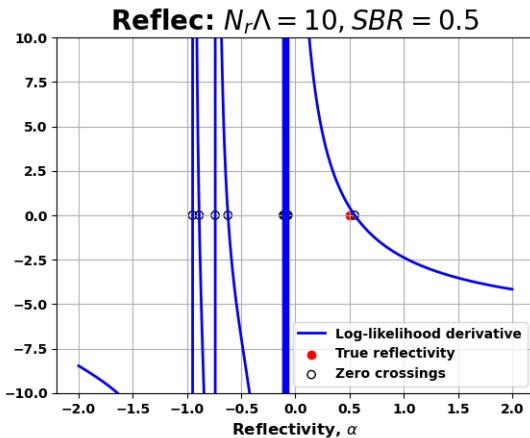


Fig. 16: Depth estimation with and without the help of reflectivity.



(a) The trajectory of the first derivative $\frac{d}{d\alpha} L_t(\alpha)$ for a high SBR value of 10.



(b) The trajectory of the first derivative $\frac{d}{d\alpha} L_t(\alpha)$ for a low SBR value of 0.5.

Fig. 17: Visualization of $\frac{d}{d\alpha} L_t(\alpha)$ for two different SBR values: one high and one low.

The local minimum problem in the reflectivity estimation is more challenging than the depth because the zero crossings in Fig. 17 are even closer to each other compared to those in Fig. 15. However, we recognize that the $\frac{d}{d\alpha} L_t(\alpha)$ is monotonically decreasing within $[0, \infty)$, and the optimization is also constrained in this region, indicating that the negative zeros should be ignored. Inspired by this, we propose an algorithm based on the bisection method for the reflectivity estimation, as illustrated in Algorithm 2.

First, we examine $\frac{d}{d\alpha} L_t(\alpha)|_{\alpha=0}$. If it is not greater than zero, according to the monotonicity of $\frac{d}{d\alpha} L_t(\alpha)$ in $[0, \infty)$, no zero crossing will exist in this region. Therefore, we assign zero to the constrained optimization without any calculations. When $\frac{d}{d\alpha} L_t(\alpha)|_{\alpha=0} > 0$, one unique solution exists in $[0, \infty)$, and we can apply a bisection method to solve it.

Reflectivity estimation results. Using Algorithm 2, we draw the scatter plot to evaluate the reflectivity estimation performance with and without knowledge of depth in Fig. 18. According to the figure, the reflectivity estimation is much better with the help of depth, resulting in low-variance estimates,

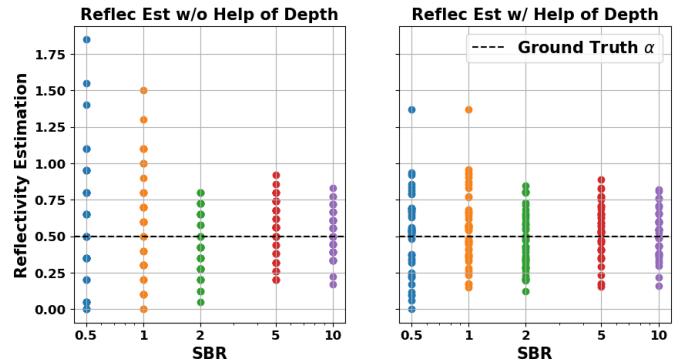


Fig. 18: Reflectivity estimation with and without the help of depth.

Algorithm 2 Robust Reflectivity Estimation Algorithm

Ensure: Estimated reflectivity $\hat{\alpha}$

- 1: Initialize b as a large positive value
 - 2: Define $L'_t(\alpha)$ ▷ Derivative of likelihood function
 - 3: **if** $L'_t(0) \leq 0$ **then**
 - 4: $\hat{\alpha} \leftarrow 0$
 - 5: **else**
 - 6: Solve $\hat{\alpha}$ such that $L'_t(\hat{\alpha}) = 0$ using the bisection method on the interval $(0, b)$
 - 7: **end if**
 - 8: **Return** $\hat{\alpha}$
-

particularly when the SBR is low.

X. SIMULATION PIPELINE

In this section, we provide a detailed description of how the SP-LiDAR simulation is conducted. The arrangement of a typical SP-LiDAR setup is illustrated in Fig. 19. As outlined in the manuscript, we adopt typical SP-LiDAR parameter settings, assuming no multi-path interactions, pulse elongation, crosstalk, or dead time. Under these assumptions, we can presume the independence of the captured photons.

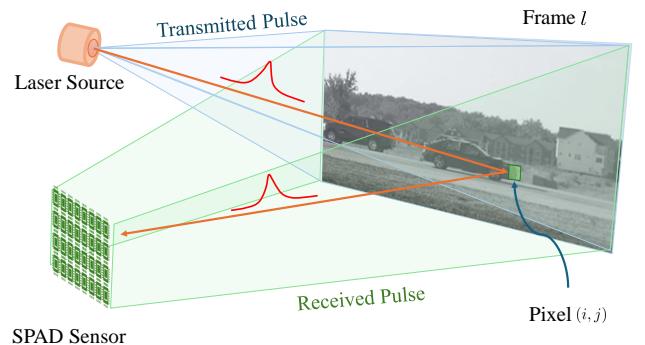


Fig. 19: The single photon LiDAR imaging setup showcases the main components of the arrangement. The laser source emits pulses in rapid succession, whereas the SPAD sensor records the time of arrival of photons.

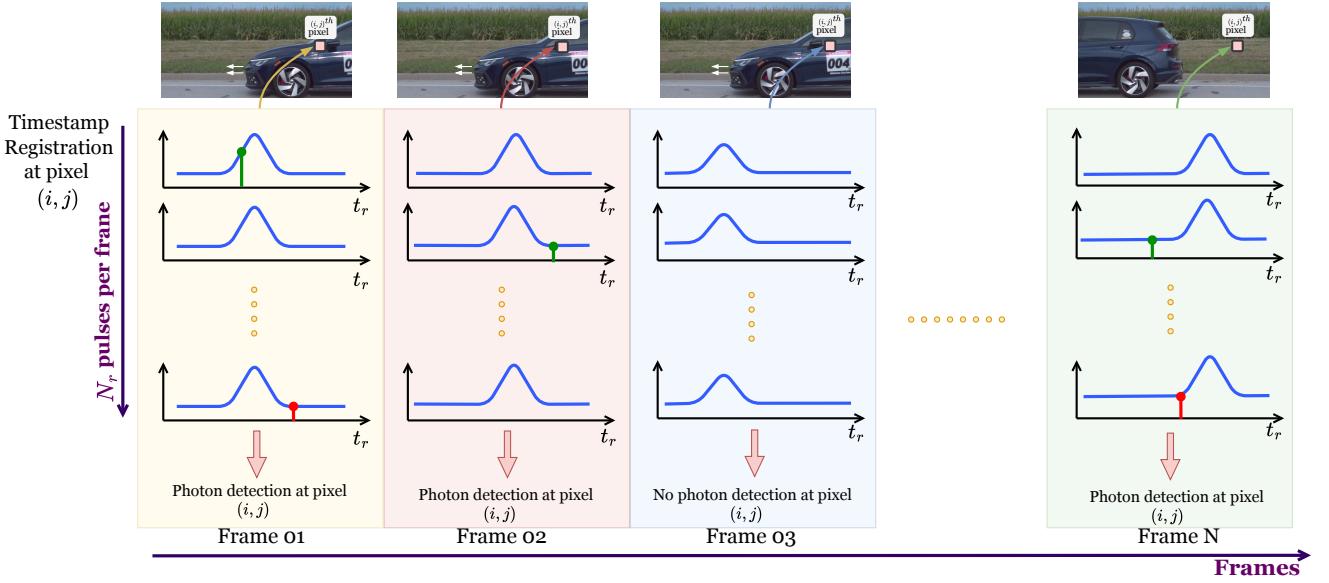


Fig. 20: The timestamp registration of the SPAD sensor captures the behavior of the first photon for the pixel (i, j) . During the exposure time $t_{\text{exp}} = N_r t_r$, only the first photon is detected. Successful registrations are marked in green, while missed photons are marked in red. In some frames, there may be no photon registrations due to the random nature of photon arrivals. Timestamps are recorded for all pixels, resulting in the generation of timestamp frames.

Furthermore, in our simulation, we assume that the scene is quasi-static, meaning it remains stationary during each exposure time. We replicate the operation of the SPAD sensor described in [4], which is capable of registering the first photon during each readout cycle, following the first-photon behavior illustrated in Fig. 20.

The figure illustrates how the first photon arriving at the sensor from the $(i, j)^{\text{th}}$ pixel is captured during each exposure time. Since the laser operates faster than the SPAD sensor's timestamp readout capability, there are N_r laser repetitions within each exposure time. Due to the random nature of photon arrivals, some exposure intervals may result in no detections, while others may register multiple photons. The number of photon detections per exposure time at pixel (i, j) is given by:

$$p_{M_{i,j}}(m_{i,j}) = \frac{e^{-N_r \Lambda_{i,j}} [N_r \Lambda_{i,j}]^{m_{i,j}}}{m_{i,j}!}, \quad (25)$$

where $\Lambda_{i,j} = \eta \alpha_{i,j} S + B_{i,j}$. However, in cases of multiple detections, only the timestamp corresponding to the first photon is recorded—a phenomenon known as first-photon behavior. Note that since the scene is dynamic from frame to frame, the underlying photon arrival distribution at pixel (i, j) varies over time.

Given that at least one photon is registered during each exposure time (i.e., $M > 0$), the timestamp distribution of the first photon detected at each pixel (i, j) can be modeled as a mixture distribution, as follows [1], [5]:

$$p(t_{i,j}) = \frac{\eta \alpha_{i,j} S}{\eta \alpha_{i,j} S + B_{i,j}} \left(\frac{s \left(t_{i,j} - \frac{2z_{i,j}}{c} - t_{\text{jit}} \right)}{S} \right) + \frac{B_{i,j}}{\eta \alpha_{i,j} S + B_{i,j}} \left(\frac{1}{t_r} \right), \quad (26)$$

where we model timing jitter as $t_{\text{jit}} \sim \mathcal{N}(0, \sigma_j)$, a temporal error interval, caused by the temporal response of the SPAD setup. Without loss of generality, we assume that the laser pulse $s(t)$ is a Gaussian pulse $\mathcal{N}(0, \sigma_t)$, yielding $S = \int_{-t_r/2}^{t_r/2} s(t) dt = 1$. Therefore, the distribution can be further simplified to:

$$p(t_{i,j}) = \frac{\eta \alpha_{i,j}}{\eta \alpha_{i,j} + B_{i,j}} \left(s \left(t_{i,j} - \frac{2z_{i,j}}{c} - t_{\text{jit}} \right) \right) + \frac{B_{i,j}}{\eta \alpha_{i,j} + B_{i,j}} \left(\frac{1}{t_r} \right). \quad (27)$$

To calculate the reflectivity of each pixel $\alpha_{i,j}$, i.e., the average number of photons per pixel, we use:

$$\alpha_{i,j} = \frac{E_0}{hc/\lambda} \cdot \frac{10^{-\alpha_{\text{atm}} \cdot 2R} \Gamma_{i,j}}{8f_{\#}^2} \cdot \frac{W_p H_p}{A_{\text{illum}}}, \quad (28)$$

and the average number of photons from the background $B_{i,j}^{\text{bck}}$ and the dark current $B_{i,j}^{\text{dc}}$ are modeled, respectively, as:

$$B_{i,j}^{\text{bck}} = \frac{W_p^{\text{bck}}}{hc/\lambda} \cdot \frac{10^{-\alpha_{\text{atm}} \cdot R} \Gamma_{i,j}}{8f_{\#}^2} \cdot W_p H_p \cdot t_r, \quad (29)$$

$$B_{i,j}^{\text{dc}} = C^{\text{dc}} t_r, \quad (30)$$

where $B_{i,j} = \eta B_{i,j}^{\text{bck}} + B_{i,j}^{\text{dc}}$, and A_{illum} indicates the area of illumination by the source. In Eq. (28), E_0 denotes the energy of the pulse, while W_p and H_p represent the width and height of a SPAD pixel. The variables R and α_{atm} correspond to the range and atmospheric attenuation, respectively. The above equations inherently follow the models introduced in [6]. All the notations and the corresponding parameter values used for this simulation are listed in Tab. VI.

To obtain high-speed depth maps, we utilized an RGB 2000-FPS video dataset [7], which was subsequently processed using the monocular depth estimation algorithm *Depth Anything*

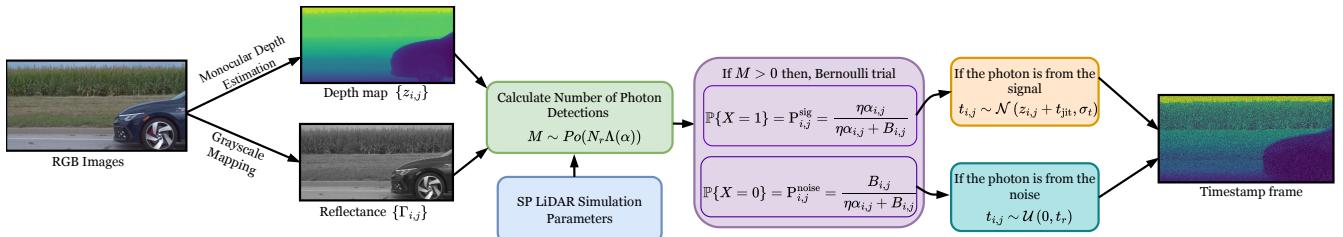


Fig. 21: An overview of the single photon LiDAR timestamp simulation process demonstrates the key components of the pipeline. First, the number of photon detections at each pixel (i, j) is obtained using a Poisson distribution. If there is at least one detection, a Bernoulli trial is then carried out to determine whether the detected photon originates from the signal or the noise. Once we sample from the corresponding distributions for all pixels, we can construct the timestamp frame.

TABLE VI: SP-LiDAR Simulation Parameters

Parameter	Symbol	Value
Dark counts	C^{dc}	126 Hz
Planck constant	h	$6.626 \times 10^{-34} \text{ m}^2\text{kgs}^{-1}$
Wavelength	$λ$	671 nm
Attenuation coefficient	$α_{atm}$	0.7dB km ⁻¹
f-number	$f_{\#}$	2.0
Reflectance	$Γ_{i,j}$	[0.0, 1.0]
Height of an effective pixel	H_p	9.2 μm
Width of an effective pixel	W_p	9.2 μm
Exposure time	t_{exp}	1000 μs
Jitter variation	$σ_j$	220 ps
Range	R	30 m
Pulse width	$σ_t$	1 ns
Energy per pulse	E_0	1.219 nJ
Background radiation	W^{bck}	0.0002 W
Repetition rate	$1/t_r$	2.25 MHz
Depth variation	$z_{i,j}$	[2 m, 60 m]
Efficiency of the sensor	$η$	0.18
Speed of Light	c	$3 \times 10^8 \text{ ms}^{-1}$

V2 [8] to generate ground truth depth maps. Using pre-trained models, we could only acquire relative depth variations. Therefore, these depth map values were converted into absolute depth values, denoted as $z_{i,j}$. For reflectance $Γ_{i,j}$, we used the grayscale values of the RGB images, a common practice in SP-LiDAR literature.

Once all parameters were obtained, we sampled timestamps from the PDF provided in Eq. (27). Since the corresponding process is a mixture of Gaussian and Uniform distributions, we could separate the PDF to determine whether a detected photon originated from the signal or noise based on the outcomes of a Bernoulli distribution defined as follows:

$$\mathbb{P}\{X=x\} = \begin{cases} P_{i,j}^{sig} = \frac{\eta\alpha_{i,j}}{\eta\alpha_{i,j} + B_{i,j}}, & \text{if } x=1, \\ P_{i,j}^{noise} = \frac{B_{i,j}}{\eta\alpha_{i,j} + B_{i,j}}, & \text{if } x=0, \end{cases} \quad (31)$$

where $X=1$ denotes the event that the detected photon originates from the signal, with corresponding probability $P_{i,j}^{sig}$, and $X=0$ denotes the event that the detected photon originates from noise, with corresponding probability $P_{i,j}^{noise}$. If the outcome of the Bernoulli trial is signal (i.e., if $X=1$), the timestamp is sampled from the corresponding Gaussian distribution. Otherwise (i.e., if $X=0$), the timestamp is sampled from the uniform distribution $\mathcal{U}(0, t_r)$. This sampling mechanism is applied to all pixels in each frame, across

all frames in each video. The overall simulation process is illustrated in Fig. 21.

XI. VISUALIZATION OF COMMON FEATURES

In this section, we provide the visualization of the common features from the experiment conducted in Sec. III-C, focusing on the latent space of both reflectivity and depth. The input images for depth and reflectivity have a size of $128 \times 128 \times 1$. The feature map size is gradually reduced through a series of convolutional layers, resulting in a latent feature map of $8 \times 8 \times 128$, corresponding to a $2 \times$ compression ratio for each branch. One-third of the latent feature map, i.e., a feature map of size $8 \times 8 \times 43$, is trained to extract the common features shared by both depth and reflectivity.

To visualize these common features in a low-dimensional space, we use Principal Component Analysis (PCA), a dimensionality reduction technique that linearly transforms the data into a new coordinate system, making it easier to identify the principal components (PCs) that capture the most significant variations [9]. Using PCA, the original 128×43 unrolled common feature maps of reflectivity and depth, denoted as $\mathcal{H}_{lsFeat} = [\mathcal{R}_{lsFeat}, \mathcal{D}_{lsFeat}]$, are projected onto a 128×3 low-dimensional feature space. Mathematically, the projection is formulated as

$$\mathcal{H}_{lsFeat}^{pca} = \mathcal{H}_{lsFeat}^c W, \quad (32)$$

where $W = [v_1 \ v_2 \ \dots \ v_k] \in \mathbb{R}^{p \times k}$ is the projection matrix, constructed from the eigenvectors corresponding to the largest k eigenvalues of the covariance matrix of $\mathcal{H}_{lsFeat}^c \in \mathbb{R}^{n \times p}$, the centered latent feature maps. Here, $p = 43$ and $n = 128$ denote the number of features and the number of samples, respectively, while $k = 3$ represents the number of PCs. The resulting transformed feature map is $\mathcal{H}_{lsFeat}^{pca} \in \mathbb{R}^{n \times k}$. Similarly, to showcase independent latent features of depth and reflectivity, the same procedure is followed.

Fig. 22 illustrates the distribution of common and independent features in the low-dimensional space. Across the feature distributions of the three scenarios, we observe that the common features are co-located, while the independent features are clearly separated. These results further support the presence of shared features between the depth and reflectivity inputs.

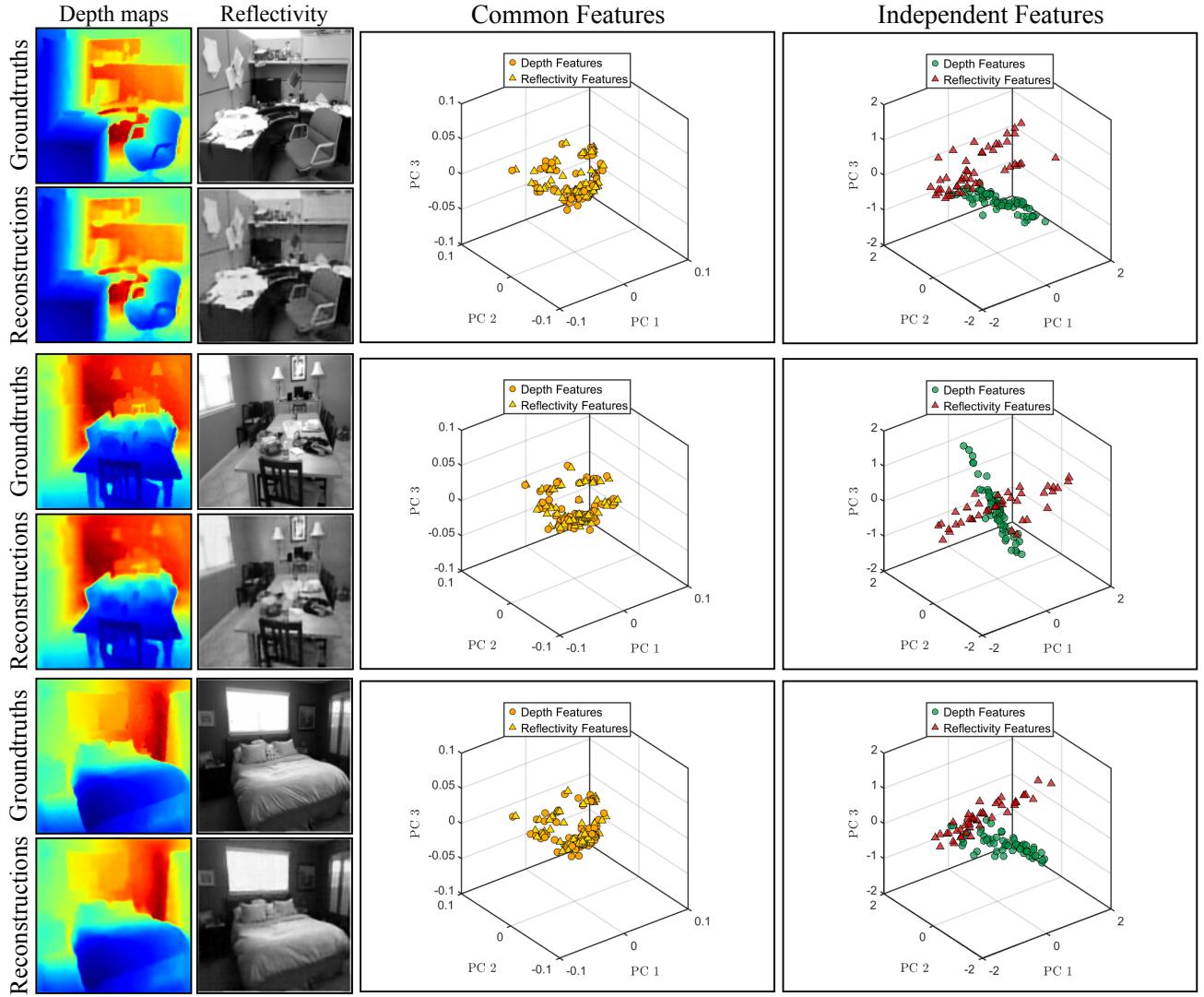


Fig. 22: The feature distribution in the low-dimensional space for four scenes shows that the common features are co-located, while the remaining features are dispersed. “PC 1”, “PC 2” and “PC 3” on the axes represent the 1th, 2nd and 3rd principal components.

REFERENCES

- [1] D. Shin, A. Kirmani, V. K. Goyal, and J. H. Shapiro, “Photon-Efficient Computational 3-D and Reflectivity Imaging With Single-Photon Detectors,” *IEEE Transactions on Computational Imaging (TCI)*, vol. 1, no. 2, pp. 112–125, 2015.
- [2] J. Rapp and V. K. Goyal, “A Few Photons Among Many: Unmixing Signal and Noise for Photon-Efficient Active Imaging,” *IEEE Transactions on Computational Imaging (TCI)*, vol. 3, no. 3, pp. 445–459, 2017.
- [3] R. Kitichotkul, J. Rapp, and V. K. Goyal, “The Role of Detection Times in Reflectivity Estimation with Single-Photon LiDAR,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 30, no. 1, pp. 1–14, 2024.
- [4] R. K. Henderson, N. Johnston, F. Mattioli Della Rocca, H. Chen, D. Day-Uei Li, G. Hungerford, R. Hirsch, D. Mcloskey, P. Yip, and D. J. S. Birch, “A 192 × 128 Time Correlated SPAD Image Sensor in 40-nm CMOS Technology,” *IEEE Journal of Solid-State Circuits*, vol. 54, no. 7, pp. 1907–1916, 2019.
- [5] S. H. Chan, H. K. Weerasooriya, W. Zhang, P. Abshire, I. Gyongy, and R. K. Henderson, “Resolution Limit of Single-Photon LiDAR,” in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024, pp. 25 307–25 316.
- [6] S. Scholes, G. Mora-Martin, F. Zhu, I. Gyongy, P. Soan, and J. Leach, “Fundamental Limits to Depth Imaging with Single-photon Detector Array Sensors,” *Nature Scientific Reports*, vol. 13, no. 1, p. 176, 2023.
- [7] P. Chennuri, Y. Chi, E. Jiang, G. M. D. Godaliyadda, A. Gnanasambandam, H. R. Sheikh, I. Gyongy, and S. H. Chan, “Quanta Video Restoration,” in *European Conference on Computer Vision (ECCV)*, 2024, pp. 152–171.
- [8] L. Yang, B. Kang, Z. Huang, Z. Zhao, X. Xu, J. Feng, and H. Zhao, “Depth Anything V2,” in *Advances in Neural Information Processing Systems (NeurIPS)*, 2024, pp. 21 875–21 911.
- [9] I. T. Jolliffe and J. Cadima, “Principal Component Analysis: A Review and Recent Developments,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, p. 20150202, 2016.