

Opportunistic Single-Photon Time Of Flight : Supplemental Document

Sotiris Nousias^{1*} Mian Wei^{1*} Howard Xiao¹ Maxx Wu¹ Shahmeer Athar¹
Kevin J. Wang¹ David A. Barmherzig¹ Anagh Malik¹ David B. Lindell^{1, 2}
Kiriakos N. Kutulakos^{1, 2}

*Joint first authors

¹Dept. of Computer Science, University of Toronto

²Vector Institute

Contents

A	Probabilistic Models and Derivations	3
A.1	Harmonic probing estimator	3
A.2	Harmonic probing estimator distribution	3
A.3	CFAR sinc comb detector	6
A.3.1	Definition of sinc comb detector	6
A.3.2	Derivation of CFAR sinc detector threshold	6
B	Algorithms	7
B.1	Laser discovery & synchronization	7
B.1.1	Additional implementation details	11
B.2	Geometric optimization	11
B.2.1	Space of solutions	11
C	Opportunistic ToF in Presence of Multi-Path Transport	13
C.1	Impact of multi-path transport on sinc comb model and laser discovery	13
C.1.1	Impact of 2-bounce specular transport on laser discovery and pulse train estimation	13
C.2	Impact of multi-path transport on pulse-delay estimation and geometry optimization	13

D	Natural Asynchrony of Pulsed Lasers	16
E	Acquisition Setup	17
E.1	Experimental hardware	17
E.2	Comparison to commercial lidars	17
E.3	Capture procedure	18
F	Experiments I: Additional Details for Figure 1 Experiment	19
F.1	Laser discovery and computational synchronization	19
F.2	Pulse-delay map estimation	20
F.2.1	Pulse-delay map estimation & frequency detection error	20
F.3	Opportunistic ToF over room-scale scenes with shadows & occlusions	20
G	Experiments II: Quantitative Evaluation of Opportunistic ToF over Room-Scale Scenes	25
G.1	3D-printed scene	25
G.2	Accuracy versus number of lasers for Figure 1 scene	25
H	Experiments III: Opportunistic ToF under Strong Ambient Light	31
H.1	Emulating ambient light	31
H.2	Precision of frequency localization as a function of SBR	31
H.3	mm-accurate 3D reconstruction under strong ambient light	31
H.4	Low-SBR experiment for scene in Figure 1	32
I	Experiments IV: Dynamic Acquisition	35
J	Simulation-Based Evaluation of Opportunistic ToF	36
J.1	Simulation details	36
J.2	3D reconstruction and localization consistency	36
J.3	Frequency localization & laser detection accuracy	37
J.3.1	Accuracy as a function of signal strength	37
J.3.2	Accuracy as a function of exposure time	37
J.3.3	Two-laser detection test	38
J.4	Pulse delay error as a function of frequency error	40
J.5	3D reconstruction accuracy	41

A Probabilistic Models and Derivations

We provide details of the derivations associated with the probabilistic models of Section 5 of the main paper. For convenience, we clarify the following notation:

- t_k : the k^{th} timestamp,
- n : harmonic order,
- \mathcal{T} : the set of timestamps,
- N : total number of harmonics used.

A.1 Harmonic probing estimator

We start by defining the periodic flux estimator, which is the value of the estimated flux using harmonic probing at $t = t_k$.

Definition 1 (Periodic flux estimator). *The periodic flux estimator $\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T})$ at a timestamp t_k is*

$$\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T}) \stackrel{\text{def}}{=} \sum_{n=-N}^N \hat{\Phi}(\mathbf{q}, nf, \mathcal{T}) \exp(j2\pi n f t_k). \quad (\text{A.1})$$

A.2 Harmonic probing estimator distribution

We start by characterizing the distribution of $\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T})$ for frequencies f that do not correspond to a laser pulse train. Under low-flux conditions, Proposition 1 tells us that $\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T})$ follows a normal distribution, with its mean and variance determined by the number of photons and number of harmonics. Intuitively, if the set of harmonics does not align with a pulse train's frequencies, summing over more harmonics adds more noise and increases the variance of the result.

Proposition 1 (Distribution of periodic flux estimator). *Under low-flux conditions, if $\Phi(\mathbf{q}, f) = 0$, $t \in \mathcal{T}$, and $|\mathcal{T}| \gg 1$, then:*

$$\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T}) \sim \mathcal{N}\left(\frac{2N + |\mathcal{T}|}{t_{\text{exp}}}, \frac{2N|\mathcal{T}|}{t_{\text{exp}}^2}\right). \quad (\text{A.2})$$

Proof Sketch of Proposition 1.

The proof proceeds in three steps:

Step 1: We show that $\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T})$ is the result of flux probing [16] with the following probing function $p(t)$:

$$p(t) = \frac{2}{t_{\text{exp}}} \sum_{n=1}^N \cos(2\pi n f (t - t_k)). \quad (\text{A.3})$$

Step 2: We characterize the distribution $\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T})$ using the following lemma:

Lemma 1 (Proposition 3 in [16]). *Let $p(t)$ be a deterministic function of time. The probing measurements $p(\mathcal{T}) = \sum_{\tau \in \mathcal{T}} p(\tau)$ are normally distributed with mean $\langle p, \phi \rangle$ and variance $\langle p^2, \phi \rangle$.*

Step 3: We derive the mean and variance of the estimator in step 1.

□

Proof of Proposition 1.

Step 1: Let us consider the k^{th} timestamp in \mathcal{T} and denote it as t_k . Let $T = |\mathcal{T}|$ be the total number of timestamps and let $\mathcal{T}_- \stackrel{\text{def}}{=} \mathcal{T} - \{t_k\}$ be the set of timestamps excluding t_k . Then we have

$$\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T}) = \sum_{n=-N}^N \hat{\Phi}(\mathbf{q}, nf, \mathcal{T}) \exp(j2\pi n f t_k) \quad (\text{A.4a})$$

$$= \sum_{n=-N}^N \hat{\Phi}(\mathbf{q}, nf, \mathcal{T}_-) \exp(j2\pi n f t_k) + \sum_{n=-N}^N \underbrace{\hat{\Phi}(\mathbf{q}, nf, \{t_k\})}_{\substack{\text{equal to} \\ \frac{\exp(-j2\pi n f t_k)}{t_{\text{exp}}}}} \exp(j2\pi n f t_k) \quad (\text{A.4b})$$

$$= \sum_{n=-N}^N \hat{\Phi}(\mathbf{q}, nf, \mathcal{T}_-) \exp(j2\pi n f t_k) + \sum_{n=-N}^N \frac{1}{t_{\text{exp}}} \quad (\text{A.4c})$$

$$= \sum_{n=-N}^N \hat{\Phi}(\mathbf{q}, nf, \mathcal{T}_-) \exp(j2\pi n f t_k) + \frac{2N+1}{t_{\text{exp}}} \quad (\text{A.4d})$$

$$= \sum_{n=1}^N \hat{\Phi}(\mathbf{q}, nf, \mathcal{T}_-) \exp(j2\pi n f t_k) + \sum_{n=-N}^{-1} \hat{\Phi}(\mathbf{q}, nf, \mathcal{T}_-) \exp(j2\pi n f t_k) + \underbrace{\frac{T-1}{t_{\text{exp}}}}_{\text{term for } n=0} + \frac{2N+1}{t_{\text{exp}}}. \quad (\text{A.4e})$$

Let us consider the first two terms in the summation of Eq. (A.4e), corresponding to the n^{th} -order harmonic, nf :

$$\hat{\Phi}(\mathbf{q}, nf, \mathcal{T}_-) \exp(j2\pi n f t_k) + \hat{\Phi}(\mathbf{q}, -nf, \mathcal{T}_-) \exp(-j2\pi n f t_k) = \frac{1}{t_{\text{exp}}} \sum_{t \in \mathcal{T}_-} \exp(j2\pi n f (t_k - t)) + \exp(-j2\pi n f (t_k - t)) \quad (\text{A.5a})$$

$$= \frac{2}{t_{\text{exp}}} \sum_{t \in \mathcal{T}_-} \cos(2\pi n f (t - t_k)). \quad (\text{A.5b})$$

Combining Eqs. (A.4a), (A.4e) and Eq. (A.5b) we obtain:

$$\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T}) = \frac{2N+T}{t_{\text{exp}}} + \sum_{t \in \mathcal{T}_-} p(t). \quad (\text{A.6})$$

Step 2: Since we assume that $|\mathcal{T}| \gg 1$, removing t_k from \mathcal{T} affects its underlying estimated flux function $\hat{\phi}$ negligibly. From Lemma 1, we know that:

$$\sum_{t \in \mathcal{T}_-} p(t) \sim \mathcal{N}(\langle p, \phi \rangle, \langle p^2, \phi \rangle). \quad (\text{A.7})$$

Substituting Eq. (A.7) into Eq. (A.6) gives us:

$$\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T}) \sim \mathcal{N}\left(\frac{2N+T}{t_{\text{exp}}} + \langle p, \phi \rangle, \langle p^2, \phi \rangle\right). \quad (\text{A.8})$$

Equation A.8 tells us therefore that we can model $\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T})$ as a normal distribution.

Step 3: In general, when the frequency f is not a harmonic of the laser repetition frequency, we expect that $\Phi(\mathbf{q}, nf) = 0$ for all $n \in \mathbb{N}$.

In low-flux regimes [16], where SPAD dead time can be ignored, we can derive closed-form expressions for the mean and variance of $\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T})$. We start by simplifying $\langle p, \phi \rangle$:

$$\langle p, \phi \rangle = \int_0^{t_{\text{exp}}} p(t) \phi(t) dt \quad (\text{A.9a})$$

$$= \frac{2}{t_{\text{exp}}} \sum_{n=1}^N \int_0^{t_{\text{exp}}} \cos(2\pi n f(t - t_k)) \phi(t) dt \quad (\text{A.9b})$$

$$= \frac{1}{t_{\text{exp}}} \sum_{n=1}^N \Phi(\mathbf{q}, nf) + \Phi(\mathbf{q}, -nf) \quad (\text{A.9c})$$

$$= 0. \quad (\text{A.9d})$$

Now, we simplify $\langle p^2, \phi \rangle$:

$$\langle p^2, \phi \rangle = \int_0^{t_{\text{exp}}} p^2(t) \phi(t) dt \quad (\text{A.10a})$$

$$= \frac{4}{t_{\text{exp}}^2} \sum_{n_1=1}^N \sum_{n_2=1}^N \int_0^{t_{\text{exp}}} \underbrace{\cos(2\pi n_1 f(t - t_k)) \cos(2\pi n_2 f(t - t_k))}_{\text{use trig identity for } \cos A \cos B} \phi(t) dt \quad (\text{A.10b})$$

$$= \frac{2}{t_{\text{exp}}^2} \sum_{n_1=1}^N \sum_{n_2=1}^N \int_0^{t_{\text{exp}}} \underbrace{\cos(2\pi(n_1 - n_2)f(t - t_k)) \phi(t) + \cos(2\pi(n_1 + n_2)f(t - t_k)) \phi(t)}_{= 0 \text{ when } n_1 \neq n_2} dt \quad (\text{A.10c})$$

$$= \frac{2}{t_{\text{exp}}^2} \sum_{n=1}^N \underbrace{\int_0^{t_{\text{exp}}} \phi(t) dt}_{\approx T} \quad (\text{A.10d})$$

$$= \frac{2NT}{t_{\text{exp}}^2}. \quad (\text{A.10e})$$

Therefore, we have:

$$\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T}) \sim \mathcal{N}\left(\frac{2N+T}{t_{\text{exp}}}, \frac{2NT}{t_{\text{exp}}^2}\right). \quad (\text{A.11})$$

□

In high-flux conditions, where dead time can be comparable or larger than the photon interarrival time, the noise model in Eq. (A.8) still holds. While there is no closed-form expression for the mean and variance in that case, it may still be possible to leverage the sample mean and variance expressions of Eq. (A.2) for the CFAR sinc-comb detection. In both our experiments and our simulations we find that even with an SBR of 0.01 and high-flux conditions, our analytical noise model of Eq. (A.2) is effective in detecting the presence of periodic flux signals.

A.3 CFAR sinc comb detector

A.3.1 Definition of sinc comb detector

In this subsection, we derive the CFAR bound for the sinc comb detector described in Section 5 of the main paper. Proposition 1 informs us that the flux estimated for a specific time instant from frequencies not present in the true flux follow a normal distribution. Leveraging this insight, we identify pulse repetition frequencies that yield reconstructed pulse trains as opposed to reconstructed signals that are statistically indistinguishable from Gaussian noise. Mathematically, this can be expressed as:

Definition 2 (CFAR sinc comb detector). To achieve a constant probability of false alarm p , a candidate frequency f is detected if there is at least one photon timestamp $t_k \in \mathcal{T}$ for which the estimated instantaneous flux at t_k is statistically significant, *i.e.*,

$$\hat{\phi}(t_k) > \underbrace{\frac{2N + |\mathcal{T}|}{t_{\text{exp}}} + \text{CDF}_{\mathcal{N}}^{-1}(1 - p) \frac{\sqrt{2N|\mathcal{T}|}}{t_{\text{exp}}}}_{\text{CFAR sinc detection threshold } \Delta} \quad (\text{A.12})$$

where N is the number of harmonics, and t_{exp} is the exposure time. We derive the detection threshold Δ below.

A.3.2 Derivation of CFAR sinc detector threshold

The main intuition of Definition A.12 is that a frequency f is only detected when it is statistically significant, *i.e.*, f is a noise frequency under our null hypothesis. Given that we know the distribution of the estimated instantaneous flux at timestamp t_k under the null hypothesis from Eq. (A.11), we set our threshold Δ according to the z-score needed for a one-sided Z-test [15] to reject our null hypothesis with a significance level of $1 - p$. In our detection scheme, we set $p = 1/(|\mathcal{T}||\mathcal{F}_{\text{cand}}|)$, *i.e.*, the likelihood of observing at least one outlier among the total number of timestamps and candidate frequencies. For instantaneous flux estimates that do not correspond to actual laser repetition frequencies, this can be expressed as follows:

$$P(\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T}) > \Delta) = p \quad (\text{A.13a})$$

$$P(\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T}) \leq \Delta) = 1 - p \quad (\text{A.13b})$$

$$P\left(\frac{\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T}) - \mu}{\sigma} \leq \frac{\Delta - \mu}{\sigma}\right) = 1 - p, \quad (\text{A.13c})$$

where $\mu = E[\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T})]$ and $\sigma^2 = \text{var}[\hat{\phi}_f(\mathbf{q}, t_k, \mathcal{T})]$. Substituting Eq. (A.11) into Eq. (A.13) and inverting both sides gives:

$$\Delta = \frac{2N + |\mathcal{T}|}{t_{\text{exp}}} + \text{CDF}_{\mathcal{N}}^{-1}(1 - p) \frac{\sqrt{2N|\mathcal{T}|}}{t_{\text{exp}}}, \quad (\text{A.14})$$

where $\text{CDF}_{\mathcal{N}}^{-1}$ denotes the inverse cumulative distribution function of the standard normal distribution.

B Algorithms

B.1 Laser discovery & synchronization

Below we provide pseudocode for the individual components of the approach described in Section 5.

Algorithm 1: Initial repetition frequency candidates

Input: \mathcal{T} – set of timestamps from pixel \mathbf{q} , t_{exp}
Output: $\mathcal{F}_{\text{cand}}$ – set of candidate repetition frequencies
 $\mathcal{F}_{\text{cand}} \leftarrow \emptyset$
sweep frequencies up to 50 MHz with step size $0.6/t_{\text{exp}}$
 $\Delta_f \leftarrow 0.6/t_{\text{exp}}$
 $\mathcal{F} \leftarrow \{f \mid f < 50 \text{ MHz and } f = n\Delta_f \text{ for } n \in \mathbb{N}\}$
 $p \leftarrow \frac{1}{|\mathcal{F}|}$
for f **in** \mathcal{F} :
 # Equation 3 of main paper
 $\hat{\Phi}(\mathbf{q}, f, \mathcal{T}) \leftarrow \frac{1}{t_{\text{exp}}} \sum_{t \in \mathcal{T}} \exp(-j2\pi ft)$
 # Equation 4 of main paper
 if $\|\hat{\Phi}(\mathbf{q}, f, \mathcal{T})\|^2 \geq \text{CDF}_{\chi^2}^{-1}(1-p) \frac{|\mathcal{T}|}{2t_{\text{exp}}^2}$:
 $\mathcal{F}_{\text{cand}} \leftarrow \mathcal{F}_{\text{cand}} \cup \{f\}$
return $\mathcal{F}_{\text{cand}}$

Algorithm 2: Frequency localization by high-res scanning

Input: \mathcal{T} – set of timestamps from pixel \mathbf{q} , t_{exp} , $\mathcal{F}_{\text{cand}}$, $\Delta_{f_{\text{fine}}}$
Output: \mathcal{F}_{loc} – set of localized repetition frequencies
 $\mathcal{F}_{\text{loc}} \leftarrow \emptyset$
 $k_{\text{max}} \leftarrow \lfloor \frac{0.6}{t_{\text{exp}} \Delta_{f_{\text{fine}}}} \rfloor$
for f **in** $\mathcal{F}_{\text{cand}}$:
 # Frequency scanning with fine frequency steps in the neighborhood
 # of the first harmonic of candidate frequency
 $k_{\text{loc}} \leftarrow \underset{-k_{\text{max}} < k < k_{\text{max}}}{\text{argmax}} \|\hat{\Phi}(\mathbf{q}, f + k\Delta_{f_{\text{fine}}}, \mathcal{T})\|^2$
 $\mathcal{F}_{\text{loc}} \leftarrow \mathcal{F}_{\text{loc}} \cup \{f + k_{\text{loc}}\Delta_{f_{\text{fine}}}\}$
return \mathcal{F}_{loc}

Algorithm 3: Frequency pruning by second harmonic detection

Input: \mathcal{T} – set of timestamps from pixel \mathbf{q} , t_{exp} , \mathcal{F}_{loc} , $\Delta_{f_{\text{fine}}}$

Output: \mathcal{F}_{loc} – set of localized repetition frequencies

$k_{\text{max}} \leftarrow \lfloor \frac{0.6}{t_{\text{exp}} \Delta_{f_{\text{fine}}}} \rfloor$

for f in \mathcal{F}_{loc} :

 # Frequency probing with fine step around the second harmonic of candidate frequency

$k_{\text{loc}} \leftarrow \underset{-k_{\text{max}} < k < k_{\text{max}}}{\text{argmax}} \quad \|\hat{\Phi}(\mathbf{q}, 2f + 2k\Delta_{f_{\text{fine}}}, \mathcal{T})\|^2$

if $\|\hat{\Phi}(\mathbf{q}, 2f + 2k_{\text{loc}}\Delta_{f_{\text{fine}}}, \mathcal{T})\|^2 < \text{CDF}_{\chi^2}^{-1}(1-p) \frac{|\mathcal{T}|}{2t_{\text{exp}}^2}$:

$\mathcal{F}_{\text{loc}} \leftarrow \mathcal{F}_{\text{loc}} - \{f\}$

else:

$f_{\text{loc}} \leftarrow f + k_{\text{loc}}\Delta_{f_{\text{fine}}}$

$\mathcal{F}_{\text{loc}} \leftarrow (\mathcal{F}_{\text{loc}} - \{f\}) \cup \{f_{\text{loc}}\}$

return \mathcal{F}_{loc}

Algorithm 4: mHz frequency localization by harmonic hopping

Input: \mathcal{T} – set of timestamps from pixel \mathbf{q} , t_{exp} , \mathcal{F}_{loc} , $\Delta_{f_{\text{fine}}}$

Output: \mathcal{F}_{loc} – set of localized repetition frequencies

$k_{\text{max}} \leftarrow \lfloor \frac{0.6}{t_{\text{exp}} \Delta_{f_{\text{fine}}}} \rfloor$

for f in \mathcal{F}_{loc} :

for $i = 3$ to 10 :

$n \leftarrow 2^i$

 # Frequency probing with fine step around the n-th harmonic of candidate frequency

$k_{\text{loc}} \leftarrow \underset{-k_{\text{max}} < k < k_{\text{max}}}{\text{argmax}} \quad \|\hat{\Phi}(\mathbf{q}, nf + nk\Delta_{f_{\text{fine}}}, \mathcal{T})\|^2$

if $\|\hat{\Phi}(\mathbf{q}, nf + nk_{\text{loc}}\Delta_{f_{\text{fine}}}, \mathcal{T})\|^2 \geq \text{CDF}_{\chi^2}^{-1}(1-p) \frac{|\mathcal{T}|}{2t_{\text{exp}}^2}$:

$f \leftarrow f + k_{\text{loc}}\Delta_{f_{\text{fine}}}$

else:

break

return \mathcal{F}_{loc}

Algorithm 5: Pulse train reconstruction by harmonic probing

Input: \mathcal{T} – set of timestamps from pixel \mathbf{q} , t_{exp} , f
Output: $\hat{\phi}_f(\mathbf{q}, t, \mathcal{T})$ – reconstructed pulse train at each timestamp t in \mathcal{T}
maximum frequency detectable by SPAD
 $f_{\text{max}} \leftarrow 15 \text{ GHz}$
 $N \leftarrow \max\{n \in \mathbb{N} \mid n < \lfloor \frac{f_{\text{max}}}{f} \rfloor\}$
for $n = -N$ **to** N :
 # Probe the n -th harmonic of f
 $\hat{\Phi}(\mathbf{q}, nf, \mathcal{T}) \leftarrow \frac{1}{t_{\text{exp}}} \sum_{t \in \mathcal{T}} \exp(-j2\pi nft)$
for $t \in \mathcal{T}$:
 # Reconstruct pulse train at each timestamp
 $\hat{\phi}_f(\mathbf{q}, t, \mathcal{T}) \leftarrow \sum_{n=-N}^N \hat{\Phi}(\mathbf{q}, nf, \mathcal{T}) \exp(j2\pi nft)$
return $\hat{\phi}_f(\mathbf{q}, t, \mathcal{T})$

Algorithm 6: CFAR sinc comb detection and synchronization

Input: \mathcal{T} – set of timestamps from pixel \mathbf{q} , t_{exp} , \mathcal{F}_{loc} , p – probability of false alarm
Output: $\mathcal{F}_{\text{laser}}$ – set of detected pulse repetition frequencies
 $\mathcal{F}_{\text{laser}} \leftarrow \emptyset$
for f in \mathcal{F}_{loc} :
 let $\hat{\phi}_f(\mathbf{q}, t, \mathcal{T})$ be the output of Pseudocode 5 with inputs \mathcal{T} , t_{exp} , and f
 # CFAR threshold for pulse train
 $z_p \leftarrow \text{CDF}_{\mathcal{N}}^{-1}(1 - p)$
 $\Delta \leftarrow \frac{2N + |\mathcal{T}|}{t_{\text{exp}}} + z_p \frac{\sqrt{2N|\mathcal{T}|}}{t_{\text{exp}}}$
 $\mathcal{T}_{\text{laser}} \leftarrow \left\{ t \in \mathcal{T} \mid \hat{\phi}_f(\mathbf{q}, t, \mathcal{T}) > \Delta \right\}$
 # Frequency detected if pulse train exceeds CFAR threshold for at least one timestamp in \mathcal{T}
 if $\mathcal{T}_{\text{laser}} \neq \emptyset$:
 $\mathcal{F}_{\text{laser}} \leftarrow \mathcal{F}_{\text{laser}} \cup \{f\}$
return $\mathcal{F}_{\text{laser}}$

Algorithm 7: Pulse-delay map estimation

Input: \mathcal{T} – set of timestamps from pixel \mathbf{q} , $\mathcal{F}_{\text{laser}}$

Output: $o_f + \tau_f(\mathbf{q})$ – pulse delay for $f \in \mathcal{F}_{\text{laser}}$ at pixel \mathbf{q}

maximum frequency detectable by SPAD

$f_{\text{max}} \leftarrow 15 \text{ GHz}$

for f in $\mathcal{F}_{\text{laser}}$:

$N \leftarrow \max\{n \in \mathbb{N} \mid n < \lfloor \frac{f_{\text{max}}}{f} \rfloor\}$

for $n = -N$ to N :

 # Probe the n-th harmonic of f

$\hat{\Phi}(\mathbf{q}, nf, \mathcal{T}) \leftarrow \frac{1}{t_{\text{exp}}} \sum_{t \in \mathcal{T}} \exp(-j2\pi nft)$

 # Period-wrapped counterpart of pulse train $\hat{\phi}_f(\mathbf{q}, t, \mathcal{T})$

 compute $\hat{\phi}_f(\mathbf{q}, (t \bmod 1/f), \mathcal{T}) = \sum_{n=-N}^N \hat{\Phi}(\mathbf{q}, nf, \mathcal{T}) \exp(j2\pi nf(t \bmod 1/f))$

$o_f + \tau_f(\mathbf{q}) \leftarrow \text{argmax}_t \hat{\phi}_f(\mathbf{q}, (t \bmod 1/f), \mathcal{T})$

return $o_f + \tau_f(\mathbf{q})$ for each f in $\mathcal{F}_{\text{laser}}$

B.1.1 Additional implementation details

Initial candidate frequencies. We identify candidate repetition frequencies by frequency-scanning the DC-50 MHz range with step size $0.6/t_{\text{exp}}$. 50 MHz is chosen as an upper bound for repetition frequencies of typical lasers used in lidar [2]. $0.6/t_{\text{exp}}$ is chosen according to Section B.10 of [16].

Frequency non-maximum suppression. CFAR sinc comb and synchronization may result in a set $\mathcal{F}_{\text{laser}}$ that contains multiple laser frequencies clustered within a sinc lobe; higher-order harmonics of the same fundamental frequency; and frequencies that lie within the span of the sinc lobes of a laser’s repetition frequency. To remove such spurious frequencies, we perform non-maximum suppression by clustering all laser frequencies that lie within the same sinc lobe, keeping only the laser frequency f that maximizes $\hat{\phi}_f(\mathbf{q}, t, \mathcal{T})$. Then we remove higher-order harmonics by projecting them to first-order laser frequencies. Finally we perform another round of non-maximum suppression by probing the frequency intervals $(f - 1.5t_{\text{exp}}, f - 0.5t_{\text{exp}})$ and $(f + 0.5t_{\text{exp}}, f + 1.5t_{\text{exp}})$ with a step size of 1 mHz.

Boosting SNR by local timestamp aggregation. We randomly choose a single 10 by 10 neighborhood of pixels to aggregate timestamps and then perform probing operation on this “superpixel” in all our experiments. This aggregation boosts SNR by up to 10x, and reduces the effect of dead-time under strong ambient light conditions while preserving memory and runtime efficiency.

Pulse-delay map estimation. We discretize the period-wrapped counterpart $\hat{\phi}_f(\mathbf{q}, (t \bmod 1/f), \mathcal{T})$ using 10000 evenly spaced samples over the domain $[0, 1/f)$. We estimate the pulse delay by first identifying the time at which the discretized flux reaches its maximum. We further refine this estimate by using it as the starting point for Newton’s algorithm, running it for 100 iterations to compute the pulse delay by locating the global maximum of $\hat{\phi}_f(\mathbf{q}, (\tau \bmod 1/f), \mathcal{T})$.

B.2 Geometric optimization

We rewrite Eq. (5) of the main paper to formulate the passive 3D geometry reconstruction task as the following optimization problem:

$$\underset{o_l, \mathbf{l}_l, d(\mathbf{q})}{\text{minimize}} \sum_{\mathbf{q}} \rho_{\alpha, c} \{ \|\mathbf{l}_l - d(\mathbf{q})\mathbf{v}(\mathbf{q})\| + d(\mathbf{q}) - c(\tau_l(\mathbf{q}) + o_l) \} \quad (\text{B.15})$$

$$\text{subject to } d(\mathbf{q}) \geq 0 \text{ for all } \mathbf{q} \text{ and } \alpha \geq 0, \quad (\text{B.16})$$

where $\rho_{\alpha, c}$ is Barron’s adaptive robust loss defined by adaptive parameters α, c with constraint $\alpha \geq 0$ [4]. We randomly initialize our estimation parameters $o_l, \mathbf{l}_l, d(\mathbf{q})$ in the range between $[0, 1)$, and use Adam optimizer [10] with 0.001 learning rate to run optimization until convergence. Convergence is defined as the point at which the average change of $o_l, \mathbf{l}_l, d(\mathbf{q})$ across all pixels \mathbf{q} falls below 10^{-3} . Typical number of iterations required for convergence ranges from 100K to 1M.

B.2.1 Space of solutions

As mentioned in Section 6 of the main paper, given L lasers and N pixels, the solution to the system of algebraic equations of Eq. (5) lies in \mathbb{R}^{4L+N} ($3L$ for the light source positions, L for the sync offset, and N dimensions for the depth of N pixels). Since Eq. (5) can be expressed as a polynomial, we can associate an algebraic manifold of dimension $4L + N - 1$ to each pulse delay measurement. For general laser and scene point positions, the algebraic manifolds associated with any two instances of Eq. (5) will intersect transversally [9]. This results in a discrete solution space with no continuous deformation

ambiguities for $L=2$, $N=8$ and $L=3$, $N=6$. Counting these solutions is beyond the scope of this paper; see [7] for a possible approach.

In practice, we find that (a) 3 lasers are sufficient for reliable reconstruction of complex scenes and (b) 2-laser reconstructions are very sensitive to outliers/shadows (see Section G.2 and Figure G.14 for further details). Lastly, when the scene is a single plane, light sources can only be localized up to a reflection ambiguity as placing light sources at their virtual positions with the scene acting as the plane of reflection preserves the ToF measurements.

C Opportunistic ToF in Presence of Multi-Path Transport

In this section, we discuss the effect of indirect light transport to our method and we expand on how it affects (a) our sinc comb model, laser discovery and synchronization, and (b) pulse-delay estimation, source localization and depth estimation.

C.1 Impact of multi-path transport on sinc comb model and laser discovery

When multi-path light transport is not negligible, the impulse $\delta(t - \tau_l(\mathbf{q}))$ in Eq. (1) becomes a function $h_{\mathbf{q}}(t)$ describing the scene’s temporal impulse response for pixel \mathbf{q} (Figure C.2). Despite this, *the sinc comb structure of incident flux in the Fourier domain is preserved* because it depends only on the laser’s frequency comb and the exposure crop (Eq. (2)). Thus, Figure 3 (row 2, col 4) and Figure 4 still apply. Consequently, laser discovery and synchronization, which depend only on the sinc comb structure, are also invariant to multi-path transport. For concreteness, we focus on some specific cases of two-bounce transport below.

C.1.1 Impact of 2-bounce specular transport on laser discovery and pulse train estimation

Now consider the case where the scene consists of a specular and a diffuser surface arranged as shown in Figure C.1. The flux incident at \mathbf{q} and \mathbf{q}' from \mathbf{l}_1 results in the detection of a *single* laser in both cases. Specifically, the train $\hat{\phi}_f(\mathbf{q}, \mathcal{T})$ reconstructed for \mathbf{l}_1 in Figure C.1a will contain *two* pulses for every pulse emitted by the laser, as shown in Figure C.2c. This treatment is precisely what allows contributions from multiple bounces—diffuse, specular, etc—to be associated to the laser that actually produced them: in the supplementary video, distant diffuse inter-reflections (as in Figure C.2b can be seen appearing and disappearing on the male mannequin’s cheek over a few consecutive frames, due to 2-bounce light from the “magenta” (Figure C.3a) and “green” (Figure C.3b) laser, respectively.

C.2 Impact of multi-path transport on pulse-delay estimation and geometry optimization

Pulse delays may be estimated erroneously for some lasers (*e.g.*, the paths $\mathbf{l}_1 \rightarrow \mathbf{x} \rightarrow \mathbf{y} \rightarrow \mathbf{q}$ and $\mathbf{l}_1 \rightarrow \mathbf{y} \rightarrow \mathbf{q}$ in Figure C.1a may result in an estimated pulse similar to that of Figure C.2c) or for all of them (*e.g.* due to path $\mathbf{l}_1 \rightarrow \mathbf{z} \rightarrow \mathbf{x} \rightarrow \mathbf{q}'$ in Figure C.1b). We handle such errors in the geometric optimization stage (Section 6), where *global constraints across all pixels and lasers can be utilized*. Specifically, our use of a robust loss function in Eq. (B.15) serves to reduce the influence of pulse delays that are inconsistent with a global solution for scene depths and laser positions *by treating them as outliers*. As can be seen from the estimated pulse delays and final influence maps of the *bathroom1* scene in Figure C.4, pixels receiving only multi-bounce light from the green laser (red arrow) are successfully reconstructed by ignoring that laser’s corrupted delays, whereas pixels corresponding to the case shown in Figure C.1b (mirror’s surface) are automatically identified as unreconstructible as shown by the influence maps in Figure C.4. Thus, as long as enough pixels in the camera’s field of view have pulse delay measurements that are not uncorrupted by such multi-path transport, the scene’s lasers can be localized in 3D and the non-specular surfaces can be reconstructed accurately. In the synthetic scene shown in Figure C.4 for example, the total RMSE was 7.2 mm without considering errors on the Venetian blinds and 5.3 cm with the errors included.¹

¹The increased depth errors on the Venetian blinds occurred because the tiny gaps between the blinds’ blades resulted in transient rendering artifacts and thus were unrelated to our method.

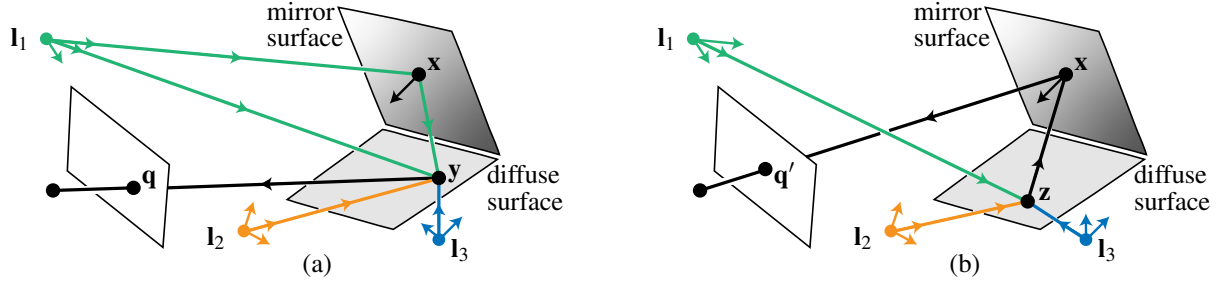


Figure C.1: Failure cases of pixelwise pulse-delay estimation. (a) Laser I_1 emits a pulse that reaches the diffuse point y directly and simultaneously reflects off the mirror-like (specular) surface at point x before also reaching the same diffuse point y . As a result, two pulses—one direct and one specular-reflected—arrive at the point y . These pulses create two identical peaks at pixel q , as illustrated previously in case P3 (Figure C.2). (b) The pulses from light sources I_1 , I_2 and I_3 reach the diffuse point z which is not directly visible from pixel q' . Instead, the light scattered at z reflects off the mirror surface at x before reaching q' . As a result, the estimated pulse delay for each light source will be incorrect since there is not direct reflection from point z to q' .

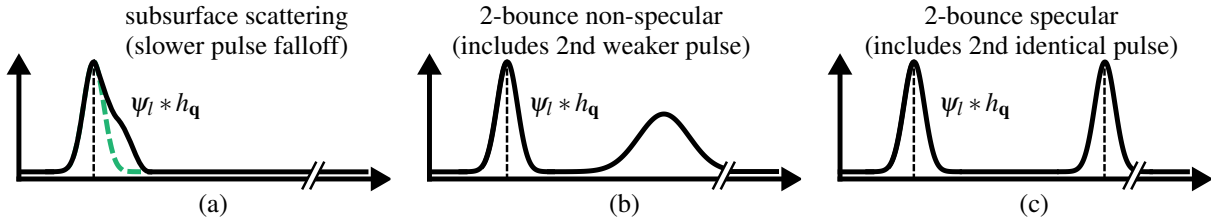


Figure C.2: Effect of multi-path light transport to the incident flux at a pixel q . (a) Subsurface scattering broadens the pulse shape, resulting in a slower falloff. (b) Non-specular two-bounce reflections introduce additional pulse with lower amplitude. (c) Specular two-bounce reflections produce secondary pulses that are potentially identical in shape and amplitude to the directly-reflected laser pulses. All of these phenomena can be modelled as a convolution of the laser pulse ψ_l with the pixel's transient response h_q , leaving the underlying sinc comb model unaffected.



Figure C.3: 2-bounce indirect reflections in Figure 1 scene. (a) Indirect light transport on the mannequin face from the magenta laser. (b) Indirect light transport on the mannequin face from the green laser.

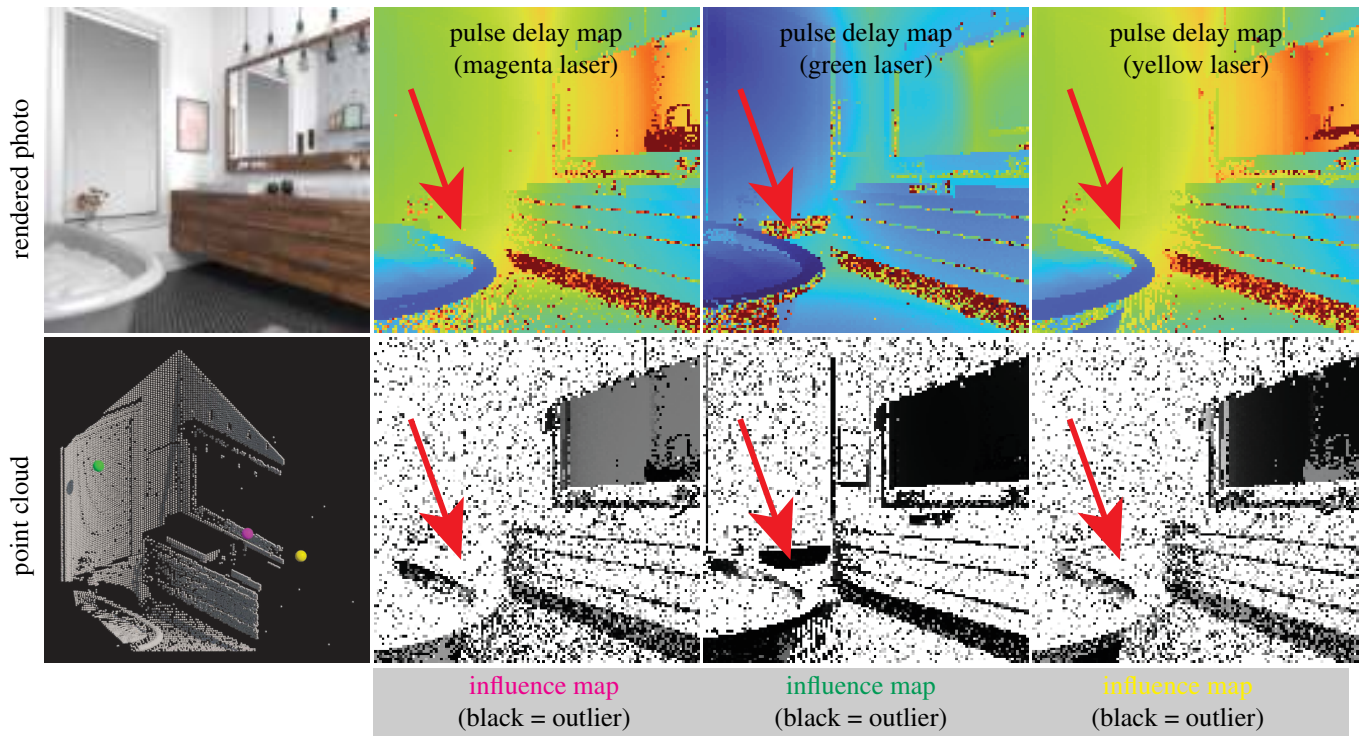


Figure C.4: Simulation results on the *bathroom1* scene from the dataset of Bitterli [5], containing specular surfaces. Specular reflections result in additional peaks, causing erroneous pulse-delay maps that do not provide correct geometric constraints. Such delays are identified and handled as outliers shown in the laser-specific influence maps on the second row. The laser-specific influence maps are computed per-pixel upon convergence of geometric optimization, by taking the derivative of the adaptive loss function, according to Eq. (9) on Barron [4]. See Section J for further details on our simulations.

D Natural Asynchrony of Pulsed Lasers

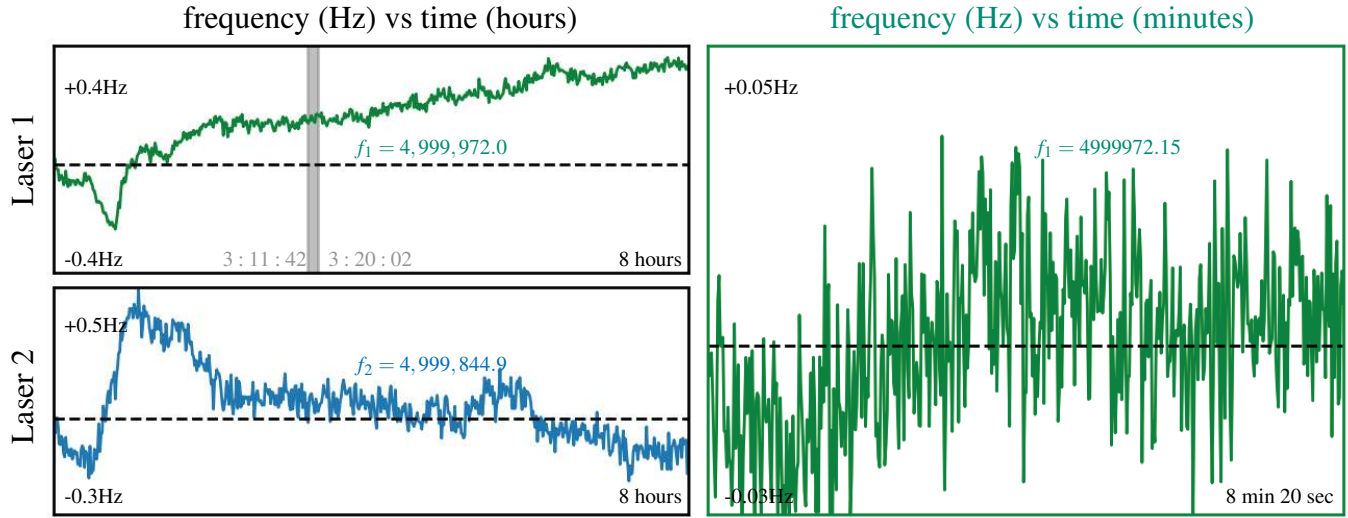


Figure D.5: Left: Frequency estimates of two laser generators operating at 5 MHz over an 8 hour period. **Right:** Zoomed-in frequency estimates for a duration of 8 minutes and 20 seconds.

We assess the degree of synchronization between two identical picosecond lasers with identical but individual driver devices as follows. We set the nominal frequency of two Alphas lasers to 5 MHz, which is well within their nominal frequency range of 1 MHz to 40 MHz. We then use two channels of our time-to-digital converter (TDC) to timestamp the sync signal output by their individual drivers. The TDC generated a sequence of time-stamped markers, one per period, for each of the lasers. We use the markers to estimate the lasers' frequency for each one-millisecond interval during an acquisition as the inverse of the average timespan between consecutive markers. Figure D.5 shows how these frequencies varied over an acquisition of 8 hours (left) and 8 minutes (right). Despite being set at the exact same frequency, the actual frequencies of the lasers initially differed by 128 Hz and exhibited additional drift of nearly 1 Hz over a period of 8 hours of constant operation. This suggests that while Hz-scale frequency stability over that interval was maintained, the absolute difference between the frequencies of those two identical laser devices was much larger (over 100 Hz in our lasers).

E Acquisition Setup

E.1 Experimental hardware

The following sensor and laser source were used in our experiments:

- **Free-running SPAD:** We use a single-pixel PDM Series Fast-Gated SPAD from Micro Photon Devices, operated in asynchronous mode. The SPAD has a pixel pitch of 50 micrometers, a timing jitter of 68 picoseconds, a 231 nanosecond dead time, and a quantum efficiency of 30% at 532 *nm*.
- **Time-to-digital converter (TDC):** The SPAD's output is converted into a stream of timestamps using a Swabian TimeTagger Ultra. The TimeTagger Ultra's timing resolution is 1 picosecond and its RMS timing jitter is 8 picoseconds. The TDC features four independent channels, one of which was used for the SPAD.
- **Galvo mirrors:** For experiments where 2D imaging was required, we used a pair of galvo mirrors from Thorlabs (Model GVS012) to acquire measurements over a 2D field of view with our single-pixel SPAD.
- **Ultrafast high-power pulsed laser:** We used a class-4 picosecond laser with a wavelength of 532 *nm* (NKT Photonics Katana 05HP). We vary its repetition rate from 9.993 MHz to 10 MHz during our acquisition and set its power to 0.2W. Its FWHM was 110 ps under the experimental conditions we employed, resulting in a flux function with frequency support up to 10 GHz.
- **Mirrors:** We used Thorlabs E01-BB02 mirrors to redirect the collimated laser beam to various points in the scene.
- **Picosecond laser drivers:** We used two Alphas Lasers PLDD-100M picosecond drivers as reference clocks for measuring frequency drift.
- **Pulsed laser illumination:** We use mirrors (Thorlabs E01-BB02) to redirect the laser beam towards a piece of A4 paper taped onto the wall. The A4 paper acts as a diffuser, scattering laser light onto our scene.

E.2 Comparison to commercial lidars

We compare the laser power we use in our experiment to a commercially available automotive flash lidar, Lumentum's M53-100 [12].

The M53-100 reports a 400 W typical peak optical power with a 0.1% duty cycle at 905 *nm* wavelength (NIR). This results in a 0.4 W average optical power. We operated the Katana HP-5 at 0.2W average power and the wavelength of the laser beam is 532 *nm*. Since photon energy is inversely proportional to its wavelength (Einstein-Planck equation [1]), photons at 905 *nm* wavelength carry approximately $1.7\times$ less energy than photons at 532 *nm* wavelength. We compute the ratio in photon flux between our light source and the M53-100:

$$0.294 \approx \underbrace{\frac{0.2 \text{ W}}{0.4 \text{ W}}}_{\text{ratio in optical power}} \times \underbrace{\frac{532 \text{ nm}}{905 \text{ nm}}}_{\text{ratio in photon energy}}. \quad (\text{E.17})$$

Therefore using a sensor with a comparable quantum efficiency at 905 *nm*; replacing our laser with the M53-100; and increasing its beam divergence from its native 19° to the 180° hemisphere used in our

experiments would result in 3.4x more photons detected per pixel on average.

E.3 Capture procedure

Since we had access only to a single-pixel SPAD, the only way to emulate concurrent imaging over a 2D field of view of multiple asynchronous lasers is to (1) scan the field of view pixel by pixel, (2) repeat this 2D scan for each laser position, (3) compute the union of the photon timestamp streams due to the pulsed lasers and ambient light and (4) account for the SPAD's dead time by removing timestamps from this union that occurred during the SPAD's dead time intervals. Specifically:

1. **Temporal pixel alignment:** In order to emulate concurrent imaging, the photon timestamps of all pixels must be computationally shifted relative to each other to create the effect of a common absolute clock. We use the laser synchronization markers to provide temporal alignment of the photon timestamps within the laser repetition period.
2. **Scanning and timestamp stream collection:** We used the galvo mirrors to scan the scene one pixel at a time, with each pixel collecting photons for the same time duration (100 milliseconds unless otherwise stated).
3. **Concurrent imaging of multiple asynchronous pulsed lasers:** Since we had access to only a single pulsed laser source, we repeat the previous steps as many times as the number of light sources (typically three, unless stated otherwise). The final timestamp stream for each pixel is obtained by combining the individual timestamps from each pulsed laser source along with ambient timestamps, while accounting for dead time. Specifically, if two photons from different streams were detected within the dead time after alignment, one of them is removed.
4. **Total acquisition time:** Total acquisition times vary by experiment and ranged from 2 hours for a 128×128 scan to 10 hours for a 512×512 scan. See Section [F](#) for the details.

F Experiments I: Additional Details for Figure 1 Experiment

F.1 Laser discovery and computational synchronization

We provide additional details about the experiment shown in Figure 1 (middle).

Scene: The scene itself is shown in Figure 1 (left, top). It consisted of two mannequins separated by a board covered with black fabric. The board was intentionally placed between the mannequins to cast shadows from at least one laser. In this scene, the single-photon camera is positioned 1.12 m from right wall, and 2.99 m from the background wall.

Illumination: We used the procedure outlined in Section E.3 to acquire timestamps corresponding to 3 pulsed lasers outside the SPAD’s field of view and approximately 2.16 m away from each other. The pulse repetition frequency of the lasers was set to 9.998 MHz, 9.999 MHz, and 10.000 MHz, respectively. The frequencies were intentionally chosen to differ at kHz scales because those are the *hardest* to distinguish (Section J.3, Figure J.20).

Scanning and timestamp stream collection: We used the galvo mirrors to scan the field of view with a spatial resolution of 0.38 cm, resulting in a 512×512 image. We used an exposure time of $t_{\text{exp}} = 0.1$ seconds for timestamp stream acquisition at each galvo position. Average photon counts at every pixel acquired is 11040 photons. Photon counts at individual pixels vary significantly due to shading and shadows: differences in photon counts between brightest and dimmest pixels are 15159, 9581, 9270 for laser 1, 2, and 3, respectively.

Frequency scanning and detection: We aggregate timestamps from a 10×10 patch. We then probe from 0 Hz to 50 MHz with a step size of 6 Hz. We set the constant probability of false alarms to 1 over the number of probed frequencies for the initial candidate frequencies, *i.e.* $0.12 \cdot 10^{-7}$. From this step, 362 candidate frequencies were above the CFAR bound which were then refined as part of our frequency detection method (Algorithms 2—6). Three laser frequencies passed the CFAR sinc comb detector. The whole process from scanning to detection takes 1.9 hours.²

Laser separation: Computational laser synchronization is essential for passive 3D imaging, as it reconstructs the pulse train of each laser, enabling the generation of pulse-delay maps used as inputs for the geometric optimization. For each detected laser frequency, we reconstruct its corresponding pulse train by performing harmonic probing (Algorithm 5) up to the 1500th harmonic for all pixels in the image. In this scene, reconstructing the pulse train for each detected laser frequency involves computing only 1500 Fourier coefficients per detected laser frequency, resulting in a total of 4500 coefficients for the three detected lasers.

Visualization of laser wavefronts propagating asynchronously: We reconstruct the flux by integrating the incident pulse train at each pixel over a 125-picosecond exposure. This corresponds to a frame rate of 8 billion frames per second. Figure F.6 shows the transient before (left) and after (right) laser separation, demonstrating the flux function decomposition into pulse trains from the detected pulsed lasers. In Figure F.6 (right), we represent each laser with a distinct color for visualization purposes. Notably, the separation of laser pulses reveals intriguing phenomena occurring on the picosecond timescale. For instance, in the top row of Figure F.6, the top wall (annotated with a white ellipse) appears illuminated by a single wavefront in the transient image of unseparated pulsed lasers. However, our method reveals that this effect actually arises from the simultaneous illumination of the wall by wavefronts from two different pulsed lasers. This phenomenon is even more prominent in Figure F.6 (row 2, left). This is consistently observed across the remaining rows in Figure F.6.

²For comparison, our implementation of [16] takes approximately 60 hours.

Computational syncing: Our approach makes it possible to automatically sync-lock to one of the pulsed lasers, effectively freezing it in time, and allowing us to observe how the other pulsed lasers drift relative to it. Figure F.7 (row 1) illustrates this first-of-its-kind visualization, where the “magenta” pulsed laser is frozen, while the other two lasers move forward in space as time progresses. To freeze a pulse, we select the same relative time across different periods, aligning with the laser’s cycle. However, the other pulses, having different periods, appear to propagate at varying speeds relative to the synchronized laser. An intriguing effect occurs when we sync-lock to the pulsed laser with the shortest period. Since this pulse repeats more frequently, it appears stationary while the slower pulses propagate in the opposite direction. This phenomenon offers a novel visualization of aliasing in the transient domain, showing an optical illusion similar to how a fan can appear to spin backward.

F.2 Pulse-delay map estimation

We provide additional details and results for the pulse-delay maps for the scene in Figure 1.

F.2.1 Pulse-delay map estimation & frequency detection error

We examine the sensitivity of our estimated frequency on the pulse delay estimation. Figure F.8 compares changes in the pulse-delay estimation resulting from frequency estimation changes ranging from 1 mHz to 10 Hz for frequencies 9.998, 9.999 and 10.0 MHz. Figure F.9 shows the mean and median pulse delay error versus the frequency errors. While we do not have access to ground truth frequencies and pulse delays, we observe that even a small change in the estimated frequency on the order of a few hertz can cause nanosecond-scale shifts to the pulse delay map, which corresponds to several metres in the path length. This highlights the importance of requiring mHz-level accuracy in estimating the pulse repetition frequencies of the pulsed lasers for our approach, otherwise the pulse-delay maps will contain significant errors. Finally, this range of errors is incorporated into the simulations in Section J to evaluate the robustness of the geometric optimization.

F.3 Opportunistic ToF over room-scale scenes with shadows & occlusions

Metrics for assessing 3D geometry recovery: Since we did not have access to ground truth geometry measurements, we evaluated the accuracy of the 3D reconstruction with other well-established geometric primitives. Specifically, we fit planes to the top and right walls in the scene by segmenting pixels based on their estimated depths. Pixels with depths greater than 3 meters in the left portion of the scene are extracted to represent the front wall. Similarly, pixels with depths greater than 2.1 meters in the right portion of the scene are extracted to represent the right wall. The top wall and right wall contain 66,827 and 22,442 segmented pixels, respectively (Figure F.10). We apply RANSAC [8] with an inlier threshold of 1 cm and a maximum of 1000 iterations to both sets of segmented pixels to compute the inlier ratio and plane-fitting RMSE. The angle between the two walls is calculated as the dot product of the normals of the fitted planes.

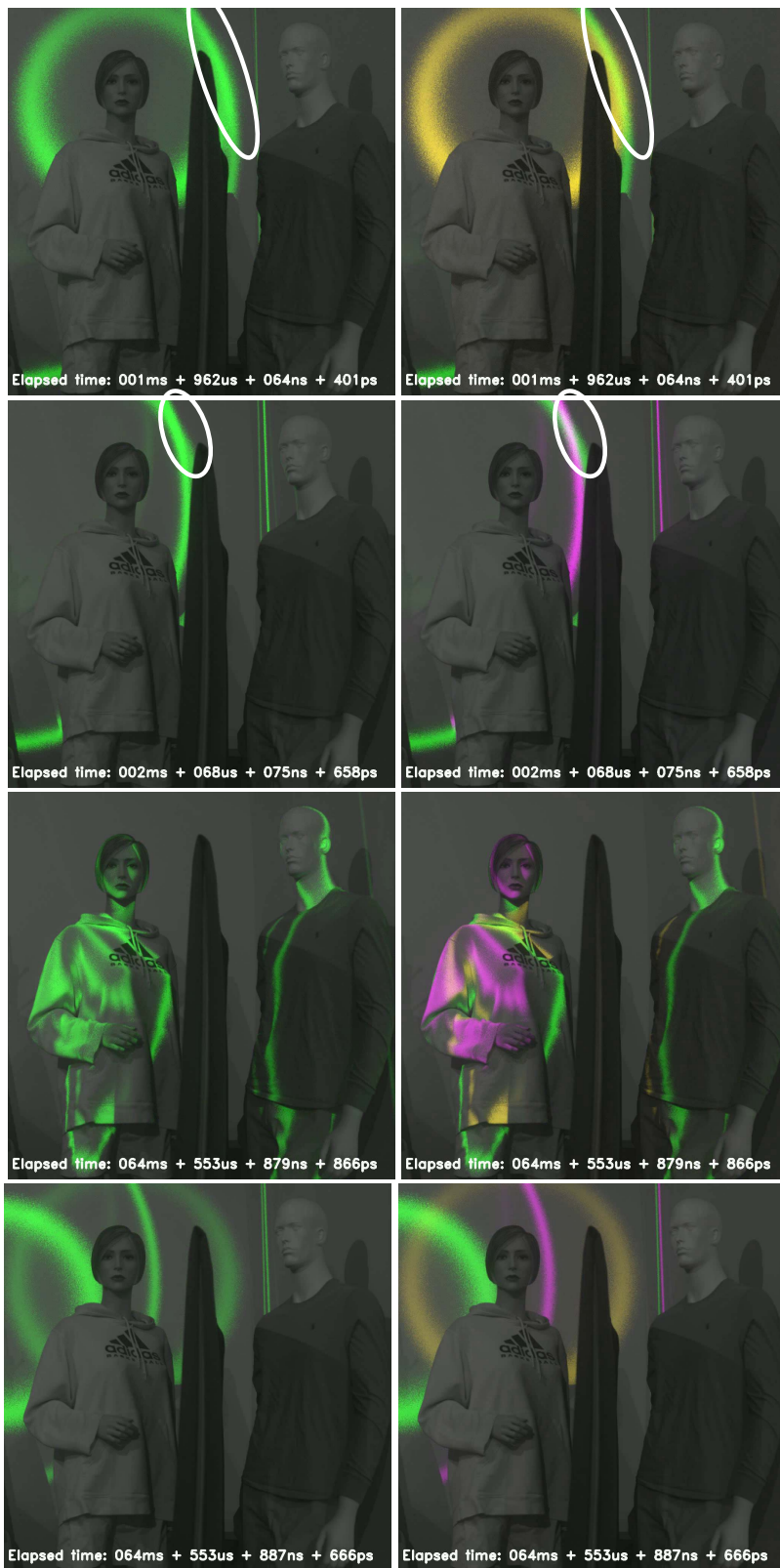


Figure F.6: Post capture laser separation. **Left:** Transient images without separated pulsed laser contributions from our approach. **Right:** Transient images with separated pulsed laser contributions from our approach.

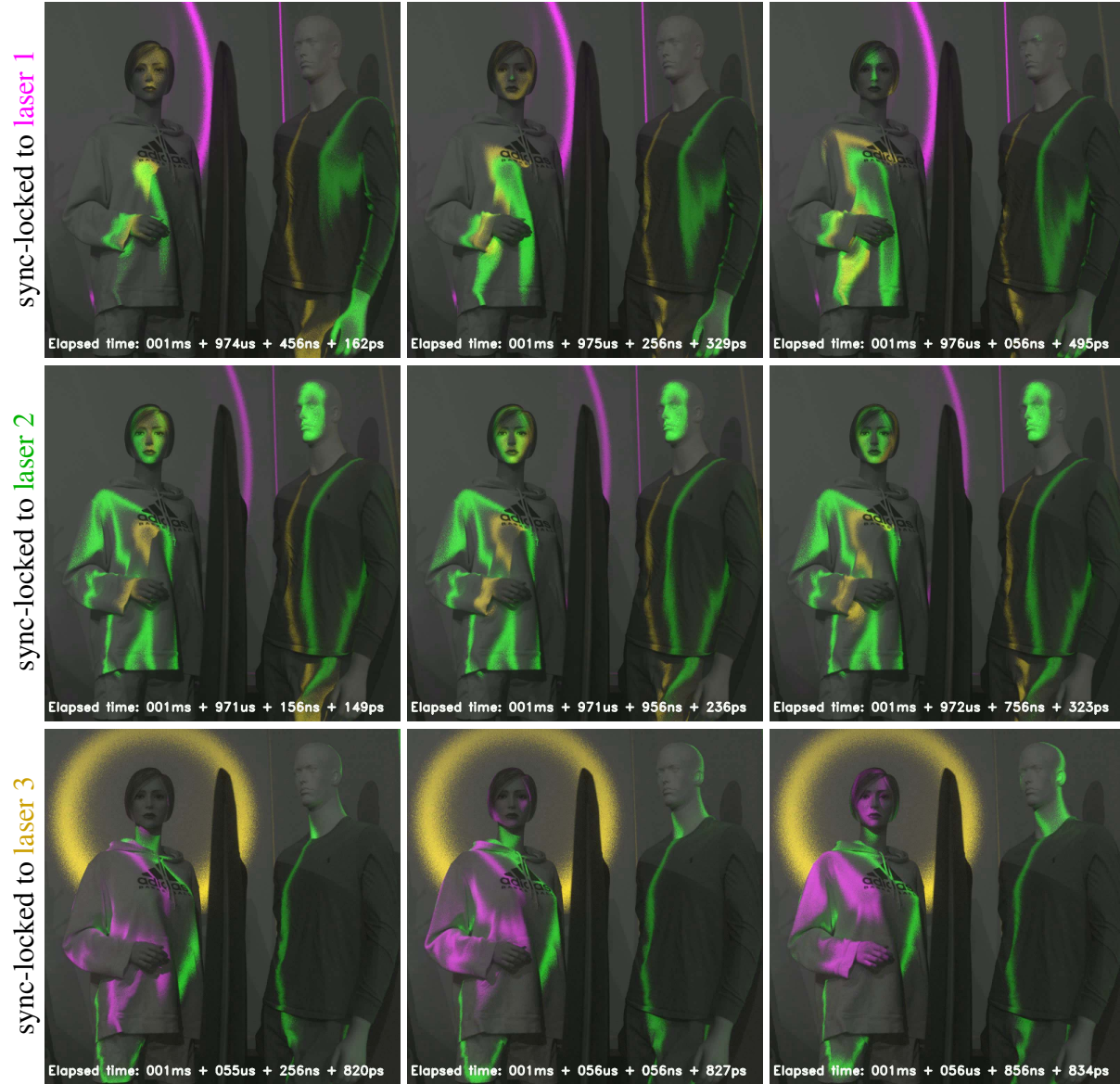


Figure F.7: Computational Sync-Lock. **Row 1:** Sync-locked to laser 1 (“magenta”), freezing its pulse in time while the other two lasers propagate forward. **Row 2:** Sync-locked to laser 2 (“green”), causing the “magenta” pulse to appear as though it is moving backward due to its longer period. **Row 3:** Sync-locked to laser 3 (“yellow”), making the other two lasers appear to move backward because of their longer periods.

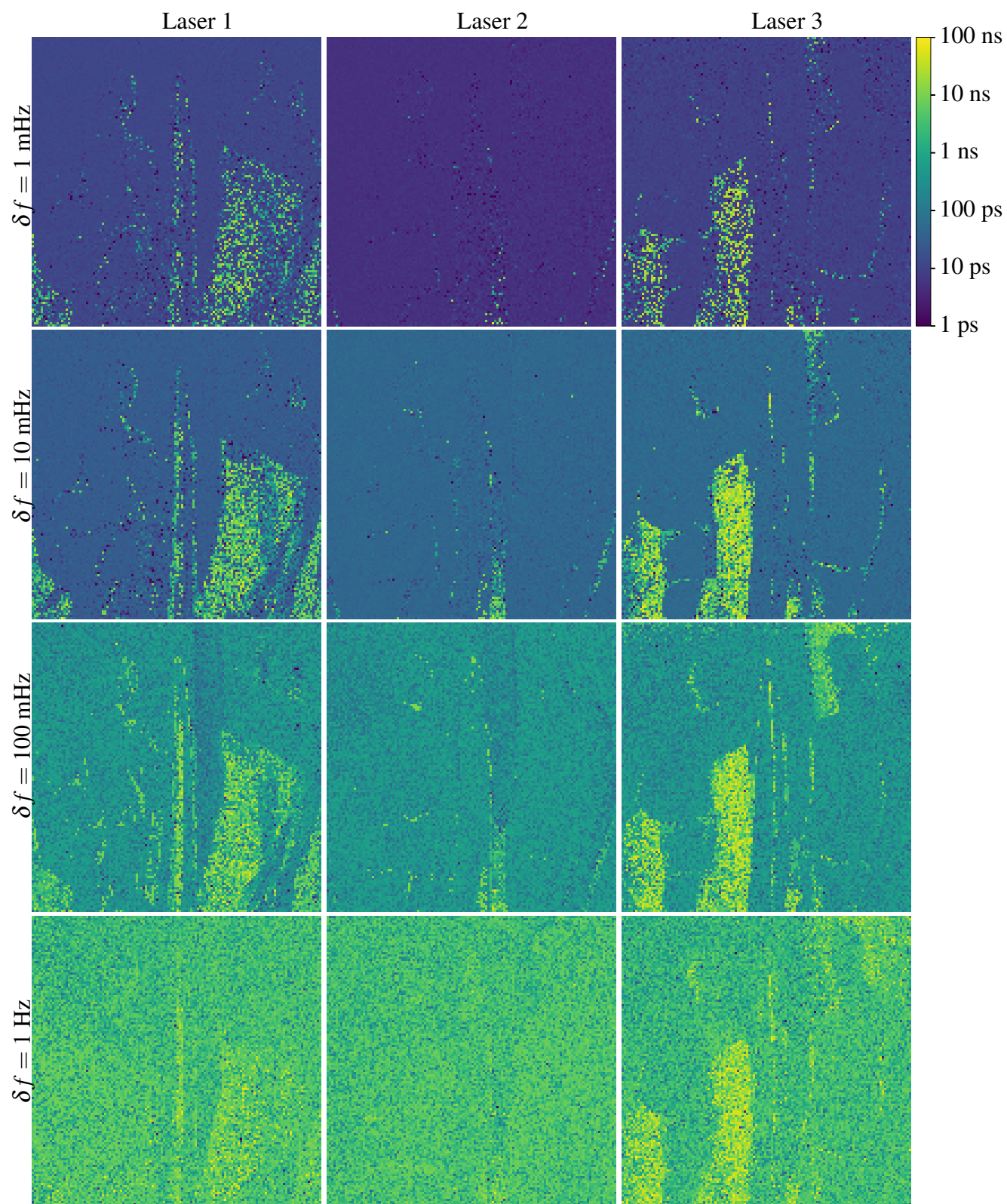


Figure F.8: Pulse-delay error for varying frequency estimation errors for three different frequencies.

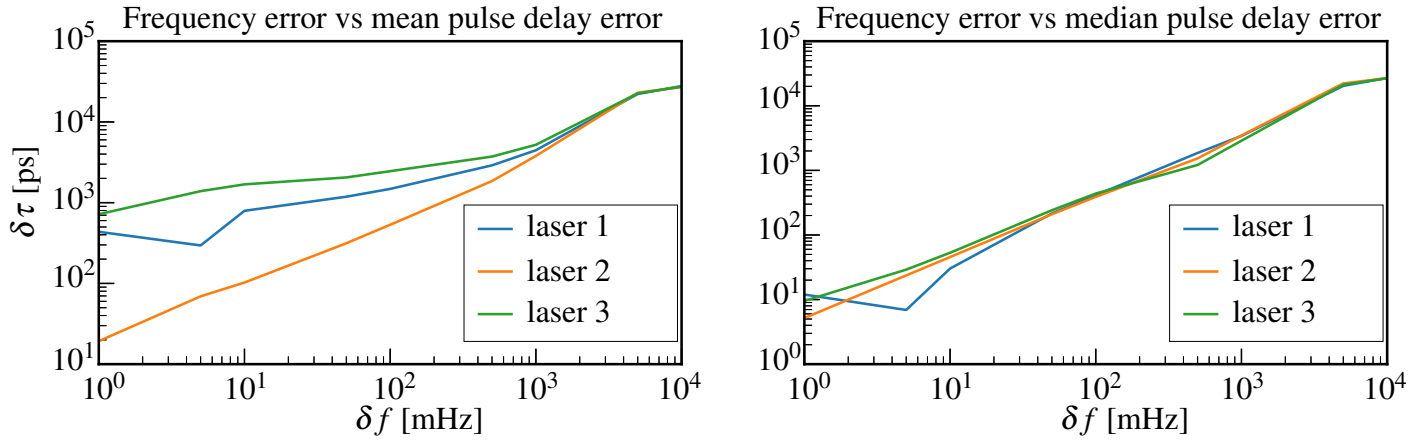


Figure F.9: Mean (left) and median (right) pulse-delay error versus frequency estimation error.

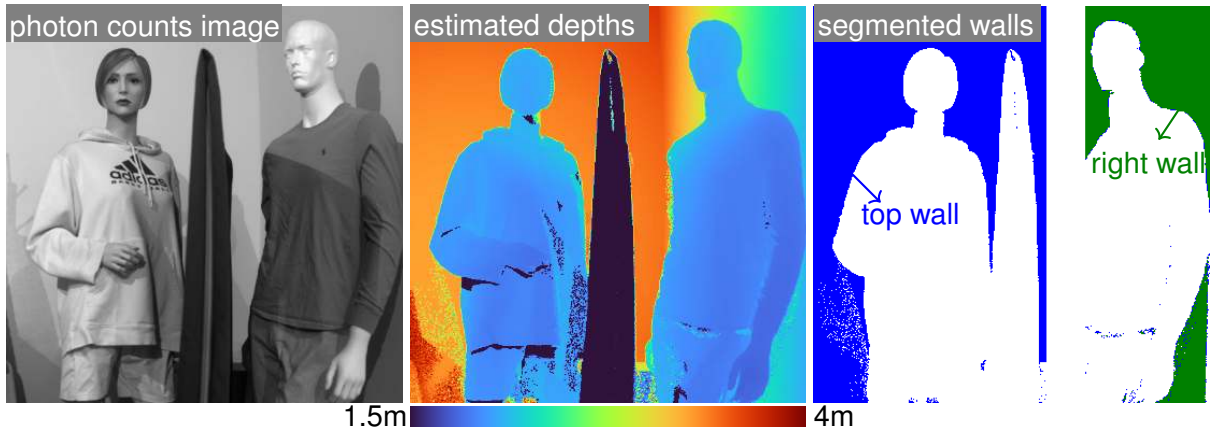


Figure F.10: **Left:** Intensity image of the scene, gamma-corrected. **Middle:** Estimated depth map from our approach. **Right:** Wall segmentation using depth-based thresholding.

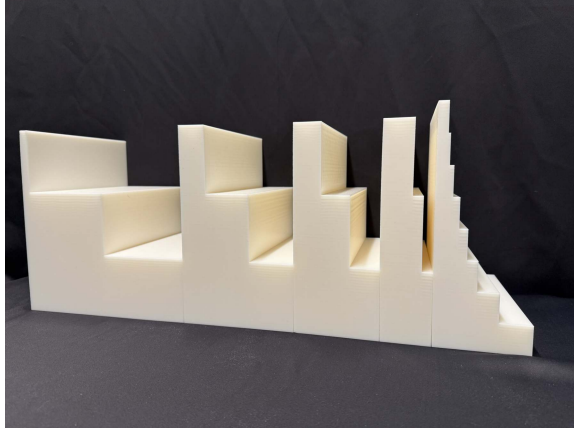


Figure G.11: From left to right, the first four staircases have step sizes 10 cm, 5 cm, 3 cm and 1 cm respectively. The last one has steps of 3 cm, 2 cm, 1 cm, 8 mm, 6 mm, 4 mm, and 2 mm.

G Experiments II: Quantitative Evaluation of Opportunistic ToF over Room-Scale Scenes

G.1 3D-printed scene

Ground truth 3D objects: We 3D-printed four staircases with equal step sizes of 10 cm, 5 cm, 3 cm, and 1 cm. Each camera-facing staircase face measures 20 cm x 7.5 cm. Additionally, we printed another staircase with step sizes of 3 cm, 2 cm, 1 cm, 8 mm, 6 mm, 4 mm, and 2 mm, with camera-facing faces measuring 20 cm x 3 cm. Figure G.11 shows the printed objects. The 3D printer had a precision of 0.2 mm.

Scene: The scene is shown in Figure 7 (row 3, top). It consists of five 3D printed models side-by-side at approximately 3 meters away from the SPAD. We scanned 114×407 points with $t_{\text{exp}} = 0.1$ s.

3D accuracy metrics: We segment each individual staircase based on the photon count image of the scene. We then register the reconstructed point cloud for each staircase with the ground truth staircase point clouds. The ground-truth point clouds are sampled from the CAD model of the staircases and contain the same number of points as the reconstructed staircases. Registration is performed using ICP with maximum correspondence distance of 0.01, voxel size of 0.001, initialization set to the identity matrix, tolerance of 10^{-9} and maximum number of iterations set to 5000. We summarize the overall inlier ratio and RMSE of 5 staircase registrations from 3 to 8 light sources in Table G.1. We show pixel-wise RMSE error maps in Figure G.12. Our reconstructions consistently achieve millimeter-level accuracy, regardless of the number of light sources used.

Additional results: We captured the same scene using a Kinect Azure [3]. Figure G.13 shows the reconstructed point cloud from Kinect, which is unable to accurately reconstruct the staircase with varying step sizes. Our method consistently outperforms the Kinect at both SBR levels (Figure 7 row 3).

G.2 Accuracy versus number of lasers for Figure 1 scene

We performed an additional experiment in which we add two more pulsed lasers to the scene with repetition frequencies of 9.996 MHz and 9.997 MHz. Using these, we perform frequency detection (Algorithms 1—6), pulse-delay estimation (Algorithm 7), and geometric optimization (Section B.2) to

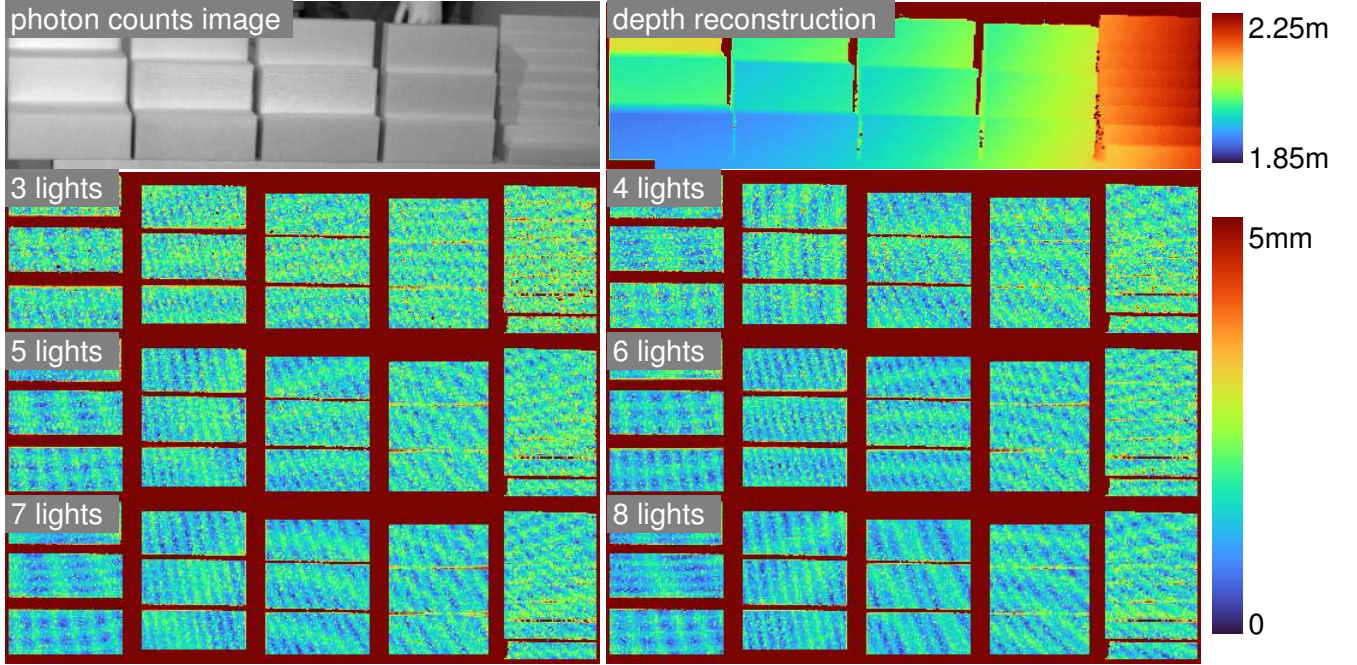


Figure G.12: Pixel-wise depth error maps for 3 to 8 light sources.

	10cm block	5cm block	3cm block	1cm block	mixed block	overall
3 pulsed lasers	2.8mm	2.5mm	2.4mm	2.1mm	2.5mm	2.3mm
4 pulsed lasers	2.0mm	2.1mm	2.0mm	1.9mm	2.2mm	2.1mm
5 pulsed lasers	2.0mm	2.0mm	1.9mm	1.8mm	2.1mm	2.0mm
6 pulsed lasers	2.2mm	2.0mm	1.9mm	1.7mm	2.1mm	2.0mm
7 pulsed lasers	1.8mm	1.8mm	1.8mm	1.7mm	2.0mm	1.9mm
8 pulsed lasers	1.7mm	1.8mm	1.8mm	1.6mm	2.0mm	1.8mm

Table G.1: RMSE of 5 staircase registrations from 3 to 8 light sources.

# of pulsed lasers	Top wall plane fitting		Right wall plane fitting		Angle between fitted planes [°]
	RMSE [mm]	%Inlier	RMSE [mm]	%Inlier	
3	3.4	93.2	3.6	93.2	90.42
4	3.0	89.1	3.4	89.4	90.20
5	3.0	87.8	2.7	89.4	90.20

Table G.2: Geometry reconstruction accuracy under different number of light sources evaluated with plane fitting.

recover depth, 3D locations of the pulsed lasers, and clock offsets.

We report quantitative metrics for plane fitting and the angle between fitted planes with varying numbers of pulsed lasers in Table G.2. As the number of pulsed lasers increases, the RMSE values for the top wall and right wall show slight improvement, particularly for the right wall, where the RMSE

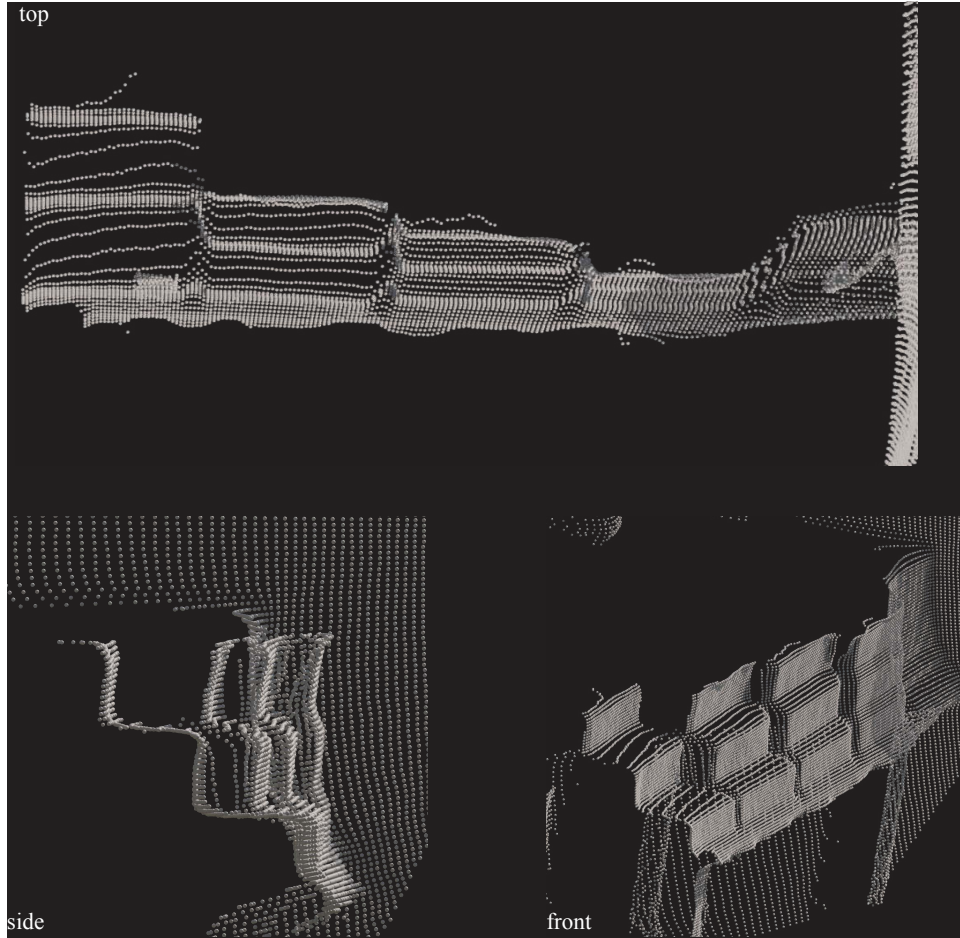


Figure G.13: 3D reconstruction from Kinect Azure [3] for the scene in Figure 7 (row 3).

decreases from 3.6 mm to 2.7 mm. The angle between the fitted planes remains consistent across all configurations, maintaining a value close to 90.2 degrees.

Optimization fails with only two light sources because occlusions result in an insufficient number of pixels observing both light sources (Figure G.14). Figure G.15 shows the quality of the reconstructed point cloud under varying numbers of pulsed lasers in the scene. Table G.3 shows the RMSE values for registering the reconstructed point clouds. Overall, our reconstructions remain consistent even with one-tenth of the exposure time. For comparison, Figure G.16 (right) shows the point cloud captured by a Kinect Azure from a slightly different viewpoint, but at approximately the same distance from the scene (Figure G.16 (left)).

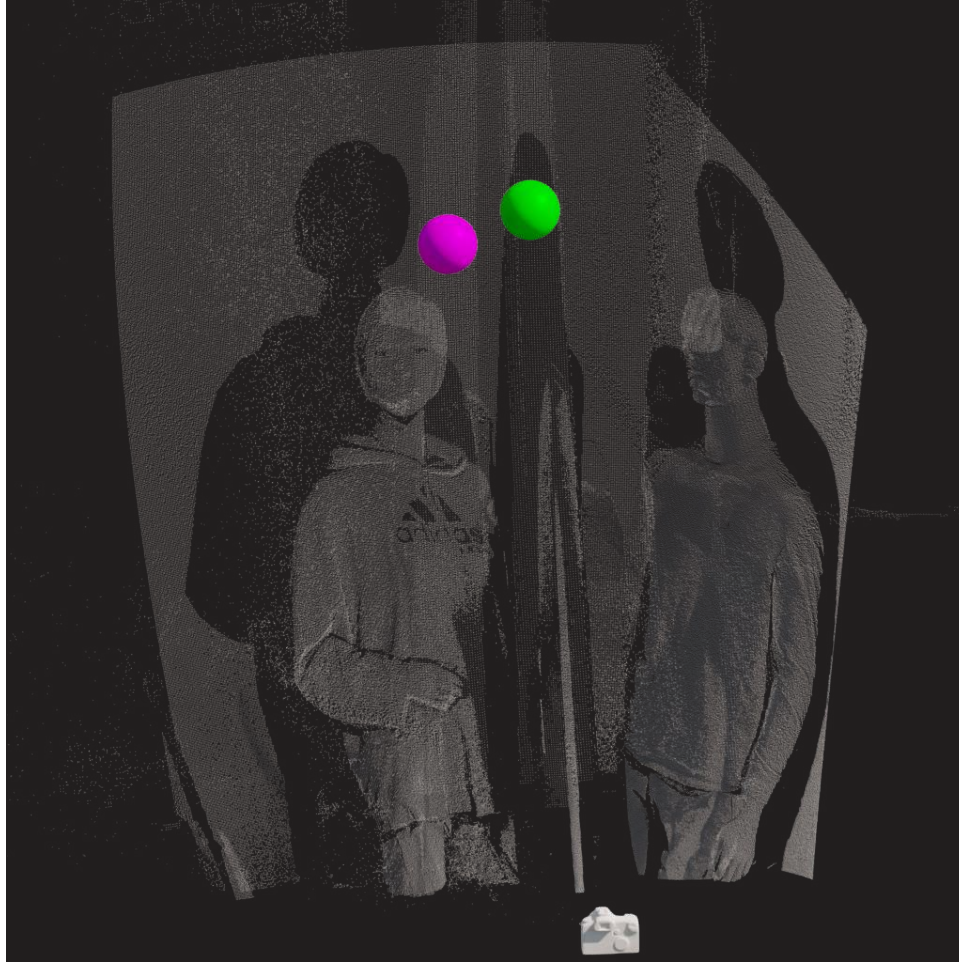


Figure G.14: Optimization fails to converge with two pulsed lasers, as shown by the estimated pulsed lasers being too close in 3D space and the background surfaces appearing curved.

Source point cloud	Target point cloud	%Inlier	RMSE[mm]
3 pulsed lasers, 0.01s exposure	3 pulsed lasers, 0.1s exposure	95.52	2.6
4 pulsed lasers, 0.1s exposure	3 pulsed lasers, 0.1s exposure	91.9	3.6
5 pulsed lasers, 0.1s exposure	3 pulsed lasers, 0.1s exposure	90.1	3.0
5 pulsed lasers, 0.1s exposure	4 pulsed lasers, 0.1s exposure	93.1	2.6

Table G.3: Registration of reconstructed point clouds under varying number of light sources and exposures using ICP algorithm.

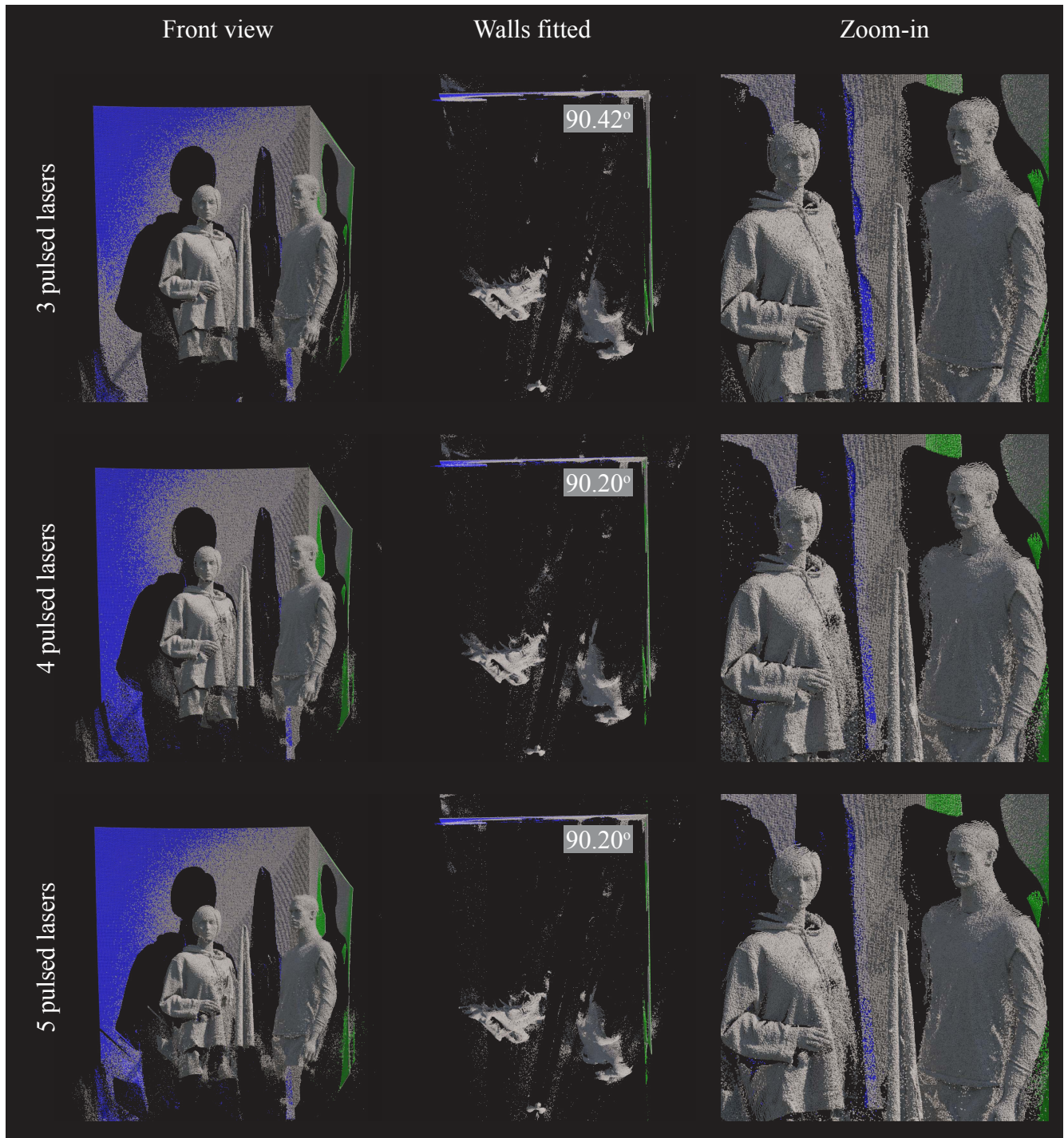


Figure G.15: 3D reconstruction with a varying number of pulsed lasers. The fitted top wall is shown in blue, while the fitted right wall is shown in green.

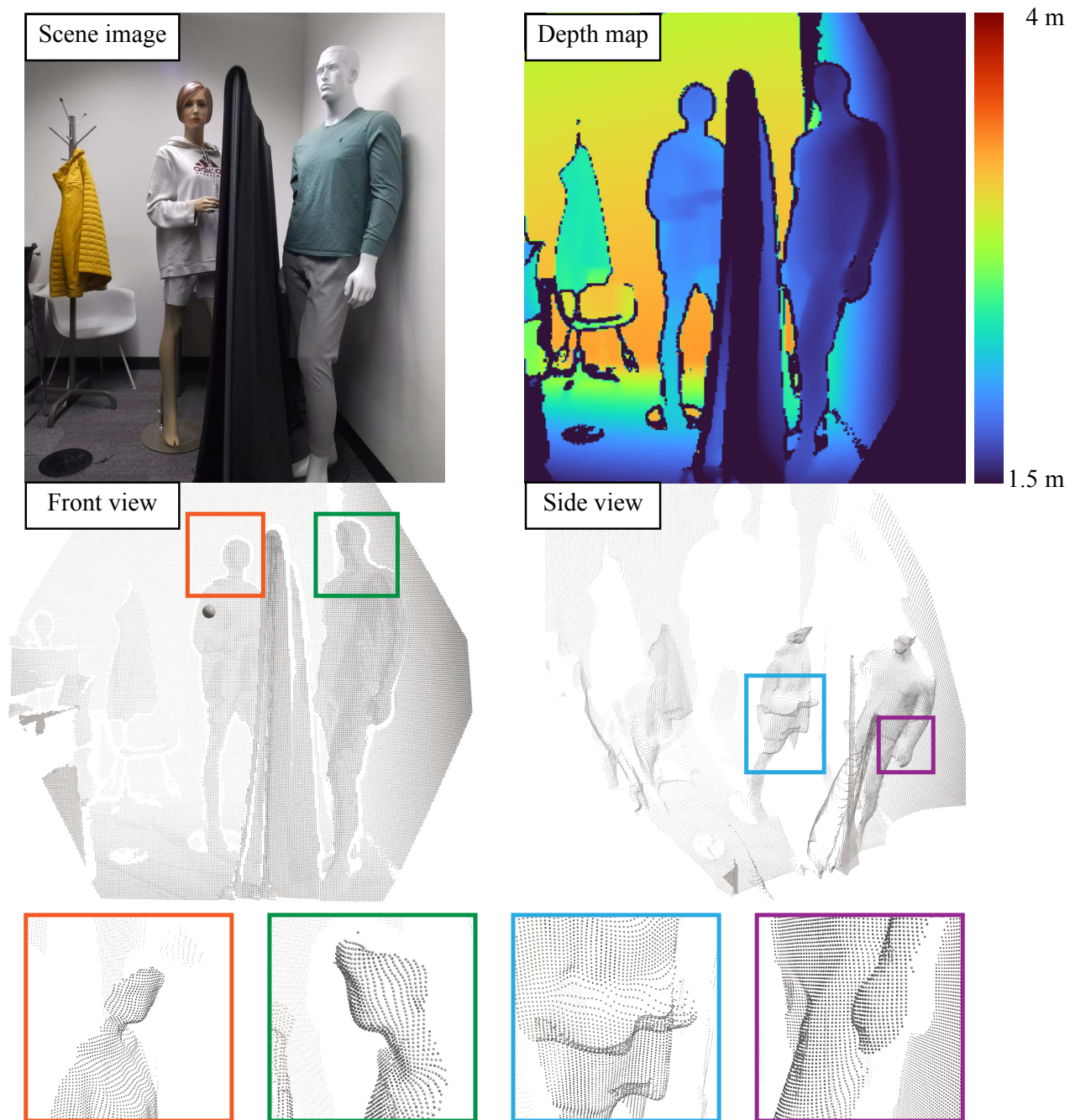


Figure G.16: 3D reconstruction by Kinect Azure [3] for the scene in Figure 1.

H Experiments III: Opportunistic ToF under Strong Ambient Light

H.1 Emulating ambient light

SBR calculation: We define the SBR β to be the ratio of the number of signal photons to background photons. For the case of three lasers and no ambient light, the SBR is 0.5. When emulating ambient light, we report the desired SBR level (i.e., the incident flux) rather than the actual SBR, as SPAD dead time can introduce non-linear effects that may slightly impact the final SBR level. Such effects were minor in our experiments: a desired SBR level of 0.01 resulted in an average SBR of 0.0103.

Controlled low-SBR conditions: In order to evaluate performance under many levels of ambient light for low SBR and dead-time impacted conditions, we combined experimentally-acquired timestamp data with simulated timestamps from ambient light photons. Instead of capturing ambient photons with ceiling lights turned on, we simulate ambient light corresponding to a DC point light source in the scene. Specifically, to simulate a desired SBR level of β , we compute the average flux $\frac{1}{L} \sum_{l=1}^L \phi_l(\mathbf{q})$ at each pixel, i.e., the per-pixel laser photon count divided by the exposure time. We simulate ambient timestamps from a homogeneous Poisson process with rate function $\frac{\beta}{L} \sum_{l=1}^L \phi_l(\mathbf{q})$. We then superimpose the laser photon stream with the ambient photon stream and enforce a dead time of 231 ns. By default, the laser photons have an SBR of around 0.5 without the inclusion of ambient photons. Our results are expressed in terms of the average SBR across all pixels.

H.2 Precision of frequency localization as a function of SBR

We conduct an experiment in which we assess the effect of signal-background ratio (SBR) on our detection method (Algorithms 1–6).

Laser details: The laser frequencies are 9,998 MHz, 9,999 MHz, and 10.000 MHz. Our exposure time t_{exp} was 100 milliseconds.

Experimental procedure: We randomly select 30 pixels from the scene of Figure 1 and aggregate timestamps from a 10-by-10 neighbourhood at each pixel. For each neighbourhood, we evaluate our ability to detect the laser frequencies for SBR levels of 0.5, 0.1, 0.01, and 0.003, resulting in 120 random trials in total. For each trial, we ran our detection method (Algorithms 1–6) and consider a laser frequency to be successfully detected if our method returns a frequency within a conservative bound of 500 Hz of the laser’s frequency.

Discussion: We report the percentage of trials for which the laser frequencies were estimated within 1 mHz, 10 mHz, and 100 mHz in Table H.1 for each light source. We observe that in high-SBR scenarios (0.5 and 0.1), we are able to estimate the laser frequencies to within sub-mHz for a majority of the trials, with at least 80% of the trials achieving mHz precision. In the low-SBR case, the accuracy suffered, but the majority of trials still returned mHz-precise frequency estimates. No spurious frequencies were returned by our method in any of the 120 trials.

H.3 mm-accurate 3D reconstruction under strong ambient light

We evaluate our method’s 3D recovery accuracy in 0.01 SBR level. We emulated ambient photons for the scene in Figure 7 (row 3) according to Section H.1. Figure H.17 shows the percentage of pixels with depth error less than 1, 2, and 3 mm respectively. Our method consistently reconstructs the same number of points with an error of less than 3 mm, even when the SBR is reduced by a factor of 50.

	SBR	%<1 mHz	%<10 mHz	%<100 mHz	% detected
Laser 1	0.5	63.3	80	100	100
	0.1	56.7	93.3	96.7	100
	0.01	36.7	66.7	83.3	86.7
	0.003	33.3	60	66.7	66.7
Laser 2	0.5	96.7	100	100	100
	0.1	90.0	100	100	100
	0.01	50	100	100	100
	0.003	36.7	90	100	100
Laser 3	0.5	73.3	86.7	100	100
	0.1	46.7	86.7	100	100
	0.01	16.7	53.3	70	80
	0.003	6.7	20	40	46.7

Table H.1: Frequency estimation accuracy across different SBR levels for the scene in Figure 1. Our method consistently detects the pulse repetition frequencies of the pulsed lasers with mHz precision even in 0.003 SBR.

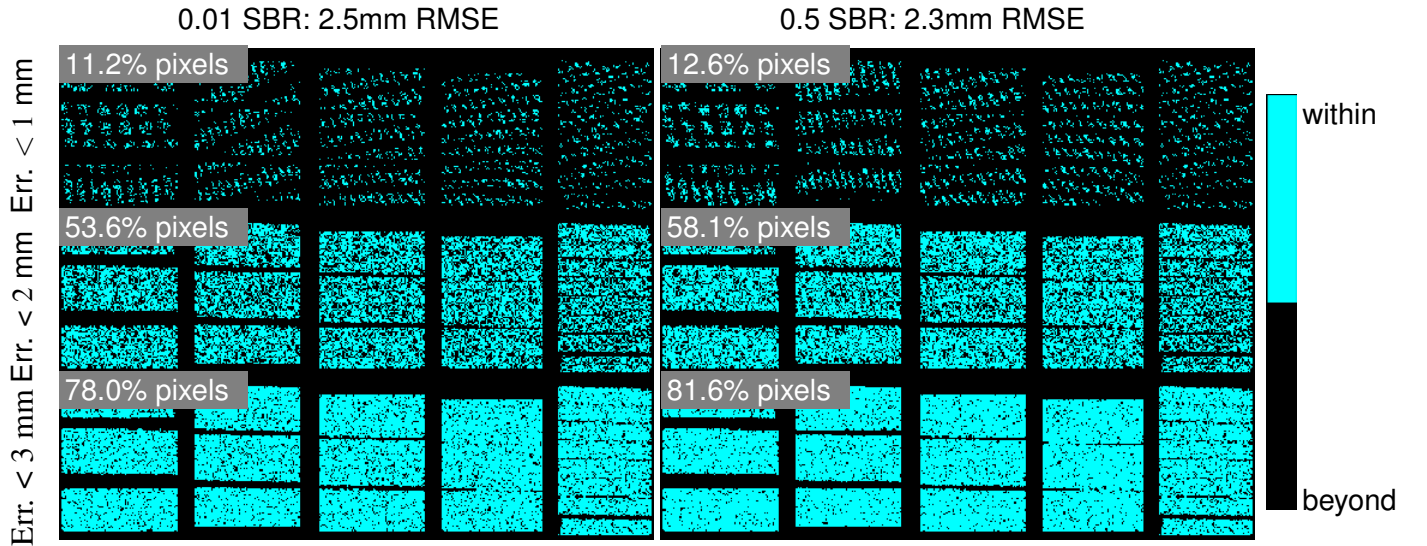


Figure H.17: Pixel-wise depth error maps for 0.01 and 0.5 SBR with thresholds 1 mm, 2 mm and 3 mm.

H.4 Low-SBR experiment for scene in Figure 1

We additionally apply our method to the scene in Figure 1 under low SBR settings.

Scene and illumination: We use the same scene and illumination conditions outlined in Section F.1.

Timestamp stream generation: We use the procedure outlined in Section H.1 to generate timestamps of the Figure 1 scene with a desired SBR of 0.01. The average photon count at every pixel after adding ambient photons was 175384 photons, which is about 40.5% of the maximum photon count for a dead time of 231 ns. The average laser photon count per pixel was 5264. Compared to the high-SBR case, our timestamp streams contained on average 17 times more photons in total—but half as many laser photons, due to SPAD dead time effects.

Frequency scanning and detection: We aggregate timestamps from a single 10×10 patch and run our frequency detection method (Algorithms 1–6) on the aggregated timestamp stream. 80 candidate frequencies were above the CFAR bound and three laser frequencies passed the CFAR sinc comb detector. The whole process from frequency scanning to detection takes 10.3 hours.³

Pulse-delay map estimation: We use the pulse-delay map estimation method (Algorithm 7) to recover the pulse delay per-pixel for each laser. For efficiency, we thin our timestamp data by culling 90% of the photons, resulting in average photon count of 17.5k during harmonic probing. Additionally, we recover the photon count images from the recovered pulse trains.

Figure H.18 shows the photon count images and the pulse-delay maps.

Discussion: We observe that the low-SBR photon count images (Figure H.18 (row 1)) depict good separation of the laser photon counts, similar to the ones shown in Figure 1 in the main paper. However, these images are noisier as a result of the increased ambient light level, and we also observe some artifacts resulting from the ambient photons bleeding into the photon count images, highlighted in Figure H.18 (top row, red ellipses). Since the laser frequencies are well-estimated within a few mHz even at an SBR of 0.01 (Table J.4), we observe that the pulse-delay maps (Figure H.18 (row 2)) are estimated with high accuracy, differing from the high-SBR pulse-delay maps by just tens of picoseconds (Figure H.18 (row 3)) in well-lit regions. We do observe that low-flux regions as well as shadowed regions have significantly larger pulse delay errors (> 3 ns), likely due to the reduction in signal photons in those areas.

Please refer to the supplementary video for a fly-by of the scene in Figure 1 captured under low-sbr conditions.

³The increase in processing time compared to our experiment in Section F is due to the much larger number of photons detected.

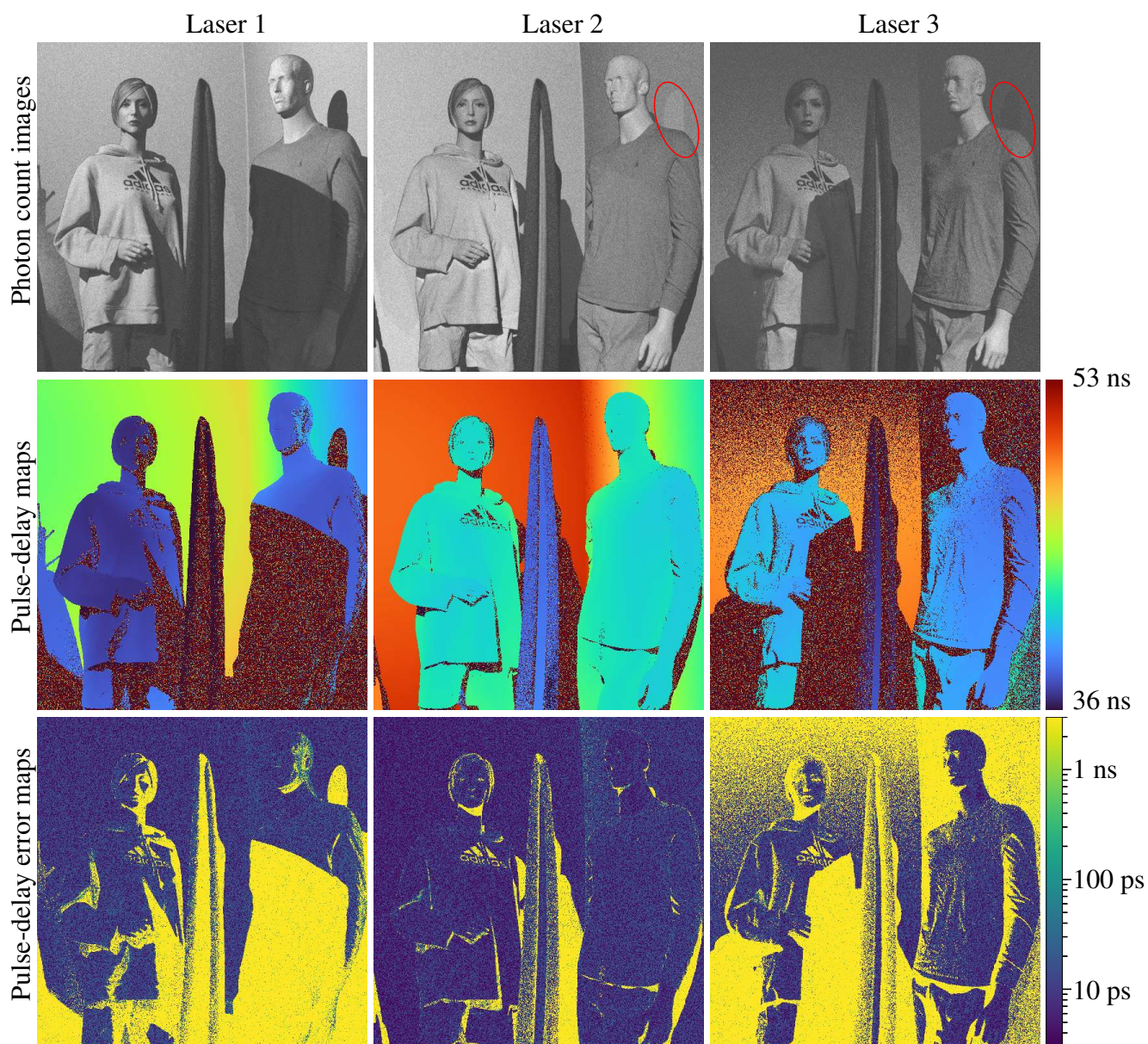


Figure H.18: Figure 1 scene under SBR of 0.01. **Row 1:** Photon count images corresponding to each pulsed laser. **Row 2:** Pulse-delay maps. **Row 3:** Pulse-delay error map between high-SBR pulse delays and low-SBR pulse delays (log-scale).

I Experiments IV: Dynamic Acquisition

In this section, we provide additional details and results for the experiment shown in Figure 7 (row 4) where both the camera and the laser sources change position from one snapshot to the next.

Scene: The scene is shown in Figure 7 (row 4, left) and includes a mannequin wearing a jacket with millimeter-scale surface features, a figure holding a bowl, and 3D-printed staircases with step sizes 1 cm and 3 cm. We captured six different snapshots following the acquisition procedure described in Section E.3. The camera moved approximately 0.3 meters between snapshots, and the pulsed lasers changed positions across snapshots as well.

Accuracy assessment: Since we have access to the ground truth staircase models, we register the reconstructed point cloud of each staircase in each view with the corresponding ground truth shape as described in Section G.1. The average registration RMSE error for the 1 cm and 3 cm staircase models is 2.6 mm and 2.0 mm respectively, across all views. The average distance between estimated camera positions from ICP is 32.7 cm.

J Simulation-Based Evaluation of Opportunistic ToF

In this section we use simulated scenes to evaluate the performance of individual components of opportunistic ToF against known ground truth.

J.1 Simulation details

Simulation scenes and renderings: Our simulations used 6 different room-scale scenes in the Mitsuba format from the collection of Bitterli [5], shown in Figure J.19.⁴ The depth range of the scenes varied from 4 m to 15 m. We use Mitsuba version 3.1.1 to extract the ground-truth 3D positions of each pixel, camera calibration parameters, depth maps, and pulse-delay maps for various randomly-sampled three-laser configurations in each scene. To these pulse-delay maps, we manually add random clock offsets for each light source, with the offsets uniformly sampled from the range [0, maximum pulse delay at each source]. Unless otherwise stated, we use 128×128 pixels and $t_{\text{exp}} = 0.1$ s.

Photon timestamp generation of 2D room scenes: For realistic rendering of the 2D room scenes, we convert the Mitsuba 3 files into a Mitsuba-2-compatible format and use a transient renderer [14] for Mitsuba 2 to render a 4,000 bin histogram at every pixel for each laser. For each pixel, we then use the histogram to sample a timestamp stream for each laser using the thinning method [11]. We set the average flux to be 10 kilophotons per second, similar to the lighting conditions of our real experiments. We combine the timestamp stream of each pulsed laser and enforce a dead time of 231 ns to remove timestamps that arrive during the SPAD’s dead time period. The remaining timestamps are quantized to a resolution of four picoseconds.

Pulsed laser model from [14]: We set the laser pulse profile from [14] to a FWHM of 235 ps. The laser pulse profile is approximately Gaussian, but our use of the transient renderer enables us to capture multi-path transients. In our room-scene simulations, we use 7.499, 7.5, and 7.501 MHz as the pulse repetition frequencies of the three lasers, respectively.

Photon timestamp generation for simulations involving a single pixel: In the simulations of Section J.3, which involve timestamps collected at a single pixel, we model a laser pulse as a Gaussian with a desired FWHM. Since we have an analytical expression for the pulse, we use Cinlar’s method [6] to simulate the timestamps for greater efficiency. To simulate timestamps for L lasers, we generate timestamps for each laser pulse train separately and then combine them into a single stream. We apply a dead time of 231 ns, the nominal dead time of our system, to remove photons that arrive within the SPAD’s dead time period. We then apply a timing jitter of 8 ps and quantize the timestamps to 1 ps to match the nominal specs of our experimental time-to-digital converter.

J.2 3D reconstruction and localization consistency

We evaluate the consistency of opportunistic ToF by generating timestamps for 50 trials per simulated scene for a total of 300 trials. In each trial, we vary the pulsed laser positions and clock offsets while keeping the pulse repetition frequency and average photon flux constant. We then use opportunistic ToF to detect frequencies (Algorithms 1–6), estimate pulse-delay maps (Algorithm 7), and estimate depth, laser 3D coordinates, and clock offsets (Section 6). Table J.1 reports the success rate of trials where our method converged, as well as the average error and standard deviation for the estimated variables in

⁴The scenes we use are “bathroom1” (Contemporary Bathroom by Marek), “livingroom1” (The White Room by Jay-Artist), “livingroom2” (The Modern Living Room by Wig42), “bedroom” (Bedroom by SlykDrako), “dining” (The Breakfast Room by Wig42), and “greywhiteroom” (The Grey & White Room by Wig42).

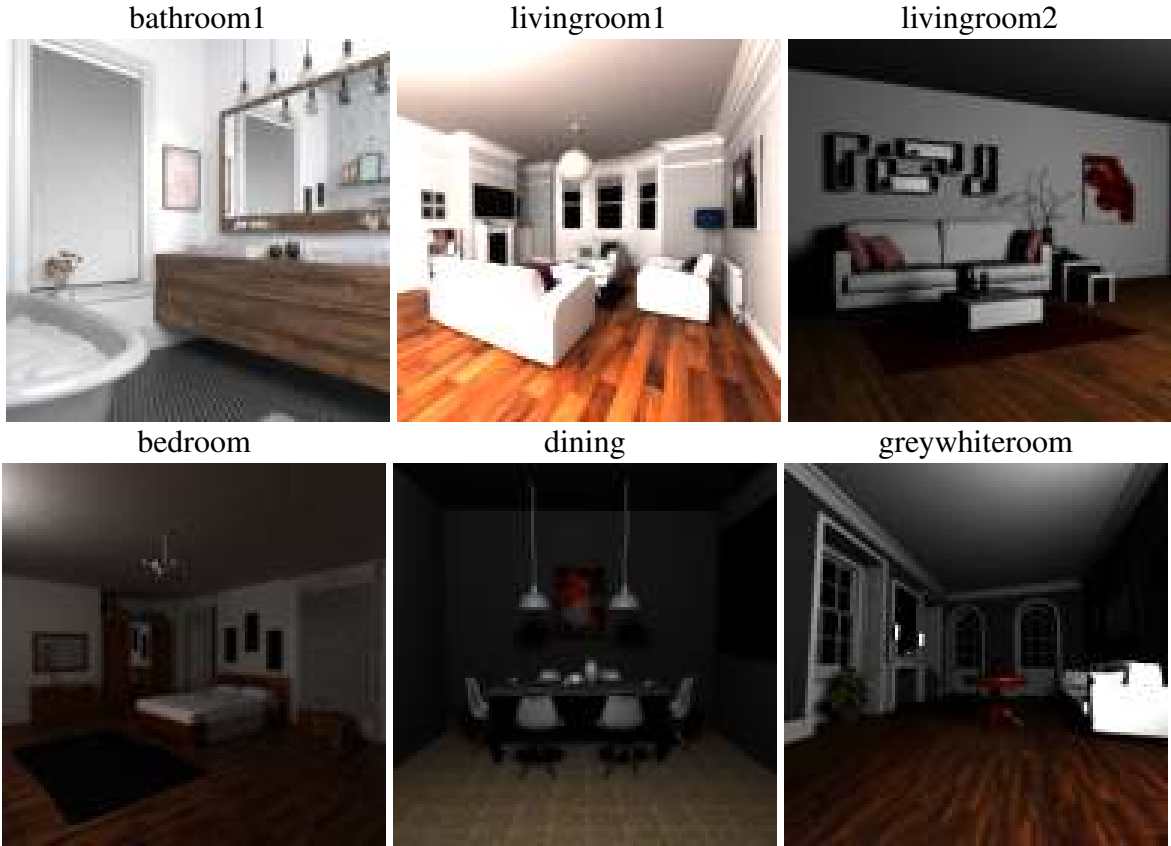


Figure J.19: Blender scenes used in simulations.

those cases. We observe that the method succeeds in roughly 94% of all trials. Failure cases occur when the lasers are positioned in locations that create significant occlusions and shadows.

% Success cases	Depth error		Pulsed laser localization error		Clock offset error	
	Average [mm]	Std. [mm]	Average [mm]	Std. [mm]	Average [mm]	std. [mm]
93.67	3.5	4.3	16.2	11.2	9.4	8.9

Table J.1: Simulation success rate and accuracy.

J.3 Frequency localization & laser detection accuracy

J.3.1 Accuracy as a function of signal strength

We assess the performance of frequency detection and estimation across different flux levels in two simulated scenes. Keeping $t_{\text{exp}} = 0.1$ s constant, we vary the average flux levels from 500 photons per second to 100 kphotons per second. Frequency detection (described in Algorithms 1–6) is performed on three distinct "superpixels" in each scene. Table J.2 shows the frequency estimation errors for the three pulsed lasers in these scenes.

J.3.2 Accuracy as a function of exposure time

As described in Section 3 of the main paper, reducing the exposure time causes blurring of each harmonic, which in turn reduces the accuracy of frequency estimation. To analyze the effect of exposure

Photons per sec.	Bathroom1 scene			
	Laser 1 error [Hz]	Laser 2 error [Hz]	Laser 3 error [Hz]	% Freq. detected
100K	0.001	0.005 (two patch detected)	0.004	88.9
50K	0.001	0.001 (two patch detected)	0.001	88.9
10K	0.005	0.019	0.004	100
5K	0.041	0.021	0.004	100
1K	0.671	0.094	0.063	100
0.5K	0.06	0.12	0.069	100

Photons per sec.	Livingroom2 scene			
	Laser 1 error [Hz]	Laser 2 error [Hz]	Laser 3 error [Hz]	% Freq. detected
100K	0.004	6e-4	3e-4	100
50K	0.003	0.002	0.003	100
10K	0.003	0.001	0.004	100
5K	0.011	0.002	0.004	100
1K	0.027	0.025	0.048	100
0.5K	0.060	0.084	0.162	100

Table J.2: Effects of varying light levels in frequency detection and estimation.

time on repetition frequency detection and estimation, we simulate a scene with three distinct pulsed lasers. Timestamps are generated following the procedure in Section J.1 for $t_{\text{exp}} = 0.1$ s. Shorter exposure times are simulated by cropping the timestamps. Frequency detection and estimation (Algorithms 1–6) is then performed on each set of timestamps. Table J.3 summarizes the frequency estimation errors for the repetition frequencies of the different pulsed lasers. Extremely short exposure times impact the detection rate of our approach. However, when a frequency is successfully detected, even at $t_{\text{exp}} = 1$ ms, the frequency error remains as low as 0.3 Hz. For reference, the expected blur of each harmonic at this exposure is 600 Hz.

J.3.3 Two-laser detection test

We perform simulations to determine the minimum separation between resolvable laser pulse repetition frequencies.

Flux function: We define the flux function to be the sum of two laser pulse trains with pulse repetition frequencies of 10 MHz and $10 + \Delta f$ MHz, respectively. We model the laser pulse as a Gaussian with 110 ps FWHM and set the amplitude of each laser pulse to be 10 kilophotons per second. The pulse delay of each laser pulse train was sampled uniformly randomly from $[0, 1/f]$ where f is its pulse repetition frequency. We use the photon timestamp generation of pulsed laser trains procedure described in Section J.1 to generate timestamp streams.

Experimental procedure: We vary the frequency separation Δf from $0.2/t_{\text{exp}}$ to $4/t_{\text{exp}}$ in increments of $0.2/t_{\text{exp}}$ and use exposure times t_{exp} of 1 second, 100 milliseconds, and 10 milliseconds. For each Δf and t_{exp} , we generated 50 timestamp streams, resulting in 1000 random trials in total. We run our detection method (Algorithms 1–6) on each timestamp stream and consider a repetition frequency to be successfully detected if the method estimates a frequency within $0.01/t_{\text{exp}}$ of the corresponding frequency. For

Exposure time	Bedroom scene			
	Laser 1 error [Hz]	Laser 2 error [Hz]	Laser 3 error [Hz]	% Freq. detected
100ms	3e-4	0.001	0.003	100
50ms	0.003	0.003	0.003	100
10ms	0.024	0.020	0.020	100
5 ms	0.025	0.173	0.068	100
1 ms	4.90 (two patch detected)	No patch detected	400 (one patch detected)	33.33
0.5ms	4.00 (one patch detected)	No patch detected	No patch detected	11.11

Exposure time	Livingroom1 scene			
	Laser 1 error [Hz]	Laser 2 error [Hz]	Laser 3 error [Hz]	% Freq. detected
100ms	0.001	0.003	3e-4	100
50ms	0.002	0.004	0.001	100
10ms	0.018	0.014 (two patch detected)	0.029	88.9
5 ms	0.073	0.060 (two patch detected)	0.029 (two patch detected)	77.8
1 ms	No patch detected	0.069 (two patch detected)	0.277 (one patch detected)	33.3
0.5ms	No patch detected	2.491 (one patch detected)	No patch detected	11.11

Table J.3: Effects of exposure time in frequency detection and estimation.

each Δf and t_{exp} , we compute the two-frequency detection probability across all trials.

Figure J.20 shows the simulation results.

Discussion: We observe that when the frequency separation is less than or equal to $1.4/t_{\text{exp}}$, our method consistently fails to detect both frequencies. Additionally, our simulations suggest that there is a phase transition where the detection probability goes from 0% to 100% and the duration of this phase transition decreases as the exposure time increases, likely due to the increased accuracy in the CFAR frequency detector as photon counts increase. These observations suggest that our frequency detection method has a minimum two-frequency resolution between $1.4/t_{\text{exp}}$ to $1.6/t_{\text{exp}}$ as we vary Δf in steps of $0.2/t_{\text{exp}}$.

We also corroborate this limit using the Rayleigh criterion limit [13]. Since the Rayleigh criterion is defined for Airy disks and not a sinc, we fit a sinc to the Airy disk by matching the FWHMs of both functions.⁵ Doing so yielded a minimum frequency separation of $1.43/t_{\text{exp}}$, which is consistent with our observed limit of $1.4/t_{\text{exp}}$ to $1.6/t_{\text{exp}}$.

We also observe that our detection method has an increased probability of detecting one of the two laser frequencies when the frequency separation is at most $0.4/t_{\text{exp}}$. For frequency separations less than $1/t_{\text{exp}}$, the two sinc lobes of the laser frequency merge into one sinc lobe centred at $10 + \Delta f/2$, *e.g.*, the mean of the two laser frequencies. In this case, while we are able to localize laser frequencies, they will be roughly $\Delta f/2$ away from the peak of the sinc lobe and therefore removed by our frequency detection method. However as $\Delta f \rightarrow 0$, the laser frequencies get closer to the peak of the sinc lobe, allowing our method to detect them, which explains the increase in successful trials.

⁵We found that the Airy disk $(2kJ_1(x/k)/x)^2 \approx \text{sinc}(x/t_{\text{exp}})$ around their lobes when $k = 3.733$ and $t_{\text{exp}} = 0.1$; J_1 is the Bessel function of the first kind of order one. Under Rayleigh criterion, $\Delta f \approx 3.8317k = 1.43/t_{\text{exp}}$ since 3.832 is the location of the first zero of J_1 .

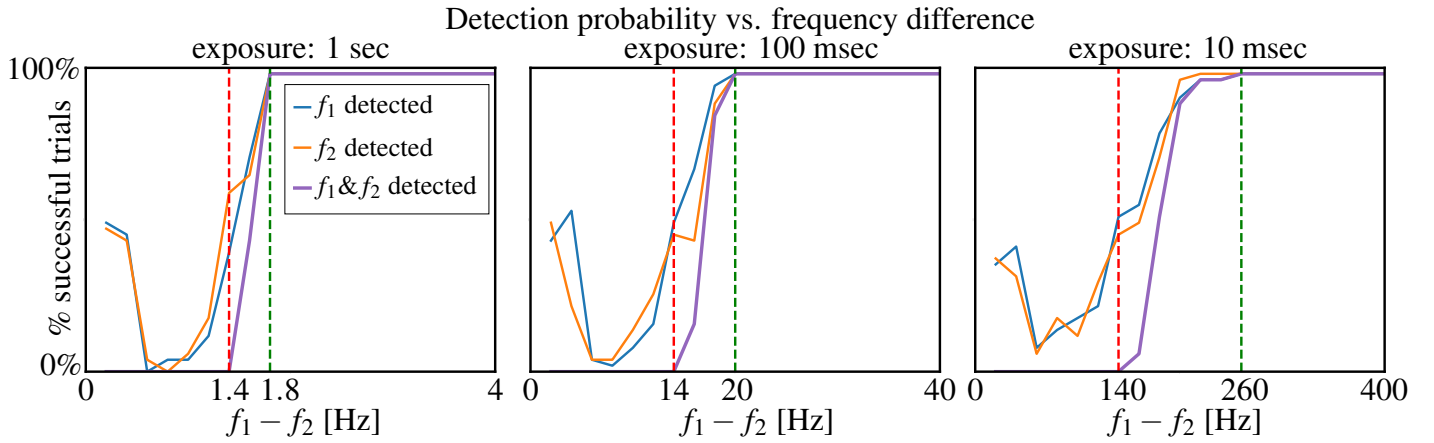


Figure J.20: Two-laser detection probability versus difference of their repetition frequencies for three exposure times.

J.4 Pulse delay error as a function of frequency error

Accurate frequency estimation is a key contribution of our approach, essential for reconstructing pulsed laser trains, estimating delay maps, and enabling passive 3D imaging. To evaluate the effect of frequency estimation errors, we set the repetition rate of the pulsed laser to 7.5 MHz and reconstruct the pulse profile using harmonic probing at each provided frequency for the scenes shown in Figure J.21. Table J.4 shows the average errors across three different scenes for repetition frequency errors ranging from 0 Hz to 1 Hz. As the error in the repetition frequency increases, the pulse FWHM becomes larger, indicating that the pulse is increasingly blurred, and the pulse-delay shifts. This occurs because even a small error in the repetition frequency introduces a shift of tens of Hz in the higher harmonics, causing them to misalign. Consequently, the harmonic components fail to “line up” properly, leading to distortions in the reconstructed pulse profile and delay map.

Freq. estimation error	Pulse FWHM [ps]			Avg. pulse-delay error $\times c$ [mm]		
	livingroom2	greywhiteroom	dining	livingroom2	greywhiteroom	dining
0	252	251	272	8.8	9.0	11.5
100 μ Hz	252	254	273	8.8	9.0	11.5
500 μ Hz	255	255	275	9.0	9.0	13.5
1 mHz	256	256	276	9.4	9.2	13.4
5 mHz	264	278	302	13.0	12.7	18.8
10 mHz	278	333	340	17.4	15.5	23.6
50 mHz	501	755	554	58.9	58.9	69.2
100mHz	1351	1753	1256	115	110	132
500mHz	16140	19981	9535	794	569	664
1Hz	28733	38349	21368	1715	1079	1328

Table J.4: Pulse FWHM and pulse-delay map error for varying pulse repetition frequency errors for three different scenes.

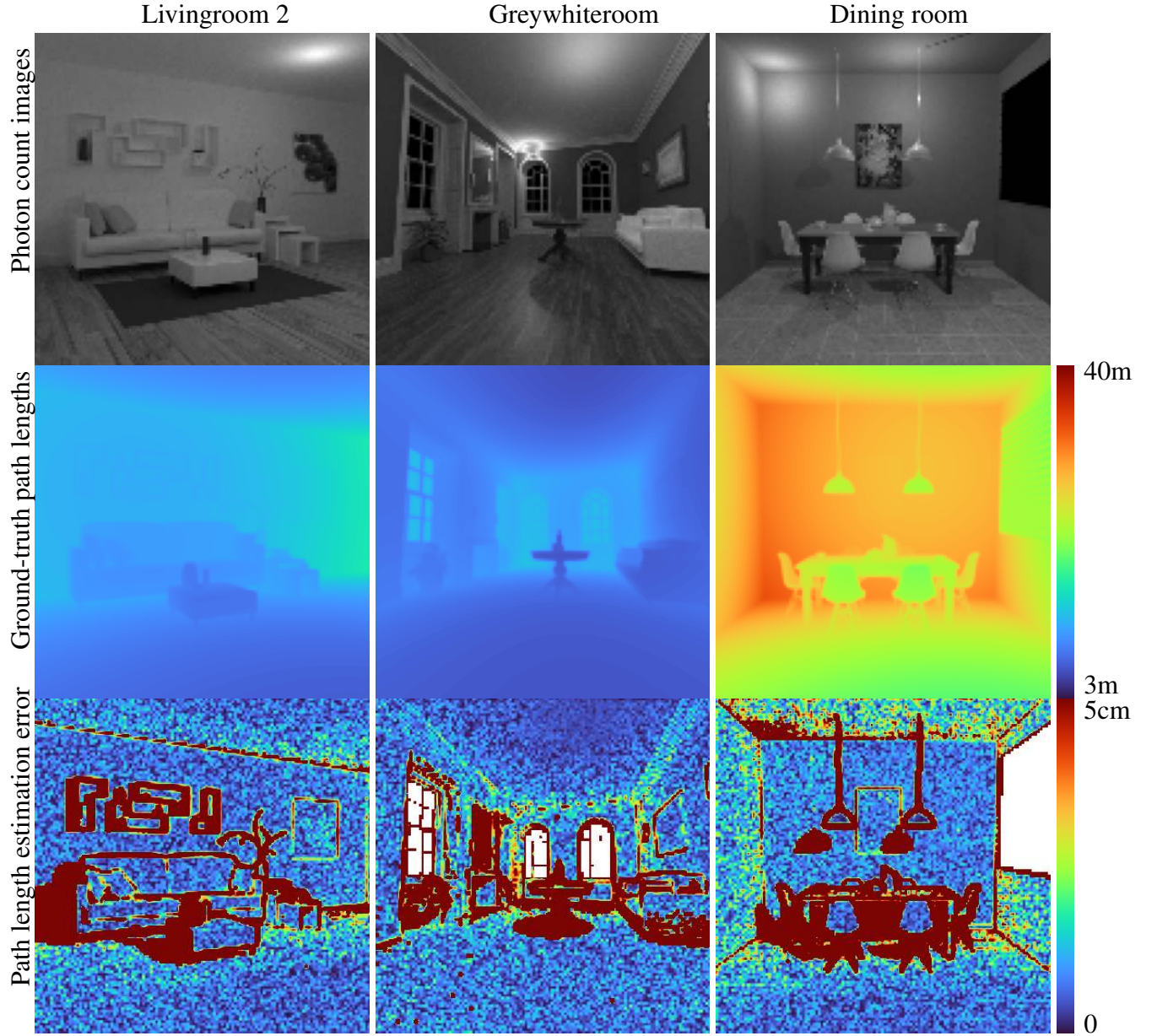


Figure J.21: Path length estimation errors in livingroom2, greywhiteroom and dining room scenes. **Top:** Intensity images (gamma-corrected). **Middle:** Ground-truth path lengths from rendering. **Bottom:** Path length estimation error against ground truth.

J.5 3D reconstruction accuracy

We assess the impact of pulse-delay errors on 3D reconstruction, pulsed laser localization, and clock offset accuracy. To do this, we perturb the ground truth simulated pulse-delay maps by adding zero-mean Gaussian noise with standard deviations ranging from 0 ns to 33 ns, as determined from the analysis in Section F.2.1, which correspond to depth errors of 0 mm to 1000 mm. Then, we run geometric optimization (Section B.2) to recover depth, pulsed-laser 3D locations and sync offsets. Table J.5 summarizes the average depth error, pulsed-laser localization error, and clock offset error across all 8 simulated scenes. Thanks to the adaptive loss function and the overdetermined nature of our system of equations, our

geometric optimization framework effectively handles significant pulse-delay errors, even those in the range of hundreds of millimeters. As noted in Section F.2.1, such pulse-delay errors are primarily caused by frequency estimation errors of hundreds of millihertz. However, as demonstrated in our real-world experiments (Section 7 of the main paper), our frequency estimation errors are much smaller, ensuring robust and accurate performance in practical scenarios.

Avg. pulse-delay error $\times c$	Avg. depth error	Average source localization error	Avg. clock offset error $\times c$
0	0.38mm	0.61mm	0.32mm
5mm	1.4mm	2.4mm	1.6mm
1 cm	2.7mm	0.82mm	2.2mm
5 cm	13mm	6.0mm	2.4mm
10 cm	29.8mm	41.7mm	16.7mm
50 cm	304mm	607mm	2153mm

Table J.5: Impact of pulse-delay estimation error in passive 3D reconstruction accuracy.

References

- [1] ISO 80000-7:2019(E). Quantities and units. Part 7: Light. Technical report, International Organization for Standardization, Geneva, CH, 2019. [17](#)
- [2] ams OSRAM Group. TMF8820/21/28 Time-of-Flight Sensor Datasheet, 2024. Accessed: 2024-11-14. [11](#)
- [3] Cyrus S Bamji, Swati Mehta, Barry Thompson, Tamer Elkhatib, Stefan Wurster, Onur Akkaya, Andrew Payne, John Godbaz, Mike Fenton, Vijay Rajasekaran, et al. Impixel 65nm bsi 320mhz demodulated tof image sensor with $3\mu\text{m}$ global shutter pixels and analog binning. In *IEEE Int. Solid-State Circuits Conf. (ISSCC)*, pages 94–96. IEEE, 2018. [25](#), [27](#), [30](#)
- [4] Jonathan T Barron. A general and adaptive robust loss function. In *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 4331–4339, 2019. [11](#), [15](#)
- [5] Benedikt Bitterli. Rendering resources, 2016. <https://benedikt-bitterli.me/resources/>. [15](#), [36](#)
- [6] Erhan Cinlar. *Introduction to stochastic processes*. Courier Corporation, 2013. [36](#)
- [7] Timothy Duff, Kathlen Kohn, Anton Leykin, and Tomas Pajdla. PLMP-point-line minimal problems in complete multi-view visibility. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 1675–1684, 2019. [12](#)
- [8] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981. [20](#)
- [9] Victor Guillemin and Alan Pollack. *Differential topology*. American Mathematical Soc., 2010. [11](#)
- [10] Diederik P Kingma. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. [11](#)
- [11] PA W Lewis and Gerald S Shedler. Simulation of nonhomogeneous poisson processes by thinning. *Naval research logistics quarterly*, 26(3):403–413, 1979. [36](#)
- [12] Lumentum. 400 W 905 nm Multi-Junction VCSEL Array, 2025. Accessed: 2025-03-13. [17](#)
- [13] Rayleigh. Xxxi. investigations in optics, with special reference to the spectroscope. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 8(49):261–274, 1879. [39](#)
- [14] Diego Royo, Jorge García, Adolfo Muñoz, and Adrian Jarabo. Non-line-of-sight transient rendering. *Comput. Graph.*, 107:84–92, 2022. [36](#)
- [15] Richard C Sprinthal. Basic statistical analysis. (*No Title*), 2011. [6](#)
- [16] Mian Wei, Sotiris Nouisias, Rahul Gulve, David B Lindell, and Kiriakos N Kutulakos. Passive ultra-wideband single-photon imaging. In *Proc. IEEE/CVF Int. Conf. on Computer Vision (ICCV)*, pages 8135–8146, 2023. [3](#), [4](#), [5](#), [11](#), [19](#)