

Depth estimation in SPAD-based LIDAR sensors

Chen, Mingzhe; Rao, Padmakumar R.; Venialgo, Esteban

DOI

[10.1364/OE.507975](https://doi.org/10.1364/OE.507975)

Publication date

2024

Document Version

Final published version

Published in

Optics Express

Citation (APA)

Chen, M., Rao, P. R., & Venialgo, E. (2024). Depth estimation in SPAD-based LIDAR sensors. *Optics Express*, 32(3), 3006-3030. <https://doi.org/10.1364/OE.507975>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.



Depth estimation in SPAD-based LIDAR sensors

MINGZHE CHEN,^{1,2} PADMAKUMAR R. RAO,² AND ESTEBAN VENIALGO^{3,*}

¹Silicon Integrated B.V., High Tech Campus 9, 5656 AE Eindhoven, The Netherlands

²Faculty of Electrical Engineering, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands

³Faculty of Mechanical, Maritime and Materials Engineering, Department of Precision and Microsystems Engineering, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands

*e.venialgo@tudelft.nl

Abstract: In direct time-of-flight (D-TOF) light detection and ranging (LIDAR), accuracy and full-scale range (FSR) are the main performance parameters to consider. Particularly, in single-photon avalanche diodes (SPAD) based systems, the photon-counting statistics plays a fundamental role in determining the LIDAR performance. Also, the intrinsic performance ultimately depends on the system parameters and constraints, which are set by the application. However, the best-achievable performance directly depends on the selected depth estimation method and is not necessarily equal to intrinsic performance. We evaluate a D-TOF LIDAR system, in the particular context of smartphone applications, in terms of parameter trade-offs and estimation efficiency. First, we develop a simulation model by combining radiometry and photon-counting statistics. Next, we perform a trade-off analysis to study dependencies between system parameters and application constraints, as well as non-linearities caused by the detection method. Further, we derive an analytical model to calculate the Cramér–Rao lower bound (CRLB) of the LIDAR system, which analytically accounts for the shot noise. Finally, we evaluate a depth estimation method based on artificial intelligence (AI) and compare its performance to the CRLB. We demonstrate that the AI-based estimator fully compensates the non-linearity in depth estimation, which varies depending on application conditions such as target reflectivity.

Published by Optica Publishing Group under the terms of the [Creative Commons Attribution 4.0 License](#). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

1. Introduction

Depth sensing, defined as the method of obtaining a measure of the distance between a sensor and a target surface, has been an active field of research in automotive and consumer electronics. It plays a fundamental role in providing input information to systems, such as autonomous driving and face recognition. The intrinsic performance in depth sensing is only determined by the quality of input data obtained with the ranging hardware. However, the best achievable performance is also determined by the utilized depth estimation algorithm and noise rejection method.

There are several competing or complementary technologies available for depth sensing, such as light detection and ranging (LIDAR), radio detection and ranging (RADAR), sound navigation and ranging (SONAR), among others. Autonomous driving has been the leading force to develop LIDAR systems that are more compact, have lower costs, and achieve improved reliability as well as full-scale range (FSR). The main advantage of LIDAR is the ability to resolve targets with excellent angular resolution. On the other hand, the main disadvantages of LIDAR are its limited FSR, and difficulty to work under unfavorable environmental conditions, such as rain and fog. In order to build advanced driver assistance systems (ADAS), a sensor fusion approach is preferred so the disadvantages of either technique are compensated [1–5]. Also, LIDAR sensors

are utilized in the smartphone industry to perform tasks such as face recognition, scene modeling, auto-focus assistance, etc. [6–10].

Essentially, a LIDAR sensor emits light and some portion of it is reflected back from a target surface to the sensor. Typically, the light source is a laser and the detection part is composed of a photodetector array or integrated camera [11–14]. The depth information is calculated directly or indirectly from the round-trip time-of-flight (TOF) of the detected light. LIDAR can be classified in two main sub-categories, namely indirect TOF (I-TOF) and direct TOF (D-TOF), depending on how the actively generated light is modulated and measured [15]. Other techniques that are based on light interferometry, such as frequency-modulated continuous wave (FMCW) LIDAR, also show potential for depth sensing [16].

In D-TOF sensors, the laser light is generated following a pulse-width modulation (PWM) scheme with a very low duty cycle, so the spatial resolution is improved and optical peak power is increased. There are mainly two techniques in D-TOF LIDAR, namely flash and scanning LIDAR [16]. Flash LIDAR is utilized in short range application such as in smartphones [10]. On the other hand, scanning LIDAR can achieve a larger FSR and therefore shows more potential for automotive applications [17,18]. In both cases, the FSR of the LIDAR system is limited by the irradiance of the background light, maximum laser power required by safety regulations, as well as the optical sensitivity.

In this paper, we specifically select the case of smartphone LIDAR to investigate effects such as non-linearities, trade-offs between parameters and constraints, etc. For this purpose, we develop a custom pseudo ray-tracing algorithm that is coupled to a Monte Carlo (MC) code to simulate the LIDAR system. Most important, we evaluate depth estimation methods, such as deep-learning based methods, in terms of statistical efficiency. For such evaluation, we calculate the Cramér–Rao lower bound (CRLB) by utilizing an analytical model, and compare it to the performance of the depth estimation methods. It is important to mention that the analytical model utilized to calculate the CRLB accounts for the shot noise.

2. LIDAR system modeling

In order to model general scenarios in SPAD-based D-TOF LIDAR, we propose a generic D-TOF LIDAR sensor that is shown in Fig. 1(a). The generic sensor is composed of a vertical-cavity surface-emitting laser (VCSEL), and a SPAD array that is connected to several time-to-digital converters (TDCs) [12–14,18–20].

We optimize the system parameters considering smartphone applications. However, the presented model and equations are valid for any D-TOF application in which correlation-based noise filtering methods, such as coincidence filters, are not utilized. It is important to note that correlation-based methods require complex hardware implementations [20].

The sensor's principle of working is simply based on emitting a light pulse with the VCSEL, which arrives at a target surface, part of the light is reflected back and captured by the SPAD array. The distance is calculated by measuring the time difference between the emitted pulse and detected pulse. Normally, the system performs multiple measurements based on time-correlated single-photon counting (TCSPC) cycles, in order to form a histogram of the recorded time differences or timestamps. The timestamps are calculated by the TDCs (see Fig. 1(b)), and we consider an architecture where multiple SPADs share a single TDC through timelines. The shared timelines allow to have a lower number of TDCs with respect to the number of SPADs [21].

To study the proposed D-TOF LIDAR sensor performance and evaluate several depth estimation methods, we propose a simulation and signal processing flow that is shown in Fig. 2. The optical modeling stage estimates the detected number of VCSEL photons per TCSPC cycle S_p , and the noise rate N_r in terms of counts per second. The noise events can include spurious counts due to background light as well as dark counts. Next, the time statistics modeling stage simulates the random and temporal behavior of the photon-counting process. For signal generation, we

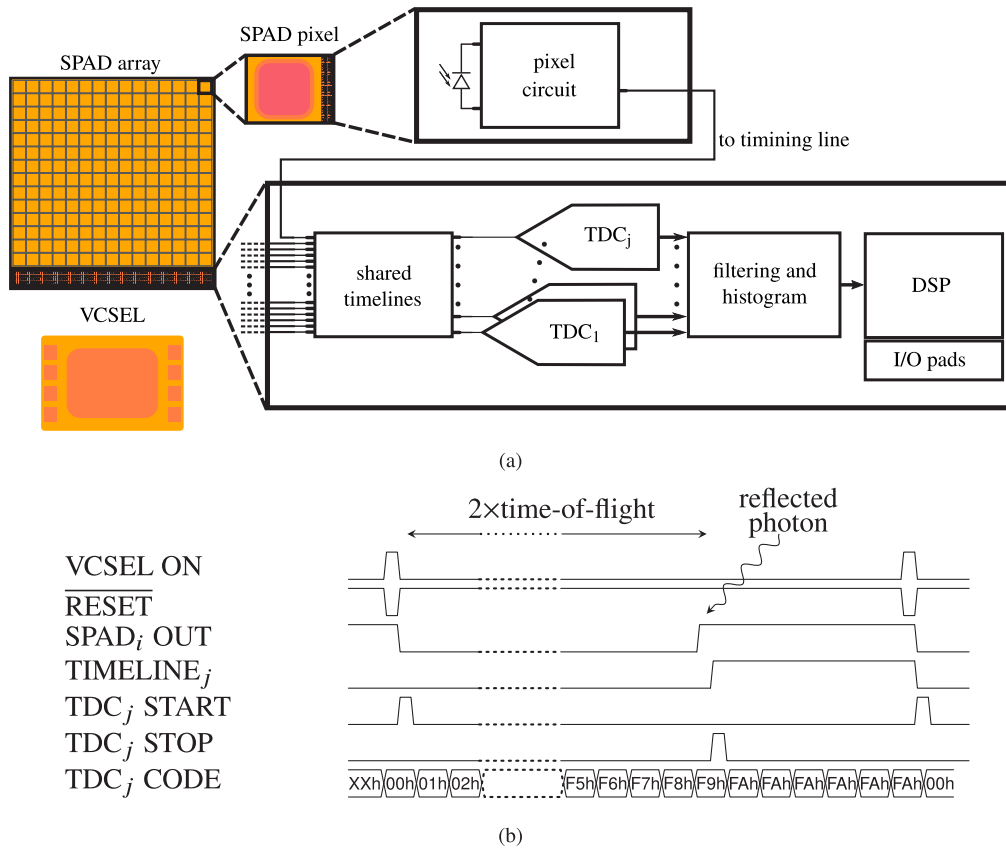


Fig. 1. Diagram and operation of a generic LIDAR sensor. (a) sensor block diagram. (b) generic timing diagram.

consider that the SPADs are recharged at the beginning of the TCSPC cycle. Also, when a detection occurs, the fired SPADs are kept quenched until the next recharge phase (see Fig. 1(b)).

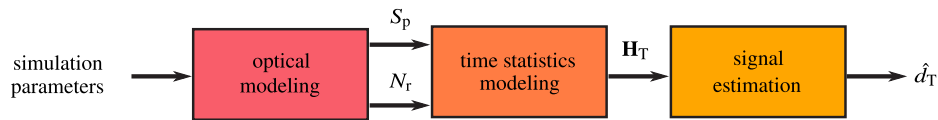


Fig. 2. Simulation and signal processing flow

The simulation and signal processing flow can simulate a single D-TOF LIDAR scenario with fixed simulation parameters. As a result, it outputs a single value that corresponds to the estimated target depth $\hat{d}_T$, which is calculated based on a timestamp histogram H_T . However, in practice, we simulate several scenarios in the same run by sweeping the simulation parameters.

In the following subsections, details about the optical and time-statistics modeling stages are described and discussed.

2.1. Optical model

Typically, in D-TOF LIDAR, the target surface is considered to be an ideal light diffuser. In order to simplify the simulation model, previous works assume that the incidence angle between a

light ray from the laser to the target surface is always perpendicular [18,20]. However, when the D-TOF LIDAR sensor moves closer to the target, the incidence angle of the light ray cannot be longer considered perpendicular (see $\theta_{e,ij}$ in Fig. 3(a)).

Therefore, we propose a model in which the target surface is divided into sub-elements ΔT_{ij} , which are considered as individual Lambertian reflectors. Also, this pixelated model can consider situations in which the laser field-of-illumination (FOI) exceeds the target surface area (see Fig. 3(a)). In this model, the detector consists of a lens, an optical bandpass filter and a SPAD array. The detector collects the summation of the diffused power from the pixelated sub-elements whose locations are inside the FOI of the VCSEL, and the field-of-view (FOV) of the detector (see Fig. 3(b)).

The optical model is based on a pseudo-sequential ray tracing algorithm that first calculates light paths between the VCSEL and the target surface. Next, it calculates the light paths between the sub-elements of the pixelated target surface and the SPAD array.

2.1.1. Signal event

A signal event is defined as a VCSEL photon detection that generates avalanche current in a SPAD. The signal light path is divided into two sub-paths: the emission path in which the VCSEL emits photons to the target (see Fig. 3(a)); and the reflection path in which the target reflects back photons onto the sensor lens (see Fig. 3(b)).

In the first sub-path, the amount of received VCSEL light by a sub-element of the pixelated target surface is calculated. The VCSEL and the SPAD array are placed at the beginning of the d -axis and are denoted as $d = 0$. And we define d_T as the distance from the VCSEL surface to the target surface.

The coverage sphere of the VCSEL is determined by a cone with its apex angle $2\theta_e$, which is equal to the FOI (see Fig. 3(a)). The total solid angle that corresponds to the FOI, Ω_{FOI} , can be approximated as

$$\Omega_{FOI} = \int_0^{2\pi} \int_0^{\theta_e} \sin \theta d\theta d\phi \approx 4\pi \sin^2 \frac{\theta_e}{2} = 4\pi \sin^2 \frac{FOI}{4}. \quad (1)$$

At every sub-element ΔT_{ij} , we define $\theta_{e,ij}$ as the incident angle between a VCSEL light ray and the normal vector to the surface of ΔT_{ij} (see Fig. 3(a)). And we define d_T as the distance from the VCSEL surface to the target surface. So, the distance $d_{e,ij}$ from the VCSEL to ΔT_{ij} is given by

$$d_{e,ij} = \frac{d_T}{\cos \theta_{e,ij}}. \quad (2)$$

And the solid angle $\Delta\Omega_{ij}$ from ΔT_{ij} is

$$\Delta\Omega_{ij} = \frac{\Delta\hat{T}_{ij}}{d_{e,ij}^2}, \quad (3)$$

where $\Delta\hat{T}_{ij}$ is the area of ΔT_{ij} .

A received radiant flux $\Delta\Phi_{e,ij}$ by ΔT_{ij} , when it is located inside the FOI, is calculated as the ratio between Ω_{FOI} and $\Delta\Omega_{ij}$; and multiplied by the total VCSEL optical power P_s , as a dot product. This calculation is expressed as

$$\Delta\Phi_{e,ij} = \begin{cases} \frac{P_s \Delta\hat{T}_{ij}}{4\pi \sin^2(\frac{FOI}{4}) d_{e,ij}^2} \cos \theta_{e,ij} & |\theta_{e,ij}| \leq \frac{FOI}{2} \\ 0 & |\theta_{e,ij}| > \frac{FOI}{2} \end{cases}. \quad (4)$$

The second sub-light-path of the ray tracing algorithm corresponds to the detected light by the SPAD array, which is back-reflected from the target surface (see Fig. 3(b)). The radiant flux

$\Delta\Phi_{s,ij}$ emitted by the $\Delta T_{i,j}$ is the product of the received radiant flux and the reflectivity ρ of the target surface, and is defined by

$$\Delta\Phi_{s,ij} = \rho\Delta\Phi_{e,ij}. \quad (5)$$

The detected radiant intensity $\Delta I_{e,ij}$, emitted from $\Delta T_{i,j}$, is proportional to the product of the peak radiant intensity $I_{0,ij}$ and cosine $\theta_{s,ij}$, which is the angle between the sensor lens $\hat{S}$ and the normal vector of $\Delta T_{i,j}$ (see Fig. 3(b)). $\Delta I_{e,ij}$ is given by

$$\Delta I_{e,ij} = I_{0,ij} \cos \theta_{s,ij}. \quad (6)$$

A plane radiator or reflector that is perfectly diffusive emits light in all directions. Further, the total emitted power is contained within half sphere with respect to the normal of that plane (see Fig. 3(b)). Therefore, the radiant flux $\Delta\Phi_{s,ij}$ diffused from $\Delta T_{i,j}$ of the target is calculated as the surface integral of the diffused radiant intensity $\Delta I_{e,ij}$ over the solid angle Ω of the half sphere

$$\Delta\Phi_{s,ij} = \int_{\Omega} \Delta I_{e,ij} d\Omega = \pi I_{0,ij}. \quad (7)$$

Approximating the illumination with the inverse-square law, the irradiance $\Delta E_{e,ij}$, measured at the sensor lens and emitted by $\Delta T_{i,j}$, is expressed as

$$\Delta E_{e,ij} = \frac{I_{e,ij}}{r^2} = \frac{\Delta\Phi_{s,ij} \cos \theta_{s,ij}}{\pi d_{s,ij}^2}, \quad (8)$$

where $d_{s,ij}$ is the distance from $\Delta T_{i,j}$ to the D-TOF LIDAR sensor.

The total radiant flux Φ_s sampled at the lens, whose area is A_l , is the summation of the radiant flux $\Delta\Phi_{s,ij}$ of every $\Delta T_{i,j}$, but only if they are inside the FOV. Φ_s is calculated as follows

$$\Phi_s = \sum_{i,j} \begin{cases} \frac{A_l \Delta\Phi_{s,ij} (\cos \theta_{s,ij})^2}{\pi d_{s,ij}^2} & |\theta_{s,ij}| \leq \frac{\text{FOV}}{2} \\ 0 & |\theta_{s,ij}| > \frac{\text{FOV}}{2} \end{cases}. \quad (9)$$

Finally, the number of photons per TCSPC cycle S_p , is calculated as

$$S_p = \Phi_s \eta_l \eta_f \frac{2\lambda}{\pi f h c} \text{PDE}, \quad (10)$$

where h is the Plank's constant, c is the light speed, f is the VCSEL repetition frequency, and λ is the VCSEL center wavelength. The photon detection efficiency (PDE) is defined as the product of photon detection probability (PDP) and fill factor (FF) of the SPAD array. Also, it accounts for a reduction ratio $2/\pi$ since the light-sensitive area of the SPAD array is represented as a square, and it is inscribed in a circle that corresponds to the light projected by the lens [20]. An optical bandpass filter is placed between the lens and the SPAD array to reduce background ambient light (see Fig. 3(b)). So, the transmittance of the optical bandpass filter η_f is also considered as a power loss.

2.1.2. Noise event

A noise event is defined as any SPAD avalanche that is not triggered by a photon emitted from the VCSEL. Noise events come from artificial light or natural light in the two different light paths: photons emitted from a noise source to the target and the target reflects photons to the sensor lens; or direct emission of photons to the sensor lens. We only consider the first light path (see

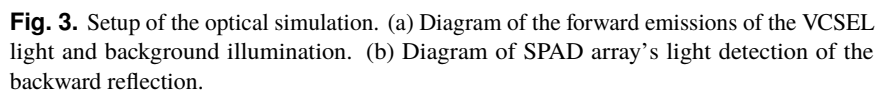


Fig. 3). Additionally, dark counts can trigger SPAD avalanches, but in practice the dark count rate (DCR) is significantly smaller than the noise rate produced by background ambient light.

We use the equivalent light irradiance, E_n , at sea level to estimate the background noise power. Similar to the signal event calculation, we utilize the same ray tracing equations but by replacing $\Phi_{s,ij}$ in Eq. (9) by

$$\Delta\Phi_{n,ij} = \rho E_n \Delta\hat{T}_{ij}. \quad (11)$$

Therefore, we obtain an expression for the total radiant flux measured at the sensor lens, due to noise photons only, expressed as

$$\Phi_n = \sum_{ij} \begin{cases} \frac{A_l \Delta\Phi_{n,ij} (\cos \theta_{s,ij})^2}{\pi d_{s,ij}^2} & |\theta_{s,ij}| \leq \frac{\text{FOV}}{2} \\ 0 & |\theta_{s,ij}| > \frac{\text{FOV}}{2} \end{cases}. \quad (12)$$

Similarly as in Eq. (10), we calculate the noise rate per time unit N_r as

$$N_r = \Phi_n \eta_l \eta_f \frac{2\lambda}{\pi h c} \text{PDE}. \quad (13)$$

2.2. Statistical model

The output of the time statistics modeling stage is a single histogram represented by the vector $\mathbf{H}_T$, when simulating fixed optical parameters. The timestamps of signal and noise events are generated using S_p and N_r as input information, which are calculated from Eqs. (10) and (13) respectively (see Fig. 2).

In a single TCSPC cycle, the number of detected VCSEL photons X is a random variable (RV) that follows a Poisson process. Its probability density function (PDF) is defined as

$$P(X = k) = \frac{S_p^k}{k!} e^{-S_p}, \quad (14)$$

where S_p is the expected number of signal events per TCSPC cycle. We assume that the VCSEL light pulse has a Gaussian shape and neglect any dispersion effects. Therefore, the timestamp of a signal event is represented by a RV with a Gaussian PDF given by

$$f(t | 2\frac{d_T}{c}, \sigma_1) = \mathcal{N}(t | 2\frac{d_T}{c}, \sigma_1), \quad (15)$$

where σ_1 relates to the VCSEL pulse width, and $\mathcal{N}$ is the Gaussian function.

The noise events are uniformly distributed over time. Their timestamps are represented by an RV with exponential PDF, which is defined by

$$g(t | N_r) = \begin{cases} N_r e^{-N_r t} & t \geq 0 \\ 0 & t < 0 \end{cases}. \quad (16)$$

Based on the Monte Carlo method, random signal timestamps are generated following Eqs. (14) and (15), and noise timestamps are generated following Eq. (16). Next, the signal and noise timestamps are grouped together and sorted. In the case of having a single-TDC system, only the first timestamp is added to the $\mathbf{H}_T$. In the case of multiple TDCs, this process is repeated for every subgroups composed of many SPADs and a single TDC, and the first timestamp of every TDC is added to the $\mathbf{H}_T$. The timestamp generation process and update of $\mathbf{H}_T$ is repeated for every TCSPC cycle.

Additionally, we consider that the dominant factor in the timing uncertainty is the rounding effect produced by TDC's least significant bit (LSB) of size Δ_{TDC} . Therefore, we perform a timestamp rounding by considering Δ_{TDC} as bin size when calculating $\mathbf{H}_T$. In our model, the sensor is able to detect several photons per laser pulse. However, we consider that a single SPAD is working in photon-starving mode and assumed that the SPAD timing jitter is much smaller than σ_l . When SPADs are not working in photon-starving mode, the so-called multi-photon distortion affects the SPAD timing jitter [22].

3. System trade-off analysis

In this section, we perform a system trade-off analysis by using the models described in the previous section. We define two criteria to evaluate the system performance, namely the depth range and depth resolution.

We define parameters that are swept during the analyses and categorize them into two subsets, namely variable and fixed. The variable parameters can be modified by the system during operation, and the fixed parameter are constrained by the application. The variable parameters are P_s , total number of TCSPC cycles N_{TCSPC} , PDP, and the TDC bin size Δ_{TDC} . The fixed parameters are E_n and ρ . It is important to mention that in this section we consider a system with a single TDC. Also, appendix C explains how the models and results of single-TDC systems can be re-adjusted to be re-interpreted for multiple-TDC systems.

Also, appendix A shows additional static parameters that are common for all simulations.

3.1. Depth range

The aim of the depth range analysis is to find the maximum distance that the system can detect under limiting conditions. We utilize the find peaks function (FPF) from MATLAB as a reference algorithm to estimate d_T [23]. In this analysis, we gradually increase d_T until the FPF is unable to locate a signal peak, and the final valid value of d_T is considered as the depth range.

Figure 4 shows the maximum depth under different parameter values. The system fails when N_{TCSPC} is lower than 500. Furthermore, the depth range decreases with the increase of E_n .

We observe that ρ and ΔE_n are highly dependent parameters when determining the depth range. Therefore, we perform a depth range analysis with two extreme values of ρ , and sweeping ΔE_n as well as P_s (see Fig. 4(d)). It can be found that for the ρ value of 60 % the depth range is slightly higher for low P_s values. Also, it reaches a faster saturation when increasing the P_s . On the other hand, the 8 % ρ has a slight lower depth range for low P_s values, but its depth range continues to increase as P_s increases. The explanation to this observation that in the curve with a ρ value of 8 %, the depth range is restricted by the VCSEL power, only. In the curve with higher ρ , the depth range is restricted by the amount of reflected background light.

Also, we further analyze the effect of PDP in the depth range in appendix B.

3.2. Depth resolution

In the depth resolution analysis, we study the mean-square-error (MSE) of the system when estimating d_T . In this section, we also utilize the FPF as our estimation method to locate the VCSEL peak position in $\mathbf{H}_T$. Further, $\hat{d}_T$ is simply calculated as the time of the peak found by FPF from $\mathbf{H}_T$, and converted into the corresponding half distance. For a better understanding of the analysis, we report the $\sqrt{\hat{\text{MSE}}}$ in order to have consistency in the depth unit. Also, we study the variance σ_{MSE} and bias component B_{MSE} of the MSE, separately.

In this analysis, we select a standard value of P_s equal to 7.36 mW, which allows a depth range of 0.5 m when PDP = 2 % (see Fig. 11 in appendix B), and is below the maximum safety value when utilizing $N_{\text{TCSPC}} = 10000$ [24].

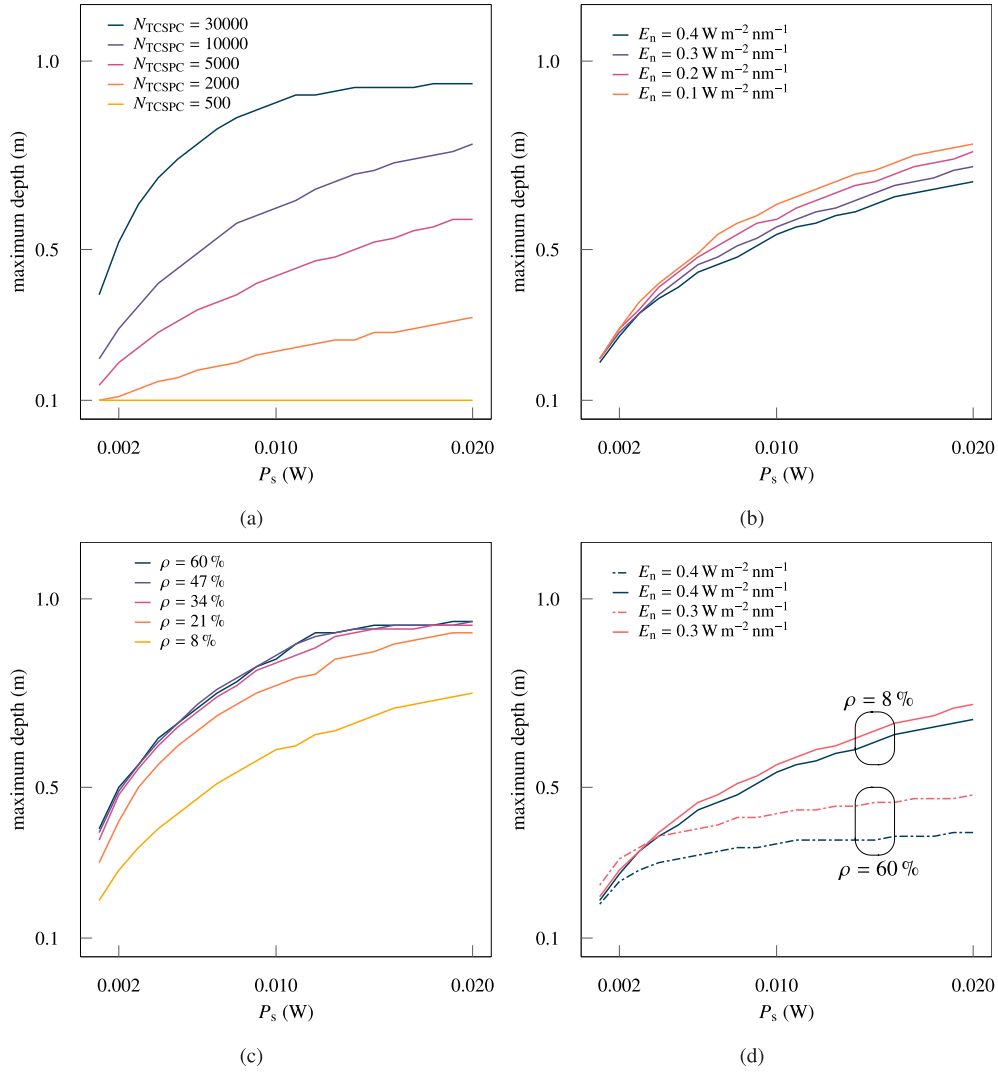


Fig. 4. Depth range under variable and fixed parameter conditions, each individual shows the depth range as a function of P_s . (a) depth range with N_{TCSPC} as parameter. (b) depth range with E_n as parameter. (c) depth range as ρ as parameter. (d) depth range under extreme conditions of ρ and E_n . When the corresponding parameters are not swept, they are set as $N_{TCSPC} = 30000$, $\Delta_{TDC} = 25$ ps, $E_n = 0.1 \text{ W m}^{-2} \text{ nm}^{-1}$, and $\rho = 8\%$.

We sweep the variable parameters, and for every point in the sweep we generate 1,000 histograms to calculate the MSE. Fig. 5 shows the results when sweeping the variable parameters and Fig. 6 shows the results when sweeping the fixed parameters.

First, we simulate the depth resolution while sweeping N_{TCSPC} (see Fig. 5(a)). We find that the MSE remains almost unchanged with respect to N_{TCSPC} (see Fig. 5(a)). However by analyzing B_{MSE} and σ_{MSE} separately, we find that the MSE value is mainly determined by B_{MSE} . It is important to note that B_{MSE} is an order of magnitude larger than σ_{MSE} . Also, the value of B_{MSE} does not depend on N_{TCSPC} ; and as expected, σ_{MSE} decreases as N_{TCSPC} increases (see Fig. 5(b)). The large value of the B_{MSE} is explained by order statistics [25,26]. This effect is further elaborated in appendix D.

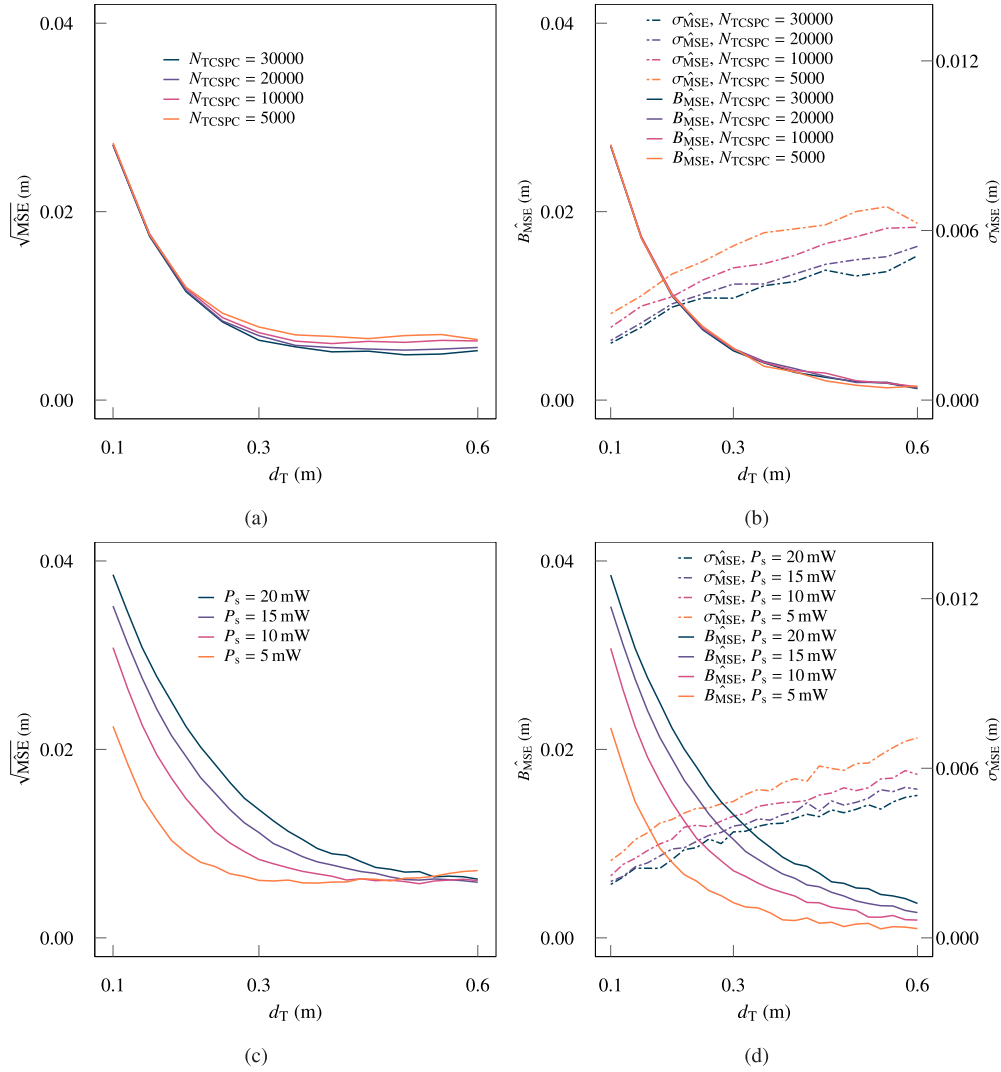


Fig. 5. Depth resolution under variable parameter conditions, and subdivided into B_{MSE} and σ_{MSE} . (a) $\sqrt{\text{MSE}}$ with N_{TCSPC} as parameter. (b) B_{MSE} and σ_{MSE} with N_{TCSPC} as parameter. (c) $\sqrt{\text{MSE}}$ with P_s as parameter. (d) B_{MSE} and σ_{MSE} with P_s as parameter. When the corresponding parameters are not swept, they are set as $P_s = 7.36$ mW, $N_{\text{TCSPC}} = 10000$, $\Delta_{\text{TDC}} = 25$ ps, $E_n = 0.1 \text{ W m}^{-2} \text{ nm}^{-1}$, and $\rho = 8\%$.

Next, we simulate the MSE as function of P_s (see Figs. 5(c) and 5(d)). And again, the change in B_{MSE} is explained by order statistics; as the total number of VCSEL photon arriving at the detector changes with respect to P_s . The tendency in Figs. 6(c) and 6(d) is also explained by a bias shift.

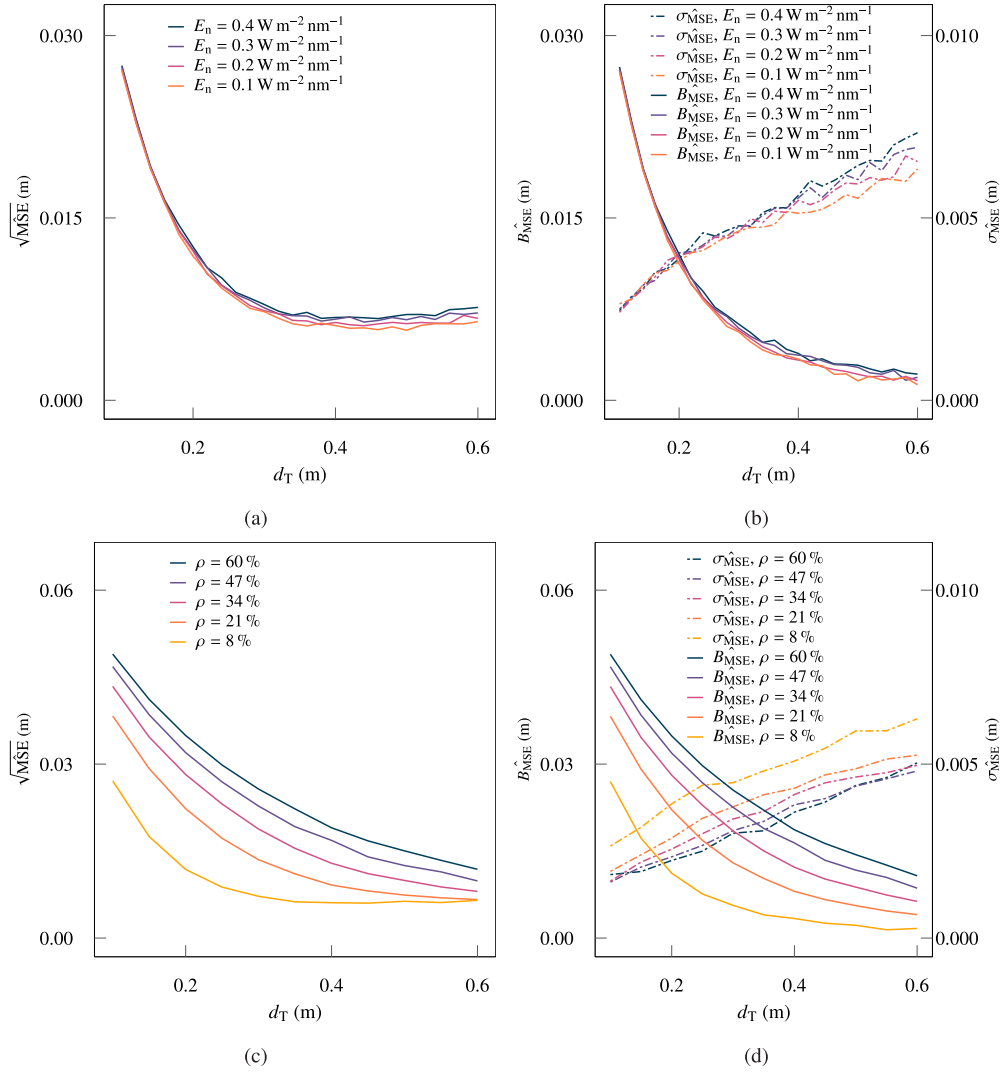


Fig. 6. Depth resolution under fixed parameter conditions, and subdivided into B_{MSE} and σ_{MSE} . (a) $\sqrt{\text{MSE}}$ with E_n as parameter. (b) B_{MSE} and σ_{MSE} with E_n as parameter. (c) $\sqrt{\text{MSE}}$ with ρ as parameter. (d) B_{MSE} and σ_{MSE} with ρ as parameter. When the corresponding parameters are not swept, they are set as $P_s = 7.36 \text{ mW}$, $N_{\text{TCSPC}} = 10000$, $\Delta_{\text{TDC}} = 25 \text{ ps}$, $E_n = 0.1 \text{ W m}^{-2} \text{ nm}^{-1}$, and $\rho = 8\%$.

4. Cramér–Rao lower bound

The CRLB predicts the optimum performance that can be achieved by any unbiased estimator, and is only determined by the estimator's input data. It is calculated from a likelihood function that relates the estimator's input data and the parameter to be estimated, which are $\mathbf{H}_T$ and d_T , respectively, in this case. In this section, we describe in detail the steps to calculate the CRLB of the 1D-TOF LIDAR system (see Fig. 1(a)).

First, we define a timestamp as an RV that corresponds to photon detection time with respect to an uncorrelated signal (see Fig. 1(b)). Next, we continue with the calculation of the timestamp PDF of background photons, only. We assume that the emission process of the background

light source is Poissonian. Therefore, the timestamp PDF, when the sensor is only exposed to background photons is given by Eq. (16), which is derived from random Poisson points [27].

So far, we consider that our system has a single TDC and so it can only record the first photon detection. Therefore, we can derive Eq. (16) by utilizing order statistics with first order, as an alternative way to random Poisson points. To do so, we propose to define the PDF of the unsorted background photons' timestamps as follows

$$s_N(t) = \lim_{T \rightarrow +\infty} \begin{cases} 0 & \text{for } t < 0 \\ \frac{1}{T} & \text{for } t \geq 0 \end{cases}. \quad (17)$$

Also, the size of the unsorted timestamp set is given by

$$R = \lim_{T \rightarrow +\infty} N_r T. \quad (18)$$

T must tend to infinity in order to achieve the definition of random Poisson points [27]. Next, we calculate the PDF of the first background photon's timestamp. We define it as the PDF of the first order statistic of set composed of R independent and identically distributed (IID) RVs with PDF equal to s_N , and is given by

$$\begin{aligned} n_1(t) &= \lim_{T \rightarrow +\infty} R[1 - S_N(t)]^{R-1} s_N(t) \\ &= \lim_{T \rightarrow +\infty} R[1 - \frac{t}{T}]^{R-1} \frac{1}{T} H(t) \\ &= \lim_{T \rightarrow +\infty} N_r [1 - \frac{t}{T}]^{N_r T-1} H(t) \\ &= N_r e^{-N_r t} H(t), \end{aligned} \quad (19)$$

where $H(t)$ is the unit step function. It is important to note that $n_1(t)$ considers the influence of the shot noise. Also, Eq. 19 express the relationship between order statistics and random Poisson points. The next step is to calculate the PDF of the timestamp that corresponds to the first photon detection, but when the VCSEL is turned on. Therefore, we define the PDF of the unsorted photons' timestamps, which includes VCSEL and background photons, as

$$s_{N+P}(t | d_T) = \lim_{T \rightarrow +\infty} \begin{cases} 0 & \text{for } t < 0 \\ \frac{1}{T} + \frac{s_p}{R} \mathcal{N}(t | 2\frac{d_T}{c}, \sigma_1) & \text{for } t \geq 0 \end{cases}. \quad (20)$$

Similarly as in Eq. 19, we define the first detected timestamp as the first order statistic of a set composed of R IID RVs with PDF equal to s_{N+P} . Subsequently, the PDF of first timestamp, which corresponds to VCSEL and background photons, is given by

$$p_1(t | d_T) = \lim_{T \rightarrow +\infty} R[1 - S_{N+P}(t)]^{R-1} s_{N+P}(t). \quad (21)$$

At this point, we can consider that p_1 approximates a PDF that corresponds to the time between random Poisson points with nonuniform density [27], as far as T tends to infinity. Next, we define the likelihood function, which relates the estimator's input data to the parameter to be estimated, as follows

$$l(t | d_T) = p_1(t | d_T) * U(t | [0, \Delta_{TDC}]). \quad (22)$$

The uniform distribution $U(t | [0, \Delta_{TDC}])$ models the rounding effect due to the finite size of the TDC's LSB. Also, we assume that the single-shot resolution of the TDC is only limited by the LSB size. We do not consider differential nonlinearity (DNL) or integral nonlinearity (INL) effects in the TDC. Timing jitter, DNL, or INL effects can be modeled by replacing $U(t | [0, \Delta_{TDC}])$ by a customized PDF (see appendix E).

Finally, we rescale the likelihood function as follows

$$l'(t | d_T) = \frac{l(t | d_T)}{\int_0^{T_h} l(t | d_T) dt}, \quad (23)$$

where T_h is the maximum time range of the timestamp histogram $\mathbf{H}_T$ or TDC full scale range. In all calculations, we consider an 8-bit TDC, so T_h is equal to 6.4 ns when $\Delta_{\text{TDC}} = 25$ ps.

The Fisher information $\mathcal{I}_1(d_T)$ observed in one single detection is given by

$$\mathcal{I}_1(d_T) = \int_0^{T_h} \left[\frac{\partial}{\partial d_T} l'(t | d_T) \right]^2 \frac{1}{l'(t | d_T)} dt. \quad (24)$$

Since every TCSPC cycle is independent from each other, assuming that the target is static during the detection process, the Fisher information observed in $\mathbf{H}_T$ is calculated as

$$\mathcal{I}_h(d_T) = N_h \mathcal{I}_1. \quad (25)$$

N_h represents the total number of counts contained in $\mathbf{H}_T$ and is subject to shot noise. However, the shot noise influence in N_h can be neglected since the histogram typically contains several thousands of counts. Finally, the CRLB is defined as the inverse of $\mathcal{I}_h(d_T)$.

We numerically calculate $p_1(t | d_T)$ for two difference ρ values with a time step of 1 ps (see Fig. 7). In this calculation, we also consider two cases per ρ value: $T = T_h$ and $T = T_h \cdot 10^4$. Obviously, the second case aims to approach that T tends to infinity, numerically. In addition, we calculate $\mathbf{H}_T$ using the MC code of subsection 2.2 but we reduce Δ_{TDC} to 1 ps and increase N_{TCSPC} to 10^7 .

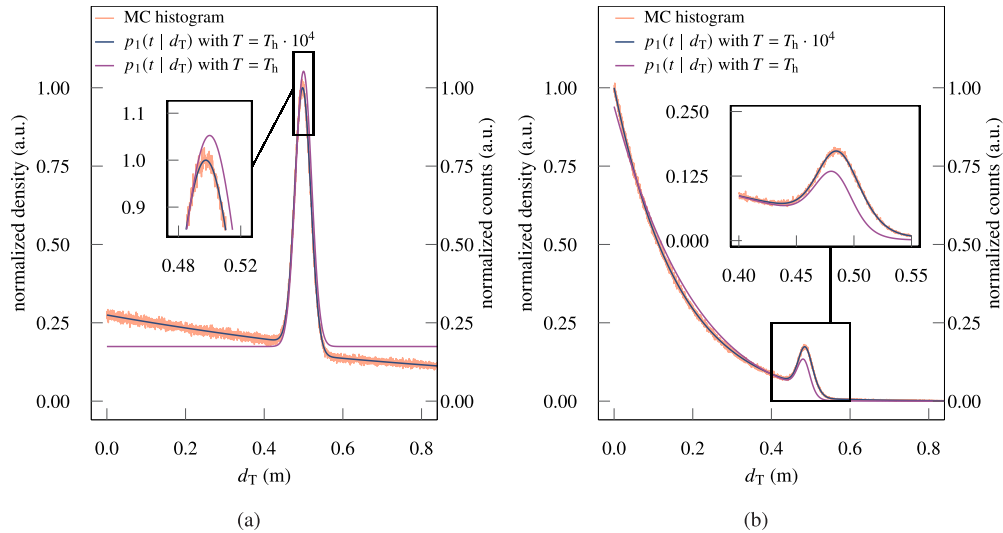


Fig. 7. $p_1(t | d_T)$ calculated with the analytical and MC models for two different reflectivity conditions and two different T values. The system parameters are $P_s = 7.36$ mW, $N_{\text{TCSPC}} = 10^7$, $\Delta_{\text{TDC}} = 1$ ps, $E_n = 0.4$ W m⁻² nm⁻¹, and $d_T = 0.5$ m. In (a) ρ is 8 % and in (b) ρ is 60 %.

As observed in Fig. 7, $p_1(t | d_T)$ is perfectly matching $\mathbf{H}_T$ generated by the MC code when $T = T_h \cdot 10^4$. It is worth noticing that the analytical model includes the shot noise influence in the detection process as long as T tends to infinity. Previous analytical models utilized to calculate the CRLB in estimation problems with single-photon detectors did not account for influence of

the shot noise [28,29]. However, in a LIDAR case, the shot noise has a significant influence since the average number of detected photons per TCSPC cycle is rather low.

Finally, we calculate the $\sqrt{\text{CRLB}}$ with respect to d_T for the two different ρ values (see Fig. 8), as well as for $T = T_h$ and $T = T_h \cdot 10^4$. In the numerical calculation, R is rounded to the nearest integer number (see Eq. (21)), which causes numerical artifacts in the $\sqrt{\text{CRLB}}$ due to large discontinuities in $l(t | d_T)$. This effect is minimized as T tends to infinity (see Fig. 8).

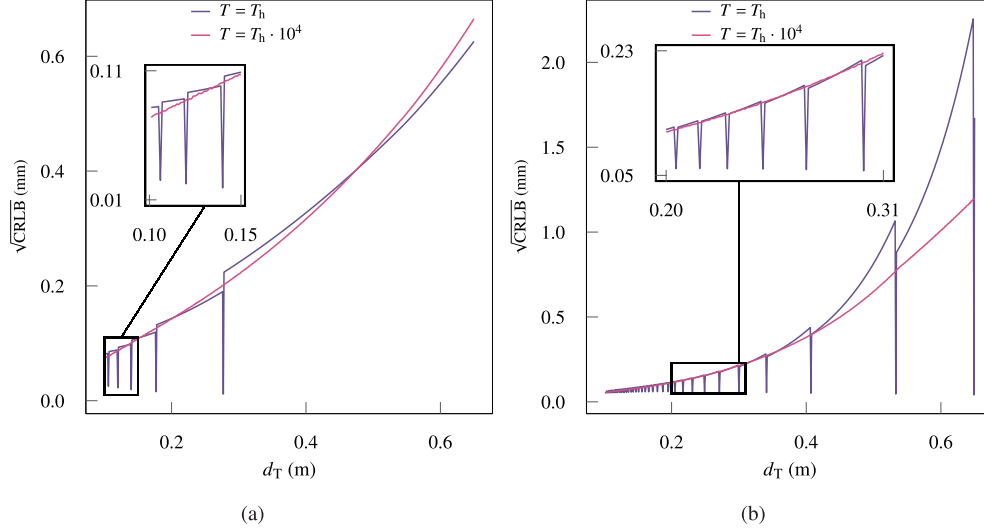


Fig. 8. CRLB calculated for two reflectivity conditions and two different T values. The system parameters are $P_s = 7.36$ mW, $N_{\text{TCSPC}} = 30000$, $\Delta_{\text{TDC}} = 25$ ps, and $E_n = 0.4 \text{ W m}^{-2} \text{ nm}^{-1}$. In (a) ρ is 8 % and in (b) ρ is 60 %.

5. Full system performance

So far, we utilize the FPF as our reference estimation method. However, as explained in section 3.2, this method is biased. Also, the behavior of the FPF depends on its input parameters, such as the minimum peak prominence, which requires to be adjusted depending on the application conditions [23]. Therefore, we propose to utilize artificial neural networks (ANNs) as a robust and unbiased depth estimator, which can be trained to adjust its coefficient, automatically.

In this section, we evaluate the performance of an improved version of the FPF (see appendix D) and compare it to the ANN-based depth estimator's performance. In addition, we benchmark the performance of both methods against the CRLB that is calculated in section 4.

5.1. Artificial neural network

Depth estimation based on a feedforward ANNs can directly calculate the depth and internally compensate for the non-linearities of the system [30]. To clarify, the ANN is not used as a classifier since it has a single output which directly gives $\hat{d}_T$. The ANN inputs are the counts of the 256 bins of single histogram $\mathbf{H}_T$, and the ANN has single hidden layer with 8 neurons. The output neuron's activation function is linear and the hidden neurons' activation function is a tansig function.

To obtain a representative training set, we generate many $\mathbf{H}_T$ by linearly sweeping the parameters of the model. We assume that the parameters of the system are already fixed, and swept only the fixed parameters (see Tables 3, 4 and 5). Thus, E_e , ρ , and d_T are the three parameters that are swept to obtain the training set (see Table 1). It is important to mention that

at every parameter condition, the simulation is repeated for 5 times, in order to account for the statistical effects.

Table 1. Range of parameter sweeps in ANN training and testing sets. The variable parameters are: $N_{\text{TCSPC}} = 30000$, $\Delta_{\text{TDC}} = 25$ ps, and $P_s = 7.36$ mW.

parameter	start	end	step training	step test
E_n ($\text{W m}^{-2} \text{ nm}^{-1}$)	0	0.4	0.04	0.04
ρ (%)	8	60	1	1
d_T (m)	0	0.6	0.0025	0.01

5.1.1. ANN performance evaluation

We evaluate the performance of the ANN to input histograms that are not used during training. Therefore, we create a testing set in addition to the training set (see Table 1). We combine all patterns that belong to the same value of d_T and evaluate the ANN performance by creating a distribution of $\hat{d}_T$ per d_T value. Fig. 9 shows the $\hat{d}_T$ at every d_T of the testing set as an error bar plot, where the bar width represents in interquartile range. As observed in Fig. 9, the non-linearities of the signal peak in $\mathbf{H}_T$ are perfectly compensated by the ANN. In addition, the selection of 8 hidden neurons is not arbitrary, and it was chosen after observing that a further increase of hidden neurons do not improve significantly the estimation performance, for the utilized dataset size.

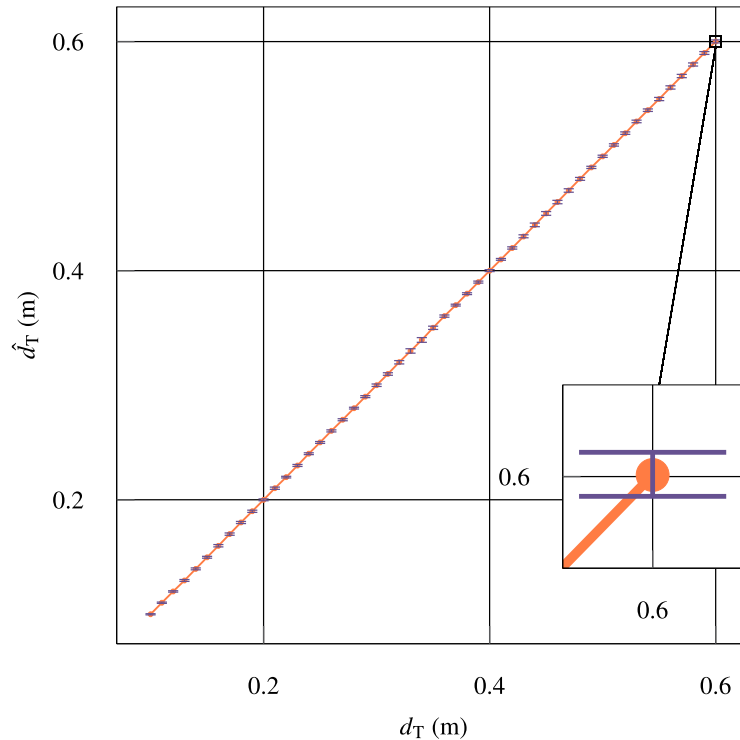


Fig. 9. Linearity evaluation of the ANN for the testing set.

Table 2. Performance summary of the estimated depth shown as $\hat{d}_T \pm 2\sigma_{\hat{d}_T}$.

depth: d_T (m)	improved FPF: $\hat{d}_T$ (m)	ANN: $\hat{d}_T$ (m)
0.1	0.0943±0.0120	0.1000±0.0016
0.2	0.2022±0.0102	0.1999±0.0028
0.3	0.3039±0.0096	0.3000±0.0033
0.4	0.4036±0.0100	0.4000±0.0029
0.5	0.5007±0.0100	0.4998±0.0035
0.6	0.5977±0.0095	0.6001±0.0049

5.2. Performance comparison

In this subsection, we compare the performance between the improved FPF (see appendix D) and ANN. We select six distributions from the linear sweep, calculate $\hat{d}_T$ minus d_T , and show the results as a set of boxplots (see Fig. 10). It can be observed that the error distribution of the improved FPF is more irregular with respect to the ANN. Additionally, it appears that the linearity of the ANN is significantly better compared to the improved FPF. Also, Fig. 10 compares the $\pm 3\sqrt{\text{CRLB}}$ to the error distribution of the estimation methods.

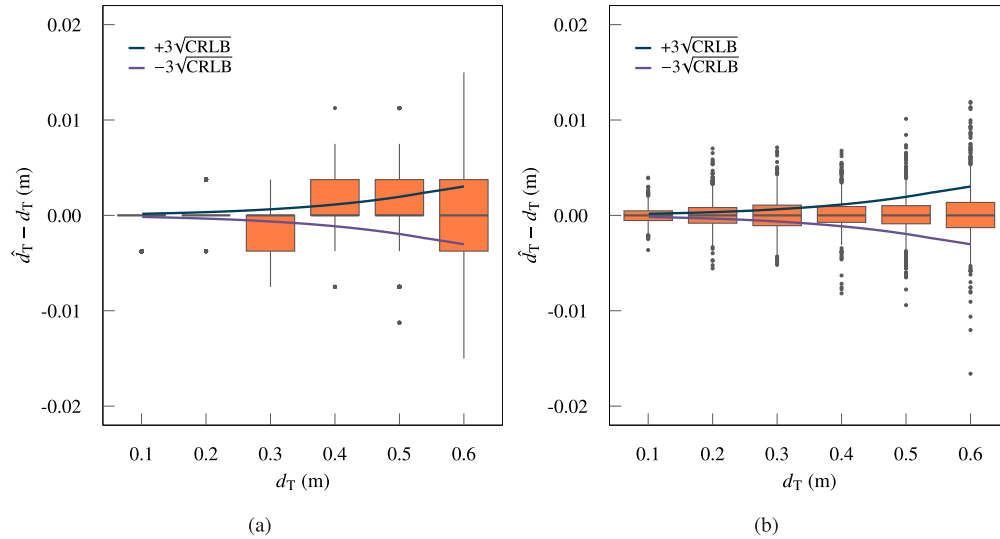


Fig. 10. Distribution of the estimation error for the testing set of the improved FPF and the ANN depicted as a boxplot and compared to the CRLB. (a) improved FPF performance; (b) ANN performance. The CRLB is calculated with $P_s = 7.36$ mW, $N_{\text{TCSPC}} = 30000$, $\Delta\text{TDC} = 25$ ps, $E_n = 0.4$ W m⁻² nm⁻¹, ρ is 60 %, and $T = T_h \cdot 10^4$.

In addition, Table 2 shows a comparison summary expressed as sample mean and standard deviation of the $\hat{d}_T$ distribution. It can be observed that the standard deviation achieved by the ANN is significantly lower compared to the improved version of the FPF.

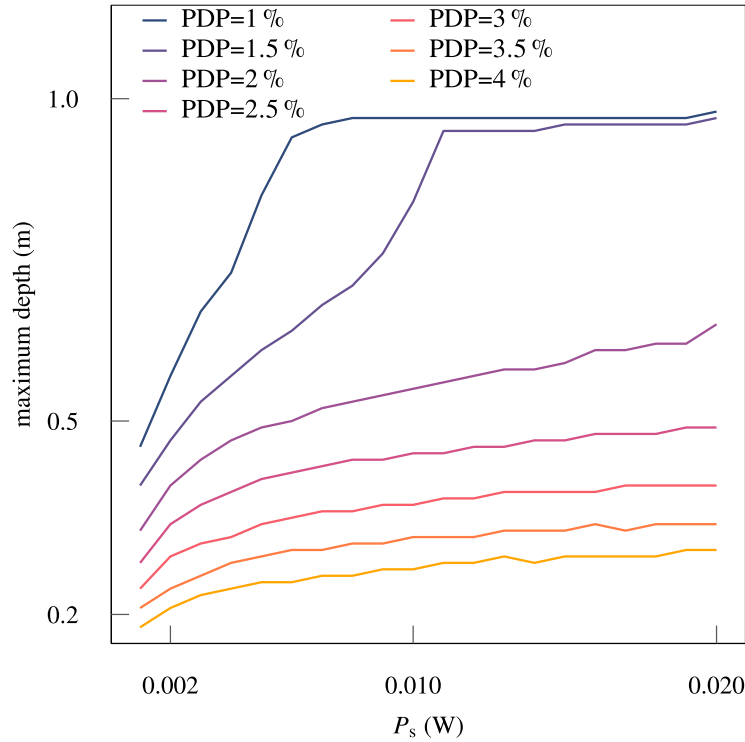


Fig. 11. Depth range under different PDP conditions. The simulation parameters are set as: $N_{\text{TCSPC}} = 30000$, $\Delta_{\text{TDC}} = 25$ ps, $E_n = 0.4 \text{ W m}^{-2} \text{ nm}^{-1}$, and $\rho = 60\%$.

6. Discussion and conclusions

The proposed simulation model based on radiometry and photon-counting statistics of section 2 is a simple method, which is not only limited to simulate D-TOF LIDAR systems for smartphone applications. The simple approach of pixelating the target surface overcame the limitation of previous models [18,20]. In addition, this approach allows to accurately simulate conditions when the target surface is close to the sensor, as well as conditions when the target surface is smaller than the FOI or FOV.

In subsection 3.1, we determine that the main parameter that limits the FSR of the LIDAR system is N_r , either if the target is highly reflective (high ρ values) or under high background irradiance (E_n). The reason behind this effect is that N_r determines the chance that a TDC is available for a laser photon detection. Also, in appendix C, we show that the influence of N_r on the entire system is reduced by increasing the total number of the TDCs; which effectively reduces the background photon rate per individual TDC. In addition, we showed that a longer FSR is obtained by increasing N_{TCSPC} , however this decreases the maximum estimation per unit time that the system can achieve. In addition, we also show in appendix C that a single-TDC system can model a multiple-TDC system by readjusting the simulation parameters, accordingly.

In subsection 3.2, the linearity of the system is mainly influenced by the PDF shape of the first photon timestamp $p_1(t|d_T)$, which is left-skewed with respect to the VCSEL temporal intensity $f(t)$. This effect of a shifted histogram peak even if the target is at the same d_T depends on S_p , which subsequently depends on P_s and ρ . Also, we conclude that if the B_{MSE} is compensated, the reminder σ_{MSE} is improved if the total number of detected laser photons per histogram is incremented. This is achieved by increasing P_s , ρ , or N_{TCSPC} .

In section 4, we analytically calculate the CRLB that corresponds to the depth estimation of a given D-TOF LIDAR system. We approximate the condition of random Poisson points with nonuniform density by simply extending the calculation time of $p_1(t|d_T)$ to infinity, and therefore account for the shot noise. It is important to mention that in this application of single-photon detection, the shot noise has a fundamental influence since the average number of detected photons, per TCSPC cycle, could be lower than the unity.

Finally, we propose the utilization of AI based methods for depth estimation, in section 5. We determine that the ANN can automatically compensate for the systematic based produced by the single-photon detection method. In addition, this method has a significantly improved depth resolution in comparison to the reference method, which is based on finding peaks in the timestamp histogram.

Nomenclature

$\Delta\Phi_{n,i,j}$ radiant flux emitted by $\Delta T_{i,j}$, due to background light photons only

$\Delta\Phi_{s,i,j}$ radiant flux emitted by $\Delta T_{i,j}$, due to VCSEL photons only

$\Delta I_{e,i,j}$ detected radiant intensity by the D-TOF LIDAR sensor, due to VCSEL photons only

$\Delta T_{i,j}$ sub-element of the target surface

$\Delta\Omega_{i,j}$ solid angle between $\Delta T_{i,j}$ and the VCSEL

$\Delta\Phi_{e,i,j}$ radiant flux received at $\Delta T_{i,j}$, due to VCSEL photons only

ΔT_{DC} TDC bin size

η_f optical bandpass filter transmittance at λ

η_l lens transmittance at λ

$\Delta\hat{T}_{i,j}$ area of $\Delta T_{i,j}$

$\hat{d}_T$ estimated target depth

λ VCSEL center wavelength

$\mathbf{H}_T$ timestamp histogram

$\mathcal{I}_1(d_T)$ Fisher information observed in one single detection

$\mathcal{I}_h(d_T)$ Fisher information observed in a $\mathbf{H}_T$

Φ_n radiant flux received at the lens of the D-TOF LIDAR sensor, due to background light photons only

Φ_s radiant flux received at the lens of the D-TOF LIDAR sensor, due to VCSEL photons only

ρ reflectivity of the target surface

σ_1 VCSEL pulse width in terms of standard deviation

σ_{MSE} estimator's standard deviation

$\theta_{e,i,j}$ angle between a VCSEL ray trajectory and the normal to the surface of $\Delta T_{i,j}$

$\theta_{s,i,j}$ angle between the sensor and the normal vector of $\Delta T_{i,j}$

ξ equivalence constant between a systems with single and multiple TDCs

A'_1 equivalent area of the receiver lens for a subgroup of SPADs and single TDC

A_1 area of the receiver lens

B_{MSE} estimator's bias

$d_{e,ij}$ the distance from the VCSEL to ΔT_{ij}

$d_{s,ij}$ distance from ΔT_{ij} to the D-TOF LIDAR sensor

d_T target depth

E_n equivalent solar irradiance at sea level

f VCSEL repetition frequency

$f(t | 2\frac{d_T}{c}, \sigma_1)$ normalized VCSEL intensity over time or timestamp PDF of unsorted photons

$g(t | N_r)$ timestamp PDF of the first detected background photon

$l'(t | d_T)$ normalized likelihood function with respect to T_h

$l(t | d_T)$ likelihood function

N_h Total number of counts contained in $\mathbf{H}_T$

N_r noise rate per time unit

N_{TCSPC} total number of TCSPC cycles

N'_{TCSPC} equivalent TCSPC cycles of a single-TDC sensor utilized for modeling a multiple-TDC sensor

$p_1(t | d_T)$ timestamp PDF of the first detected photon

P_s VCSEL optical power

R timestamp set size

$S_{N+P}(t)$ unsorted timestamp CDF of background light and VCSEL photons

$s_{N+P}(t)$ unsorted timestamp PDF of background light and VCSEL photons

$S_N(t)$ unsorted timestamp CDF of background light photons

$s_N(t)$ unsorted timestamp PDF of background light photons

S_p detected number of VCSEL photons per TCSPC cycle

T maximum time utilized to calculate $p_1(t | d_T)$

T_h maximum time range of the timestamp histogram $\mathbf{H}_T$

$U(t | [0, \Delta_{\text{TDC}}])$ Uniform distribution with support $[0, \Delta_{\text{TDC}}]$

ANNs Artificial neural networks

CRLB Cramér–Rao lower bound

DCR dark cont rate

FF fill factor

FOI field-of-illumination

FOV field-of-view

FPF find peaks function

FSR full scale range

M total number of SPADs

MC Monte Carlo

MSE mean-square-error

PDE photon detection efficiency

PDF probability density function

PDP photon detection probability

SPAD single-photon avalanche diode

TCSPC time-correlated single-photon counting

TDC time-to-digital converter

VCSEL vertical-cavity surface-emitting laser

Appendix A System parameters

Since, we focus our analysis on smartphone LIDAR applications, we define our systems parameters based on the characteristics of commercially available sensors [31–34]. Tables 3, 4 and 5 show parameters that are not classified as either variable nor fixed parameter, and are used in the simulations.

Appendix B PDP influence in depth range

In the depth range analysis, we find out that PDP has a drastic impact on N_r , which is one of the main factor that limits the depth range. Therefore, we evaluate the effect of PDP, separately, by sweeping its value from 1% to 4%. In this simulation, we choose the highest background light condition and target reflectivity (see Fig. 11). The value of E_n is selected according to [32], and the PDP sweep range is derived from [37].

From Fig. 11, it can be observed at high background noise, the increase of PDP decreases the depth range. The reason is that the probability of detecting background photon before the arrival of VCSEL photons increases exponentially with respect to PDP. Therefore, for high values of E_n and PDP, the VCSEL photons are missed. We consider that a maximum distance of 0.5 m is sufficient for a smartphone application [31–34], and so we fixed the PDP value to 2 % in all of the simulations.

We consider that P_s is the easiest controllable parameter in the system. Thus, each time we simulate the depth range, we also sweep the VCSEL power up a maximum value of 20 mW. We consider that if a target is too close to the sensor, for eye safety reason, the sensor can immediately turn off the VCSEL after one complete measurement cycle. Therefore, the exposure time is equal to the inverse of the repetition rate multiplied by N_{TCSPC} . For example, the maximum value of P_s for a class 1 sensor with $N_{\text{TCSPC}} = 30000$ is equal to 12.77 mW [24].

Table 3. SPAD array parameters.

definition	symbol	value
lens area (derived from [31–34])	A_l	0.54 mm ²
lens transmittance (according to [35])	η_l	94 %
bandpass transmittance (according to [36])	η_f	94 %
bandpass width (according to [36])	B_O	10 nm
field-of-view (derived from [31–34])	FOV	21°
total number of SPADs (derived from [31–34])	M	196
fill factor (derived from [12,13,18–20])	FF	25 %

Table 4. VCSEL parameters.

definition	symbol	value
VCSEL center wavelength (according to [31])	λ	940 nm
VCSEL repetition freq. (derived from [31–34])	f	40 MHz
VCSEL pulse width (std) (derived from [31–34])	σ_1	127.65 ps
field-of-illumination (according to [32])	FOI	21°

Table 5. Target surface parameters.

definition	symbol	value
target length (according to [32])	L	260 mm
target width (according to [32])	W	200 mm

Appendix C System with Multiple TDCs

In this section we expand the equations and models of section 2 in order to simulate systems composed of multiple TDCs. As shown in Fig. 1(a), the SPADs are shared in subgroups that are connected to a single TDC. For example, if we have 14 TDCs in the system, the total number of 196 SPADs are shared in subgroups of 14 SPADs per TDC.

By modifying the values of N_{TCSPC} and A_l , we can approximate the condition of having several smaller and statistically independent SPAD arrays or subgroups connected to a single TDC. Therefore, assuming that we have ξ total TDCs in the system, we can recalculate an equivalent and smaller lens area as follows

$$A'_l = \frac{A_l}{\xi}. \quad (26)$$

By replacing A_l and A'_l in Eqs. (9) and (12), we can obtain a TOF histogram of a subgroup of SPADs connected to a single TDC. In this calculation, we assume that the light power is distributed uniformly across the SPAD array's active area. Since we have multiple subgroups, the overall effect is accounted by increasing N_{TCSPC} , as follows

$$N'_{\text{TCSPC}} = \xi N_{\text{TCSPC}}. \quad (27)$$

We re-run the MC code and fully account for a system composed of multiple TDCs, and compare to the model of a single TDC with modified A'_l and N'_{TCSPC} . The output TOF histogram for both cases are shown in Fig. 12, and obviously, they are statically equivalent.

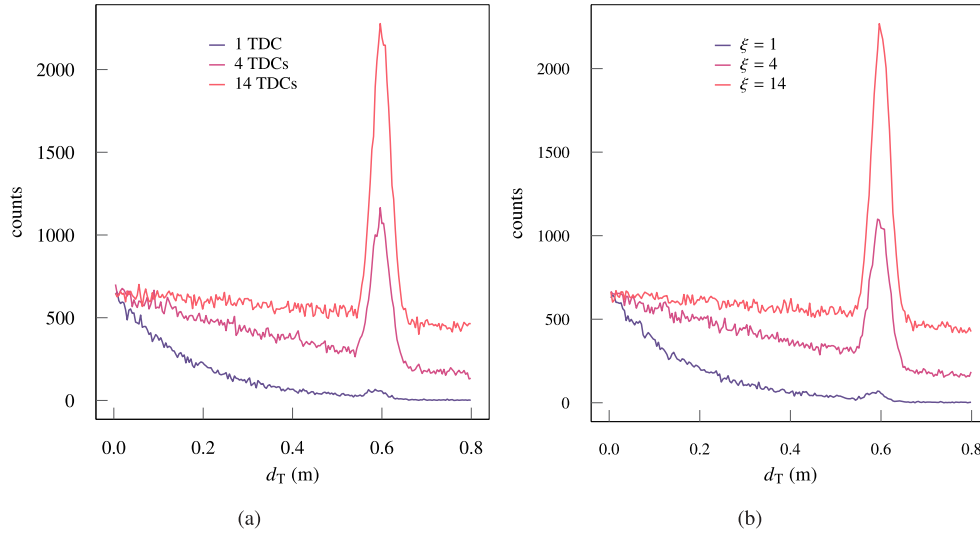


Fig. 12. Comparison between fully modeled multiple-TDC system and equivalent single-TDC system. The system parameters are $P_s = 7.36$ mW, $N_{\text{TCSPC}} = 30000$, $\Delta_{\text{TDC}} = 25$ ps, $E_n = 0.4$ W m⁻² nm⁻¹, $\rho = 60$ %. (a) TOF histogram obtained with fully modeled a system with 14 TDCs. (b) TOF histogram obtained with modeling the equivalent single-TDC system.

Also, it is importance to note that just by increasing the number of TDCs in the system, without using any background-light rejection method, the noise counts are reduced and signal counts are increased.

Appendix D Improved find peak function

According to the analysis in section 3.2, the signal peak found in $\mathbf{H}_T$ is shifted with respect to the VCSEL intensity peak. The reason is that the shape of the signal peak is left-skewed with respect to the VCSEL pulse, because the system can only detect the first photon (see Fig. 13). So, in order to compensate the bias we utilize a simple linear regression, which is expressed as follows

$$\hat{d}_T' = K\hat{d}_T + B, \quad (28)$$

where K and B are coefficients of the linear regression. These two coefficients are calculated based on fitting the typical obtained depth with the target depth.

We add another improvement to the FPF that is a simple background subtraction algorithm, which is performed in three steps. First, background-light timestamps are measured with the VCSEL intentionally turned off, and we save the corresponding timestamp histogram. Next, we turn on the VCSEL, capture timestamps and generate a new timestamp histogram. Finally, we input to the FPF the subtraction between the two previous histograms. In this way, we remove the exponential shape contained in the $\mathbf{H}_T$.

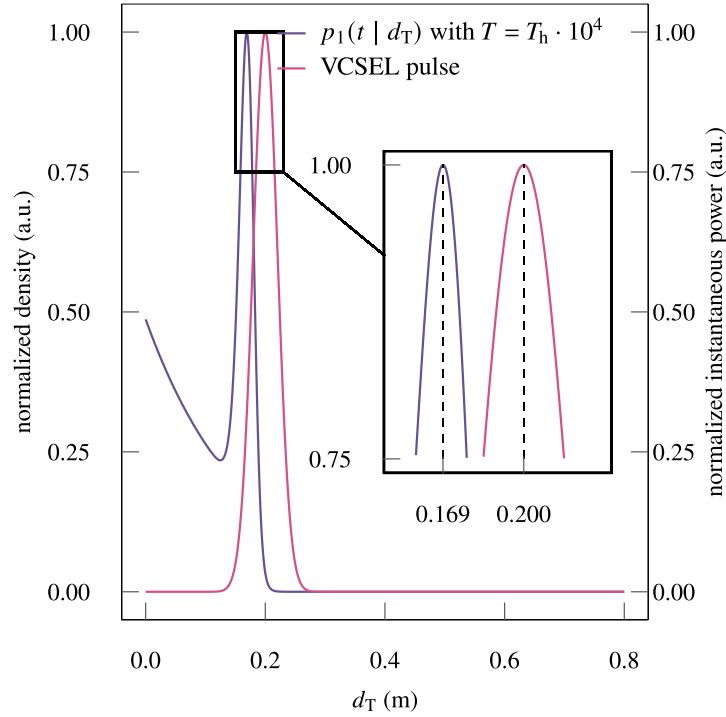


Fig. 13. $p_1(t|d_T)$ against VCSEL pulse. The system parameters are $P_s = 7.36$ mW, $E_n = 0.04$ W m⁻² nm⁻¹, $d_T = 0.2$ m, and ρ is 8 %.

Appendix E Influence of TDC timing jitter

We perform an extra MC simulation to depict the influence of the TDC timing jitter on the system performance. It is important to note that the SPAD jitter, VCSEL pulse width, and TDC timing jitter cannot be added in quadrature. The reason behind this is that the photons go through a sorting process before they are detected by the TDC. Also due to this reason, the laser peak width in the timestamp histogram can be narrower than the addition in quadrature of the SPAD jitter and VCSEL pulse width; given that the number of average detected photons is larger than the unity [38].

Figure 14 shows $\mathbf{H}_T$ when we model the TDC timing jitter as a Gaussian R.V., which is added to the timestamps after the sorting process. We intentionally reduce Δ_{TDC} to 1 ps and increase N_{TCSPC} to 10^7 , to observe the influence of the timing jitter in $\mathbf{H}_T$. As shown in Fig. 14, a TDC timing jitter up to 100 ps FWHM has not significant influence on the peak width in $\mathbf{H}_T$ that corresponds to the laser photons.

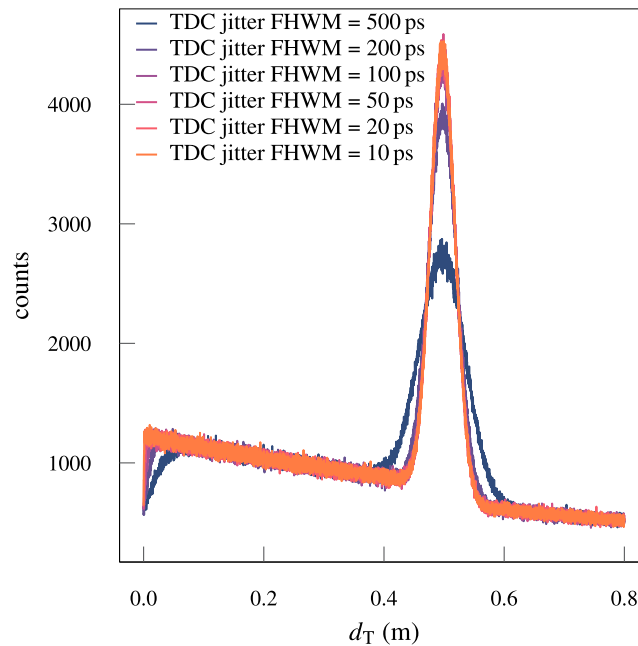


Fig. 14. H_T calculated with the MC model for several TDC timing jitter values. The system parameters are ρ is 8 %, $P_s = 7.36$ mW, $N_{\text{TCSPC}} = 10^7$, $\Delta_{\text{TDC}} = 1$ ps, $E_n = 0.4$ W m⁻² nm⁻¹, and $d_T = 0.5$ m.

Acknowledgments. Silicon Integrated co-funded this research.

Disclosures. The authors declare no conflict of interests.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. D. Göhring, M. Wang, M. Schnürmacher, *et al.*, "Radar/lidar sensor fusion for car-following on highways," in *The 5th International Conference on Automation, Robotics and Applications*, (2011), pp. 407–412.
2. H. Cho, Y.-W. Seo, B. V. Kumar, *et al.*, "A multi-sensor fusion system for moving object detection and tracking in urban driving environments," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, (2014), pp. 1836–1843.
3. J. Kocic, N. Jovicic, and V. Drndarevic, "Sensors and sensor fusion in autonomous vehicles," in *2018 26th Telecommunications Forum (TELFOR)*, (2018), pp. 420–425.
4. M. Mahlich, R. Schweiger, W. Ritter, *et al.*, "Sensorfusion using spatio-temporal aligned video and lidar for improved vehicle detection," in *2006 IEEE Intelligent Vehicles Symposium*, (2006), pp. 424–429.
5. R. Halterman and M. Bruch, "Velodyne HDL-64E lidar for unmanned surface vehicle obstacle detection," in *Unmanned Systems Technology XII*, vol. 7692 G. R. Gerhart, D. W. Gage, and C. M. Shoemaker, eds., International Society for Optics and Photonics (SPIE, 2010), pp. 123–130.
6. cnet, "Lidar is one of the iphone and ipad's coolest tricks. here's what else it can do," <https://www.cnet.com/tech/mobile/lidar-is-one-of-the-iphone-ipad-coolest-tricks-its-only-getting-better/>. (accessed January 17, 2022).
7. J. H. Gao and L.-S. Peh, "A smartphone-based laser distance sensor for outdoor environments," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, (2016), pp. 2922–2929.
8. J. A. P. Y. Phun and C. Safitri, "Smartphone authentication with hand gesture recognition (HGR) using LiDAR," in *2021 5th International Conference on Informatics and Computational Sciences (ICICoS)*, (2021), pp. 93–98.
9. T. Wu, J. Liu, Z. Li, *et al.*, "Accurate smartphone indoor visual positioning based on a high-precision 3D photorealistic map," *Sensors* **18**(6), 1974 (2018).
10. G. Luetzenburg, A. Kroon, and A. A. Bjørk, "Evaluation of the apple iphone 12 pro LiDAR for an application in geosciences," *Sci. Rep.* **11**(1), 22221 (2021).

11. A. Payne, A. Daniel, A. Mehta, *et al.*, “7.6 a 512×424 CMOS 3D time-of-flight image sensor with multi-frequency photo-demodulation up to 130mhz and 2gs/s ADC,” in *2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, (2014), pp. 134–135.
12. M. Perenzoni, D. Perenzoni, and D. Stoppa, “A 64 × 64-pixels digital silicon photomultiplier direct TOF sensor with 100-MPhotons/s/pixel background rejection and imaging/altimeter mode with 0.14% precision up to 6 kmh<https://www.overleaf.com/project/61dea127f975132f45f2ef37> for spacecraft navigation and landing,” *IEEE J. Solid-State Circuits* **52**(1), 151–160 (2017).
13. C. Niclass, M. Soga, H. Matsubara, *et al.*, “A 100-m range 10-frame/s 340 × 96-pixel time-of-flight depth sensor in 0.18- μ m CMOS,” *IEEE J. Solid-State Circuits* **48**(2), 559–572 (2013).
14. S. W. Hutchings, N. Johnston, I. Gyongy, *et al.*, “A reconfigurable 3-D-stacked SPAD imager with in-pixel histogramming for flash LIDAR or high-speed time-of-flight imaging,” *IEEE J. Solid-State Circuits* **54**(11), 2947–2956 (2019).
15. A. Süß, V. Rochus, M. Rosmeulen, *et al.*, “Benchmarking time-of-flight based depth measurement techniques,” in *Smart Photonic and Optoelectronic Integrated Circuits XVIII*, vol. 9751 S. He, E.-H. Lee, and L. A. Eldada, eds., International Society for Optics and Photonics (SPIE, 2016), pp. 199–217.
16. F. Villa, F. Severini, F. Madonini, *et al.*, “SPADs and SiPMs arrays for long-range high-speed light detection and ranging (LiDAR),” *Sensors* **21**(11), 3839 (2021).
17. R. Roriz, J. Cabral, and T. Gomes, “Automotive lidar technology: A survey,” *IEEE Trans. Intell. Transport. Syst.* **23**(7), 6282–6297 (2022).
18. P. Padmanabhan, C. Zhang, and E. Charbon, “Modeling and analysis of a direct time-of-flight sensor architecture for LiDAR applications,” *Sensors* **19**(24), 5464 (2019).
19. C. Niclass, M. Soga, H. Matsubara, *et al.*, “A 0.18- μ m CMOS SoC for a 100-m-range 10-frame/s 200×96-pixel time-of-flight depth sensor,” *IEEE J. Solid-State Circuits* **49**(1), 315–330 (2014).
20. A. Ronchini Ximenes, “Modular time-of-flight image sensor for light detection and ranging: A digital approach to lidar,” <https://doi.org/10.4233/uuid:c434368a-9a67-45de-a66f-f5dc30430e03> (2019). (accessed April 03, 2023).
21. A. Ronchini Ximenes, P. Padmanabhan, and E. Charbon, “Mutually coupled time-to-digital converters (tdcs) for direct time-of-flight (dtof) image sensors,” *Sensors* **18**(10), 3413 (2018).
22. M. W. Fishburn, *Fundamentals of CMOS single-photon avalanche diodes* (Delft University of Technology, 2012), pp. 55–68.
23. Mathworks, “findpeaks function reference,” <https://www.mathworks.com/help/signal/ref/findpeaks.html>. (accessed April 03, 2023).
24. M. Chen, *ID-TOF: system modeling, processing algorithm design and implementation* (Delft University of Technology, 2022), pp. 11–13, Msc. Dissertation.
25. B. Arnold, N. Balakrishnan, and H. Nagaraja, *A First Course in Order Statistics*, Classics in Applied Mathematics (Society for Industrial and Applied Mathematics, 1992).
26. M. Fishburn and E. Charbon, “System tradeoffs in gamma-ray detection utilizing spad arrays and scintillators,” *IEEE Trans. Nucl. Sci.* **57**(5), 2549–2557 (2010).
27. A. Papoulis, *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, 1984), pp. 55–58, 76–77, McGraw-Hill series in electrical engineering. Communications and Information theory.
28. S. Seifert, H. van Dam, and D. Schaart, “The lower bound on the timing resolution of scintillation detectors,” *Phys. Med. Biol.* **57**(7), 1797 (2012).
29. E. Venialgo, S. Mandai, T. Gong, *et al.*, “Time estimation with multichannel digital silicon photomultipliers,” *Phys. Med. Biol.* **60**(6), 2435 (2015).
30. M. Hagan and M. Menhaj, “Training feedforward networks with the marquardt algorithm,” *IEEE Trans. Neural Netw.* **5**(6), 989–993 (1994).
31. AMS, “TMF882X datasheet,” https://ams.com/documents/20143/6015057/TMF882X_DS000693_5-00.pdf. (accessed April 03, 2023).
32. AMS, “TMF8701 datasheet,” https://ams.com/documents/20143/4409399/TMF8701_DS000602_8-00.pdf. (accessed April 03, 2023).
33. STmicroelectronics, “VL53L3CX datasheet,” <https://www.st.com/resource/en/datasheet/vl53l3cx.pdf>. (accessed April 03, 2023).
34. STmicroelectronics, “VL6180 datasheet,” <https://www.st.com/resource/en/datasheet/vl6180.pdf>. (accessed April 03, 2023).
35. Thorlabs, “N-BK7 transmittance,” https://www.thorlabs.com/images/TabImages/Uncoated_N-BK7_Transmission_780.gif. (accessed April 03, 2023).
36. Thorlabs, “FBH940-10 specifications,” https://www.thorlabs.com/images/tabimages/FBH940-10_Transmission_G1-370.gif. (accessed April 03, 2023).
37. ONsemiconductor, “RB-series SiPM sensors datasheet,” <https://nl.mouser.com/datasheet/2/308/MICROB-SERIES-D-1489599.pdf>. (accessed April 05, 2023).
38. S. Mandai, E. Venialgo, and E. Charbon, “Timing optimization utilizing order statistics and multichannel digital silicon photomultipliers,” *Opt. Lett.* **39**(3), 552–554 (2014).