UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE MATHEMATIQUES, SCIENCES DE L'INFORMATION ET DE L'INGENIEUR – ED269

UMR 7357 - Laboratoire des sciences de l'Ingénieur, de l'Informatique et de l'Imagerie (ICube)

THÈSE présentée par :

Wassim KHADDOUR

soutenue le : **21 septembre 2023**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Micro et Nanoélectronique

Versatile Time-Correlated Single Photon Counting System and Applications

THÈSE dirigée par :

M. MADEC Morgan	MCF HDR, Université de Strasbourg
M. UHRING Wilfried	Professeur, Université de Strasbourg

RAPPORTEURS :

M. GINHAC Dominique	Professeur, Université de Bourgogne
M. FERUGLIO Sylvain	MDC HDR, Sorbonne Université

AUTRES MEMBRES DU JURY :

Mme. DESGREYS Patricia	Professeur, Telecom Paris
M. GRIFFITHS Andrew	Professeur, ESPCI
M. DADOUCHE Foudil	MCF, Université de Strasbourg
M. DUMAS Norbert	MCF, Université de Strasbourg

INVITÉS :

M. CLERMONT Lionel	Ingénieur de recherche, Centre Spatial de Liège

Université de Strasbourg

Wassim KHADDOUR

Versatile Time-Correlated Single Photon Counting System and Applications

École doctorale Mathématiques, sciences de l'information et de l'ingénieur | ED 269 Université de Strasbourg

Résumé

Dans cette thèse, nous présentons un système TCSPC polyvalent sur une plateforme SoC-FPGA à faible coût, capable de mesurer des signaux optiques avec haute sensibilité, résolution et précision. Nous proposons deux conceptions de TDC, un système efficace de transfert de données, et un programme C qui s'exécute sur un processeur embarqué. Nous introduisons également une nouvelle méthode de calibration et une méthode robuste pour la détection de microgouttelettes en microfluidique. Les performances du système ont été évaluées en résolution, précision, et stabilité. Le système a été utilisé dans quatre applications dans les domaines environnementale, biologique et astronomique : capteur de pollution de l'eau, capteur de turbidité, criblage et tri de microgouttelettes et caractérisation de la lumière parasite dans les optiques spatiales. Le système réalise des mesures TCSPC de haute performance et avec flexibilité pour diverses applications.

TDC, TCSPC, FPGA, FLT, tri de microgouttelettes.

Résumé en anglais

In this thesis, we present a versatile TCSPC system on a low-cost SoC-FPGA platform that can measure periodic optical signals with high sensitivity, resolution and precision. We propose two TDC designs, an efficient data transfer design, and a data processing C program that runs on an embedded processor. We also introduce a novel calibration method for asynchronous TDCs and a robust method for the detection of microfluidic droplets. The system performance was evaluated in terms of resolution, precision, accuracy and stability. The system was employed in four different applications in environmental, biological, and astronomical fields; water pollution sensor, turbidity measurement, microdroplet screening and sorting, and stray light characterization in spatial optics. The results showed that the system can achieve high-performance TCSPC measurements at low cost and high flexibility for various applications.

TDC, TCSPC, FPGA, FLT, Microfluidic microdroplet sorting.

Acknowledgement

This thesis is the result of a long and demanding journey, which I could not have accomplished without the help and support of many people.

I am extremely grateful to my thesis directors, Prof. Wilfried Uhring and Dr. Morgan Madec for their invaluable guidance, and encouragement throughout this project. My deepest and utmost gratitude goes to Prof. Wilfried Uhring who has given me the opportunity to pursue this research project. He was more than a mentor, a supporter, and a friend to me. His remarkable passion and enthusiasm for his research inspired me and his expertise and encouragement guided me through the challenges and difficulties of this work. Words fail to express my gratitude for his kindness, his confidence in me, and the friendship he extended to me since my first day in the laboratory. It was a true honor for me to be one of his students. My heartfelt gratitude also goes to to Dr. Foudil Dadouche, who opened the door for me to embark on this journey, as he was the one who first recommended me to join the laboratory in an internship; I thank him for his precious advice, invaluable guidance and continuous support. I am also profoundly grateful to Dr. Norbert Dumas for his guidance and vital help especially during the long days of manipulations and experiments.

I would also like to convey my sincere appreciation and gratitude to my defense committee, consisting of Prof. Patricia Desgreys, Prof. Dominique Ginhac, Dr. Sylvain Feruglio, Prof. Andrew Griffiths, and Dr. Lionel Clermont for their constructive and positive feedback and recommendations. I appreciate their time and expertise in evaluating my thesis.

I am also thankful to Dr. Anne Pallarès, Dr. Philippe Schmitt, and Gwenaël Pallarès from the MécaFlu research team, and Lionel Clermont and Pascal Blain from the "Centre Spatial de Liege", for their collaboration and contribution to my research.

I would like to thank my fellow PhD students and all the members of ICube laboratory, especially Nicolas Colin, Pascal Leindecker, and Jérémy Bartringer for their help and technical support.

Last but not least, I wish to express my heartfelt gratitude to my friends and family, who have always been there for me with their unconditional love and support. I am grateful to my sisters and brother for their care and encouragement, and I owe a special thanks to my sister Rana without whom this journey would not have been possible. The utmost and ultimate gratitude goes to my parents for all their hard work, dedication, and support over the years. I will be forever indebted to them, and to them I dedicate this thesis.

Abstract

Time-correlated single photon counting (TCSPC) is a powerful technique for recording periodic optical signals with high sensitivity, resolution, and precision. It is applied in a wide range of applications in physics, biology, engineering, and quantum technologies. Traditional TCSPC systems have been widely utilized in scientific and industrial research. However, these systems have several drawbacks such as high cost, bulky size, and low performance. To overcome these limitations, integrated TCSPC systems have been developed in application-specific integrated circuits (ASIC). The different components of the TCSPC system as well as the data processing electronics are integrated on a single silicon chip, enabling the design of high-performance compact TCSPC systems. However, ASIC systems also have many disadvantages such as the long development cycle, high cost for small production runs, and lack of flexibility. On the other hand, recent advances in field programmable gate array (FPGA) technology have enabled the implementation of high-performance TCSPC systems with many features such as low cost, short development cycle, flexibility, and reconfigurability. Moreover, a system-on-chip FPGA (SoC-FPGA) integrates a powerful embedded processor with the FPGA fabric on the same chip, providing an efficient platform for the realization of high-performance hardware/software co-processing applications.

This thesis presents the design and implementation of a versatile TCSPC system on a low-cost SoC-FPGA platform that can be used for different applications. The thesis proposes two time-to-digital converter (TDC) designs with a resolution of about 20 ps, one for synchronous light sources and the other for asynchronous ones. It also proposes an efficient design for high-speed data transfer between the FPGA and the embedded processor, as well as a bare-metal C program that runs on the processor and performs all the TCSPC data processing onboard. Furthermore, this thesis introduces a novel calibration method for asynchronous TDCs that enhances the linearity by a factor of 10.

The performance of the proposed system is evaluated in terms of speed, accuracy, resolution and linearity. The versatility of the system is demonstrated by employing it in four different applications: a water pollution sensor, time-resolved optical turbidity, high-throughput microfluidic droplet screening, and sorting and stray light characterization.

For the droplet sorting application, a robust method for the detection of microfluidic droplets in real time is proposed. With this method, the system is able to detect and screen droplets based on their fluorescence lifetime at a droplet rate of more than 3500 droplets/second.

The contributions of this thesis advance the state-of-the-art in FPGA-based TCSPC systems and high-throughput droplet screening and provide a low-cost, flexible, high-performance TCSPC system for various applications that require high-precision time measurement and droplet analysis.

Contents

List of Figures

List of Tables

Abbreviations

ACP	Accelerator Coherency Port
ADC	Analog-to-Digital Converter
ALM	Adaptive Logic Module
APD	Avalanche Photodiodes
API	Application Programming Interface
ASIC	Application-Specific Integrated Circuit
CDPU	Control and Data Processing Unit
CFD	Constant Fraction Discriminator
CMOS	Complementary Metal-Oxide-Semiconductor
COG	Center Of Gravity
CPU	Central Processing Unit
CSL	Center spatial de liege
DAC	Digital-to-Analog Converter
DCR	Dark Count Rate
DFF	D Flip-Flop
DMA	Direct Memory Access
DNL	Differential Nonlinearity
DOT	Diffuse Optical Tomography
DSP	Digital Signal Processing
EOM	Electro-Optic Modulator
FACS	Fluorescence-Activated Cell Sorting
FIDA	Fluorescence Intensity Distribution Analysis
FIFO	First In First Out
FLIM	Fluorescence Lifetime Imaging Microscopy
FLT	Fluorescence Lifetime
FOV	Field Of View
FPGA	Field-Programmable Gate Array
FRET	Förster Resonance Energy Transfer
FWHM	Full Width at Half Maximum

HDL	Hardware Descriptions Language
HPS	Hard Processor System
HTS	High-throughput screening
HWLIB	Hardware Library
INL	Integral Nonlinearity
IP	Intellectual Property
IRF	Instrument Response Function
LAB	Logic Array Block
LIDAR	Light Detection And Ranging
LSB	Least Significant Byte
LUT	Lookup Table
MLE	Maximum Likelihood Estimator
MMU	Memory Management Unit
MPU	Microprocessor Unit
OAS	Optical Analysis System
PAH	Polycyclic Aromatic Hydrocarbon
PBS	Phosphate-Buffered Saline
PDMS	Polydimethylsiloxane
PIO	parallel I/O
PLL	Phase-Locked Loop
PMT	Photomultiplier Tube
PVT	Process, Voltage and Temperature
QE	Quantum Efficiency
QY	Quantum Yield
RCU	Replaceable Concentrator Unit
RL	Register Line
RMS	Root Mean Square
SCU	Snoop Control Unit
SER	Single Electron Response
SNR	Signal-to-Noise Ratio
SPAD	Single-Photon Avalanche Diode
SPI	Serial Peripheral Interface
SPST	Spatial Point Source Transmittance
STD	Standard Deviation
TAC	Time-to-Amplitude Converter

TCSPC	Time-Correlated Single Photon Counting
TDC	Time-to-Digital Converter
TDL	Tapped Delay Line
TOF	Time Of Flight
TPSF	Temporal Point Spread Function
TROT	Time-Resolved Optical Turbidity
TSS	Total Suspended Solids
TTS	Transit Time Spread
UV	Ultra-Violent
VDL	Vernier Delay Line
VHDL	Very High-Speed Hardware Description Language
WFD	Water Framework Directive
WPS	Water Pollution Sensor

Chapter 1: Introduction

1.1 Background

Time-correlated single photon counting (TCSPC) is a powerful technique for the recording of periodic optical signals with extremely high sensitivity, resolution, and precision. It is suitable for applications that deal with low-level signals, such as fluorescence decay measurements, light detection and ranging (LIDAR), and diffuse optical tomography (DOT).

TCSPC originated from the 'delayed coincidence' method, which was used to measure the lifetimes of excited nuclear states [1, 2]. In the 1960s, photon counting was applied in photon correlation spectroscopy for the measurement of diffusion coefficients [3]. In that time, flashlamps with short pulse widths of around 2 ns became available, which enabled the TCSPC technique [4]. The first application of TCSPC for the spectroscopy of excited molecules was reported in the early 1970s. Since then, this technique has been widely adopted for time-resolved spectroscopy, particularly for assessing fluorescence lifetimes in solutions [5-12]. Nonetheless, due to the low repetition rate of the light sources and the limited speed of the signal processing electronics in the 1970s and 1980s, the acquisition times were extremely long to integrate a sufficient number of photon counts [5, 6, 9, 11, 13, 14]. Furthermore, TCSPC was limited to be one-dimensional, and restricted to the recording of periodic light signals. Rapid advances in electronics and laser technology during the 1990s enabled significant increases in the repetition rate and power of pulsed light sources. This in turn led to important improvements in TCSPC systems in terms of measurement speed, accuracy, and sensitivity. Furthermore, the development of laser scanning confocal microscopes and the arrays of photon detectors allowed the implementation of multidimensional TCSPC systems for further applications, such as fluorescence lifetime imaging (FLIM) [15, 16].

The general principle of TCSPC is illustrated in Figure 1.1. It is based on the detection of single photons of a periodic optical signal, the measurement of the relative arrival time of the detected photons within the signal period, and the creation of a histogram from the arrival times. By accumulating a large number of photon counts, the constructed histogram represents the waveform of the periodic signal [10, 17]. This technique requires a low intensity and a high

repetition rate of the signal so that the probability of detecting one photon in one signal period is much less than one. Typically, an average count rate of 0.01 to 0.1 photon/period is acceptable [12]. Much higher count rates would lead to distortions in the recorded signal due to the pile-up effect.

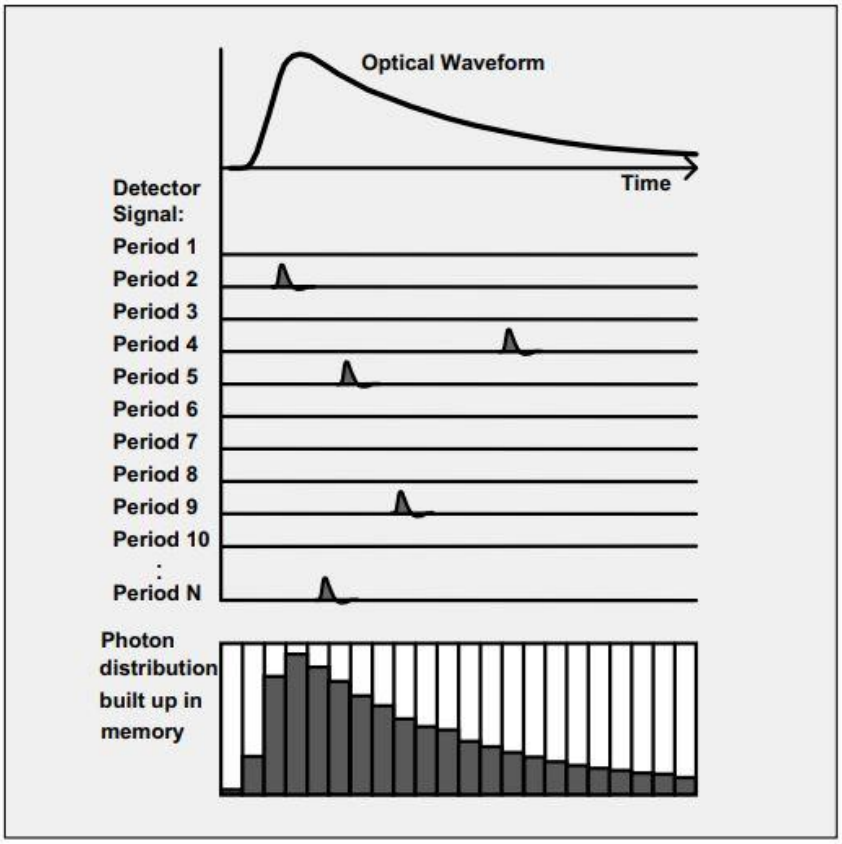


Figure 1.1 General principle of time-correlated single photon counting [14].

TCSPC has a broad range of applications in various scientific research and industrial fields such as physics, biology, and engineering. Some of the main applications are:

1- Time-resolved fluorescence lifetime measurement: fluorescence lifetime is the time required for a population of excited molecules to decrease exponentially by a factor of e as a result of fluorescence, i.e. emitting photons, and other nonradiative processes [18]. In biology, fluorescence measurement is a powerful tool to observe biological processes [19]. Fluorescence characteristics of dyes or other molecules are influenced by the reactions taking place. The fluorescence intensity depends on the reaction under study. However, it also depends on the probing volume and many other parameters such as the concentration and the excitation light intensity. In contrast, fluorescence lifetime (FLT)

is an intrinsic property that is independent of these parameters [20]. Hence, fluorescence lifetime measurement provides more reliability than intensity measurement, and it has various applications in many fields such as biochemistry, environmental sensing, and biomedical analysis [21].

TCSPC is a typical method for measuring fluorescence lifetime with high temporal resolution. It is used in a wide variety of applications, such as single-point spectroscopy [22], fluorescence lifetime imaging microscopy (FLIM) [23, 24], and Forster Resonance Energy Transfer (FRET) [25].

2- Light detection and ranging (LIDAR) applications: LIDAR is a popular technique for distance determination based on the time of flight (TOF) measurement [26, 27]. The LIDAR technique is widely applied in numerous areas, including autonomous vehicles, industrial robots, drones [28], augmented reality [29], and space science [30]. In LIDAR systems, a light source, usually a laser, emits short pulses. A portion of the emitted light is reflected back by the target object and detected by a photon detector. The time interval between the emission and detection is measured by a time-to-digital converter (TDC) and is then used to calculate the distance to the target [31]. TCSPC-based LIDAR provides high-precision depth measurement with high sensitivity to low-level signals.

3- Quantum measurements: The TCSPC technique is an essential tool for various emerging quantum technologies, such as quantum imaging and sensing, quantum-state preparations, quantum cryptography, and emission tomography [32].

Figure 1.2 shows the architecture of a classic TCSPC setup. A photon detector, commonly a single photon avalanche diode (SPAD) or a photomultiplier tube (PMT), generates a pulse for each detected photon, which is called the *STOP* signal. However, if a PMT is used as the photon detector, a constant fraction discriminator (CFD) should be used to reduce the timing jitter resulting from the amplitude jitter of the PMT [33].

A photodiode is generally used to generate the reference signal, i.e. the *START* signal, synchronous with the excitation light pulses. A second CFD is used to reduce the timing jitter of the *START* signal.

A timing channel measures the arrival time of the detected photons as follows: a time-to-amplitude converter (TAC) receives the *START* and *STOP* signals and generates an analog signal with an amplitude proportional to the time interval between these signals. The output of the TAC is then amplified and fed to an analog-to-digital converter (ADC) to be converted into

a digital value representing the photon arrival time. The digital value serves as an address for a memory unit. Each memory location corresponds to a specific time bin and stores the number of photons detected within that bin. When a photon is detected, the number of counts in the addressed memory location corresponding to its arrival time is incremented by one. Thus, a histogram representing the temporal distribution of the detected photons throughout the measurement is built up in the memory. Finally, the measurement histogram is sent to a PC for further data processing, for instance, computing the fluorescence lifetime.

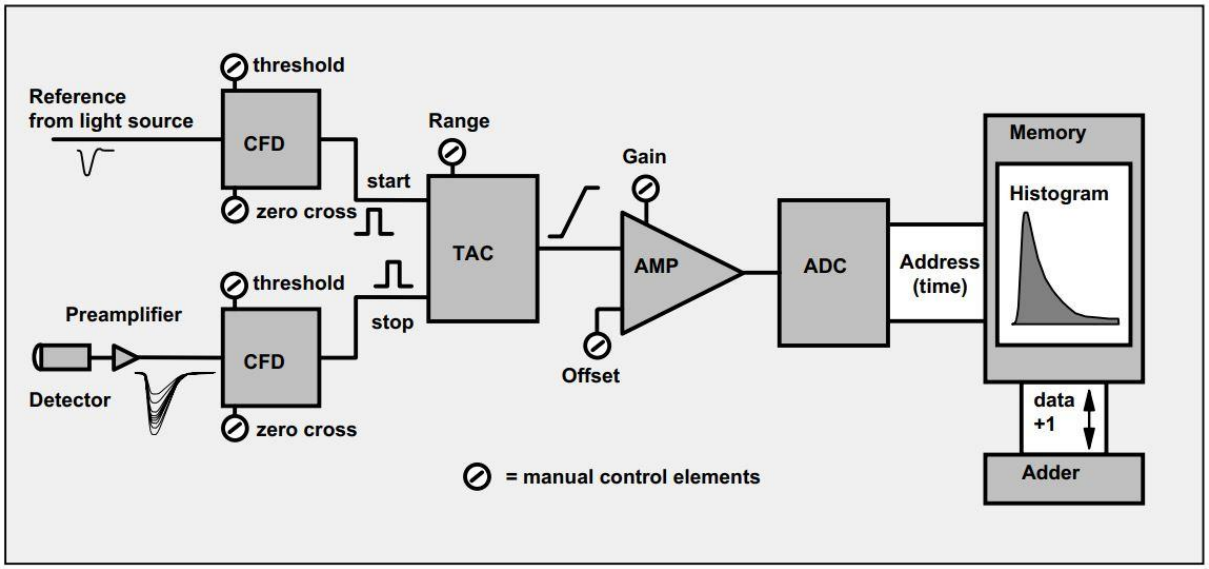


Figure 1.2 Classic TCSPC setup [14].

Traditional TCSPC systems have many limitations such as high cost, bulky size, and low pulse repetition rates and detection rates. To overcome these limitations, integrated TCSPC systems have been developed to replace the discrete electronics of classic TCSPC systems. The dead time was reduced and the count rate was improved by integrating arrays of SPADS, high-throughput fully digital TDC, and embedded signal processing on a single silicon chip as an application-specific integrated circuit (ASIC) [34]. Furthermore, recent advances in FPGA technology have enabled the implementation of high-performance TCSPC systems with high-resolution TDCs. In recent years, FPGA-based TCSPC systems have been the focus of many scientific research works [35-37].

1.2 Research Aim

The aim of this thesis is to design and implement a versatile TCSPC system on a low-cost SoC-FPGA platform that can be employed in different applications. The TCSPC technique is used

in a wide variety of applications in various fields. However, each application requires specific customizations and different post-processing of the measurement histogram data. Moreover, in typical TCSPC systems, the measurement histogram is transferred to a PC for data processing. The limited transfer rate of the interface between the system electronics and the PC creates a bottleneck that limits the system performance and prevents real-time operation. Implementing the data processing, including calibration, background suppression and post-processing, on an ASIC is a complex, inflexible, and expensive approach. On the other hand, the recent advancement in technology enables the integration of powerful hard processor systems with FPGA fabrics in a single chip, creating System-on-Chip FPGAs (SoC-FPGAs) [36]. These devices are available at low cost and offer an excellent platform for the development of high-performance hardware/software co-processing applications. This research addresses the limitations of TCSPC systems in terms of cost and versatility, by taking advantage of the features of SoC-FPGA platforms, such as low cost, flexibility, short development time, and the embedded hard processor system. The objectives of this research are:

- To design and implement two time-to-digital converters (TDCs) with high-resolution and low-nonlinearity for TCSPC systems on a low-cost SoC-FPGA kit. These TDCs should be able to measure the arrival time of photons with high accuracy and at high rates and to operate with different light sources: one for synchronous sources and the other for asynchronous ones.
- To realize a high-speed data transfer system based on direct-memory-access (DMA) for the data transfer between the FPGA and the embedded processor. This design should ensure transferring the time measurement data at a high rate that enables real-time operation.
- To develop a bare-metal C program that runs on the embedded processor and performs all the data processing, including the TDC calibration, the different histogram corrections, and the data post-processing, onboard which enables supporting applications that require real-time operation.
- To evaluate the performance of the proposed TCSPC system in terms of speed, accuracy, resolution, and linearity.
- To demonstrate the versatility of the system by deploying it in four different applications which are: a water pollution sensor (WPS), time-resolved optical turbidity (TROT), high-throughput microfluidic droplet screening and sorting, and stray light characterization.

Figure 1.3 shows a schematic diagram of the proposed co-design of the TCSPC system. The TDC design will be implemented on the FPGA logic, while the Data processing will be executed on the hard processor system of the SoC-FPGA platform. The system will also integrate a custom-built PCB motherboard that supplies stable power to the SoC-FPGA, enables USB communications with the PC, and incorporates a laser diode pulse generator, among other components.

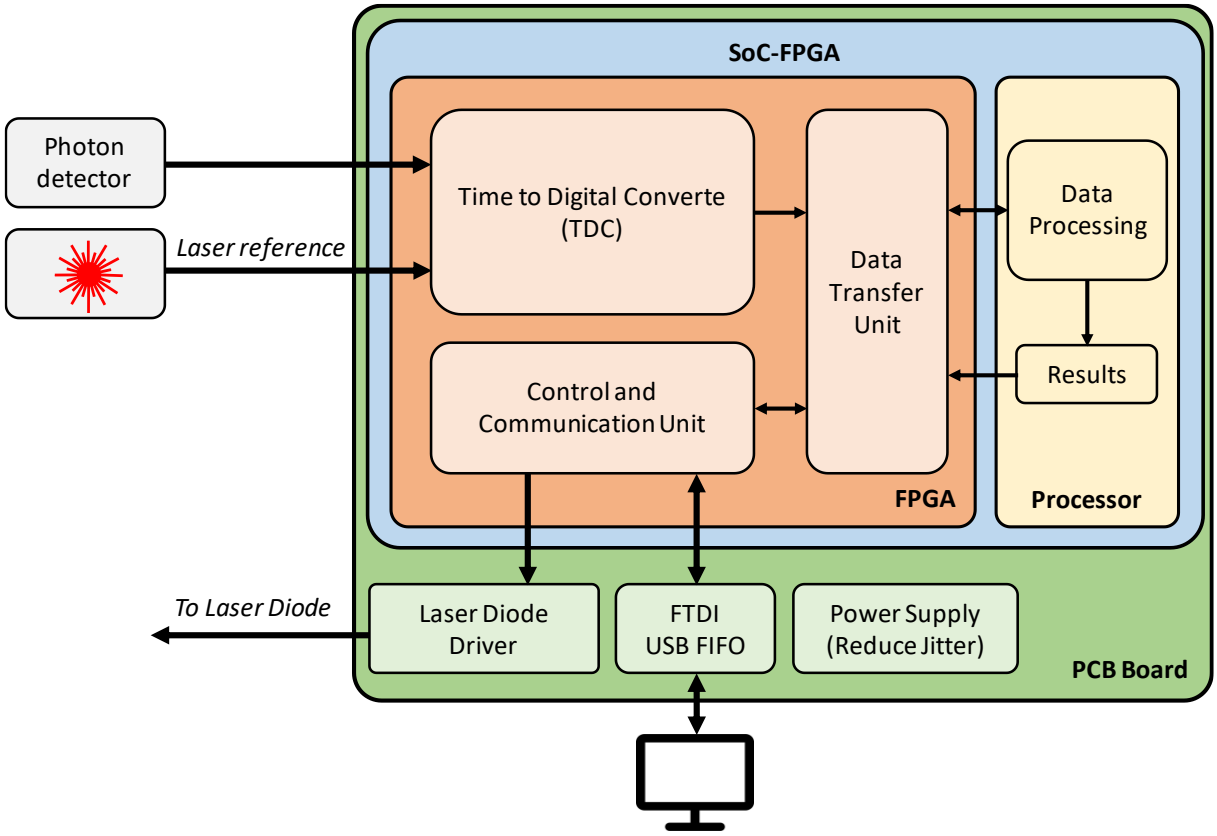


Figure 1.3 Schematic diagram of the proposed hardware/software co-design of the TCSPC system.

1.3 Contributions

The main contributions of this thesis are as follows:

- It presents the design and the implementation of two TDCs, a synchronous one and an asynchronous one, on a low-cost Cyclone V FPGA with a temporal resolution of about 20 ps and a high linearity.
- It proposes and implements an innovative calibration method for asynchronous TDCs that improves the linearity by up to an order of magnitude compared to the conventional bin-by-bin method applied to histograms.

- It develops a robust method for the detection of microfluidic droplets at high rates and in real-time. This method enables the TCSPC system to be applied for high-throughput droplet screening at more than 3500 droplets/second based on the fluorescence lifetime, which surpasses the droplet rate achieved by a more complex and costly system reported in the literature.

These contributions advance the state-of-the-art in FPGA-based TCSPC systems and high-throughput droplet screening and provide a low-cost, flexible, high-performance solution for various applications that require high-precision time measurement and droplet analysis.

1.4 Outline

This thesis presents the design and implementation of a versatile TCSPC system on a low-cost FPGA. The remainder of this thesis is organized as follows:

Chapter 2 provides a comprehensive review of the literature on the TCSPC technique and the TDCs, which are the key components of TCSPC systems. It explains the operating principle, characteristics, and the various architectures of TDCs.

Chapter 3 presents the design and implementation of a synchronous and an asynchronous FPGA-based TDC, and proposes solutions to address the challenges encountered due to the limitations of low-cost FPGAs. It also proposes methods to enhance the TDC linearity, which is one of the most critical characteristics.

Chapter 4 discusses a hardware/software co-design for the transmission of measurement data from the TDC FIFO buffer in the FPGA to the SDRAM of the embedded processor. It also discusses the different data processing algorithms implemented in the C program, including the creation of histograms, the elimination of the background signal and the other noise sources, as well as the different data post-processing algorithms.

Chapter 5 discusses in detail the calibration of the two types of TDC. In addition, it proposes, verifies, and evaluates a novel and innovative method for the calibration of asynchronous TDCs.

Chapter 6 reports the experimental results obtained for the characterization of the designed TDCs in terms of linearity, precision, accuracy, and stability. Moreover, the reliability and accuracy of the TCSPC system are assessed by performing a basic fluorescence lifetime measurement of a standard fluorophore.

Chapter 7 evaluates the versatility of the designed TCSPC system by deploying it in four applications from different fields. This chapter analyzes the specific requirements of each application and proposes optimizations and improvements that enable the system to meet these requirements.

Chapter 8 presents a conclusion and some perspectives for future research.

Chapter 2: Literature Review

The principle of the TCSPC technique is based on the detection of single photons of a periodic optical signal, the measurement of the detection time of these photons within the signal period, and finally the reconstruction of the light waveform from the statistical distribution of the individual time measurements, after repeating the measurements for a sufficient number of times. As demonstrated in the introduction, a TCSPC system includes three essential components: a pulsed light source, a single photon detector, and a time-to-digital converter (TDC). This chapter first discusses the figures of merit of the TCSPC technique, then provides a literature review of the main components of TCSPC systems, with a particular focus on the TDC, as one of the major contributions of this thesis is the design of an FPGA-based TDC.

2.1 TCSPC Technique

The TCSPC technique relies on the fact that, for a low-level high-repetition-rate signal, the photon flux is so weak that the probability of detecting a single photon during one signal period is much less than one [38]. The detector output, in this case, is a sequence of pulses randomly distributed along the experiment time corresponding to the detection instances of the individual photons. With these considerations, there would be many signal periods with no photon pulse, while others would contain one photon pulse, and the periods with more than one pulse would be extremely rare.

To appreciate the merits of TCSPC, it is compared with the direct alternative technique for the recording of optical signals, which is the analog technique. Analog recording interprets the detector signal as an analog waveform. The detector signal is first sampled at short time intervals, then digitized, and accumulated throughout many signal periods. The signal-to-noise ratio (SNR) is determined by the square root of the number of photons N within the impulse response time of the detector. At low intensity, this ratio drops far below one and the baseline instability and the electronic noise limit the number of accumulations and hence the sensitivity of the analog recording system. Obviously, analog recording is more suitable for low-repetition-rate and high-intensity signals. For instance, in fluorescence decay measurement, this can be achieved by using low-frequency high-peak-power pulsed lasers for the excitation of the studied

samples. However, this might lead to saturation effects in the sample, nonlinearity errors, and long-term degradation of the detector performance [39]. On the other hand, in the TCSPC technique, the signal intensity is determined by the pulse density of the detector signal rather than its amplitude. Each detected photon generates a pulse at the detector output and thus a photon count that contributes to the final measurement histogram. This leads to an almost ideal counting efficiency and thereby to an optimum SNR value for a given number of detected photons.

In analog recording, the width of the instrument response function (IRF) cannot be shorter than the single electron response (SER) of the detector. Thus, the bandwidth of an analog recording system is limited by that of the used detector. In contrast, in the TCSPC technique, the arrival time of the photon pulse rising edge is measured with high precision to construct the histogram. Therefore, the resolution is not limited by the width of the detector SER but rather by the accuracy of the arrival time measurement, and the bandwidth is only limited by the transit time spread (TTS) of the pulses in the detector, which is usually an order of magnitude shorter than the SER width. Consequently, for a given detector, the TCSPC technique yields a higher bandwidth and a shorter IRF than the analog recording method.

Another advantage of the TCSPC technique over analog recording is its immunity to the gain noise of the photon detector. In practice, a TCSPC system counts all the detected photons with the same weight regardless of their pulse amplitude, so the SNR is not influenced by the gain noise of the detector. In contrast, the analog recording technique is highly sensitive to this noise, as the magnitude of the detector signal is directly affected by the gain fluctuations due to the variations in the supply voltage.

2.2 Pulsed Light Sources

In early TCSPC systems, nanosecond flashlamps were used as excitation light sources [40, 41]. However, these sources had several limitations, such as low repetition rate, low excitation power, and large pulse width, which made them unsuitable for the measurement of short lifetimes and increased the measurement acquisition time. Nonetheless, flashlamps had the advantage of providing almost any excitation wavelength when coupled with a monochromator.

Modern fluorescence measurement devices use pulsed lasers as the excitation light source, such as mode-locked argon and frequency-multiplied Nd:YAG lasers. However, these sources offer a limited number of discrete wavelengths. This drawback was partly overcome by synchronously pumped jet-stream dye lasers [42]. Pulsed lasers provide short pulses with a

pulse width of less than 100 ps and a high repetition rate of up to 120 MHz. Nevertheless, these laser systems are costly, complex, and require constant maintenance.

The next advancement in pulsed light sources was the development of titanium-sapphire (Ti:Sa) lasers. These lasers offer tunable wavelengths, high repetition rates, ultra-short pulse durations in the picosecond and femtosecond range, and virtually unlimited power peaks [43]. In addition, these lasers enable the generation of UV wavelengths by frequency doubling and tripling [44]. In fact, Ti:Sa lasers are almost ideal excitation light sources for fluorescence measurement. However, the high cost of these lasers is the only drawback.

Recently, picosecond laser diodes have become the most widely used excitation light sources in TCSPC systems [45-48]. They have many advantages over other sources, such as low cost and high reliability, and require no maintenance or alignment [49]. Laser diodes are available with a wide range of discrete wavelengths that can match the absorption spectra of various fluorophores. Moreover, tunable wavelengths can be generated by swapping lasers of different wavelengths, or by combing the beams of several lasers optically and selecting the desired laser electronically.

2.2.1 Adopted Pulsed Light Sources

We used different pulsed light sources in our TCSPC system depending on the application. For the stray light characterization, a picosecond pulsed laser with a wavelength of 532 nm and a repetition rate of 21.42 MHz was used. For the other applications, pulsed laser diodes with different wavelengths were coupled with a high-repetition-rate pulse generator [50]. A 405-nm laser diode served as the light source for the demonstration of the water pollution sensor and the microfluidic droplet sorting. However, this wavelength is not optimal for the excitation of the target pollutants and fluorophores. For the time-resolved optical turbidimeter, a wavelength of 530 nm was used.

2.3 Photon Detectors

Photon detectors can be categorized into two main types based on the photoelectric effect mechanism:

1) Detectors that rely on based on the external photoelectric effect, where a photocathode emits photoelectrons upon the absorption of a photon, such as photomultiplier tubes (PMTs), microchannel plate photomultipliers (MCP-PMTs), and hybrid photomultiplier tubes (HPMTs).

2) Detectors that rely on the internal photoelectric effect, where photoelectrons are generated internally within the detector material when a photon is absorbed, such as single-photon avalanche diodes (SPADs) and superconducting nanowire single-photon detectors (SNSPDs) [51].

The choice of the photon detector depends on the specifications and requirements of the target application, such as the photon rate, the wavelength, range, and the temporal resolution. The photon detectors are usually characterized by the following key parameters:

1- Quantum efficiency (QE):

Quantum efficiency, also referred to as detection efficiency, describes the sensitivity of the detectors. It is the probability that an incident photon produces a measurable electrical pulse at the detector output. QE can also be defined as the ratio between the number of generated photoelectrons and the number of photons reaching the active area of the detector [52]. QE depends on the wavelength of the detected light [53], which is an important factor to consider for selecting an appropriate photon detector for a specific application.

2- Dark count rate (DCR):

The dark count rate of a photon detector is the rate of false pulses generated at the detector output in the absence of incident photons. It is the count rate under complete darkness. Dark counts are random events that are independent of the detected photons. They have different origins according to the type of detector, such as thermionic emission, tunneling, or impurities in the detector material [54, 55]. DCR is affected by different factors, such as the bias level and the temperature. Dark counts are a major source of noise in the photon detector that affects its sensibility and overall performance. Therefore, it is important to choose a photon detector with a low DCR, especially when the detection of weak signals is required.

3- Transit time spread (TTS) or Jitter:

The transit time is the time interval between absorbing a photon and the generation of an electrical pulse at the detector output. This interval is unstable and varies from one photon to another, resulting in the jitter of the detector. The jitter is a crucial parameter, especially for time-resolved applications, that affects the timing measurement precision and the temporal resolution of the TCSPC system [14].

4- Afterpulsing probability:

Photon detectors may generate more than one electrical pulse for a single detected photon. Afterpulsing probability is the likelihood that the detector produces a second electrical pulse within a few microseconds following the photon pulse. Hence, afterpulsing counts are false counts that are temporally correlated with true counts. Afterpulsing is an important parameter that may cause serious problems at high repetition rates because afterpulsing counts generated during many signal periods cause a considerable signal-dependent background. Usually, afterpulsing probability is expressed as a percentage, indicating the probability that the detector generates an afterpulsing event following a real event after a given time interval.

Photomultiplier tubes (PMTs) and single-photon avalanche diodes (SPADs) are the most commonly used photon detectors in TCSPC systems. In the following, we present a brief review of these detectors.

2.3.1 Photomultiplier Tube (PMT)

Photomultiplier tubes (PMTs) are conventional detectors for single-photon counting systems. PMTs were first demonstrated in the 1930s, after the development of photocathodes, secondary emissive surfaces, and electron multipliers [56]. A PMT typically consists of three main components enclosed in an evacuated glass tube: a photocathode, an electron multiplier (consisting of dynode stages), and an anode. When photons hit the photocathode, the latter emits photoelectrons into the vacuum by the external photoelectric effect. The photoelectrons are then amplified along the dynode stages by the secondary electron emission process. The photoelectron emitted by the photocathode is accelerated towards the first positively charged dynode, where it collides and releases more electrons. These electrons are then accelerated towards the next dynode, where they collide and cause the emission of more electrons, and so forth. Thus, this process significantly amplifies the signal, as the number of electrons increases with each dynode stage. The secondary electrons emitted from the last dynode stage are then collected by the anode which delivers the electron current to an external circuit. The anode current at the output of the PMT is proportional to the amount of light captured by the photocathode. At very low light intensities, the PMT output would be in the form of discrete, well-separated pulses that could be amplified and then processed by photon counting electronics.

There is a variety of dynode geometries with different gain, response time, and transit-time spread. The "linear focused dynodes" structure, depicted in Figure 2.1, is the most suitable structure for single-photon counting because it provides fast response and low jitter. The amplification process is very effective and typically yields gain values between 10^6 and 10^7 [57]. However, the dynode stages require high operating voltages of a few kV. Because of this, as well as the complex design of the multiple dynode stages, PMTs are relatively large and bulky detectors. Nevertheless, recent advancements have enabled the miniaturization of PMT designs and the integration of high-voltage power supplies into small compact units [58].

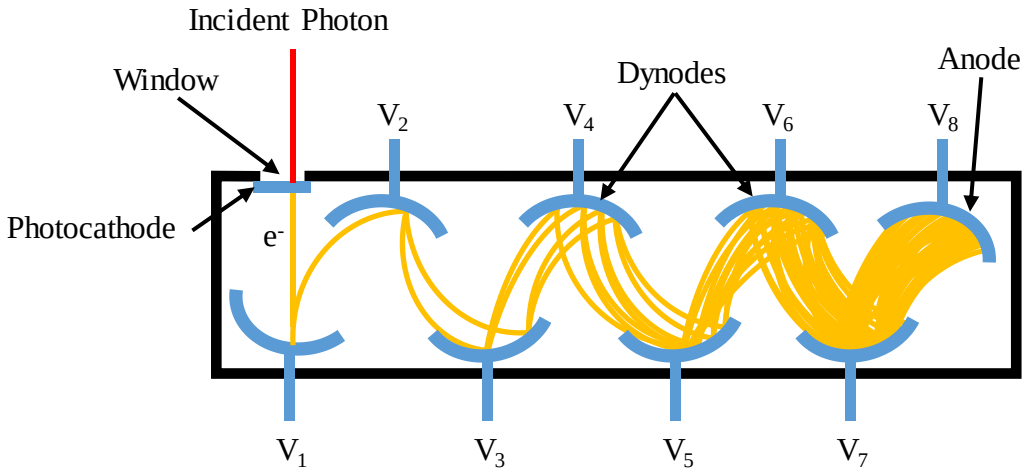


Figure 2.1 Operating principle of linear focused PMT.

In addition to their large size, PMTs typically have several drawbacks, such as the saturation at moderate light intensities, their substantial timing jitter, and their modest quantum efficiency.

- **Saturation:** PMTs are only suitable for very low-level optical signals. At high light intensities, the output pulses of individually amplified photoelectrons overlap and cannot be resolved as discrete pulses.
- **Substantial timing jitter:** the PMT output pulses have amplitude fluctuations due to variations in the amplification process of the dynodes. These fluctuations result in a timing jitter of the order of the pulse rise time, which necessitates the use of a constant fraction discriminator, as illustrated in Figure 2.2.
- **Modest quantum efficiency:** PMTs rely on the external photoelectric effect. The disadvantage of this effect is that the photoelectrons are emitted in all directions, including backward. Therefore, the quantum efficiency is always lower than 0.5. In practice, the maximum achievable quantum efficiency is about 0.4 in the wavelength range from 400 to 500 nm [59].

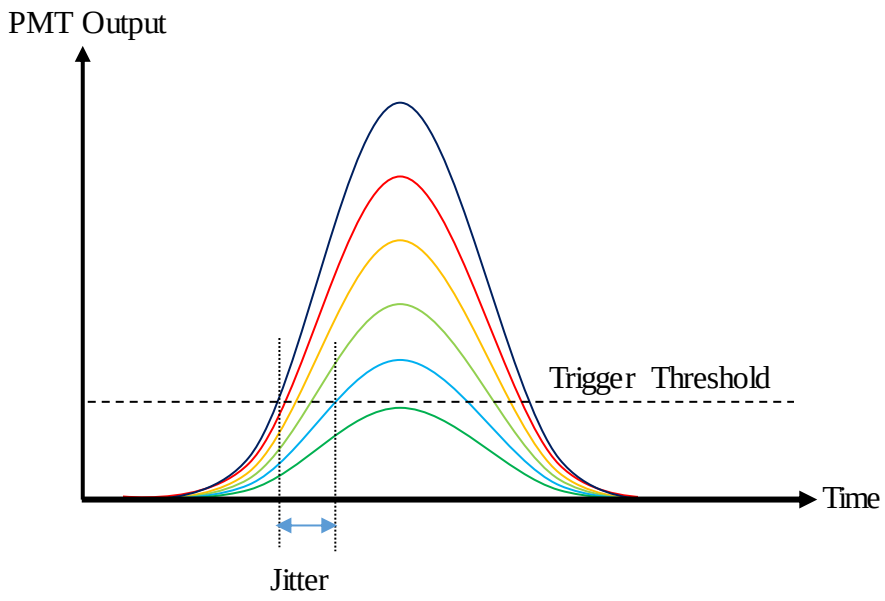


Figure 2.2 PMT timing jitter using a simple level trigger due to the amplitude fluctuation.

2.3.2 Single-Photon Avalanche Diode (SPAD)

In contrast to photomultiplier tubes, semiconductor photon detectors exploit the internal photoelectric effect, in which incident photons generate electron-hole pairs inside the semiconductor material. An example of this type of photon detector is the conventional avalanche photodiode (APD), which is a PN junction operating in reverse bias below its breakdown voltage. In this regime, the electric field is high enough that the photo-generated carriers initiate an avalanche process, where they break off new carriers from the lattice of the semiconductor material. This makes the APD function as both a photon detector and a linear amplifier with a finite gain of up to 10^3. However, this gain is insufficient to detect single photons with high temporal resolution. Furthermore, linear APDs suffer from several drawbacks: they have low gain limited to a few hundred; they exhibit high noise levels; they are highly sensitive to the variations in temperature, voltage, and doping levels, which affect the gain; and they require long integration times, which make them unsuitable for the detecting ultra-fast low-level signals. These limitations can be overcome by increasing the bias voltage above the breakdown point by several volts, which leads to a virtually infinite gain, as shown in Figure 2.3-a which illustrates the PN junction operating modes in reverse bias. A single photon avalanche diode (SPAD) is a semiconductor detector that operates in this regime, called “Geiger Mode”. In this mode, the electric field in the multiplication region is so intense that a single photon can trigger a self-sustained avalanche. The avalanche produces a large current pulse whose leading edge indicates the arrival time of the detected photon. However, the

avalanche current persists as long as the applied voltage is above the breakdown level. In this state, the SPAD will be insensitive to further arriving photons as the absorption of additional photons will not alter the signal output. Furthermore, the self-sustained avalanche will damage the SPAD by overheating. Therefore, the self-sustained avalanche must be stopped after each detection process using a Quenching circuit [60, 61]. This circuit detects the avalanche onset, then reduces the bias voltage below the breakdown level to stop the avalanche, and finally restores the voltage to its initial value so that a new photon can be detected, as illustrated in Figure 2.3-b.

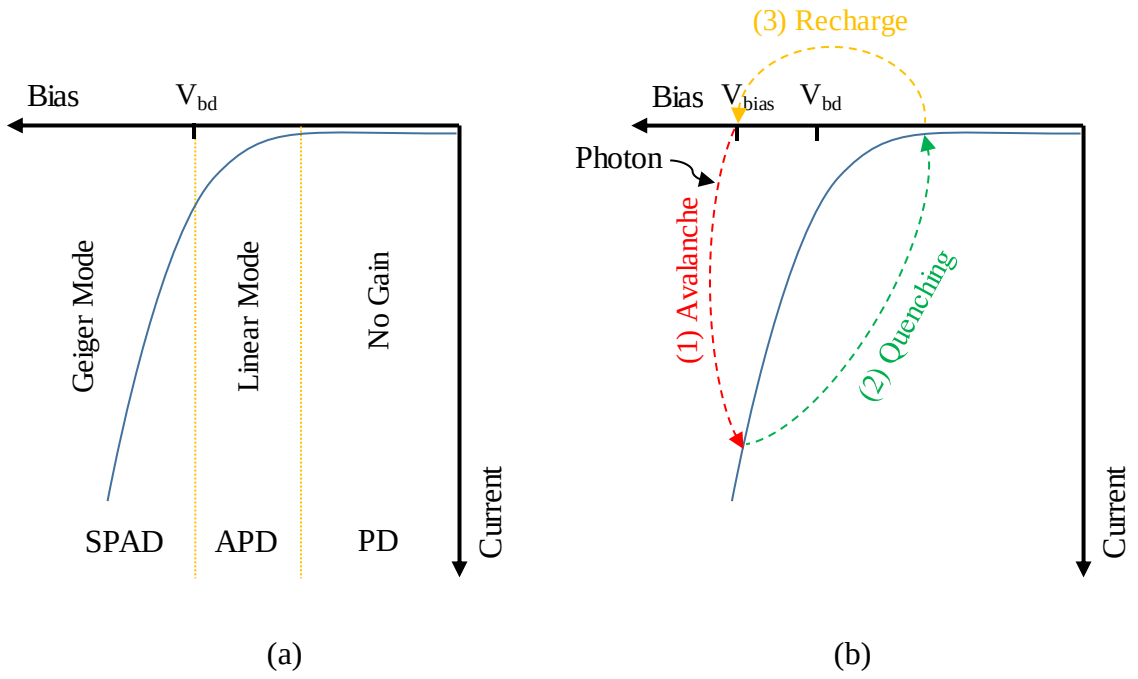


Figure 2.3 (a) P-N junction operating modes in reverse bias, (b) SPAD biasing cycle.

SPADs can be fabricated in both standard and custom CMOS technologies. However, compared to custom technology devices, standard devices have lower QE [57], significantly higher DCR per unit area, and a higher afterpulsing probability. They are usually designed with smaller active areas and low excess bias voltages to limit the detector noise.

In recent decades, significant progress has been made in the development of SPADs as good alternatives to PMTs. Novel designs and advances in CMOS technology have enabled the fabrication of SPADs with performance comparable to that of PMTs. However, unlike PMTs, which are bulky, fragile, costly, and require high voltages, SPADs are more economical, operate at moderate reverse biases, and can be integrated with their front-end circuitry on the same substrate, resulting in smaller and more compact sensors.

2.3.3 Adopted Photon Detector

For the target applications, we used a commercial SPAD as the photon detector. This detector is the "ID100 Visible Single-Photon Detector" from ID Quantique, which has the following specifications [62]:

- High sensitivity in the spectral range from 350 to 900 nm, with a peak photon detection efficiency at 500 nm.
- Fast active quenching circuit.
- High timing resolution of 40 ps.
- Low dead time of about 45 ns.
- Dark count rate as low as 5 Hz.
- Small IRF shift at high count rates.
- Active area diameter of 50 μm.
- Afterpulsing probability of about 0.5 %.
- Maximum count rate (pulsed light) of 20 MHz.

2.4 Time-to-Digital Converters (TDCs)

Time-to-digital converters (TDCs) are devices that measure precise time intervals between two signals, called *START* and *STOP* signals, with high resolution. They are used in a wide range of applications, such as time-of-flight measurement, nuclear physics, and RF data transmission.

TDCs can be implemented based on either analog or digital architectures. Analog TDCs employ a time-to-analog converter (TAC) to convert the time interval between the *START* and *STOP* pulses into a voltage signal, which is then amplified and digitized using an analog-to-digital converter (ADC). This approach yields high resolution and conversion rates [63]. On the other hand, digital TDCs are mainly implemented as application-specific integrated circuits (ASICs) or using field programmable gate arrays (FPGAs). This type of TDC has become more prevalent due to the numerous advantages of digital circuits over analog ones, especially in terms of size, efficiency, and noise immunity.

In recent decades, many studies have been conducted to enhance the performance of TDCs in terms of resolution, linearity, measurement range, and dead time. The following sections of this chapter discuss the main parameters used to evaluate the performance of a TDC and present the most common architectures of TDCs.

2.4.1 TDC Operation Parameters

The operating principle of a TDC is to map continuous time intervals into discrete time bins. Figure 2.4 compares the quantization function of an ideal TDC and a real one. The quantization process limits the TDC's accuracy and introduces conversion errors that are characterized by several key parameters such as the temporal resolution, the precision, and the differential and integral nonlinearity (DNL and INL). These parameters, together with others such as the measurement range and the dead time, evaluate the TDC performance. These parameters are defined and discussed below.

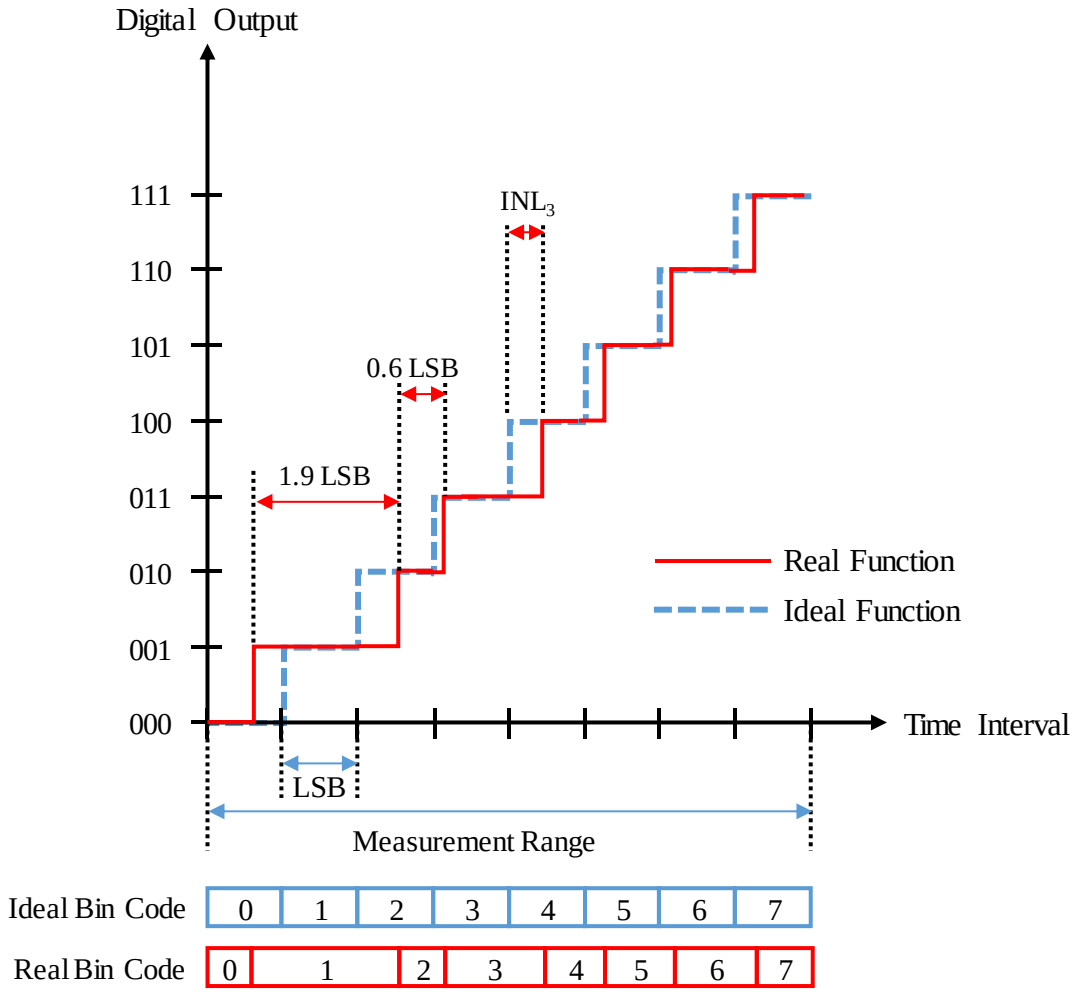


Figure 2.4 Quantization function of an ideal TDC and a real one.

- **Measurement Range (MR)**

The measurement range (MR) is the maximum time interval that can be measured by the TDC. It is determined by the sum of the widths of all the bins in the TDC. The measurement range of an N-bit TDC can be calculated by Equation 2.1.

$$MR = \sum_{b=0}^{2^N-1} T_b \tag{2.1}$$

where T_b is the width of the b-th bin.

- **Temporal Resolution or LSB**

The resolution, also termed the least significant bit (LSB) or the bin width, is the smallest change in the time interval that can be detected by the TDC. Ideally, all the bins have the same width. However, in real TDCs, the bin width varies from one bin to another. Therefore, the resolution of a TDC is defined by the average width of its bins. For an N-bit TDC, the resolution is calculated by Equation 2.2.

$$LSB = \frac{MR}{2^N} \tag{2.2}$$

where MR represents the measurement range of the TDC.

- **Differential and Integral Nonlinearity (DNL and INL)**

The nonlinearity of a TDC is described by the deviation of its transfer function from that of an ideal TDC with the same number of bits. The TDC nonlinearity is characterized by two parameters: differential and integral nonlinearity (DNL and INL). DNL is the difference between the actual width of a bin and the ideal bin width, normalized by the ideal bin width. It can be calculated by Equation 2.3.

$$DNL_b = \frac{T_b - LSB}{LSB} \tag{2.3}$$

where DNL_b is the DNL value at the b-th bin, T_b is the width of this bin and LSB is the average bin width.

INL is the deviation of the output code of the real TDC from the output of an ideal TDC expressed in LSB. It can be computed by accumulating the DNL values across the TDC bins, as shown in Equation 2.4.

$$INL_b = \sum_{i=0}^{b} DNL_i \tag{2.4}$$

The nonlinearity of the TDC is usually described by the maximum and minimum values or the peak-to-peak values of the DNL and INL.

The code density test is the most common method to evaluate the linearity of a TDC [64-66]. This method involves measuring a large number of time intervals that are uniformly distributed over the entire measurement range of the TDC. The number of counts in each bin of the resulting histogram is proportional to the width of the bin. Ideally, the number of counts should be the same for all the bins. The DNL and INL of the TDC are estimated by comparing the obtained number of counts in the TDC bins and the expected values for an ideal TDC.

- **RMS and Single-Shot Precision**

TDCs are subject to various noise sources, such as the TDC quantization noise and the clock jitter, that introduce variations in the conversion results. Thus, repeated measurements of the same fixed time interval may produce different values. The TDC precision is an important characteristic that evaluates the TDC stability and the reproducibility of the measurement result in the presence of noise. It is usually determined by the standard deviation or the root mean square (RMS) error of the measurement results of a fixed time interval.

In TCSPC systems, in addition to the TDC noise sources mentioned above, other noise sources, such as the jitter of the light source and the photon detector, contribute to the measurement uncertainty and should be taken into account in the evaluation of the overall system precision. Therefore, the precision of a TCSPC system is typically determined by the standard deviation or the full width at half maximum (FWHM) of its instrument response function (IRF). The RMS precision of the TCSPC system can be estimated by Equation 2.5 [65, 67, 68].

$$\sigma_{rms} = \sqrt{\sigma^2{}_q + \sigma^2{}_{clk} + \sigma^2{}_{INL} + \sigma^2{}_{laser} + \sigma^2{}_{detector} + \sigma^2{}_{others}} \qquad (2.5)$$

where σ_q is the TDC quantization error, σ_{clk} is the clock jitter, σ_{INL} is the standard deviation of the TDC INL, σ_{laser} is the jitter of the light source, $\sigma_{detector}$ is the jitter of the photon detector and σ_{others} is the jitter of other noise sources, such as the I/O pads and bonding wiring.

- **Dead Time and Measurement Rate**

The Dead time is the time required for the TDC to complete a measurement and to be ready to perform a new measurement [69]. The measurement rate, which is the reciprocal of the dead time, indicates the maximum conversion rate attainable by the TDC. In TCSPC systems, the maximum photon detection rate is limited by the dead time of the TDC and the detector, as well as the TDC data readout rate. A high measurement rate is crucial for high-speed TCSPC applications such as LIDAR and fluorescence lifetime measurement.

2.4.2 TDC Architectures

Different architectures of TDCs have been proposed and implemented in the literature. Analog-based TDCs have been the earliest architectures used in conventional TCSPC systems. However, with the rapid advancements in CMOS technologies, fully digital TDCs have been developed and become a more attractive option for TCSPC systems. This section presents and compares the different architectures of TDCs.

- **Analog-Based TDCs**

The operating principle of traditional analog TDCs is to use analog circuits that convert the time interval into a voltage, and then digitize this voltage using a standard analog-to-digital converter (ADC). Figure 2.5 shows a basic block and signal diagram of this architecture. First, a pulse generator generates a pulse with a width proportional to the time interval between the *START* and *STOP* signals. This pulse is then converted into a voltage by an analog integrator. Finally, an ADC converts this voltage into a digital value [70-72].

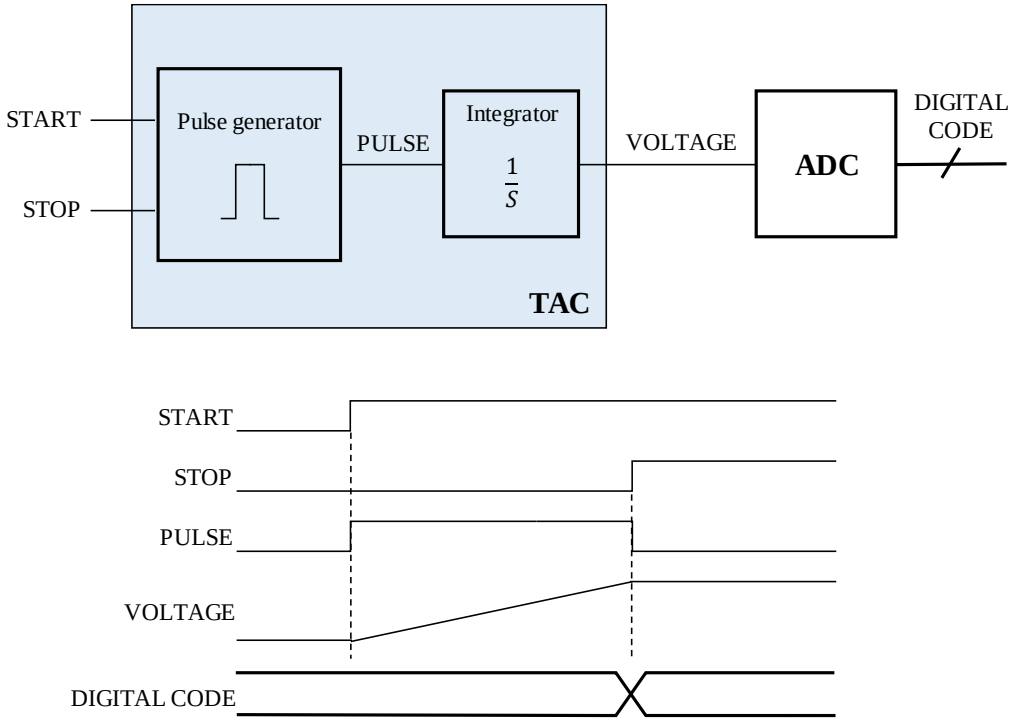


Figure 2.5 Block and signal diagram of a basic analog-based time-to-digital converter.

Analog TDCs can achieve high resolution at the cost of a limited dynamic range. A common drawback of TACs based on current integration in a capacitor is the poor linearity due to the finite output resistance of the current source, which complicates the calibration of the TDC. Moreover, the overall performance of the TDC depends on the performance of the ADC in

terms of resolution, dead time, and power consumption [73]. In addition, analog TDCs are sensitive to noise and to process and temperature variations.

- **Fully Digital TDCs**

With the rapid progress of integrated circuit technologies in the 1990s, digital solutions have become more popular and gradually replaced analog circuits, owing to the different advantages of these solutions over analog ones. This trend also affected TDCs, which evolved to fully digital architectures that provided features such as flexibility, integrability, and robustness against temperature variations and electromagnetic noise. In addition, the advancements in CMOS technology have enabled the enhancement of the temporal resolution and the overall performance of digital TDCs [74]. Several architectures of fully digital TDCs have been proposed in the literature. This section presents the main architectures and discusses their principle, advantages, and limitations.

1- Direct Counting Architecture

A straightforward method to measure a time interval is to use a counter operating at a high frequency, that counts the clock cycles between the *START* and *STOP* signals, as illustrated in Figure 2.6. The counter starts incrementing after the arrival of the *START* signal and stops after the arrival of the *STOP* signal, and the time interval (*TI*) is calculated from Equation 2.6.

$$TI = n \times T_{clk} \tag{2.6}$$

where n is the counter value and T_{clk} is the clock period. This method can achieve large measurement ranges by increasing the counter's bit width. For a counter with N bits, the measurement range is given by Equation 2.7.

$$MR = 2^N \times T_{clk} \tag{2.7}$$

However, this method suffers from measurement errors due to the asynchronous nature of the measured signals. The maximum error is: ($2 \times T_{clk}$), when both the *START* and *STOP* signals are asynchronous with the counter clock; and (T_{clk}), when only one signal is asynchronous with the clock. Nevertheless, the measurement error can be reduced by repeating the measurement multiple times and averaging the results. The resolution of the direct counting approach is limited by the clock frequency and requires high frequencies to achieve high resolutions. However, it is difficult to clock a counter at a frequency above some hundreds of megahertz in a commercial FPGA. Likewise, there will be a maximum frequency limit in ASICs depending

on the technology used [75]. Hence, this approach is suitable for applications that require wide measurement ranges without demanding high temporal resolutions.

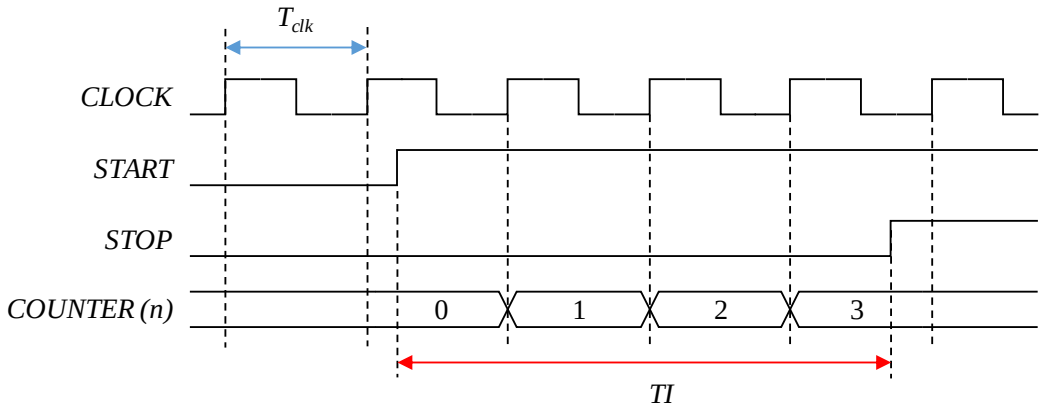


Figure 2.6 Operating principle of the direct counting TDC.

2- Tapped Delay Line (TDL) Architecture

Tapped delay line (TDL) is one of the most widely used architectures for implementing precise TDCs in both ASIC and FPGA devices [76]. This architecture has several advantages, such as simplicity, short dead time, and reliability. Figure 2.7 illustrates the basic architecture of a TDL-based TDC. The *START* signal propagates through cascaded delay elements (DEs), typically buffers, causing the output of the propagated elements to switch from one logic level to another. The outputs of the DEs are sampled by flip-flops at the rising edge of the *STOP* signal, producing a thermometer code, which is a binary vector that consists of either a series of ones followed by a series of zeros "1111…0000" or zeros followed by ones "0000…1111". The number of propagated elements corresponds to the time interval between the *START* and *STOP* signals. The thermometer code is then encoded into a binary code, and the measured time interval (*TI*) is estimated from Equation 2.8.

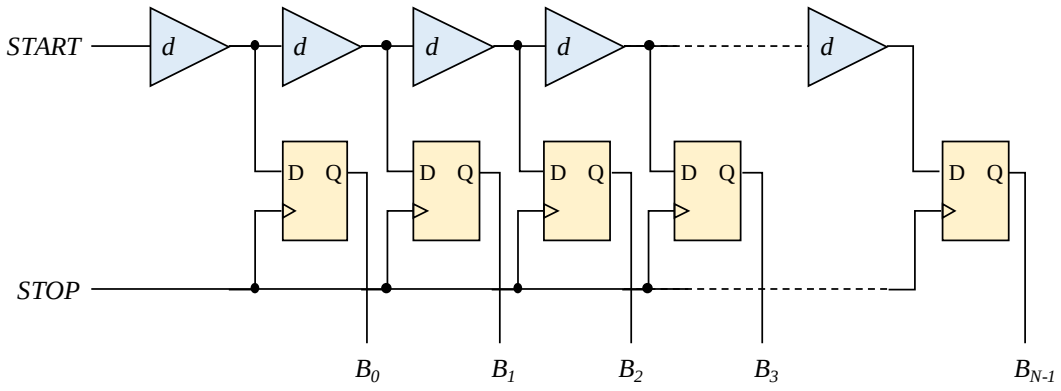


Figure 2.7 Basic architecture of a tapped delay line-based TDC.

$$TI = n \times d \tag{2.8}$$

where *n* is the binary code representing the number of propagated delay elements and *d* is the average delay of the DEs.

This structure generally provides a high resolution determined by the DE delay, but it has a short measurement range limited by the number of DEs and the average delay as shown in Equation 2.9.

$$MR = N \times d \tag{2.9}$$

where *N* is the total number of DEs. A reasonable measurement range requires a large number of DEs which increases the resource and power consumption. Furthermore, this structure has other drawbacks such as the sensitivity to process, voltage and temperature variations that affect the propagation speed through the delay elements. These variations introduce measurement errors that may also arise from the skew of the *STOP* signal that drives the registration flip-flops.

3- Vernier Method

This method is inspired by the principle of the Vernier caliper, a device for measuring lengths invented by Pierre Vernier in 1631. The Vernier caliper consists of two graduated scales: a main scale similar to a normal ruler and an auxiliary scale with ten divisions that span nine divisions of the main scale. When performing a measurement, the auxiliary scale slides in parallel to the main scale. By identifying which graduation of the auxiliary scale aligns with a graduation of the main one, a fine precision beyond that of the main scale can be attained.

Figure 2.8 illustrates the basic architecture of a Vernier TDC, which consists of two controllable oscillators, two counters, and a coincidence circuit [77, 78]. The two oscillators can be launched by trigger signals to generate two clock signals with slightly different periods T_0 and T_1. These clock signals drive the two counters. The working principle of the TDC is as follows: the *START* and *STOP* signals trigger the oscillators. The oscillators generate clock signals with an initial phase difference that equals the time interval between the *START* and *STOP* signals. Subsequently, the phase difference gradually decreases until the two clock signals coincide, as depicted in Figure 2.9. At this moment, the coincidence circuit disables the counters and the time interval (*TI*) between the rising edges of the *START* and *STOP* signals can be calculated by Equation 2.10.

$$TI = n_0 \times T_0 - n_1 \times T_1 \tag{2.10}$$

where n_0 and n_1 are the values of the two counters after the coincidence.

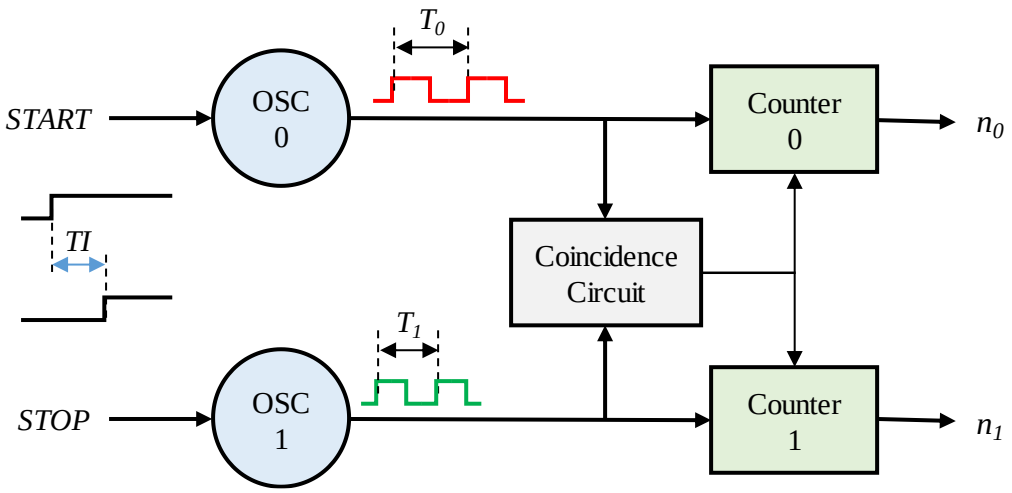


Figure 2.8 Basic architecture of Vernier TDC.

This structure has some advantages such as the immunity to process, voltage and temperature variations [75]. The temporal resolution of the Vernier TDC is determined by the difference between the oscillators' periods, while the linearity depends on the stability and accuracy of the oscillators. The main drawback of this method is the long dead time that corresponds to the duration between the start of the measurement, i.e. the activation of the first oscillator, and the coincidence moment. The dead time is conversely proportional to the resolution, so a trade-off between them should be considered according to the application. With this structure, a temporal resolution of less than 10 ps has been achieved [79-81].

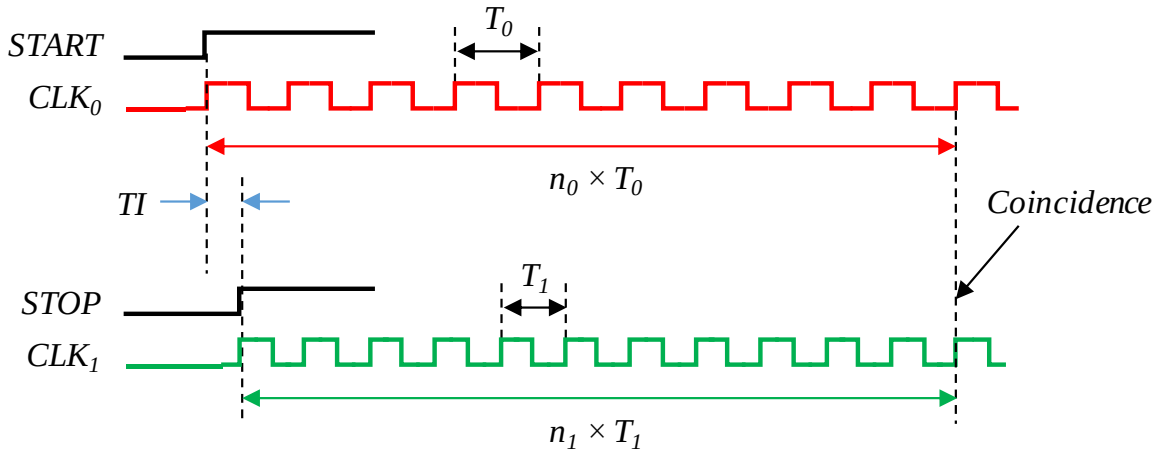


Figure 2.9 Timing diagram of Vernier TDC.

4- Vernier Delay Line (VDL)

The resolution in classical TDLs is limited by the delay of the DEs that constitute the delay line. The Vernier delay line (VDL) architecture was proposed to improve the resolution beyond the DE delay [82-84]. The VDL, also known as the differential delay line, is a hybrid architecture that combines the principles of the TDL and the Vernier structures. It consists of a TDL with delay elements in the path of the *STOP* signal, i.e. the sampling signal, as depicted in Figure 2.10. Similar to the Vernier structure oscillators. the elementary delay of the start line d_0 is slightly longer than that of the stop line d_1. The operating principle of this architecture is as follows: when the *START* signal arrives, it propagates through the slower delay line. Then, when the *STOP* signal arrives, it propagates through the faster line chasing the *START* signal. As a result, the time interval between the two signals gradually shrinks during the propagation and decreases by $(\Delta T = d_0 - d_1)$ after each DE, as illustrated in the timing diagram in Figure 2.11. The outputs of the start DEs are sampled by the flip-flops sequentially rather than simultaneously. In this case, the flip-flop line acts as an early-late arbiter line that detects the position *n* at which the *START* and *STOP* signals become in phase. The initial time interval (*TI*) between these signals can be calculated by Equation 2.11.

$$TI = n \times \Delta T \tag{2.11}$$

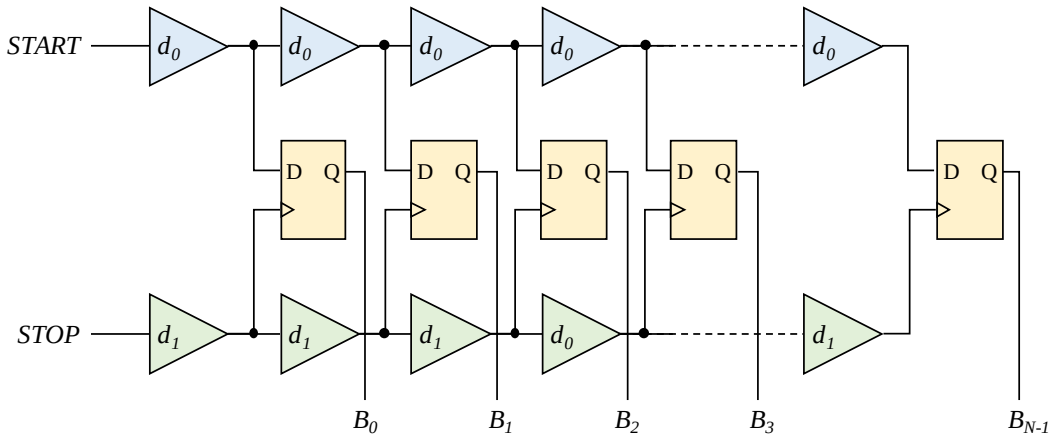


Figure 2.10 Block diagram of VDL TDC.

The resolution of this architecture depends on the delay difference ΔT, which can be minimized to achieve a very fine resolution. However, improving the resolution limits the measurement range (MR), which is proportional to ΔT and the number of DEs (*N*), as given by Equation 2.12.

$$MR = N \times \Delta T \tag{2.12}$$

Therefore, a considerably long delay line is required to obtain a reasonable MR, which leads to high resource and power consumption. Moreover, this also increases the dead time since the output becomes valid only after the STOP signal reaches the end of the VDL. However, a larger MR can be attained without increasing the dead time or the resource and power consumption by replacing the linear structure with a loop one.

Another drawback of this structure is the high sensitivity to delay mismatch and process, voltage and temperature (PVT) variations, which is a common problem for all delay line-based structures. Nevertheless, this problem is even more severe for the VDL due to its high resolution and the fact that these variations affect both of its delay lines.

To enhance the resolution of VDL TDCs, many techniques have been proposed, such as the fractional difference schema [85] and the multidimensional VDL TDC [86, 87].

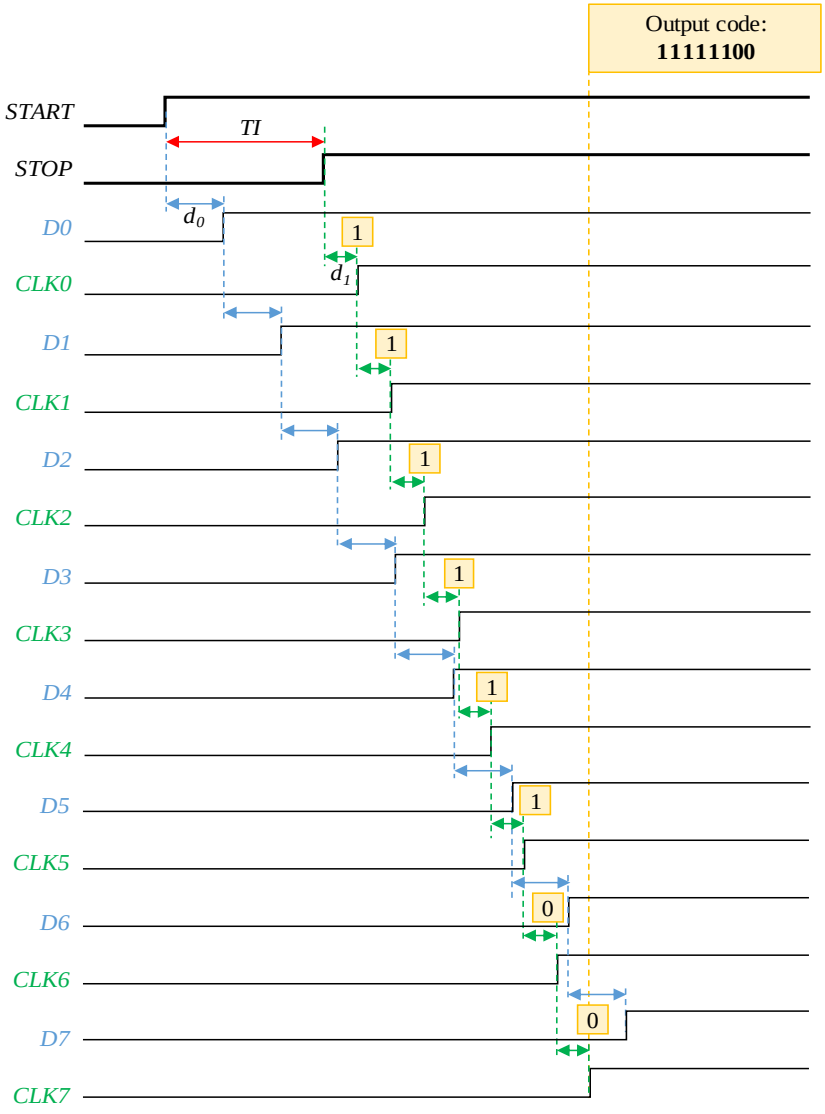


Figure 2.11 Timing diagram of VDL TDC.

5- Pulse Shrinking

This architecture employs a pulse-shrinking technique to measure the time interval (*TI*) between the *START* and STOP signals. First, the START and *STOP* signals are pre-processed to generate a pulse with a width equal to TI. This pulse then propagates through DEs that have asymmetric rise and fall times (T_r and T_f), preferably with a shorter fall time ($T_r > T_f$). As a result, the width of the pulse shrinks as it traverses the DEs until it disappears [88, 89]. The pulse width decreases by ($\Delta T_{rf} = T_r - T_f$) after each traversed DE. The number of propagated DEs (n) before the pulse vanishes is proportional to *TI* which can be calculated by Equation 2.13.

$$TI = n \times \Delta T_{rf} \tag{2.13}$$

The outputs of the DEs are used to clock a series of flip-flops that have their data inputs connected to a constant '1'. For the propagated DEs, the output of the corresponding flip-flops switches to '1', until the pulse vanishes where the output of the remaining flip-flops remains at '0'. The number of propagated DEs is detected by the position of the transition from '1' to '0', as illustrated in Figure 2.12.

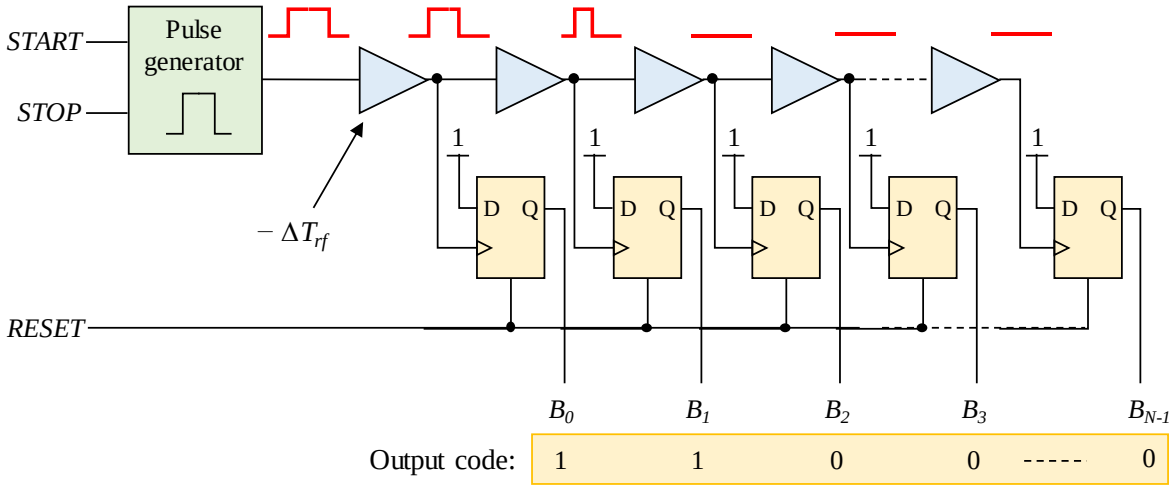


Figure 2.12 Block diagram and operating principle of a pulse shrinking TDC.

This architecture offers a high resolution, defined by ΔT_{rf}. However, it suffers from a large nonlinearity and a high sensitivity to PVT variations. Furthermore, the measurement range is limited by the number of DEs. A large number of DEs allows a large measurement range but worsens the nonlinearity problem. Therefore, a trade-off between the temporal resolution and the measurement range should be considered. The resolution also trades the dead time. Hence, this architecture is not suitable for applications that require high measurement rates.

Another pulse shrinking technique can be implemented based on loop structure [90-92]. The pulse shrinks in the loop, causing the counter to increment by one at each iteration, until it vanishes. Hence, the final value of the counter is proportional to the pulse width. However, the measurement range in the looped architecture is limited by the total delay of the loop because the entire pulse should fit inside the loop.

6- Multiphase Clock Architecture

A multiphase clock TDC consists of a phase-locked loop (PLL) and an array of D flip-flops (DFF). The PLL generates (*N*) clocks with the same frequency, the same duty cycle of 50%, and a fixed phase difference of (*180/N*) degrees among them. Figure 2.13 illustrates the schematic and the operating principle of a multiphase clock TDC with four clocks, where the phase shifts are 0°, 45°, 90°, and 135°, respectively. Hence, the clock period is subdivided into 8 identical parts. The temporal resolution is determined by the number of phase-shifted clocks and is given by Equation 2.14.

$$LSB = \frac{T_{clk}}{2N} \tag{2.14}$$

where T_{clk} is the clock period.

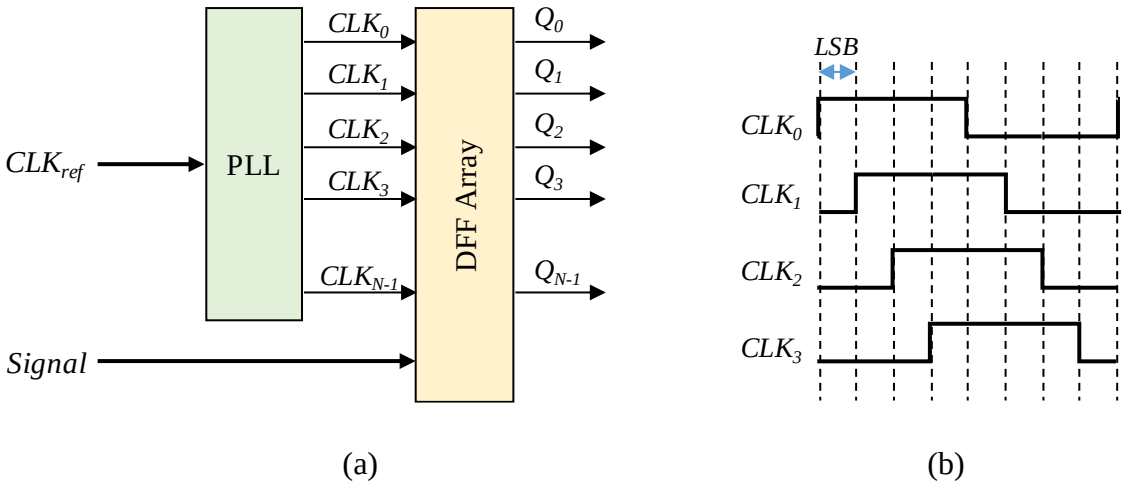
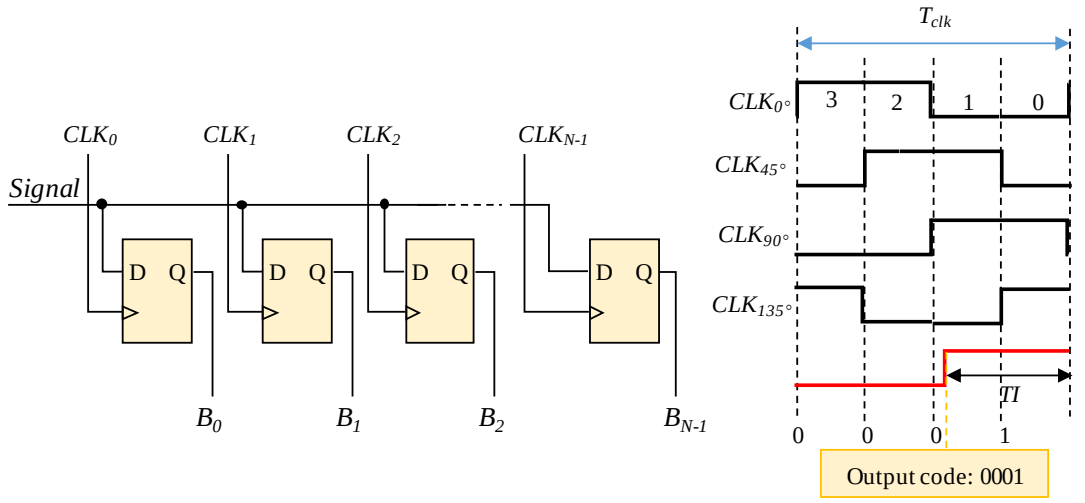


Figure 2.13 (a) General block diagram of multiphase clock architecture, (b) Timing diagram of a four-phase clock TDC.

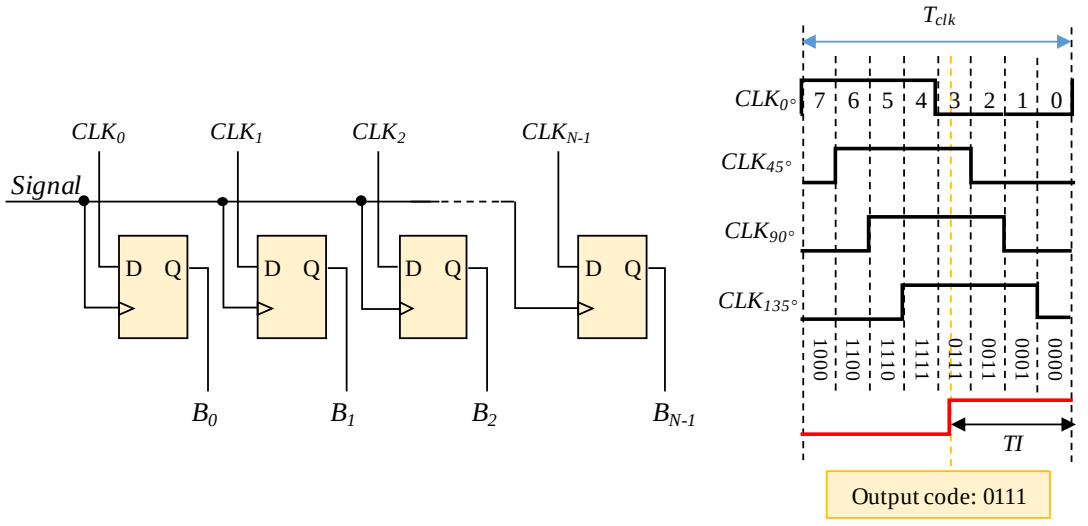
The clock signals and the signal to be measured are fed to the DFF array in two different configurations:

- Clock triggering: the clock signals are connected to the "clock" inputs of the DFFs and the measured signal is connected to their "D" inputs [93-95], as illustrated in Figure 2.14-a.
- Signal triggering: the measured signal is connected to the "clock" inputs of the DFFs and the clock signals are connected to the "D" inputs [96], as shown in Figure 2.14-b.

The code value at the DFF array output is encoded to a digital value that indicates the time interval between the measured signal and the subsequent rising edge of the clock CLK_0.



(a)



(b)

Figure 2.14 Block and timing diagram of multiphase clock TDC in two configuration modes: (a) Clock triggering, (b) Signal triggering.

7- Ring Oscillator (RO)

A ring oscillator (RO) TDC is a self-timed architecture that includes an on-chip feedback loop consisting of buffered stages to generate the clock signal, a counter, and a sampling logic [97, 98]. As depicted in Figure 2.15, the counter is driven by the ring oscillator signal and counts the number of cycles between the *START* and *STOP* signals. The ring oscillator also serves as an interpolator by providing different phases of the loop. The sampling logic records the state of these phases at the arrival moments of the *START* and *STOP* signals, and the time interval (*TI*) can be computed by Equation 2.15.

$$TI = t_1 - t_2 + C \times T_{os} \tag{2.15}$$

where t_1 and t_2 are respectively the time intervals between the *START* and *STOP* signals and the next rising edges of the clock, calculated from the sampled states of the buffered phases; *C* is the counter value; and T_{os} is the oscillation period, which is equal to the product of the phase delay and the number of stages (*N*), as shown in Equation 2.16.

$$T_{os} = d \times N \tag{2.16}$$

where *d* is the phase delay of the ring oscillator.

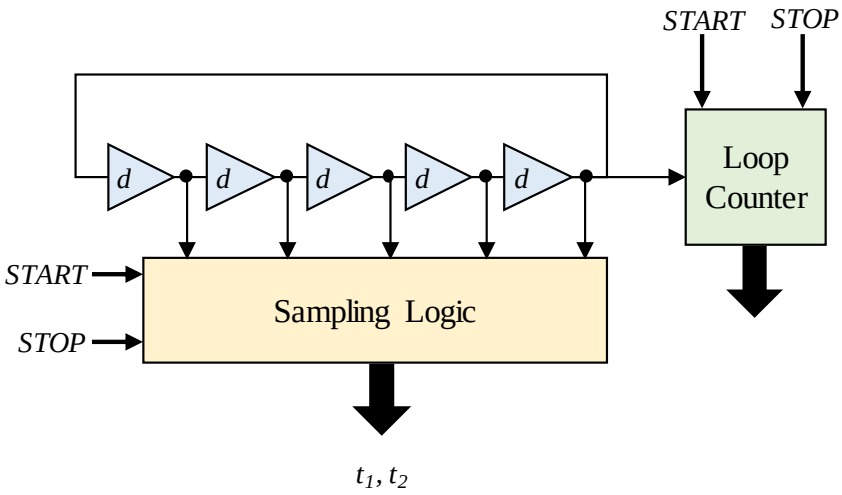


Figure 2.15 Basic architecture of a ring oscillator TDC.

RO-based TDCs should be calibrated using a reference time interval with high accuracy to determine the oscillation period. This type of TDC can achieve a high resolution limited by the phase delay, and a large measurement range by increasing the number of bits of the loop counter. Moreover, the loop architecture significantly reduces the area occupation of the RO-based TDCs, making them suitable for large-scale TDC arrays. However, this structure has

many drawbacks, such as the high power consumption in the free-running mode due to the continuous operating of the ring oscillator even when no measurement is performed, and the sensitivity to PVT variations [98].

Several variants of this structure have been proposed to improve the temporal resolution and power consumption, such as the gated RO [99], switched RO [100], Vernier RO [101], 2-D Vernier RO [102], and multipath RO [103].

8- Nutt Interpolation Method

The performance of the TDC architectures discussed above depends on several interrelated characteristics, such as temporal resolution, linearity, measurement range, dead time, and resource and power consumption. Optimizing one of these characteristics usually entails compromising one or more of the others. For example, implementing longer delay lines to increase the MR also increases the resource and power consumption, and increasing the MR of a Vernier TDC necessitates either reducing the resolution or increasing the resource consumption. An alternative solution is to employ an interpolation technique that can simultaneously achieve high temporal resolutions and large measurement ranges with minimal additional resource consumption. Nutt's interpolation method [104] is a well-established technique that has been widely applied in ADC and TDC designs. In this technique, the time interval is decomposed into three components: two fractional components and one integral component, corresponding to a coarse clock period, as illustrated in Figure 2.16.

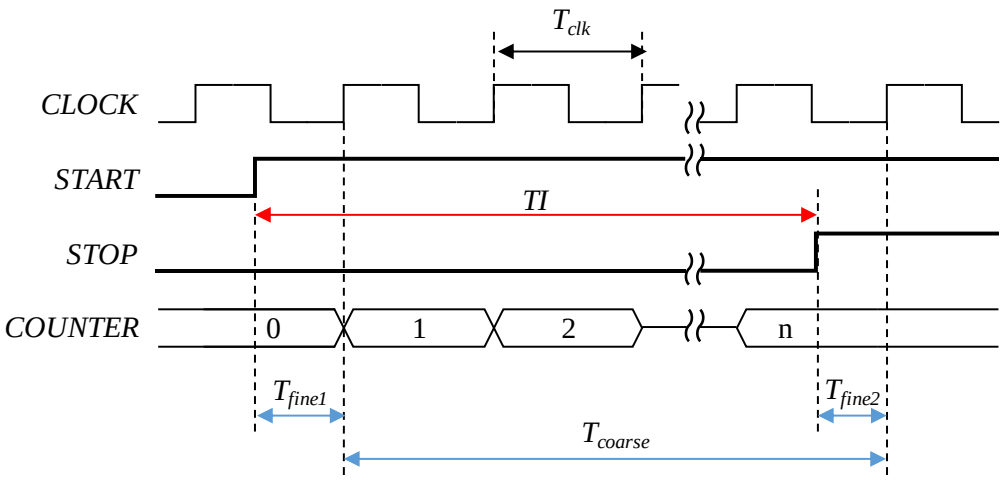


Figure 2.16 General principle of Nutt method.

The first fractional component (T_{fine1}) represents the time interval from the *START* signal to the subsequent rising edge of the coarse clock, while the second fractional component (T_{fine2})

corresponds to the time interval from the *STOP* signal to the next rising edge of the clock. The integral component (T_{coarse}) is the time interval between the endpoints of the two fractional components. T_{coarse} can be easily measured by a counter driven by the coarse clock. On the other hand, the fractional components can be measured by high-resolution TDCs functioning as interpolators. Consequently, TI can be determined by Equation 2.17.

$$\begin{aligned} TI &= T_{fine1} + T_{coarse} - T_{fine2} \\ &= T_{fine1} + n \times T_{clk} - T_{fine2} \end{aligned} \tag{2.17}$$

where n is the coarse counter value and T_{clk} is the coarse clock period.

In synchronous TCSPC systems, the *START* signal is in sync with the coarse clock. Hence, the interpolation structure can be simplified to a coarse-fine scheme, in which the time interval is decomposed into an integral component (T_{coarse}) and only one fractional component (T_{fine}). This scheme comprises a coarse counter for measuring T_{coarse} and a high-resolution interpolator for the measurement of T_{fine}, which reduces the resource and power consumption. This structure is widely adopted in TDL-based and RO-based TDCs.

2.5 TDC Implementation

Fully digital TDCs are predominantly implemented on ASIC and FPGA platforms [105-108]. ASIC TDCs can achieve superior resolution, precision, and linearity due to the full-custom design features of these circuits [109]. However, ASIC devices have several disadvantages, such as high design and production costs, long development time, and lack of flexibility. The initial development cost of ASIC devices is very high, but the unit cost decreases with increasing the production scale, which makes this solution more suitable for general-purpose and large-scale commercial TDC systems, such as LIDAR systems for autonomous drive.

On the other hand, FPGA-based TDCs have several merits, such as low cost, reconfigurability, flexibility, and short design process. Furthermore, with the continuous progress in FPGA technology, the performance gap between FPGA and ASIC TDCs has been reduced in recent years [110, 111]. Hence, implementing TDCs on FPGAs is the best solution for fast prototyping and for applications that require flexibility, reusability, and customization.

2.5.1 ASIC-Based TDCs

ASIC TDCs are mainly implemented based on the delay line and the ring oscillator architectures. The Nutt technique is usually applied in these architectures to extend the measurement range. The delay line-based architectures, such as the TDL, VDL, and pulse shrinking architectures, achieve superior temporal resolution. Some recent works have demonstrated TDC designs with sub-picosecond resolutions using these architectures [112, 113]. On the other hand, the RO-based architectures are advantageous for the implementation of large-scale TDC arrays because these architectures enable reducing the area consumption. For instance, a 192 × 128 SPAD array with in-pixel gated ring oscillator TDCs was implemented in STMicroelectronics 40 nm CMOS technology [114]. The TDC has a tunable temporal resolution from 33 ps to 120 ps and an area occupation of 9.2 × 9.2 μm.

2.5.2 FPGA-Based TDCs

A field programmable gate array (FPGA) is a reconfigurable integrated circuit that can be programmed to implement various digital systems. It comprises an array of programmable logic blocks that can perform different logic functions and a routing matrix with configurable switches that interconnect the logic blocks. The configuration of the logic blocks and the switches is determined by the content of a large number of random access memory (RAM) cells that store the configuration data. The user describes the desired functionality using a high-level hardware description language (HDL), such as VHDL and Verilog. The HDL design is then synthesized and compiled using specific software tools, such as Quartus and Vivado, to generate a bitstream file that is used to program the configuration RAM cells [115].

The first commercial FPGA device, which comprised simple configurable logic blocks and volatile SRAM cells, was introduced by Xilinx in 1985. Since then, FPGA technology has advanced rapidly by enhancing the logic density and performance and incorporating additional types of programmable resources such as memory blocks, digital signal processing (DSP) circuits, and hard processor systems [116, 117]. Moreover, many high-performance SoC-FPGA devices are currently available in the market at very low prices, such as the DE10-Nano by Terasic, a development kit based on an Intel Cyclone V SoC-FPGA that combines a dual-core ARM processor with the FPGA fabric [118].

Compared with ASIC solutions, FPGAs offer unique features, such as flexibility, low cost, and short time-to-market. FPGA designs can be easily debugged, upgraded, and integrated with

other modules. These features enable the implementation of different TDC architectures in FPGAs, with different temporal resolutions, sampling rates, and measurement ranges that meet the target application requirements. Furthermore, FPGA TDCs can be integrated into larger designs with calibration, data processing and communication modules on the same FPGA.

TDL is the most commonly used architecture for FPGA-based TDCs. The first implementation of a TDC on FPGAs was reported in 1997 by Kalisz et al [119]. They implemented a TDL-TDC on an FPGA manufactured in 0.65-µm CMOS technology and achieved a resolution of 200 ps. Since then, many studies aimed to enhance the resolution of TDL-TDCs, mainly based on the carry-chain paths provided by Xilinx and Altera, the dominant suppliers of FPGA devices in the last two decades. With the rapid advancement of FPGA technologies, the temporal resolution has been significantly enhanced. In 2016, a temporal resolution of 3.9 ps RMS was achieved by a TDL-TDC implemented on 20-nm Xilinx Kintex UltraScale FPGA using a dual-sampling technique [120].

In recent years, several approaches have been proposed to further improve the TDC resolution, such as multi-chain averaging [121], wave-union method [122], and two-stage pulse shrinking [123].

2.6 Conclusion

This chapter provided a comprehensive overview of the TCSPC technique and its main components. It evaluated and compared the various choices for each component, highlighting their strengths and limitations.

Among the excitation light sources, laser diodes are the most prevalent and preferred choice for TCSPC systems. They offer several benefits, such as low cost and high reliability, and do not need any maintenance or alignment. Laser diodes can cover a broad spectrum of discrete wavelengths. Table 2.1 summarizes the comparison of different excitation light sources.

For photon detections, PMTs and SPADs are the main photon detectors for TCSPC systems. However, SPADs have many advantages over PMTs, such as low cost, moderate biasing voltage, compact size, and integration capability with front-end circuity.

Table 2.1 Comparison of excitation light sources.

Light source	Pulse width	Repetition rate	Power	Wavelength	Cost
Flashlamps	Nano to microseconds	Low	Low	Broad spectrum	Low
Pulsed lasers	Tens of picoseconds	Up to 120 MHz	High	Discrete	High
Ti:Sa lasers	Femto to picoseconds	High	Very high	Tunable + UV	Very high
Laser diodes	Tens to hundreds of picoseconds	High	Low	Discrete	Low

This chapter also reviewed the various types of TDCs and presented the advantages, disadvantages, and the different architectures of each type. Analog TDCs are widely used in current commercial TCSPC systems due to their high resolution, linearity, reliability, and stability. However, analog TDCs have some drawbacks such as their large size and high cost. With the recent advancements in CMOS technologies, ASIC and FPGA have become the preferred platforms to implement TDCs. ASIC-TDCs offer a high level of integration, high resolution, and low power consumption. However, the long development cycle and the high cost of these TDCs limit their suitability for large-scale general-purpose applications. On the other hand, FPGA-TDCs provide high flexibility, compatibility, and low development cost. FPGA-TDCs can be easily integrated with other modules on the same device. These features make FPGAs ideal for fast prototyping, scientific experiments, and high-end instruments. The main limitation of FPGA-TDCs is the poor linearity and the moderate temporal resolution, which is constrained by the predefined logic structure. However, several solutions have been proposed to improve the linearity of these TDCs and to narrow the resolution gap with ASIC-TDCs. Table 2.2 summarizes the benefits and drawbacks of the main implementation methods of TDCs.

Table 2.2 Comparison of different implementation methods of TDCs.

	Benefits	Drawbacks
Analog	Superior resolution (sub-picosecond) High linearity Reliability and stability	Large size High complexity High cost. Long dead time (>95ns) Limited channel numbers (<10 channels)
ASIC	High resolution (~1ps) Good linearity Low dead-time Highest integration level. Suitable for large-scale fabrication of general-purpose TDCs and TCSPC arrays	Long development cycle High cost Non-updatable
FPGA	Good resolution Low dead-time Flexibility and compatibility Short development time Low development cost Good integration level Debugging and upgrading capability	Poor linearity Limited resolution

Chapter 3: Time-to-Digital Converter

TDCs are the core of TCSPC systems [124]. The role of a TDC, in general, is to measure the precise time interval between two events represented by two signals: a reference and a measured signal. In TCSPC systems, TDCs measure the arrival time of detected photons relative to the emission of the excitation light. Thus, the reference signal corresponds to the excitation light emission, while the measured signal is the output of the photon detector. According to the synchronicity of the reference signal with the system clock, TDCs can be categorized into synchronous and asynchronous types. One of the target applications of our system requires a synchronous TDC, while the other three require asynchronous ones. Therefore, we designed and implemented both types of TDCs.

Considering the features of FPGAs presented in the precedent chapter, we chose to realize the TCSPC system using a low-cost SoC-FPGA platform, namely the Intel Cyclone V DE10-Nano from Terasic. We implemented the TDC system on the FPGA part and employed the embedded hard processor system (HPS) to perform all the related data processing locally on the development platform.

In such complex and high-precision designs as TDCs, the routing topology has a critical impact on the system behavior and performance because of the dense connection and configuration network of the FPGA. However, low-cost FPGA devices such as Cyclone V do not support manual routing and placement [125]. Therefore, we encountered many problems and challenges during the design and implementation of the FPGA-TDC, which necessitated imposing many constraints to control the routing and the placement of the different parts, to manage the propagation delays that ensure the correct functionality of the system, and to enhance the performance.

This chapter introduces the architecture of the designed TDCs and covers in detail each component of these architectures. Furthermore, it discusses the problems encountered due to the limitations of low-cost FPGAs and the proposed solutions to address these problems.

3.1 Synchronous TDC

In this work, we adopted the tapped delay line (TDL) architecture for implementing the TDC on the FPGA, as it provides the best resolution. Furthermore, we combined this architecture with the Nutt method to extend the measurement range (MR). However, in synchronous TDCs, the reference signal is synchronized with the system clock, which allows for simplifying this structure into a coarse-fine scheme [126-128]. As shown in the timing diagram in Figure 3.1, the time interval (*TI*) is decomposed into two parts: T_{coarse} and T_{fine}. T_{coarse} is easily measured by a counter driven by the system clock, while T_{fine} is measured by the TDL. Although the TDL architecture provides a high resolution, the implementation of a TDL in FPGA requires addressing many timing issues related to the signal propagation at the logic cell level. In the following, we present the design of the synchronous TDC and discuss in detail the challenges encountered in each part, as well as the proposed solutions to overcome these problems.

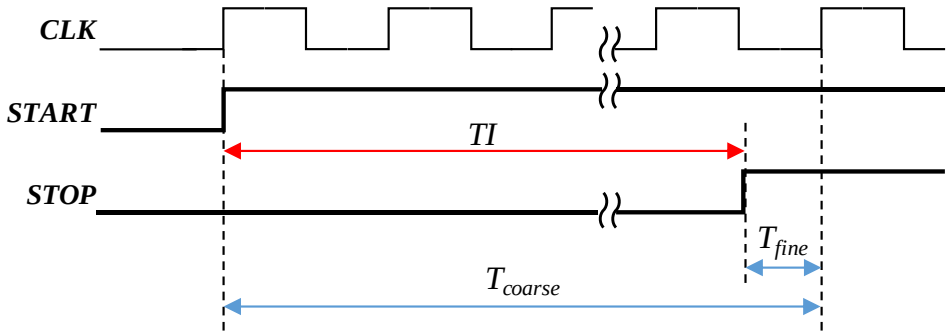


Figure 3.1 Timing diagram of Synchronous TDC. TI is decomposed into two components T_{coarse} and T_{fine}.

The global architecture of our TDC is depicted in Figure 3.2. It comprises two measurement blocks (a fine and a coarse block), a signal controller, and a data writing block. The signal controller receives the single-photon avalanche diode (SPAD) signal and generates appropriate signals for the fine and coarse blocks to measure the arrival time of the detected photon. It also sends a request signal (REQ) to the data writing controller to store the arrival time data in a first-in-first-out (FIFO) memory. The following subsections describe these blocks in detail and discuss the challenges related to their implementation and the proposed solutions to address these challenges.

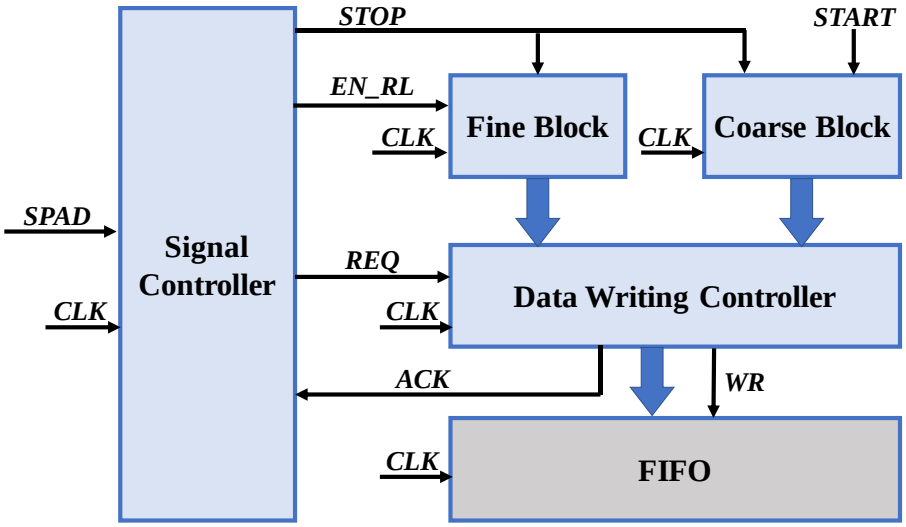


Figure 3.2 Global architecture of the synchronous TDC.

3.1.1 Fine Block

The fine block is the essential component of the TDC system, as it determines the resolution. It measures the time interval from the *STOP* signal to the next rising edge of the clock. To achieve a high resolution for our TDC, we designed the fine block based on the TDL structure. The STOP signal propagates through delay elements (DEs), whose outputs are sampled at the following clock rising edge and then encoded into a fine timestamp. The fine block consists of the TDL and an encoder to convert the TDL thermometer code into a binary code.

A. Tapped Delay Line (TDL)

The resolution of the FPGA-TDL can be optimized by utilizing high-speed paths dedicated for the carry propagation, called the carry-chains, instead of the global routing network. These chains provide a very high propagation speed that improves the TDC resolution [129, 130].

- **Cyclone V Architecture and Carry Chains**

The logical architecture of the Cyclone V FPGA used to implement the system is depicted in Figure 3.3. The Logic Array Block (LAB) is the basic building block of the Cyclone V FPGA. Each LAB comprises ten Adaptive Logic Modules (ALMs) that can be configured to perform various logic, arithmetic, and register functions. Each ALM includes four 4-input look-up tables (LUTs), two dedicated adders, and four registers. The adder sum-out output can be sampled by two registers, and the carry connections between the adders provide a fast carry chain along the LABs existing in the same column [131].

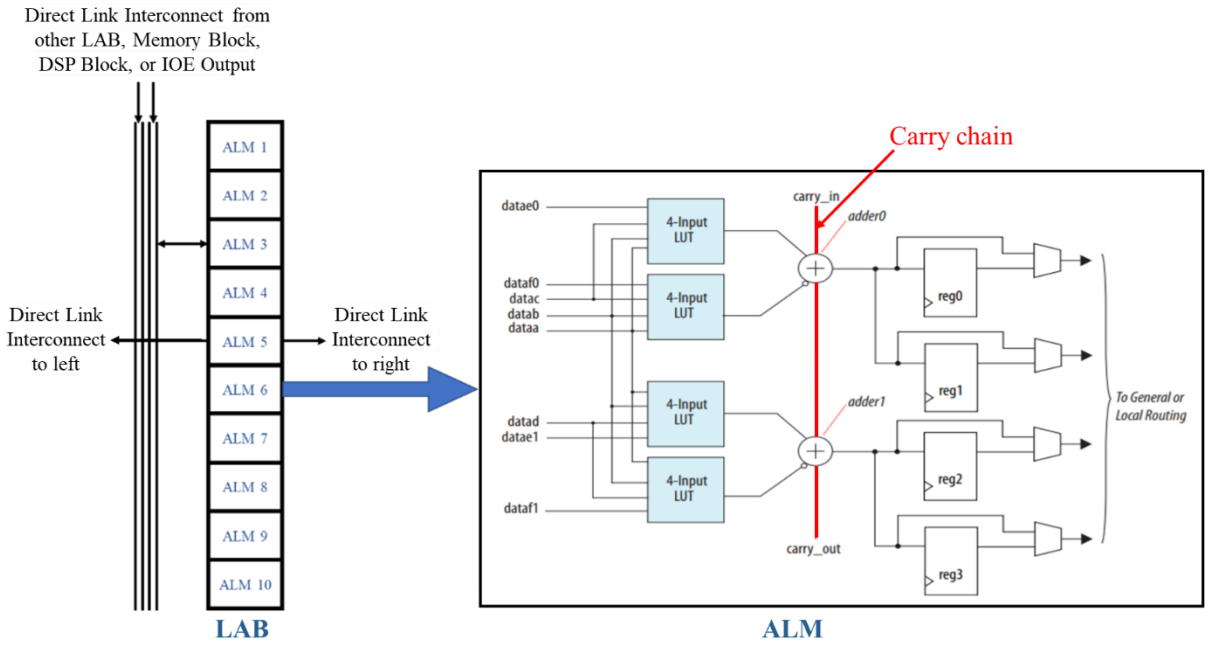


Figure 3.3 Logical architecture of Cyclone V FPGA [131].

- **TDL Design**

Since Cyclone V FPGA does not support manual routing, we employed an N-bit adder to implement the TDL based on the carry-chain of the Cyclone V FPGA [129, 132]. Each delay element (DE) consists of a one-bit adder with the sum-out output connected to the data input of a DFF. The adders form a delay line (DL) and the DFFs form a register line (RL). The first adder receives '1' and the measured signal as its operands. while each of the other adders has its two operands set to '0' and '1'. The carry-out output of each adder is connected to the carry-in input of the subsequent adder. Figure 3.4 shows the architecture of the carry-chain-based TDL.

In the absence of photon detection, the *STOP* signal is in a low state and both the sum-out and carry-out of the first adder are in a low state as well. As a result, the sum-out outputs of all subsequent adders are in a high state. Upon photon detection, the *STOP* signal changes to a high state at the first operand input of the first adder. Following the propagation time of this adder, its sum-out switches to '0' and its carry-out to '1'. After a certain delay, the carry-in of the second adder transitions to '1', causing its sum-out to switch to '0' and its carry-out to '1'. This process continues along the DL and the *STOP* signal propagates through the carry chain, resulting in zeros at the sum-out of the propagated adders, while the sum-out outputs of the non-

propagated adders remain at ‘1’. The RL samples the sum-out outputs of the adders at the next rising edge of the clock. Ideally, in perfect TDLs, the value at the RL would be a thermometer code (unary code) that can be easily encoded into a binary value.

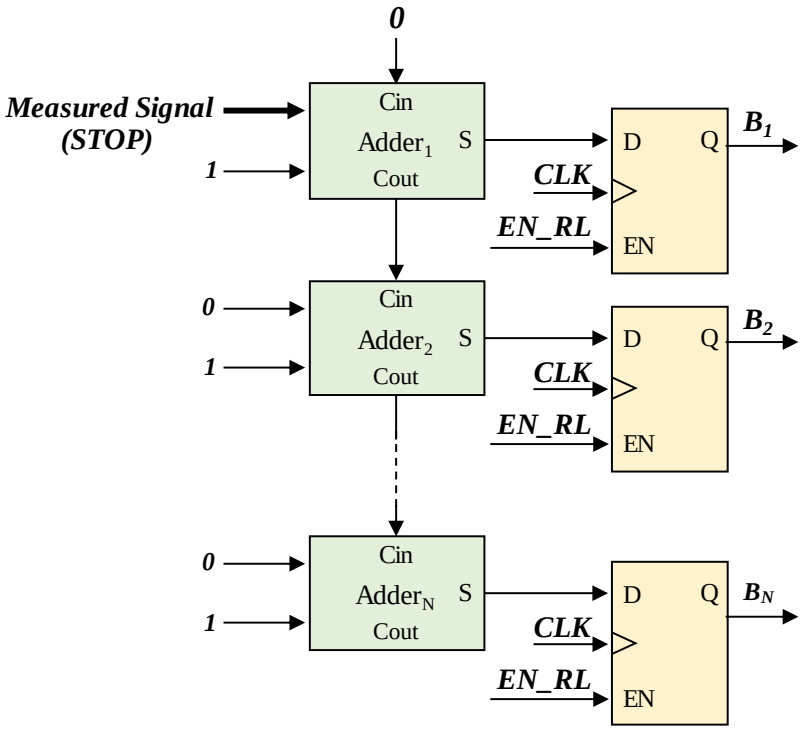


Figure 3.4 Carry-chain-based TDL architecture.

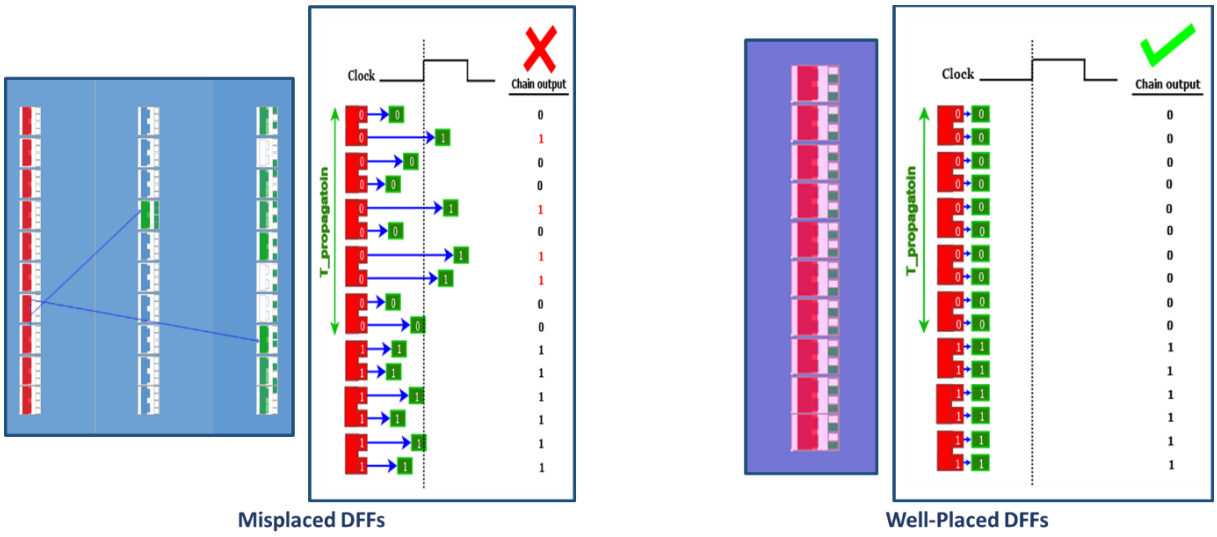
- **TDL Implementation**

We developed a VHDL code to implement a 1024-bit adder combined with a 1024-bit register line, as proposed in [129]. However, the initial compilation of our VHDL code with the Intel Quartus Prime software resulted in an unwanted optimization of the design. The synthesis tool eliminated the delay line since the adders’ operands are constant values of ‘1’ and ‘0’. To preserve the delay line, we modified the VHDL code by adding the keep attribute for the adder output signal and setting it to “true”. This prevented the synthesis tool from simplifying the design. Alternatively, this issue could be addressed by defining the TDL block, which contains the adder code, as a design partition.

When the design was recompiled, Quartus did not optimize the design and placed the TDL. However, it did not place all the registers in the same ALMs as their corresponding adders, as illustrated by the chip planner view in Figure 3.5-a. This undesirable placement resulted in large variations in the delay from the sum-out of the adders to the data input of their respective registers, and consequently in the delay of the DEs, which degraded the linearity of the TDC.

Moreover, the effect of the clock skew was aggravated, leading to erroneous thermometer codes with many bubble bits, as demonstrated in Figure 3.5-a. In fact, an ideal TDL with identical DEs generates clean thermometer codes, e.g. “0000001111111111”, with a single transition from ‘0’ to ‘1’. The position of this transition indicates the distance propagated by the measured signal at the sampling instance. Nevertheless, the discrepancies in the DEs may cause multiple transitions at the TDL output, e.g. “0000101001111111”, resulting in incorrect thermometer codes. This effect is known as the bubble problem.

In order to prevent the misalignment of registers, we instantiated both the adder delay line (DL) block and the register line (RL) block within an upper-level block, and created a logic lock region for this block using the chip planner tool. We adjusted the width of this region to 1 and its height, i.e. the number of LABs constituting this region, to a value calculated according to the required number of EDs (*N*). Given that each LAB in Cyclone V FPGA comprises 20 adders, the height of the logic lock region should be assigned to *N* divided by 20. For a TDL of 1024 adder, we set the height to 52. With these configurations, we forced the fitter to place the registers in the same ALMs as their corresponding adders, as shown in Figure 3.5-b.



(a) (b)

Figure 3.5 Registers placement problem : (a) without logic lock region, the registers are misplaced, (b) with logic lock region, the registers are placed in the same ALMs as their corresponding adders.

Another benefit of creating a logic lock region for the TDL is to control the physical placement of the latter in the FPGA chip by adjusting the origin of the region. This enables us to experimentally observe the impact of the physical placement on the TDC performance and to select the location that yields the best performance.

- **System Clock Frequency**

In TDL architecture, the TDC clock frequency should be carefully selected according to the total propagation delay of the TDL (*Tp*), which can be calculated from Equation 3.1.

$$T_p = N \times d \tag{3.1}$$

where N is the number of delay elements (DEs) and d is the DE's average delay.

To guarantee the correct operation of the TDC under temperature and voltage variations, T_p should cover a system clock period with some margin [133, 134]. This implies that T_{clk} should be slightly smaller than T_p, as expressed in Equation 3.2.

$$\begin{aligned} T_{clk} &< T_p \\ T_{clk} &< N \times d \end{aligned} \tag{3.2}$$

If T_{clk} is much shorter than T_p ($T_{clk} << T_p$), the TDL would have much more DEs than required to cover T_{clk}. The extra DEs at the end of the TDL are called non-effective DEs as they never participate in the fine time measurement. The reason is that the measured signal would not reach these DEs when the TDL output is sampled at the next clock rising edge, and the DFFs associated with these DEs would always be at logic "1", resulting in dead bins as depicted in Figure 3.6. Moreover, the non-effective DEs increase the encoding time and the resource consumption for both the TDL and the encoder.

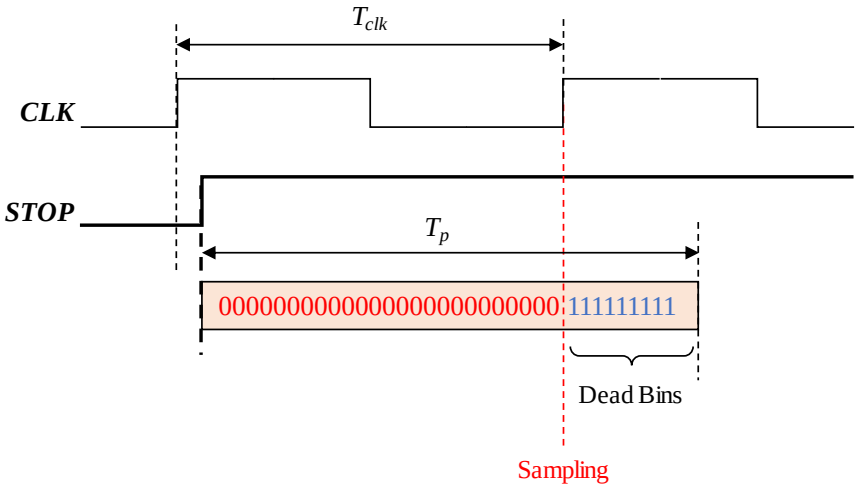


Figure 3.6 Dead bins in the case of a clock period much shorter than the TDL delay.

In contrast, if T_{clk} is longer than T_p ($T_{clk} > T_p$), the TDL would not cover an entire clock period. Hence, when the measured signal arrives at the beginning of a clock cycle, in a range from 0 to

(T_{clk} – T_p), all the DEs would be propagated at the next clock rising edge, as illustrated in Figure 3.7, giving zeros at the output of all the DEs. Consequently, the first bin (Bin_0) would have a very large width which degrades the TDC resolution.

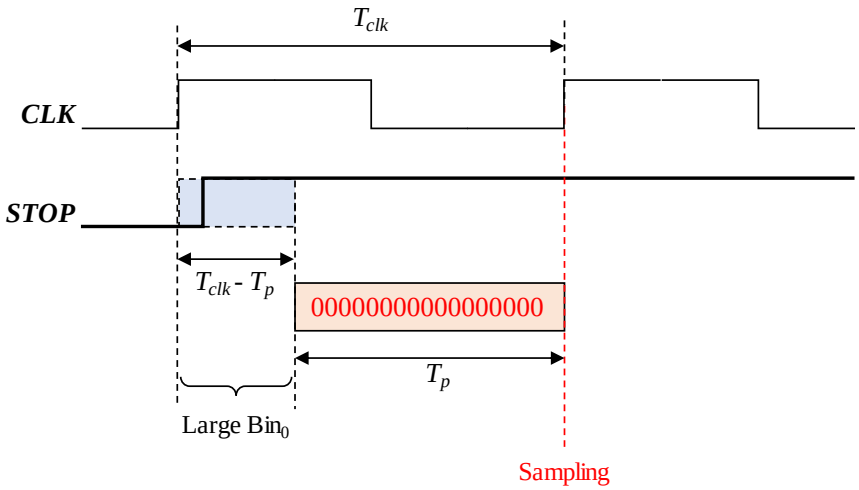


Figure 3.7 Large width of Bin_0 in the case of a clock period longer than the TDL delay.

In order to estimate T_p and the linearity of the TDL, we analyzed the propagation delays along the TDL by both Quartus Timing Analyzer tool and experimental measurements.

The timing analysis report of the carry-chain path of the TDL indicated that the propagation time from the first adder to the last one (T_p) is about 31.57 ns. Based on this value and Equation 3.2, we configured the system clock frequency at 32 MHz and compiled the design. At this frequency, the clock period (T_{clk} = 31.25 ns) would be slightly shorter than T_p. However, when we tested the TDC experimentally, the code density results showed that the first bin (Bin_0) had a very large width, which means that T_p is considerably shorter than T_{clk} (T_p << 31.25), as discussed above and illustrated in Figure 3.7.

To estimate the actual propagation time through the 1024-adder TDL, we progressively reduced T_{clk} and conducted a code density test for each T_{clk}. The findings showed that the width of Bin_0 decreased with the reduction in T_{clk}. At ($T_{clk} \approx$ 6.4 ns), all the bins of the code density histogram were filled without any large or dead bins. With T_{clk} less than 6.4 ns, we encountered dead bins, meaning that TP was longer than the clock period., as demonstrated in Figure 3.6.

Further analysis and experiments, carried out on TDLs with different numbers of DEs and different clock periods, demonstrated that the timing results obtained by the analyzing tool were

overestimated by a factor of 5 compared to the experimental results. This finding is in agreement with [135], which reported that the propagation time of 691 adders is about 4 ns. This discrepancy between the experimental and the timing analyzer results is likely due to the incapability of the timing analyzer model to assess accurate timing with such an optimized routing topology with very short propagation delays of a few picoseconds. Indeed, in a standard compilation, the logic cells composing the TDL would generally be dispersed in the LABs, resulting in longer propagation delays.

- **Temporal Resolution and Linearity**

Figure 3.8 presents a part of the carry chain path report, obtained by the Timing Analyzer tool, covering the first 7 ALMs of a LAB. The report indicates that the propagation delay is not uniform for all the adders, but varies according to the order of the adder in the ALM and the position of the ALM in the LAB. The first adder in the ALM has an estimated delay of approximately 100 ps for the first and sixth ALMs of the LAB, and of approximately 50 ps for the rest of the ALMs. These estimated delays, derived by the timing analyzer, correspond to actual delays of about 20 ps and 10 ps, respectively, as confirmed by the experimental finding. On the other hand, the second adder of all the ALMs has a neglected delay.

		Total	Incr	RF	Type	Fanout	Location	Element
ALM-1	152	2.372	0.108	RR	CELL	1	LABCELL_X22_Y51_N0	inst4|inst|inst9|inst3|_~1234|cout
	153	2.372	0.000	RR	IC	2	LABCELL_X22_Y51_N3	inst4|inst|inst9|inst3|result[81]|cin
	154	2.372	0.000	RR	CELL	1	LABCELL_X22_Y51_N3	inst4|inst|inst9|inst3|result[81]|cout
	155	2.372	0.000	RR	IC	1	LABCELL_X22_Y51_N6	inst4|inst|inst9|inst3|_~2383|cin
ALM-2	156	2.422	0.050	RR	CELL	1	LABCELL_X22_Y51_N6	inst4|inst|inst9|inst3|_~2383|cout
	157	2.422	0.000	RR	IC	1	LABCELL_X22_Y51_N9	inst4|inst|inst9|inst3|_~1870|cin
	158	2.422	0.000	RR	CELL	1	LABCELL_X22_Y51_N9	inst4|inst|inst9|inst3|_~1870|cout
	159	2.422	0.000	RR	IC	1	LABCELL_X22_Y51_N12	inst4|inst|inst9|inst3|_~1267|cin
ALM-3	160	2.472	0.050	RR	CELL	1	LABCELL_X22_Y51_N12	inst4|inst|inst9|inst3|_~1267|cout
	161	2.472	0.000	RR	IC	2	LABCELL_X22_Y51_N15	inst4|inst|inst9|inst3|result[85]|cin
	162	2.472	0.000	RR	CELL	1	LABCELL_X22_Y51_N15	inst4|inst|inst9|inst3|result[85]|cout
	163	2.472	0.000	RR	IC	1	LABCELL_X22_Y51_N18	inst4|inst|inst9|inst3|_~2398|cin
ALM-4	164	2.522	0.050	RR	CELL	1	LABCELL_X22_Y51_N18	inst4|inst|inst9|inst3|_~2398|cout
	165	2.522	0.000	RR	IC	1	LABCELL_X22_Y51_N21	inst4|inst|inst9|inst3|_~1885|cin
	166	2.522	0.000	RR	CELL	1	LABCELL_X22_Y51_N21	inst4|inst|inst9|inst3|_~1885|cout
	167	2.522	0.000	RR	IC	1	LABCELL_X22_Y51_N24	inst4|inst|inst9|inst3|_~1288|cin
ALM-5	168	2.586	0.064	RR	CELL	1	LABCELL_X22_Y51_N24	inst4|inst|inst9|inst3|_~1288|cout
	169	2.586	0.000	RR	IC	2	LABCELL_X22_Y51_N27	inst4|inst|inst9|inst3|result[89]|cin
	170	2.586	0.000	RR	CELL	1	LABCELL_X22_Y51_N27	inst4|inst|inst9|inst3|result[89]|cout
	171	2.586	0.000	RR	IC	1	LABCELL_X22_Y51_N30	inst4|inst|inst9|inst3|_~1291|cin
ALM-6	172	2.686	0.100	RR	CELL	1	LABCELL_X22_Y51_N30	inst4|inst|inst9|inst3|_~1291|cout
	173	2.686	0.000	RR	IC	2	LABCELL_X22_Y51_N33	inst4|inst|inst9|inst3|result[91]|cin
	174	2.686	0.000	RR	CELL	1	LABCELL_X22_Y51_N33	inst4|inst|inst9|inst3|result[91]|cout
	175	2.686	0.000	RR	IC	1	LABCELL_X22_Y51_N36	inst4|inst|inst9|inst3|_~2401|cin
ALM-7	176	2.736	0.050	RR	CELL	1	LABCELL_X22_Y51_N36	inst4|inst|inst9|inst3|_~2401|cout
	177	2.736	0.000	RR	IC	1	LABCELL_X22_Y51_N39	inst4|inst|inst9|inst3|_~1888|cin
	178	2.736	0.000	RR	CELL	1	LABCELL_X22_Y51_N39	inst4|inst|inst9|inst3|_~1888|cout

Delay of the first adder in the first and sixth ALM of the LAB

Delay of the first adder in the remaining ALMs of the LAB

Delay of the second adder of all the ALMs of the LAB

Figure 3.8 Carry-chain path report in the timing analyzer tool.

These variations in the propagation delay of the adders and from the adders to the flip-flops along with the clock skew cause large variations in the bin widths. This results in ultra-large and zero-width bins that degrade the linearity of the TDC. Moreover, the zero-width bins induce the bubble problem and produce erroneous thermometer codes at the output of the TDL.

- **Tuned-Downsampling**

The discrepancies in the DE propagation time could be useful to enhance the TDL resolution, but this requires complex bin realignment techniques to overcome the bubble problem [127] [135]. Another common technique to improve the linearity of the TDC and to reduce the bubble effect is downsampling. This technique consists in selectively sampling the delay elements with a fixed ratio, such as one out of every four adders. In this case, each DE would comprise several adders. However, this method improves the linearity at the cost of reducing the resolution [136, 137]. Furthermore, the large variations in the adders' delays limit the linearity improvement that can be achieved by conventional downsampling.

In this study, we propose a tuned-downsampling method that takes into account the delay variations of the adders. Instead of downsampling the TDL output with a constant ratio, we dynamically adjust the number of adders in each DE based on the delay of these adders, such that all the DEs have approximately the same delay.

Figure 3.9 demonstrates the topology of the tuned-downsampling method which is explained as follows.

- Based on the timing analysis results presented above and shown in Figure 3.8, each LAB can be divided into two halves, upper and lower, with similar propagation time characteristics. Each half consists of five ALMs and each ALM has two adders. In each half LAB, the propagation time through the two adders of the first ALM is about (20 + 0 = 20 ps), while it is about (10 + 0 =10 ps) for each of the subsequent four ALMs.
- In theory, three DEs with a propagation time of about 20 ps can be implemented in each half LAB. The first DE consists of the two adders of the first ALM. The second one comprises the four adders located in the second and the third ALMs. The third DE consists of the four adders of the fourth and fifth ALMs. This approach is applied for all the half-LABs that form the TDL.

To implement the tuned-downsampling, we developed a VHDL code that defines the interconnections between the TDL adders to the DFFs of the register line.

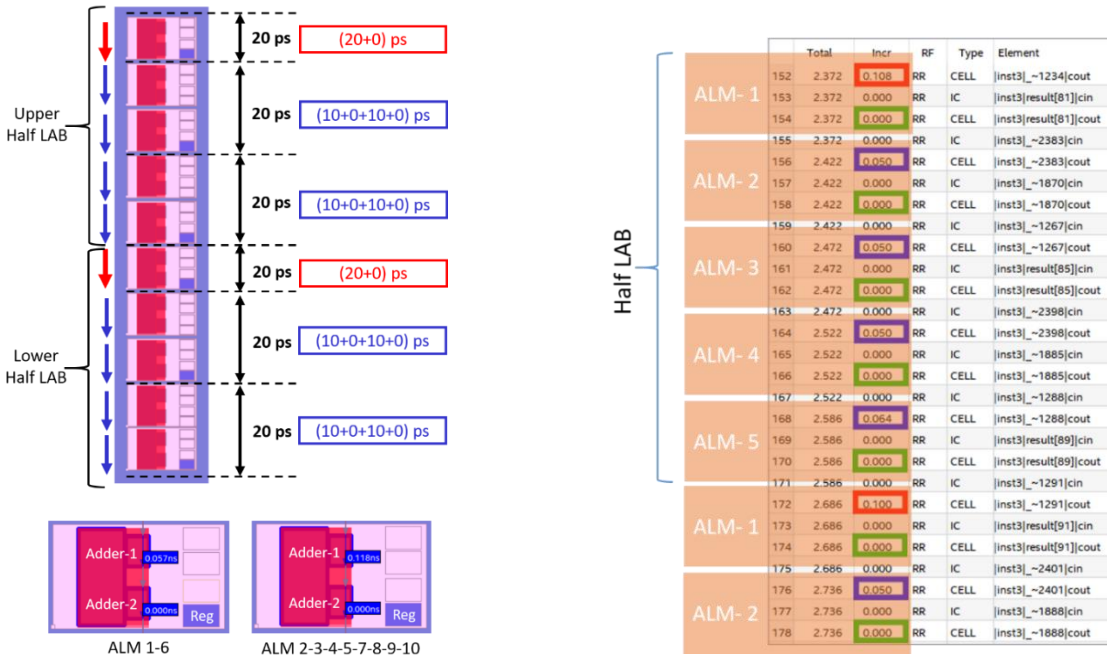


Figure 3.9 Tuned downsampling: three delay elements are implemented in each half LAB with an identical delay of about 20 ps.

The tuned-downsampling improves the linearity of the TDL and mitigates the bubble effect, albeit at the cost of reducing the resolution to about 20 ps. However, such a temporal resolution is generally sufficient for most TCSPC applications that employ a Single Photon Avalanche Diode (SPAD) as a photon detector, since the jitter introduced by the latter is on the order of some tens of picoseconds [138].

Another advantage of the tuned-downsampling technique is the reduction of resource consumption. Indeed, due to the high speed of the carry chain in Cyclone V FPGA, a high-frequency system clock is needed to drive the system. For example, a clock frequency of 667 MHz is required for a 255-DE TDL without downsampling. However, such a high frequency would not meet the timing requirements of the routing and the other blocks of the system (coarse counter, data writing controller, etc.), as these blocks are also driven by the same clock. Likewise, a very long TDL with a large number of adders is needed to cover the clock period for a reasonable frequency (855 DEs for 200 MHz). This increases the resource consumption, the encoding time, and thus the TDC's dead time. Therefore, the tuned downsampling technique provides a suitable trade-off between the temporal resolution on one hand, and the resource consumption and the dead time on the other hand.

Based on the tuned-downsampling technique, we implemented a TDL of 255 DEs with an average elementary delay of about 21.7 ps and a total propagation time (T_p) of about

(255 × 21.7 ps ≈ 5.55 ns). To respect the system clock frequency criteria given in Equation 3.2 ($T_{clk} < T_p$), we set the clock frequency to 190 MHz, corresponding to a clock period of 5.26 ns which is slightly shorter than T_p.

B. Thermometer-to-Binary Encoder

The encoder's function in a TDC is to convert the thermometer code of the TDL output into a binary value that represents the fine timestamp (T_{fine}). Various encoder architectures have been proposed in the literature, such as the ROM encoder, ones-counter, fat tree, folded Wallace tree, and multiplexer-based encoder. These architectures are comprehensively compared and evaluated by [139-143]. Among them, the ones-counter architecture is the most suitable one for TDCs as it automatically performs the bubble bits correction [127].

As shown previously, the *STOP* signal propagates through the TDL resulting in zeros at the sum-out of the propagated adders. The adders' output is sampled by the RL at the following clock rising edge, generating a thermometer code with a sequence of zeros followed by a sequence of ones. The number of zeros in this code is proportional to the time interval from the *STOP* rising edge to the next clock rising edge. Assuming that the TDL covers exactly one clock period ($T_p = T_{clk}$), the number of ones in the thermometer code would correspond to the time interval between the *STOP* signal and the previous clock rising edge, as depicted in Figure 3.10. Therefore, this interval can be quantified by counting the ones at the RL with a ones-counter to produce the fine timestamp. However, if the TDL propagation time is longer than the clock period ($T_p > T_{clk}$), the extra non-effective DEs introduce dead bins, and the fine timestamp at the encoder's output would never attain their values. Nevertheless, we propose a calibration method that is insensitive to the dead bins and implicitly resolves this issue. This method will be explained in detail in the following chapter.

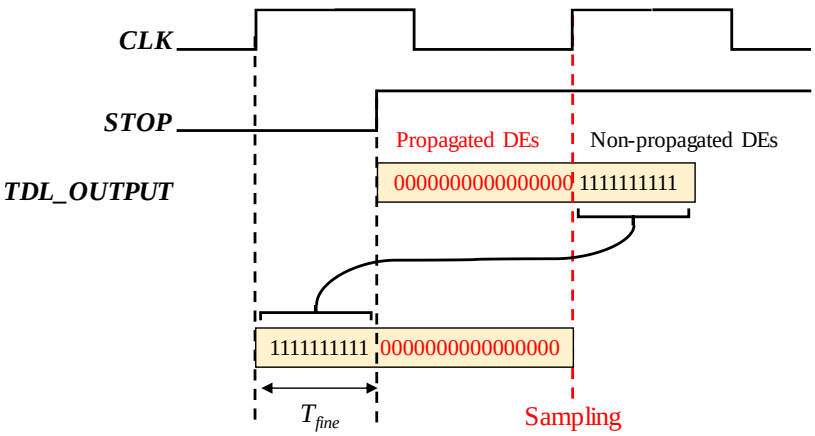


Figure 3.10 Ones-counter encoding principle.

With these considerations, T_{coarse} corresponds to the number of entire clock cycles between the *START* and *STOP* rising edges, and T_{fine} is the time interval between the *STOP* rising edge and the previous clock rising edge.

- **Ones-Counter Implementation**

This architecture consists of adders arranged in multiple stages. The number of stages is determined by the bit-width of the thermometer code. The adders count the number of logic '1' bits in the thermometer code and generate the corresponding binary at the encoder output.

For our TDL, we implemented a 255-bit to 8-bit encoder with 8 stages of adders, as depicted in Figure 3.11. The first stage comprises 128 1-bit full adders that compute the sum of two adjacent bits of the RL thermometer code. In the second stage, the 2-bit outputs of the first stage adders are fed to the inputs of 64 2-bit full adders, and so forth, the seventh stage contains two 7-bit adders. Finally, the last stage is an 8-bit adder that aggregates the output of the two adders of the seventh stage. The least significant 8 bits of the last adder constitute the binary output of the encoder. Thus, the designed encoder enables the conversion of the RL output into an 8-bit fine timestamp with automatic correction of the bubble bits.

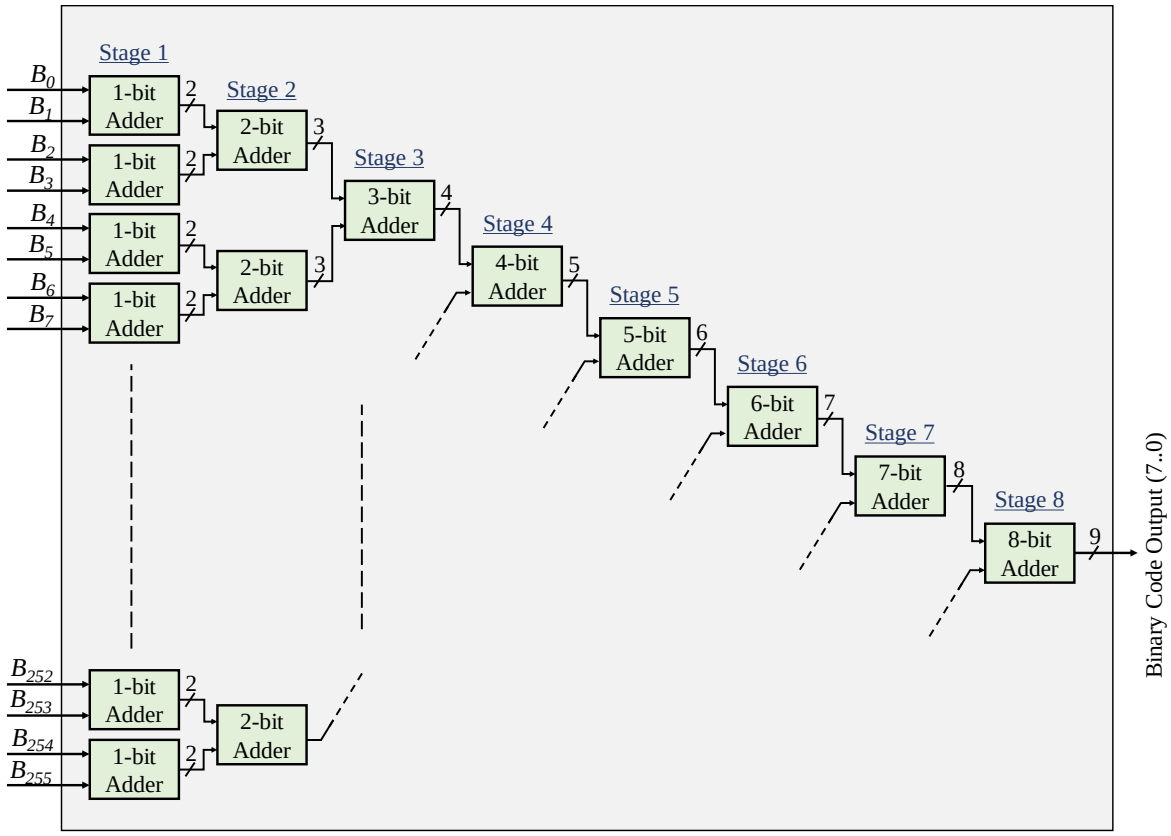


Figure 3.11 Basic architecture of a 255-to-8-bit ones-counter encoder, it consists of 8 stages of adders that count the number of ones in the thermometer code and convert it into a binary code.

- **Encoding Delay and Dead Time**

The propagation time through the encoder's cascaded stages was analyzed using the Quartus Timing Analyzer and the Signal Tap Logic Analyzer tools. The results indicated that the critical path's delay is approximately 8 ns, exceeding one clock period (T_{clk} = 5,26 ns). This propagation time was rounded up to two clock periods, and an additional clock period was added for the synchronous data writing, as will be explained later. This results in a total dead time of three clock periods (~15,8 ns) due to the encoding and data writing processes, limiting the maximum sampling rate to about 63.3×10^6 sample/s. Nevertheless, this dead time is shorter than the dead time of a commercial SPAD, which is about 40 ns for the one utilized in our system [62]. Therefore, the photon detector's dead time would be the main limitation of the sampling rate.

3.1.2 Coarse Block

The coarse block is responsible for measuring T_{coarse} as the number of clock cycles between the *START* and the *STOP* signals. A straightforward implementation of this block is a counter driven by the system clock, which starts counting at the rising edge of the *START* signal and stops at the *STOP* signal rising edge. Figure 3.12 depicts the architecture of the Coarse block that comprises an N-bit counter and a register with the same bit-width as the counter. The counter is reset to zero by the *START* signal, which is a clock-period-width pulse synchronized with the system clock, and then starts incrementing at the clock's rising edges. When the *STOP* signal arrives, the register stores the counter value which indicates the number of clock cycles between the *START* and *STOP* signals. However, this architecture suffers from a severe metastability issue due to the asynchronous sampling of the counter value as the *STOP* signal that triggers the register is not synchronized with the system clock. A solution for this problem is proposed in the following section.

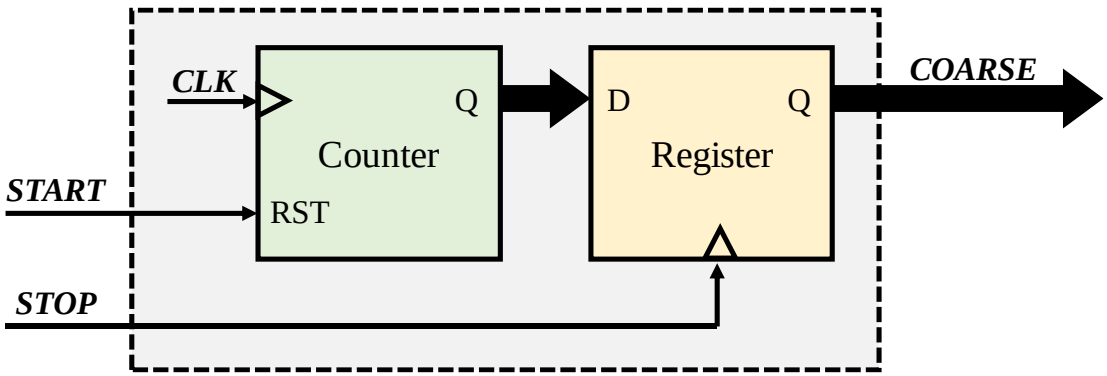


Figure 3.12 Coarse block.

- **Metastability Problem**

The registered coarse value is metastable when the sampling occurs near the transition of the counter value, i.e. when the *STOP* signal arrives close to the clock rising edge. This metastability may introduce an error of one LSB in the coarse value, i.e. an error of one clock period in the measurement.

A common approach to address this problem is often adopted in the implementation of a coarse time measurement in ASIC TDCs [144, 145]. This approach employs two independent counters operating at the two opposite edges of the clock. Since the two counters are phase-shifted with respect to each other, at any given time there will be at least one counter that is stable, as illustrated in Figure 3.13. The selection between these two counters depends on the actual phase difference between their clocks and on the fine value which indicates the arrival instance of the *STOP* rising edge within the clock cycle. The selection of the stable counter can be done using a lookup table.

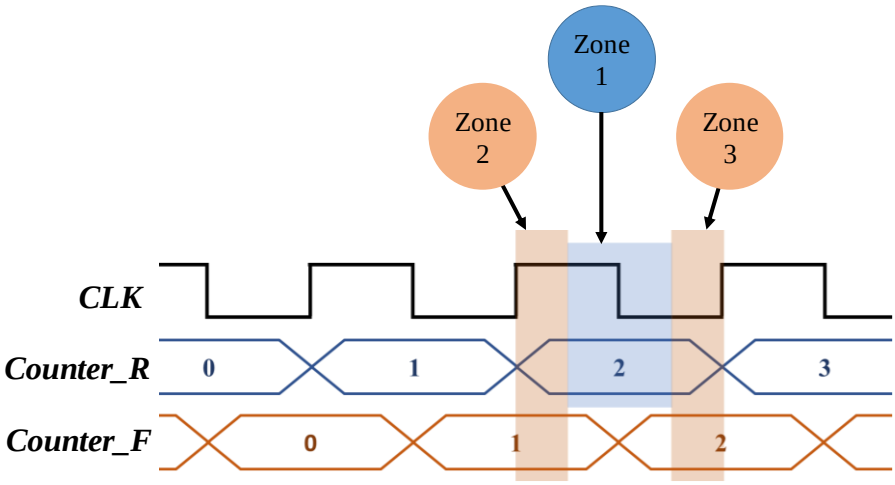


Figure 3.13 Timing diagram of the dual counter solution for the metastability issue arising from the asynchronous sampling.

A major challenge of implementing this solution on FPGA is that the relative phase between the two counters may vary with each recompilation of the TDC description code, affecting the stable timing zones of the two counters. This is because the placement and routing of the design on the FPGA chip may change with each recompilation [146]. A possible solution to this issue is to use fixed logic lock regions to constrain the placement of the counters and registers, thus forcing the fitter to preserve their placement across different compilations.

- **Coarse Block Implementation**

Figure 3.14 shows the basic architecture of the Coarse block which consists of three parts:

1- Two 8-bit counters, Counter_R and Counter_F, that increment on the rising and the falling edge of the TDC clock, respectively. Counter_R is reset by the *START* signal, which is a clock-cycle-width pulse synchronized with the TDC clock. Counter_F is reset by the *START* signal after being synchronized to the inverted TDC clock.

2- Two 8-bit DFF banks, Register_R and Register_F, that store the counters' values at the STOP signal.

3- Selection sub-block that chooses the appropriate coarse value from Register_R or Register_F.

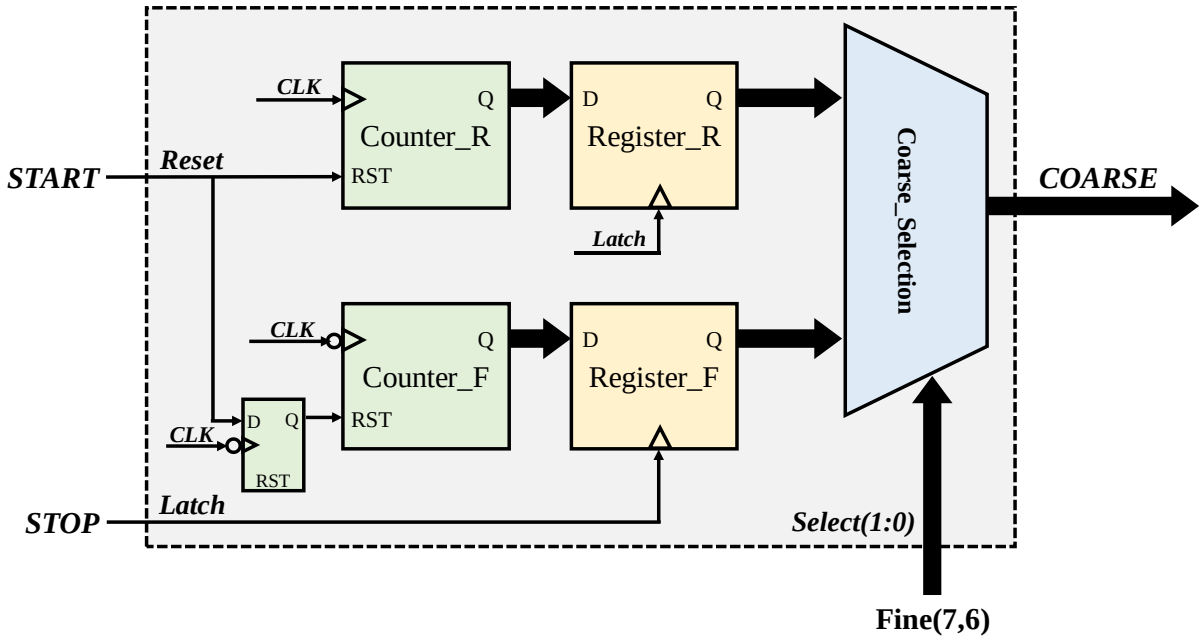


Figure 3.14 Schematic diagram of the Coarse block.

The selection sub-block determines the coarse value based on the 8-bit fine value which indicates the arrival time of the *STOP* signal within the clock. As depicted in Figure 3.15, three scenarios can be distinguished:

- If the *STOP* signal arrives in the middle of the non-phased clock cycle within a range of half period, i.e. the fine value is between 64 and 191, the selection sub-block selects the coarse value from Register_R, which is associated with the counter incrementing at the rising edge of the clock.

- If the *STOP* signal arrives at the last quarter of the clock cycle, i.e. the fine value is between 192 and 255, the selection sub-block selects the value from Register_F, which is associated with the counter incrementing at the falling edge of the clock.
- If the *STOP* signal arrives in the first quarter of the clock cycle, i.e. the fine value is between 0 and 63, the coarse value is set to Register_F incremented by one.

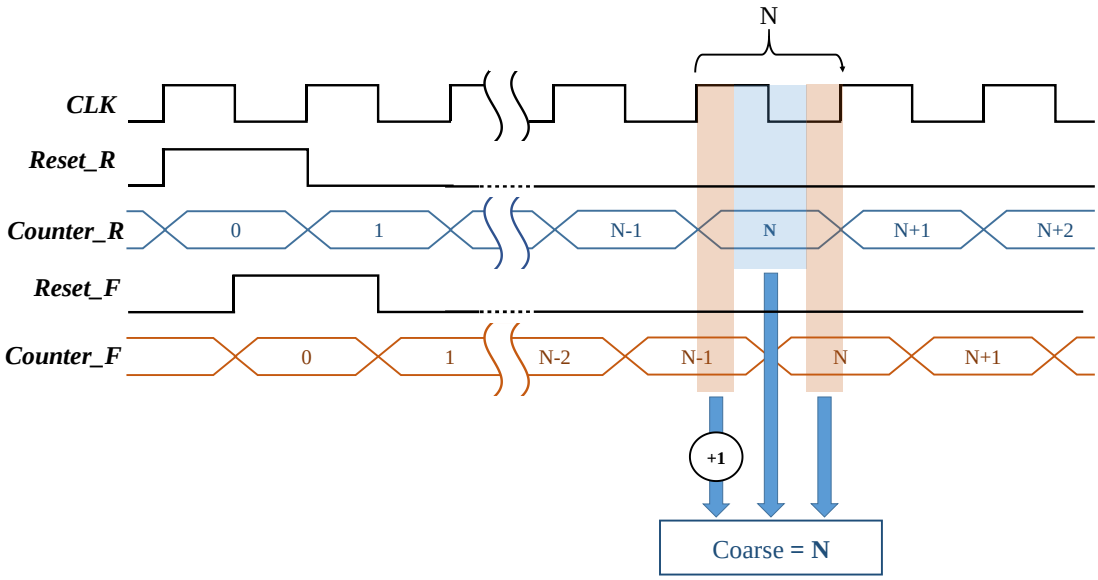


Figure 3.15 Selection sub-block operation principle; the stable coarse value is selected based on the fine value, which indicates the STOP signal's arrival time within the clock period.

The selection sub-block can be implemented using comparators and multiplexers (MUXs). However, in order to reduce the resource consumption and the selection latency, we used basic logic gates instead of comparators to generate the selection signals for the MUXs, by testing just the most significant two bits of the fine value. As shown in Table 3.1, these two bits are "00" for the first range, "11" for the last range, and "01" or "10" for the middle range. Figure 3.16 illustrates the basic architecture of the selection sub-block and the selection logic. This structure can be generalized for any bit-width of the fine value by using only the two most significant bits.

Table 3.1 Coarse value selection rules based on the fine value.

Fine value range	B7B6	Selected register	Final coarse value
0-63	"00"	Register_F	Register_F + 1
64-191	"01" or "10"	Register_R	Register_R
192-255	"11"	Register_F	Register_F

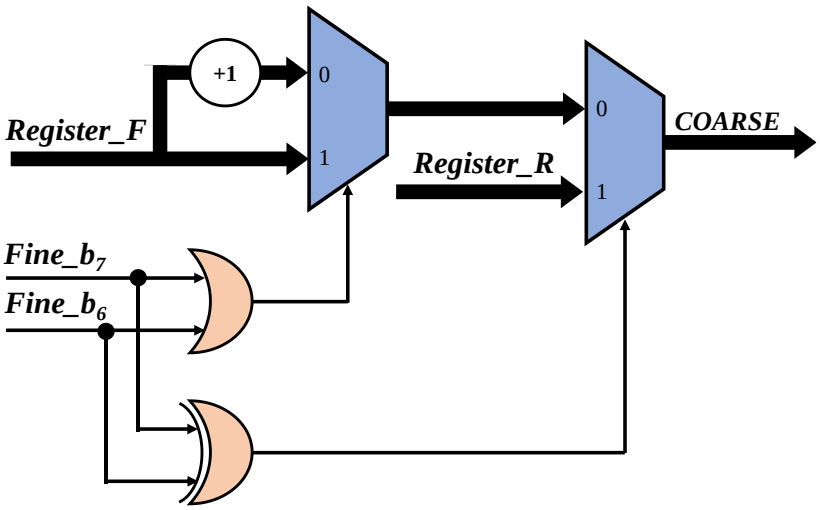


Figure 3.16 Basic architecture of the selection sub-block.

3.1.3 Signal Controller

The signal controller block is composed of three DFFs, each with a distinct configuration to perform three tasks: a) SPAD signal detection and preprocessing, b) enabling the register line (RL) of the TDL, c) generating the writing request for the data writing controller to save the measured arrival time. The signal controller architecture is shown in Figure 3.17.

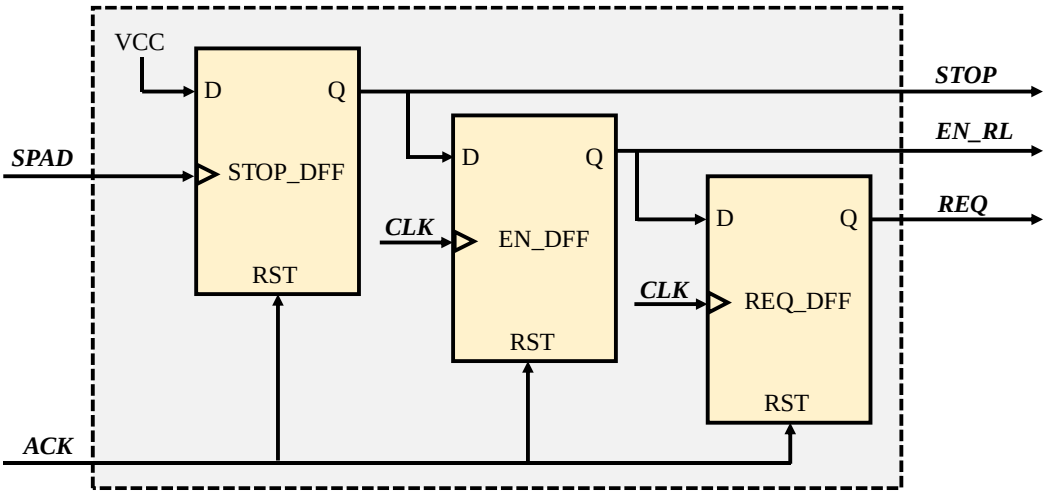


Figure 3.17 Signal controller architecture.

a) SPAD Signal Detection and Preprocessing

To ensure the proper functioning of the proposed TDL, the measured signal must have a single transition within the TDL delay time. However, the *SPAD* signal is affected by the noise

introduced by the ground bounce and the voltage sag, which can be indirectly observed as bubble bits at the TDL output when the *SPAD* signal is directly fed into the TDL without any preprocessing. The bubble bits are noise bits caused by the noise sources coupling at the TDL input. These bubble bits reduce the accuracy and reliability of the TDC measurement. can be eliminated by applying a low pass filter to the thermometer code, but this approach has several drawbacks. It increases the FPGA resource consumption and degrades the RMS accuracy of the TDC.

The first function of the signal controller is to detect the *SPAD* signal and block the fluctuations on this signal in order to generate a noise-free *STOP* signal that can be properly measured by the TDL. This function is realized by a DFF named STOP_DFF with the following configurations: the data input is fixed to "1" and the clock port is connected to the *SPAD* signal, while the reset port is connected to the acknowledgment signal (*ACK*) which is provided by the data writing controller when the arrival time measurement is completed and the measured time is stored in a memory. Hence, when a photon is detected, the rising edge of the *SPAD* signal will set the data output of STOP_DFF, i.e. the *STOP* signal, to "1". The *STOP* signal remains at "1", regardless of the noise on the *SPAD* signal and the width of the *SPAD* pulse, until the measurement of the detected photon is completed and announced by the *ACK* signal sent by the data writing controller. At this point, the *STOP* signal will be reset to '0' and the signal controller will be ready to detect another rising edge of the *SPAD* signal.

b) Enabling the TDL RL

As seen before, for the measurement of the T_{fine}, the *STOP* signal propagates through the delay line (DL) and the DL output should be registered by the register line (RL) at the subsequent clock rising edge. To that end, the DFFs constituting the RL are clocked by the system clock and enabled by a signal, denoted as *EN_RL*, that switches to '0' after the clock rising edge that follows the *STOP* rising edge.

The second function of the signal controller is to generate the *EN_RL* by synchronizing the *STOP* signal generated by the STOP_DFF, to the system clock using a DFF, denoted as EN_DFF, and then inverting the output of this DFF to generate the EN_RL signal. EN_DFF is clocked by the system clock, its data input is connected to the *STOP* signal and reset by the *ACK* signal. The output of EN_DFF is inverted by an inverter and the output of this inverter is connected to the enable port of the RL DFFs. Thus, the *EN_RL* signal is reset to '0' after the first clock edge and deactivates the RL.

We compiled the system design and conducted experimental tests. The code density test results showed that the first bin (Bin_0) had a relatively large width. Furthermore, with multiple recompilations of the design, there was always a large width either for the first or the last bin. The origin of this problem is the delay (Δt) between the arrival time of the *STOP* signal at the beginning of the TDL (t_{TDL}) and at the EN_DFF (t_{EN_DFF}). This delay may vary from one compilation to another depending on the placement of these components and the routing paths between them. Two case scenarios for this delay are illustrated in Figure 3.18.

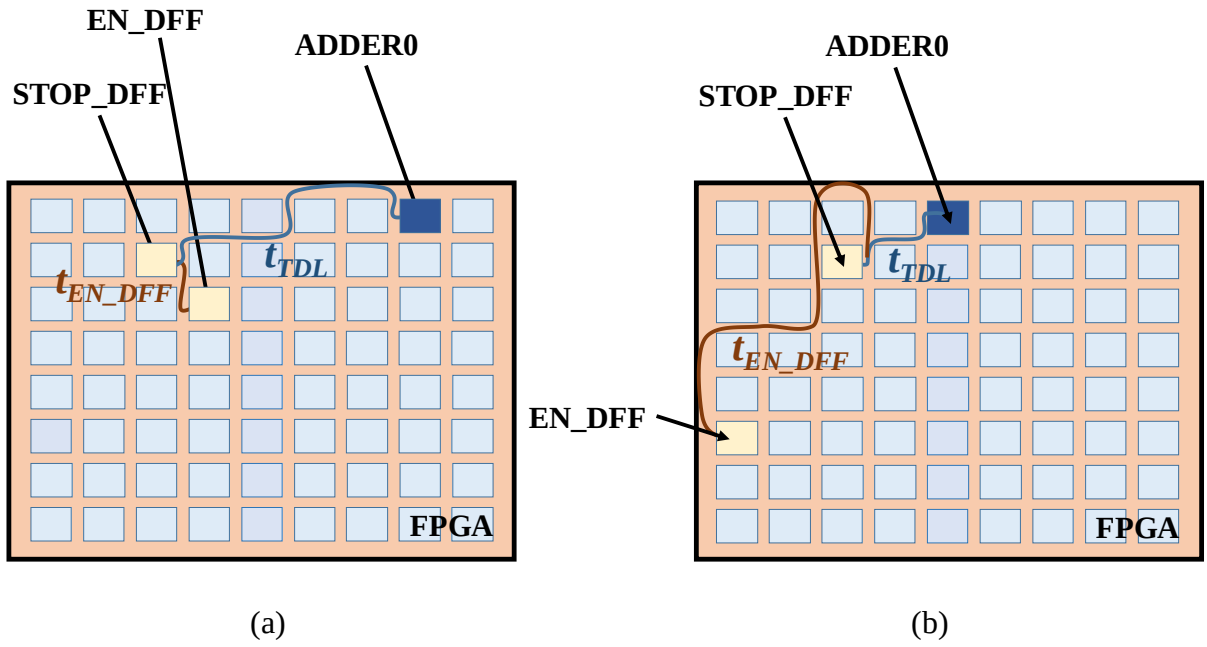


Figure 3.18 Examples of the delay (Δt) between the arrival time of the STOP signal at the EN_DFF and at the TDL for two different compilations: (a) $t_{TDL} > t_{EN\_RL}$ ($\Delta t > 0$), (b) $t_{TDL} < t_{EN\_RL}$ ($\Delta t < 0$).

In the ideal case, the delay Δt would be zero, so that the EN_DFF could sample the *STOP* signal at the first clock rising edge after the *STOP* signal reaches the TDL. However, in practice, this delay is unavoidable and can have either a positive or a negative value depending on the routing paths. There are two cases to consider:

Case 1, the *STOP* signal arrives at the TDL after it arrives at the EN_DFF ($\Delta t > 0$), as depicted in Figure 3.19-a:

In this case, if the *STOP* signal arrives at the TDL at the beginning of the clock cycle within the interval from 0 to Δt, it would have already reached the EN_DFF before the rising edge of clock of that cycle, as illustrated in Figure 3.19-b. consequently, the signal *EN_RL* would be reset to zero and disable the RL before the propagation of the *STOP* signal in the DL, resulting in ones at the output of all the RL DFFs. Thus, the encoder would produce an incorrect code of 255.

Therefore, with each event where the *STOP* signal arrives at the TDL within the range marked in red in Figure 2.19-b, the fine value would be 255. As a result, the width of this bin would be equal to Δ*t*.

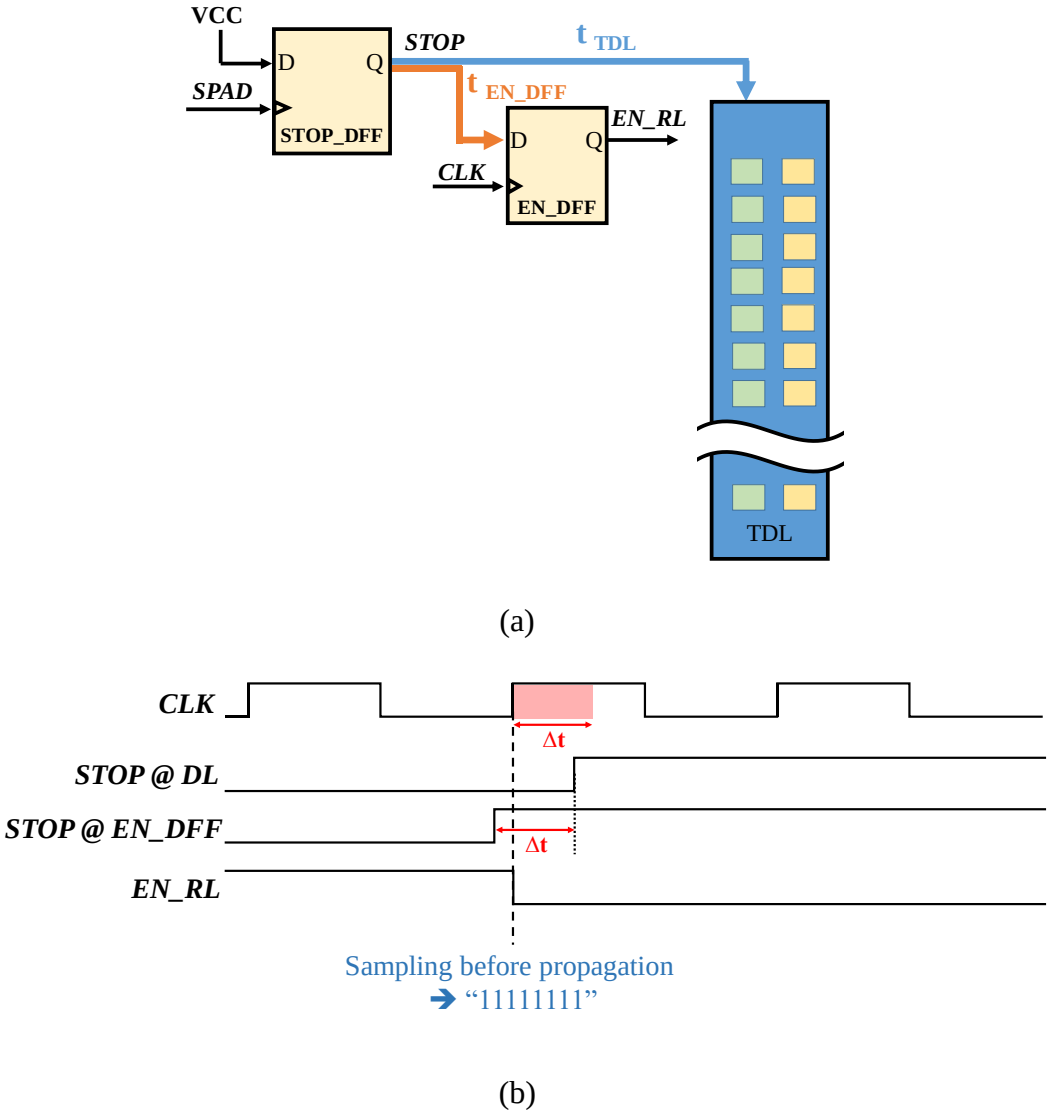


Figure 3.19 Case 1, the STOP signal arrives at EN_DFF before the TDL: (a) the STOP routing path to TDL is longer than its path to EN_DFF, (b) when the STOP signal arrives at the TDL in the red range, the DL would be sampled at the previous clock rising edge, resulting in an incorrect code of 255.

Case 2, the *STOP* signal arrives at the TDL before it arrives at the EN_DFF (Δt < 0), as shown in Figure 3.20-a:

In this case, if the *STOP* signal arrives at the TDL at the end of the clock cycle within the interval from T_{clk} - Δt to T_{clk} (marked in red in Figure 3.20-b), it would arrive at the EN_DFF after the next clock rising edge, and EN_DFF would sample the *STOP* signal at the second

following edge. Therefore, the RL output would be sampled after the arrival of the *STOP* signal at the TDL by more than one clock period, as depicted in Figure 3.20-b. At this point, the *STOP* signal would have propagated through all the DEs, resulting in zeros at their outputs. Hence, the encoder would produce an incorrect code of 0. As a result, the TDL would have a large width equal to Δt in the first bin.

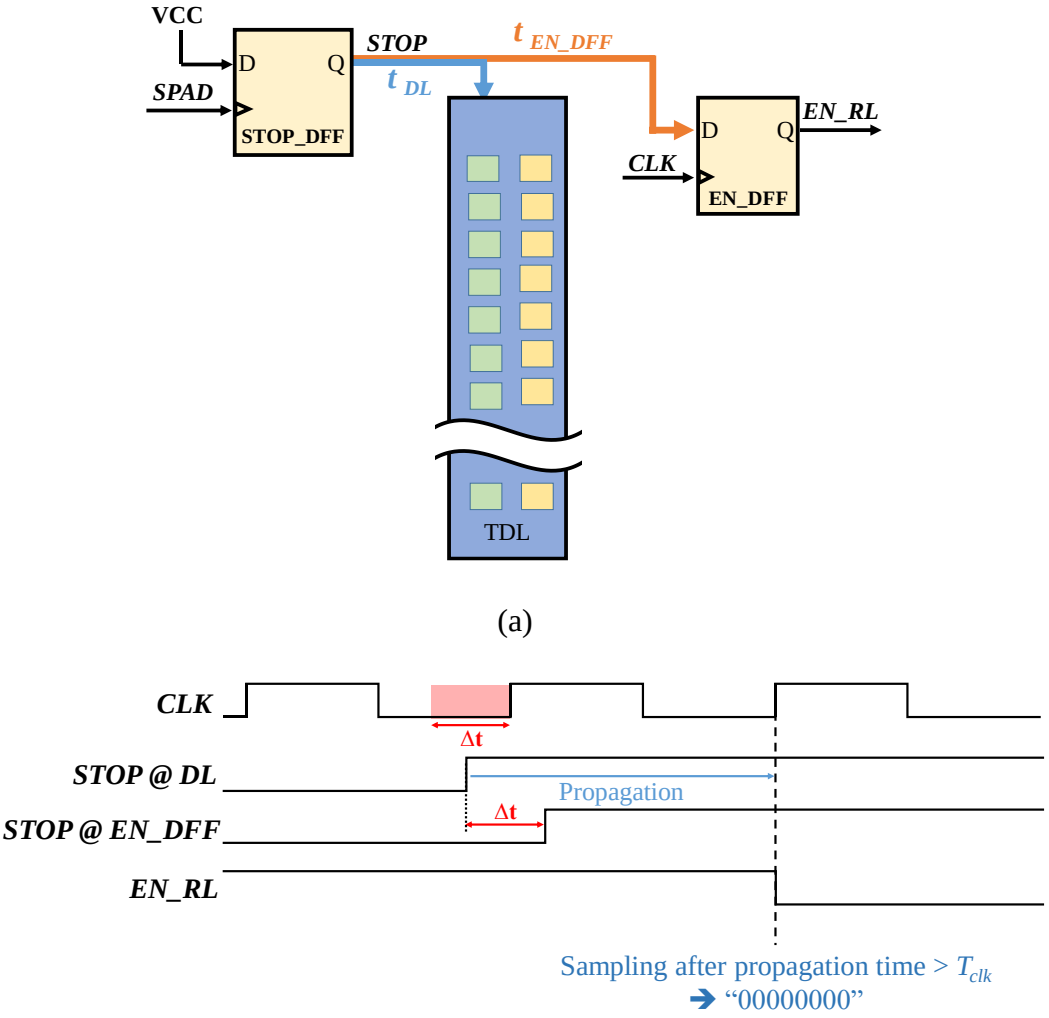


Figure 3.20 Case 2, the STOP signal arrives at the TDL before EN_DFF: (a) the STOP routing path to TDL is shorter than its path to EN_DFF, (b) when the STOP signal arrives at the TDL in the red range, the DL would be sampled at the second following clock rising edge resulting in an incorrect code of 0.

To avoid having a large width for the first/last bin, Δt should be as small as possible. However, controlling this delay is a crucial challenge in low-cost FPGAs, since manual routing is not

available in these devices. Furthermore, creating a fixed logic lock region for the EN_DFF and placing it close to the beginning of the TDL is not an effective solution, because the smallest size of a logic lock region is one LAB (1×1). Depending on the placement and routing of the EN_DFF and the inverter in this LAB, which may involve LUTs, Δt could be a few hundred picoseconds, resulting in a large width of the first or the last bin, which degrades the linearity of the TDC. However, without any constraints, Δt is unpredictable and varies with each new compilation, making the TDC implementation irreproducible.

To overcome this challenge, we integrated the EN_DFF in the TDL and placed it at its beginning, by using the DFF associated with the first adder of the TDL, as depicted in Figure 3.21. In fact, this DFF is not part of the RL as the output of the first adder is not sampled in this design due to the tuned-downsampling (the first DE consists of the first two adders with the DFF corresponding to the second one, see Figure 3.9 above). Therefore, this DFF could be used to serve as the EN_DFF, by linking the first adder to its DFF in the down-sampling VHDL code and using the data out of this DFF as the *EN_RL* signal. This way, the *STOP* signal reaches the first adder of the TDL and the EN_DFF almost at the same instant, with a negligible Δt that is equal to the delay from the first adder to its associated DFF. Moreover, when the rising edge of the *STOP* signal arrives at the first adder of the TDL, the sum-out of this adder becomes '0', as if the *STOP* signal were inverted before reaching the EN_DFF. Therefore, there is no need for an inverter after the EN_DFF.

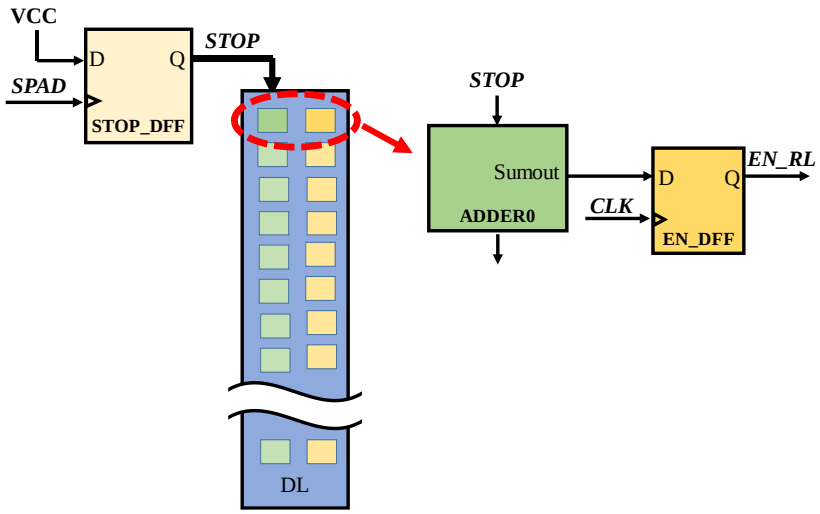


Figure 3.21 EN_DFF placement at the beginning of the TDL. The DFF associated with the first adder is used as the EN_DFF to minimize the delay.

Another benefit of this approach is that it guarantees almost the same behavior of the TDC system when the design is recompiled and enables the implementation of several TDCs in multichannel systems. Nevertheless, it is also recommended to perform an incremental compilation. This feature enables preserving the compilation outcomes of unmodified logic in the design [147].

c) Requesting Data Writing

As explained above, the encoder converts the thermometer code at the TDL output into a binary fine timestamp. However, the propagation time through the encoder may exceed one clock cycle. Hence, the system should wait for a certain delay after the photon detection before reading the accurate fine timestamp from the encoder output and writing the measurement data into a memory for subsequent processing.

The third function of the signal controller is to provide a request signal (*REQ*) to the data writing controller, which is the module in charge of storing the measurement data, after waiting for the encoder delay. This signal indicates that the fine timestamp is ready at the encoder output and can be stored in the memory.

The propagation time through the 8 stages of the encoder in our system was measured to be about 8 ns, which is longer than a clock period (T_{clk} = 5.26 ns). Therefore, the system must wait for at least two clock periods before the thermometer code is encoded to a fine timestamp. The signal controller block generates the *REQ* signal by delaying the *STOP* signal for two clock cycles using two DFFs operating as a shift register. These DFFs are clocked by the TDC clock and reset by the *ACK* signal. In practice, the EN_DFF can serve as the first DFF of the shift register, and another DFF, denoted as REQ_DFF, is added to complete the shift register, as shown in Figure 3. 17. When a photon is detected, the *STOP* signal will be shifted and after two clock cycles, the *REQ* signal at the output of REQ_DFF will be set to ‘1’. This signal triggers the data writing controller to read the fine and coarse timestamps and to write the measured time data into the memory. Once the data is stored, the *REQ* signal is reset to ‘0’ by the *ACK* signal provided by the data writing controller after completing the writing procedure.

3.1.4 Data Writing Controller

The data writing controller is a finite state machine that computes the measured time interval (the photon arrival time in TCSPC systems) from the coarse and fine values and writes the result into a FIFO memory. In its idle state, it waits for the *REQ* signal from the signal controller,

which indicates that a measurement has been performed. Upon receiving this signal, it combines the 8-bit coarse and fine values in a 16-bit word, where the coarse value occupies the most significant byte and the fine value occupies the least significant byte. It then proceeds to the next state and writes this word into the FIFO memory. To avoid overflow, it checks the FIFO status before writing the data word; if the FIFO is full, it waits until a word is read from the FIFO and a location is freed before providing the FIFO with the new word and a writing request pulse. Thereafter, it returns to the idle state and sends the acknowledge signal (*ACK*) to the signal controller to reset the *STOP*, *EN_RL*, and *REQ* signals and prepare the TDC for detecting a new photon. Therefore, when a photon is detected, the signal controller can't detect another photon until the arrival time of the current photon is already written to the FIFO. Thus, the maximum photon detection rate is limited by the memory write and read rates, among other factors, and the size of the FIFO memory should be carefully chosen to achieve the desired maximum detection rate. This issue is discussed in more detail in the chapter 4. In addition, other functions of the data writing controller related to the microfluidic droplet sorting application are discussed in Chapter 5.

3.1.5 Supplementary Blocks

In addition to the essential parts, we integrated some auxiliary components that perform additional functions such as the synchronous driving of excitation light sources, adding time-tag options, and generating uncorrelated events for the calibration process.

1- Synchronous Driving of Excitation Light Sources

As previously stated, our synchronous TCSPC system uses laser diodes, which are coupled with a pulse generator [50], as excitation light sources. The pulse generator requires a trigger signal to generate the light pulses. The synchronous trigger block provides the pulse generator with a periodic signal that is synchronized with the TDC clock. This block functions as a frequency divider. It takes the TDC clock as the reference signal and generates a clock-cycle-width pulse every N_L clock cycle, where N_L is the division factor. This factor is a configurable parameter that should be selected according to the desired repetition rate of the excitation light pulses, or the time window of the recording optical signal. For example, in LIDAR applications, the division factor N_L should be adjusted to a value that ensures a time window proportional to the maximum detected distance. The laser repetition period should exceed the time needed for the light to travel from the laser source to the detected object and back to the photon detector. For instance, if the TDC clock frequency is 200 MHz (T_{clk} = 5 ns) and the object distance is 100 m,

the time widow (T_{Laser}) should be at least 667 ns, which is the time required for the light to travel 200 m, and therefore N_L should be greater than 134.

2- Photon Time-Tag – Temporal Packetizing

Some TCSPC-based applications require not only recording the arrival time of the photon with respect to the laser pulse but also its absolute arrival time from the beginning of the experiment [14]. The former is referred to as "micro time", which corresponds to the time measured by the TDC, and the latter is denoted as "macro time". To measure the "macro time", we employ a macro time clock that drives a counter which is initialized to zero at the start of the experiment. When a photon is detected, the value of this counter, representing the "macro time", can be appended to the photon word. Moreover, the macro time clock can be used to accurately control the experiment duration. In our system, the macro time clock is derived from the TDC clock using a frequency divider. As for the light source trigger, the division factor of this signal is a configurable parameter that determines the resolution of the macro time measurement.

Another functionality of this block is the temporal packetizing of the detected photons. This functionality is useful for many applications that require an instantaneous estimation of the average light intensity detected during short time intervals. In this case, the macro time clock can be used as a temporal packet signal (*PACKET*), and the number of photons detected within a temporal packet, i.e. between two consecutive *PACKET* pulses, represents the light intensity detected during that packet.

In temporal packetizing applications, when a photon is detected, the data writing controller writes its arrival time "micro time" into the FIFO. Additionally, when a temporal packet pulse is generated, the data writing controller writes a packet word including the packet number, i.e. the current value of the packet counter, into the FIFO, as illustrated in Figure 3.22. To distinguish between the photon words and the packet words, a flag bit is added to the 16-bit data words, this bit is assigned to '1' in the packet words and to '0' in the photon arrival time data words.

3- Calibration Signal Generator

The TDC requires a calibration process to correct the nonlinearity. Calibrating the TDC involves performing a code density test which involves measuring the arrival time of uncorrelated events that have random delays with respect to the *START* reference signal.

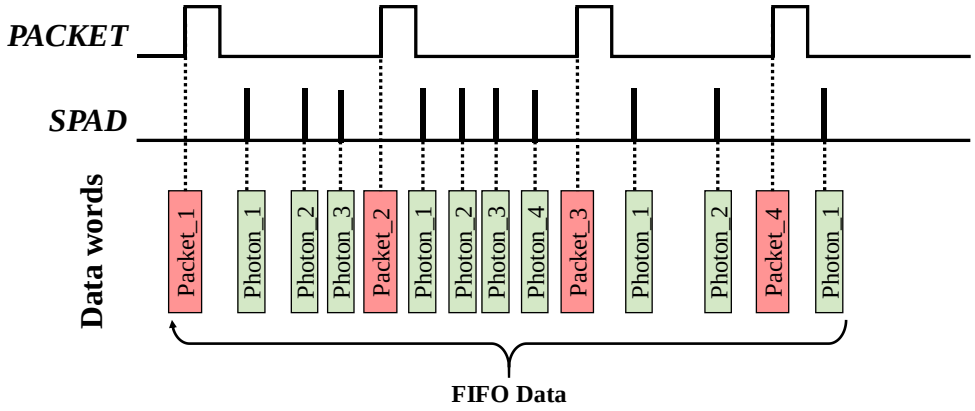


Figure 3.22 Temporal packetizing principle: at each packet pulse the data writing controller writes to the FIFO a packet word indicating the macro time of the photons detected during that packet period.

One method to generate uncorrelated events is to expose the SPAD to ambient light. In this case, the photons detected by the SPAD will have random arrival times relative to the *START* signal. However, this approach requires taking into consideration the pile-up effect of the SPAD. In fact, at high photon detection rates, the probability of detecting a photon that arrives directly after the START rising edge, i.e. at the beginning of the TDL, is higher than that of detecting a photon that arrives at the end of the TDL. Hence, the photon arrival time distribution is not uniform along the TDL. To overcome this problem, the code density test using the SPAD should be performed at a low photon detection rate, such that the pile-up effect can be neglected. For example, for a TDL with a total propagation time of 5 ns, a photon detection rate of 1 M photon/s would result in an average time interval of 1 µs between two successive photons, which is 200 times larger than the propagation time along the TDL. Consequently, the pile-up effect can be ignored at this rate. Furthermore, the code density test should integrate a large number of events which implies a relatively long calibration time at low photon detection rates.

An alternative method to conduct the code density test is to measure the arrival time of a periodic signal, termed the calibration signal, that is uncorrelated with the TDC clock with a slightly different frequency from the TDC clock or one of its harmonics. As a result, the phase difference between the TDC clock and the calibration events will gradually increase or decrease with a fine step that depends on the two frequencies. Hence, the arrival time of the calibration events covers the entire TDL delay with this fine step [148, 149].

To avoid any correlation between the TDC clock and the calibration signal, the calibration signal should be driven from a different reference clock than that of the TDC clock. It is

common to generate it by a phase-locked loop (PLL) driven by an external crystal oscillator [149]. Alternatively, a ring oscillator implemented in the FPGA can drive the PLL that generates the calibration signal.

3.2 Asynchronous TDC

As previously explained, synchronous TCSPC systems drive the light source in sync with the TDC system clock, enabling the simplification of the TDC structure into a coarse block and a single fine block. However, some applications require high-power external light sources that do not support external triggering, such as pulsed lasers, femto lasers, etc. One of our target applications, which is the stray light characterization, is an example of these applications. As will be presented in Chapter 5, this application employs an external pulsed laser that operates at predefined frequencies and cannot be triggered by the TCSPC system. In such applications, both the *START* and *STOP* signals are asynchronous to the system clock, and the measurement of the time interval (*TI*) between these two signals requires measuring their individual arrival times and subtracting these two times. To support this type of application, we upgraded our TCSPC system by adding a second time measurement channel for the *START* signal. Figure 3.23 shows the basic architecture of the designed asynchronous TDC system.

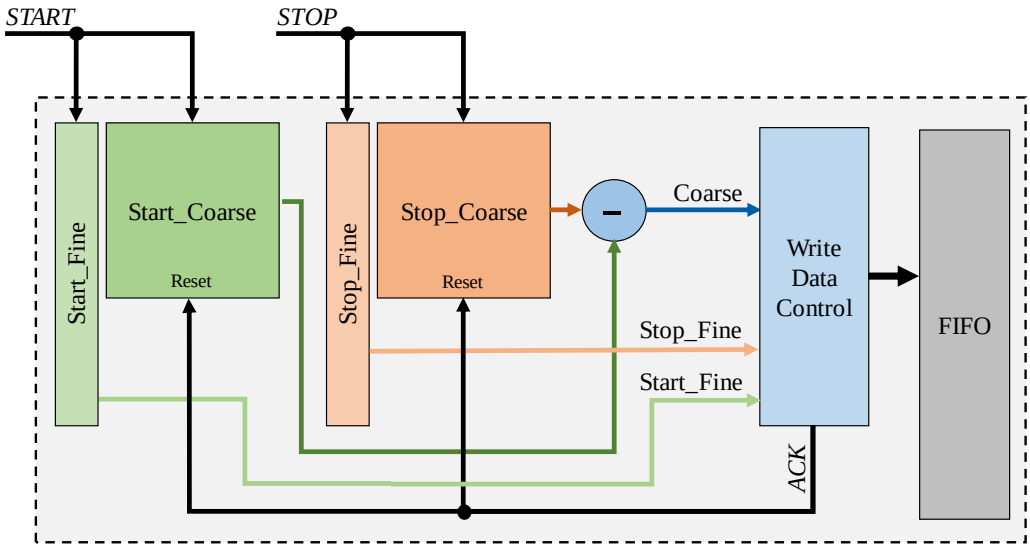


Figure 3.23 Asynchronous TDC basic architecture.

The Timing diagram in Figure 3.24 demonstrates the operating principle of the asynchronous TDC system. The arrival times of the *START* and *STOP* signals (T_{START}, T_{STOP}) are measured relative to a reference point, which corresponds to the *ACK* signal that the data writing controller sends after storing the precedent measurement data in the FIFO.

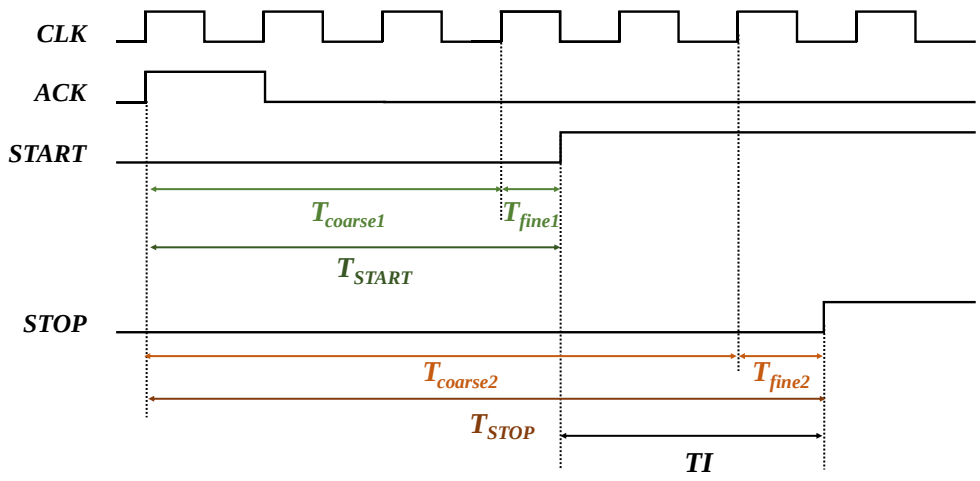


Figure 3.24 Operating principle of asynchronous TDC: two time measurement channels measure the arrival times of the START and STOP signals (T_{START}, T_{STOP}), and TI is calculated by the subtraction of these two times.

In the following sections, we present the modifications that we applied to our synchronous TDC system to implement the asynchronous system.

3.2.1 Asynchronous Coarse Block

In the synchronous TDC, the coarse block counts the clock cycles elapsed between the *START* and the *STOP* signals. This is achieved by resetting the coarse counter by the *START* signal and sampling its value at the *STOP* signal. However, this implementation suffers from a metastability problem due to the asynchronous sampling of the coarse counter value. This problem becomes more severe in asynchronous TDCs because it arises not only from the asynchronous sampling but also from the asynchronous resetting to zero of the coarse counter by the asynchronous *START* signal. Consequently, in asynchronous TDCs, the metastability problem can introduce an error of one or two LSBs in the coarse value, which corresponds to an error of one or two clock periods in the time measurement result.

As previously discussed, we addressed the metastability problem in the synchronous TDC by using two counters and a selection logic circuit to select the correct coarse value from three options (Register_R, Register_F, Register_F+1) based on the fine value. However, implementing this solution in the asynchronous TDC is more complicated because there are three scenarios for each of the reset and the sampling operations, resulting in nine different options for the final coarse value. From these options, the selection circuit should select the stable value according to the fine values of the *START* and *STOP* signals.

A simpler approach to address this problem is to use two instances of the coarse block that was used in the synchronous TDC for the separate measurements of the stable coarse values for the *START* and *STOP* signals. The final coarse value is then obtained by subtracting these two measured coarse values. However, this approach requires the simultaneous and synchronous resetting of four counters. Therefore, we used the *ACK* signal to reset the four counters and to serve as a reference point for the coarse measurement. In practice, we reset the counters that operate on the rising edge of clock directly by the *ACK* signal, and those operating on the falling edge by the *ACK* signal shifted by half a clock period. Figure 3.25 depicts the global architecture of the coarse block in our asynchronous system. For more details on the structure of the coarse sub-blocks, see Figure 3.14.

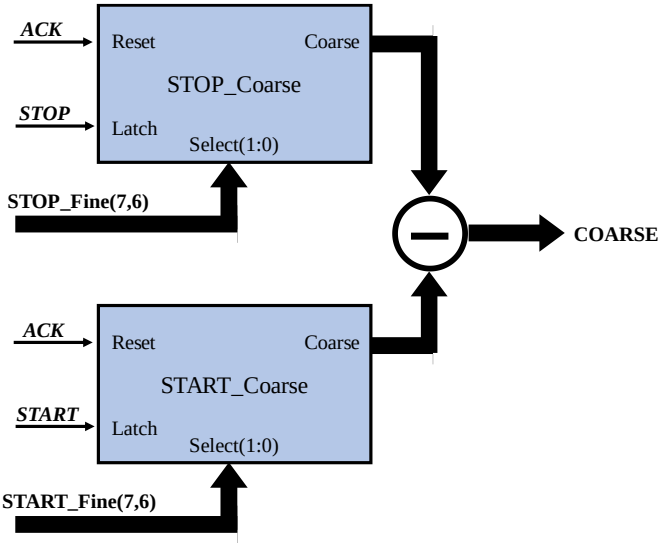


Figure 3.25 The global architecture of the Coarse block: it consists of two coarse sub-blocks for the START and STOP signals, the final coarse value is calculated by subtracting the two coarse values.

3.2.2 Asynchronous Fine Block

The various constraints that we applied in the design of our synchronous system, including assigning a fixed logic lock region for the TDL and placing the EN_DFF at the beginning of the TDL, ensure the reproducibility of the TDL while preserving the same performance. This enables the duplication of the TDL without having large variations in the characteristics of the two TDLs.

In the asynchronous TDC, we implemented a second instance of the fine block for the *START* fine measurement. This approach is applicable after addressing the large-width bin problem discussed in Section 3.1.3-b. In fact, many experiments demonstrated that without controlling

the placement of the EN_DFF, the dual-TDL asynchronous TDC would have many dead bins and large-width bins. Furthermore, creating logic lock regions for the start and stop TDLs allows placing these two TDLs far from each other, by configuring the origins of their two regions, which reduces the noise due to the interaction between the asynchronous *START* and *STOP* signals that propagate through these TDLs.

3.2.3 Asynchronous Signal Controller

In the TCSPC technique, the photon detection rate should be low enough that the probability of detecting a photon during the excitation period is much less than one. Thus, there is no need to provide the ability to detect more than one photon in the same excitation light period. Hence, we designed our TCSPC system to operate in the reverse start-stop mode, where the SPAD signal serves as the *START* signal and the light emission reference as the *STOP* signal [150]. In this mode, the measurement is launched by the photon detection, rather than by each light reference pulse, which reduces the power consumption. With this consideration, the TDC measures the time interval (*TI*) between the rising edge of the SPAD signal and the subsequent rising edge of the light reference signal, and the photon time of flight (TOF) is obtained by subtracting the measured time (*TI*) from the light emission period (T_L), as illustrated by Equation 3.3, Equation 3.4, and Figure 3.26.

$$TI = T_{STOP} - T_{START} \tag{3.3}$$

where T_{START} is the arrival time of the rising edge of the SPAD signal and T_{STOP} is the arrival time of the rising edge of the following light reference pulse.

$$TOF = T_L - TI \tag{3.4}$$

Figure 3.27 depicts the architecture of the asynchronous signal controller. In addition to the START_DFF used to detect the *SPAD* signal, we added another DFF, denoted as STOP_DFF, for the detection of the asynchronous light reference signal. The data input of this DFF is connected to the output of the START_DFF, which is the *START* signal, and the clock port is connected to the light reference signal. Hence, a *STOP* pulse is only generated when a *SPAD* pulse has been already detected by the START_DFF. The two DFFs are reset by the *ACK* signal sent by the writing data controller after the completion of the time measurement.

This block also provides the RLs of the TDLs with the enable signals. Two EN_DFFs are used to synchronize the *START* and *STOP* signals to generate the EN_RL signals for the start and stop fine blocks, as illustrated for synchronous TDCs.

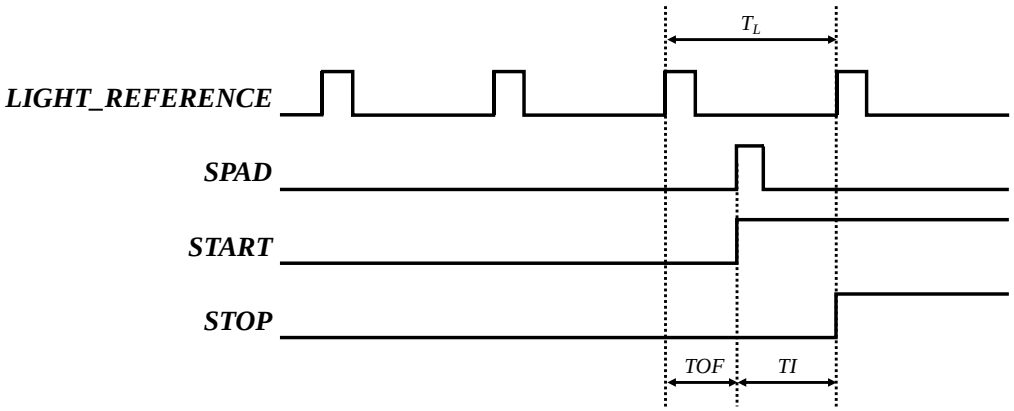


Figure 3.26 Revers start-stop mode: the SPAD signal represents the START signal and the excitation light reference represents the STOP signal, TOF is calculated by subtracting the measured time interval from the excitation light period.

The measured fine and coarse times of the *START* and *STOP* signals are available at the output of their respective fine and coarse blocks after the arrival of the *STOP* signal by at least two clock periods, which is the propagation delay through the STOP fine encoder. The signal controller shifts the *STOP* signal for two clock periods using a shift register composed of EN_DFF and REQ_DFF, as shown in Figure 3.27. The output of the REQ_DFF is the *REQ* signal that activates the data writing controller to calculate the *TOF* and to write into the FIFO memory.

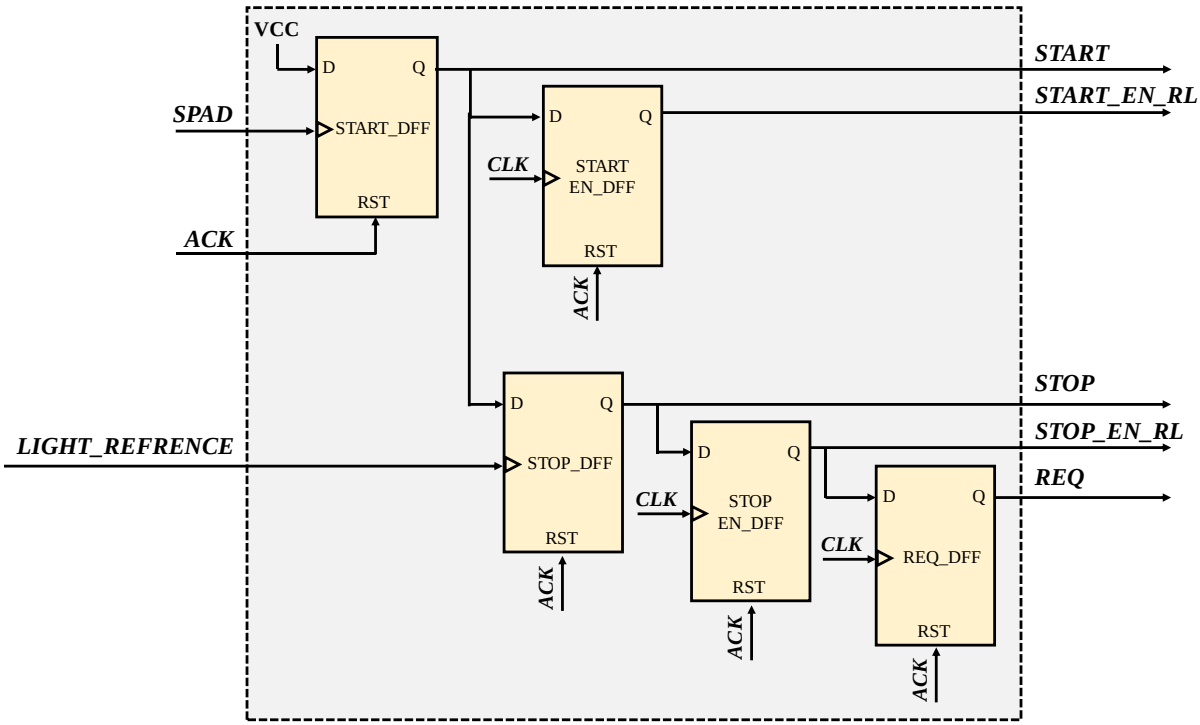


Figure 3.27 Asynchronous signal controller architecture

3.2.4 Asynchronous Data Writing Controller

As in the synchronous system, the signal controller waits for the *REQ* signal to write the time measurement result into the FIFO memory. It calculates the final coarse value by subtracting the STOP coarse value from the START coarse value. It then combines this coarse value with the fine values of the *START* and *STOP* signals and additional time tag bits, forming a 32-bit photon data word in the format shown in Figure 3.28. Finally, it writes this word to the FIFO memory. Using this data word, the final time of flight calculation will be later performed by the hard processor system, as described in the next chapter. After writing the data word to the FIFO, the data writing controller sends the acknowledgement signal (*ACK*) to the signal controller, which resets the *START*, *STOP*, *EN_RLs*, and *REQ* signals to prepare the system for the detection of a new photon.

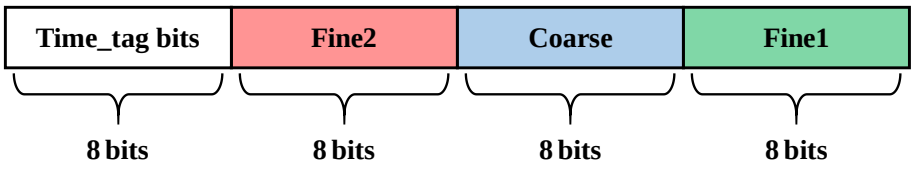


Figure 3.28 Photon word: a 32-bit word that combines the coarse value, the two fine values, and additional time-tag bits.

3.2.5 Asynchronous Driving of Excitation Light Sources

In order to support excitation light sources that require a trigger signal in our asynchronous system, we implemented an asynchronous trigger block that generates periodic trigger pulses that are asynchronous with the system clock to drive the light source.

This block employs a separate PLL that generates a clock signal, denoted as CLK_{ex}, which is uncorrelated with the TDC clock and has a frequency slightly different from that of the TDC clock. In addition, a frequency divider is added to generate the excitation light trigger signal (L_{ex}) with a configurable frequency from the CLK_{ex} by adjusting the division factor.

3.3 Conclusion

This chapter presented a synchronous coarse-fine TDC for TCSPC systems, implemented in a law-cost Cyclone V SoC-FPGA. The fine measurement employed the TDL structure, using the high-speed carry-chain paths available in this FPGA device to achieve high temporal resolution.

However, due to the lack of manual placement and routing support in such low-cost FPGAs, several timing issues were encountered in the design and implementation of the TDC. These issues were addressed by applying various constraints to ensure the correct functionality, reproducibility, and reliability of the TDC system. In addition, a tuned-downsampling technique based on the propagation time characteristics along the carry-chain was proposed, which enhanced the linearity of the TDC with a temporal resolution acceptable for the target applications. Furthermore, a robust coarse structure was proposed to avoid the metastability problem caused by the asynchronous sampling of the counter value, by using two counters operating in phase with each other and a logic circuit to select the stable coarse value.

The design was extended to support asynchronous excitation light sources, by adding another time measurement channel for the excitation light reference signal.

The synchronous and asynchronous TDC systems were successfully implemented on the Cyclone V SoC-FPGA, achieving an average temporal resolution of about 20 ps and a maximum detection rate of about 66.7 M photon/s for the synchronous system. On the other hand, the detection rate of the asynchronous system is proportional to the excitation light frequency.

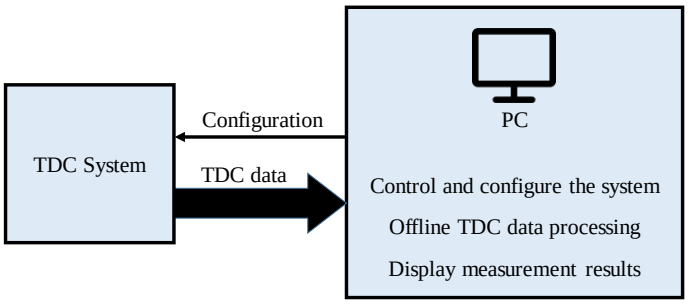
Chapter 4: Data Processing

In TCSPC systems, time measurement data of the TDC system are transferred into a computing system for processing. The data processing tasks vary according to the application. Some applications require a simple display of the recorded optical signal in the form of a histogram, while others need to extract more complex parameters such as fluorescence lifetimes, centers of gravity, relative distances, etc. However, some processing tasks are common to all the applications, such as the TDC calibration and the correction of various imperfections of the photon detector and optical and electronic components.

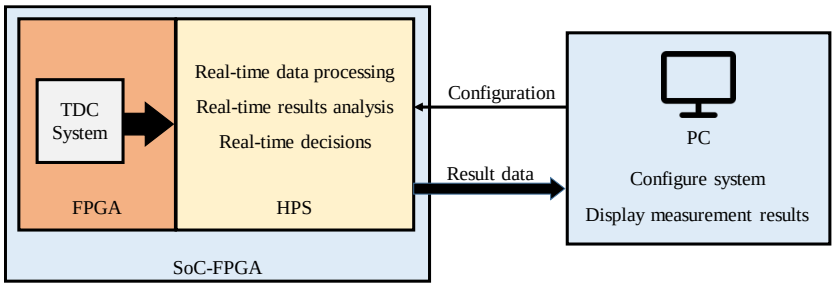
In conventional TCSPC systems, data are sent to a control PC for processing. In this approach, the PC is responsible for configuring the TCSPC system, performing offline data processing, and displaying the final results, as illustrated in Figure 4.1-a. However, our system employs the hard processor system (HPS) integrated in the SoC-FPGA board to perform all the data processing locally on board, and the communication with a control PC is limited to the system configuration and the results monitoring, as shown in Figure 4.1-b. Moreover, thanks to the high-speed interconnection buses available between the HPS and the FPGA fabric, time measurement data can be transferred to the HPS at much higher rates than when data are transferred into a PC. These high data transfer rates enable the system to operate in real time.

We implemented a hardware-software co-design. The hardware part, which is a Qsys design developed using the Quartus Platform Designer tool, configures the data buses used for the communications between the FPGA and the HPS parts. Whereas the software part is a bare-metal C program running on the ARM processor of the HPS part. This application controls the data transfer, performs online data processing, and handles the communications with the control PC. We used the ARM DS-5 Development Studio platform for the building and debugging of this application.

This chapter discusses the data transfer mechanism and the different data processing tasks required by our target applications.



(a)



(b)

Figure 4.1 Data processing modes: (a) data processing is performed by the control PC, (b) data processing is executed locally by the HPS integrated in the SoC-FPGA board and the results are transmitted to the PC.

4.1 Data Transfer

The size of time measurement data varies according to the operating mode of the TCSPC system. In this context, two principal operating modes can be distinguished:

a) Direct Histogram Mode

In this mode, the TDC system directly constructs a histogram that represents the recorded optical signal in an on-chip memory. With each detected photon, the content of the memory location that corresponds to the arrival time of this photon is incremented by one. Therefore, the data size is not related to the number of detected photons, but only determined by the total number of bins and the memory data width. In other words, the data size in this mode is constant regardless of the measurement acquisition time.

b) Individual Photon Arrival Time or Time-Tag Mode

In this mode, the TDC system writes the time measurement data, which consists of the photon and packet words, into a FIFO memory. Hence, the data size is proportional to the number of detected photons, which depends on the photon rate and the acquisition time. In general, the data size in this mode is larger than the first one.

In order to support a wider range of applications, we designed our TCSPC system to operate in the second mode. As described in the preceding chapter, the system writes the arrival time data of the photons along with the time-tag bits and the packet words into the FIFO formatted in 32-bit, i.e. each photon or packet word occupies four bytes. This leads to a high data writing rate. For instance, a photon detection rate of 10 M photon/s would result in a data writing rate of more than 40 MB/s. Moreover, if a higher dynamic range of the TDC is needed, the bit-width of the coarse counter should be increased and consequently the bit-width of the data words, which would further increase the data writing rate. For example, at a photon detection rate of 50 M photon/s with a data word width of 64 bits, the data writing rate would exceed 400 MB/s.

If data are processed by a control PC, the size of the measurement data file generated for offline processing may exceed one gigabyte, even for a short experiment duration. In addition, to ensure the TCSPC system operation at the maximum achievable photon detection rate, the time measurement data written to the FIFO should be read by the computing system at a reading rate close to the writing rate. However, achieving such a high transfer rate from the FPGA to the control PC is challenging. Therefore, we utilized the high-speed interconnection resources between the FPGA and the HPS portions, available in the Cyclone V SoC-FPGA, for transferring the TDC data to the SDRAM of the HPS where all the data processing will be performed. This approach achieves high data transfer rates that allow the system to function in real time, which is an essential requirement for some applications, such as the real-time microfluidic droplet sorting explained later in Chapter 5.

4.1.1 Cyclone V SoC-FPGA Architecture – HPS

Before presenting the proposed data transfer mechanism, we describe the basic architecture of the Cyclone V SoC-FPGA that we used for the implementation of our system. This SoC-FPGA integrates a hard processor system (HPS) with the FPGA fabric and offers high-speed bridges and interfaces for the communication between these two parts [151]. Figure 4.2 depicts a simplified diagram of the Cyclone V architecture. The HPS part includes a microprocessor unit

(MPU) with a dual-core ARM Cortex-A9 processor, flash memory controller, SDRAM controller, on-chip memory, direct memory access (DMA) controller as well as many peripherals for communication and support functions such as timers, UART, SPI, I2C, USB 2.0 controllers, etc. The MPU system consists of:

- Two ARM Cortex-A9 cores with the following characteristics: 32-kB instruction and data level 1 (L1) caches and a memory management unit (MMU).
- An accelerator coherency port (ACP) that allows peripheral masters to access coherent memory through an ACP ID mapper.
- A snoop control unit (SCU) that maintains the data coherency between the L1 caches of the two processors, and manages the processors' access to the L2 cache as well as the accesses from the ACP.

The HPS also provides a level 3 (L3) interconnect which is a partially-connected crossbar. L3 includes three switches that enable HPS and FPGA master and slave peripherals to access different HPS components.

Furthermore, Cyclone V SoC-FPGA provides high-speed buses and interfaces for the communication between the HPS and FPGA parts. These datapaths are:

1- HPS-to-FPGA bridge: a high-performance bus that enables masters in the HPS logic to access slaves in the FPGA fabric. This bridge is connected to the L3 main switch and supports configurable data width that can be set to 32-bit, 64-bit, or 128-bit.
2- Lightweight HPS-to-FPGA bridge: a lower-performance fixed-width (32-bit) bus that can be used by masters in the HPS to access slaves in the FPGA. It is mainly used to access control and status registers of FPGA peripherals. This bridge is controlled by the L3 slave peripheral switch.
3- FPGA-to-HPS bridge: a high-performance bus, connected to the L3 main switch, through which masters in the FPGA fabric can communicate with slaves in the HPS. This bridge also enables cache-coherent access to the SDRAM subsystem through the ACP of the MPU subsystem. Furthermore, the data width of this bridge is configurable to 32-bit, 64-bit, or 128-bit.
4- FPGA-to-SDRAM interface: a highly-configurable interface that enables the FPGA to directly, non-coherently access the SDRAM through the HPS SDRAM controller, without passing by the L3 interconnect and the ACP of the MPU subsystem.

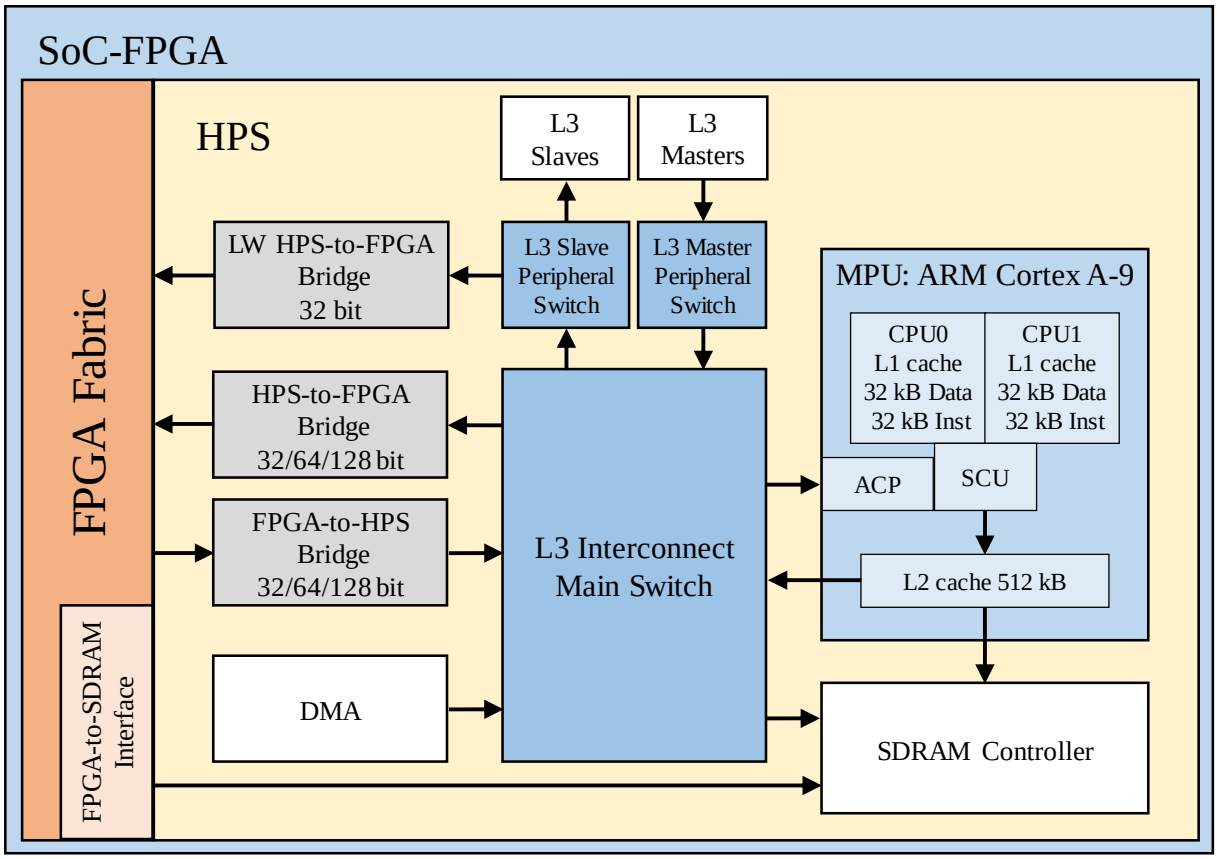


Figure 4.2 Schematic representation of the main components and interfaces of the Cyclone V SoC-FPGA architecture

4.1.2 Implemented Data Transfer Mechanism

The direct method to retrieve data from the TDC FIFO implemented in the FPGA by the processor is to access these data via a parallel I/O (PIO) or one of the dedicated bridges between the FPGA and the HPS. However, this method is suboptimal and provides transfer rates typically limited to a few tens of MB/s [152]. Given a targeted photon rate of 10 million photons per second, the data writing rate in the TDC FIFO would be about 40 MB/s, which is comparable to the data transfer rate using this method, Consequently, almost 100% of the CPU usage would be dedicated only for the data transfer, and no CPU resources would be available for data processing.

However, significantly higher data transfer rates can be attained by transferring the FIFO data into the SDRAM of the HPS in direct memory access (DMA) through the dedicated bridges and interfaces. Furthermore, this approach offloads the processor from the data transfer operations and preserves its resources for the data processing.

For the selection of the bridge, several studies compare the performance of the interconnection paths available in Cyclone V in terms of transfer rate [152-154]. These studies show that the FPGA-to-SDRAM interface provides the highest data transfer rate when no cache-coherent access is required. In our system, data should be transferred from the TDC's FIFO into the SDRAM, where the processor accesses these data to perform the required data processing, with no resource implemented in the FPGA needing to re-access the data. Hence, we chose to use the aforementioned interface for the data transfer. This interface was configured in a Qsys design implemented in the FPGA, providing direct access to the SDRAM controller of the HPS without passing by the L3 interconnect and the ACP of the MPU subsystem. In addition, we integrated in this design a DMA engine, namely the Modular Scatter-Gather DMA (mSGDMA) intellectual property (IP) core, which is freely provided by Altera. With this mechanism, we achieved a maximum data transfer rate of about 1.7 GB/s at FPGA operating frequency of 50 MHz and a DMA data size of 500 MB. However, the transfer rate varies according to the DMA data size and the FPGA operating frequency, i.e. the data writing rate into the FIFO. Figure 4.3 illustrates the adopted data transfer path from the TDC system to the SDRAM and presents the components involved in the data transfer.

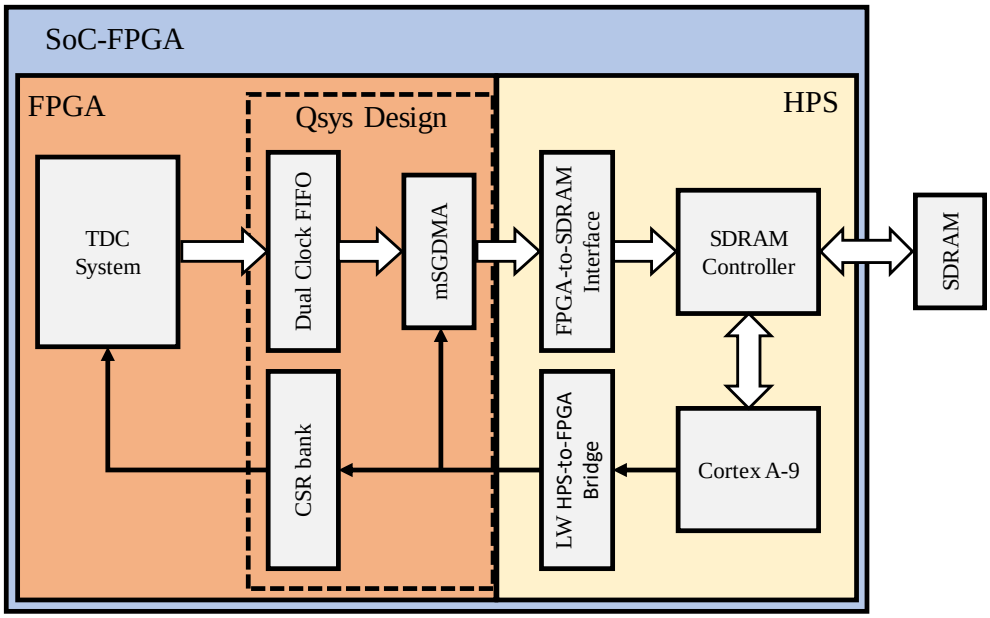


Figure 4.3 Data transfer mechanism: a Qsys design integrates an mSGDMA IP for the data transfer directly into the SDRAM through the FPGA-to-SDRAM interface.

1- Hardware - Qsys Design

We built the Qsys design using the Quartus platform designer tool. This design comprises the various components required for the implementation and operation of the data transfer channel.

Additionally, it includes several parallel I/O (PIO) ports that function as control and status registers, enabling the processor to configure the TDC system and the other FPGA. It also handles the clock crossing between different clock domains in the FPGA and HPS as well as the differences in data widths. The Qsys components involved in the data transfer process are:

1) Arria V/Cyclone V Hard Processor System IP

This IP core allows other components implemented in FPGA to interface with the hard processor logic and peripherals. It enables configuring the different interconnections between FPGA and HPS mentioned above. For our system, these interconnections are configured as follows:

- The lightweight HPS-to-FPGA bridge is enabled and its width is set to 32-bit. Via this bridge, the processor controls all the other components. For example, it configures the mSGDMA and triggers it to perform the data transactions, it also configures the blocks implemented in the FPGA via the PIOs.
- An FPGA-to-SDRAM interface is added with the type Avalon-MM Write-Only since it will be only used for the data transfer from FPGA to the SDRAM. The data width of this interface is set to 32-bit, i.e. the width of the TDC data words.
- HPS-to-FPGA bridge is enabled with a width of 32-bit. This bridge will be used to send and receive data to/from a USB interface implemented on the FPGA.

2) Modular Scatter-Gather DMA (mSGDMA)

This mSGDMA IP is a direct memory access engine that enables high-speed data transfer between different memory spaces or between a data stream and a memory space. It operates by executing preloaded instructions, known as descriptors, that specify the source and destination addresses and the data size, i.e. the number of bytes to be transferred (DMA_size). The processor only needs to provide the mSGDMA with the descriptors and initialize the data transfer, while it remains free to perform other tasks concurrently. The processor does not directly participate in the data-moving operation [155].

The mSGDMA IP uses two Avalon interface standards for the connection with the data source and destination, which are the Avalon Memory Mapped (Avalon-MM) and the Avalon Stream (Avalon-ST) interfaces. These standards enable connecting intel FPGA components and support high-speed streaming, writing, and reading data. In our design, the data source is the TDC FIFO, which is an Avalon Streaming Source, whereas the destination

is the region of the SDRAM where data will be written via the Avalon-MM interface. Hence, the mSGDMA is configured to function in the "Streaming to Memory-Mapped" mode.

The processor configures, controls, and triggers the mSGDMA via the lightweight HPS-to-FPGA bridge. Through this bridge, the processor provides the mSGDMA with the descriptors and accesses its CSRs.

3) Dual Clock FIFO

This FIFO is the buffer into which the TDC's data are written before being transferred to the SDRAM by the mSGDMA via the FPGA-to-SDRAM interface. We used a dual clock FIFO because the writing data controller writes data at the TDC clock frequency, whereas the mSGDMA reads these data at the HPS system clock frequency. Consistent with the data width of the FPGA-to-SDRAM interface and the mSGDMA, the data width of the FIFO is set to 32-bit. To avoid losing data, the FIFO data depth should be larger than a value proportional to the DMA data size and the data writing rate, i.e. the detected photon rate, as discussed later in this chapter.

2- Software – Data Transfer Algorithm

We developed a C program that runs on the HPS to control the data transfer and perform the different data processing. We opted for a bare-metal application, as it offers better performance than an operating system-based application. Bare-metal applications run at the native speed of the processor without any context switching overhead. Moreover, these applications can access the physical addresses of the HPS peripherals directly without a virtual memory system [156]. We used the Intel SoC FPGA Hardware Library (HWLIB) in the development of the bare-metal application. This library provides direct references to the HPS physical addresses, registers, and peripherals, and it deploys many macros, APIs, and drivers for more complex functionality [156, 157]. Furthermore, we configured the SoC-FPGA to boot from the microSD card on which we stored the FPGA configuration file and the C program execution file.

We developed the data transfer software based on a reference design example provided by Intel [158]. The data transfer algorithm operates as follows: first, it allocates a data buffer in the SDRAM with the same size as the data to be transferred (DMA_size). Next, it declares and configures a descriptor by assigning the destination address, which is the pointer to the data buffer, and the DMA data size (DMA_size). Then, it loads the mSGDMA with this descriptor

using the "write_extended_descriptor" API provided by Intel. Finally, the processor sets the Go bit in the mSGDMA control register to initiate the DMA transaction. At this point, the mSGDMA starts the data movement operations from the TDC's FIFO stream to the data buffer in the SDRAM via the FPGA-to-SDRAM interface, while the processor remains free to perform other tasks.

4.1.3 DMA-Based Data Transfer Problems

As mentioned above, transferring data using the mSGDMA requires the size of the data to be transferred to be known in advance, in order to configure the descriptor accordingly. Nevertheless, the size of time measurement data generated by the TDC is variable and unpredictable, as our TCSPC system operates in the time-tag mode and writes a data word for each detected photon. Since the photon rate is unknown, unstable, and may vary during the experiment, it is impossible to estimate the size of the experiment data to configure the mSGDMA with this parameter. A more critical issue occurs if the experiment time is long enough to produce data that exceeds the available memory capacity. Moreover, when real-time measurements are required, data should be processed continuously. In this case, transferring data in blocks of predefined size is complicated to implement, as many challenges arise regarding the size of data blocks and the latency. These challenges are addressed in three function modes of the TCSPC system as follows:

1- Offline Mode or Single Measurement Mode:

In this mode, the TCSPC system performs a single measurement for a predetermined duration, called the experiment time (*T_exp*), specified by the user. The system records a 32-bit word for each detected photon, thus the experiment data size (*D_exp*) is proportional to *T_exp* and to the detected photon rate during the measurement (*PhR_exp*). This size can be estimated by Equation 4.1

$$D\_exp = T\_exp \times PhR\_exp \tag{4.1}$$

- **DMA Data Size Issue:**

In reality, *PhR_exp* is unknown, unpredictable, and variable. Therefore, we cannot estimate D_exp to configure the mSGDMA with an appropriate value of *DMA_size* that corresponds to the experiment data size *D_exp*. At first glance, it seems that the simplest solution to this issue is to consider the maximum photon rate attainable by the system (*PhR_max*), and configure the

mSGDMA with the maximum data size (*D_max*) that can be generated during the experiment, as shown in Equation 4.2.

$$\begin{aligned} D\_max &= T\_exp \times PhR\_max\ (data\ word) \\ &= 4 \times\ T\_exp \times PhR\_max\ \ (byte) \end{aligned} \tag{4.2}$$

where the multiplication by four is for the conversion from data word to byte since the photon word is formatted in 32-bit.

Thus, the processor declares a memory space with the size (*DMA_size* = *D_max*) and allocates it in the SDRAM. When the user starts a measurement, the processor clears the FIFO to discard the old data, then it triggers the mSGDMA to start transferring the measurement data into the SDRAM memory space. Subsequently, the processor waits for the mSGDMA to finish the data transfer, i.e. the transfer of *D_max* data word. Thereafter, it starts processing these data and calculates the measurement final value, as illustrated in Figure 4.4. However, to generate *D_max* data words, the system should detect *D_max* photons, and the time required for that (*T_acquisition*) is proportional to the actual photon rate *PhR_exp*, as shown in Equation 4.3.

$$\begin{aligned} T\_acquisition\ &= \frac{D\_max}{PhR\_exp} \\ &= \frac{T\_exp\ \times\ PhR\_max}{PhR\_exp} \end{aligned} \tag{4.3}$$

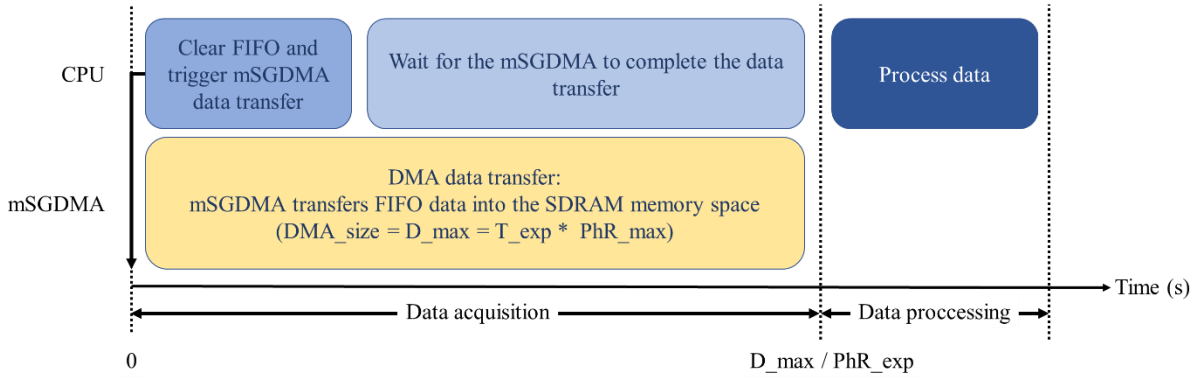


Figure 4.4 Offline mode: considering the maximum detectable photon rate, the processor should wait for the mSGDMA to transfer D_max data word before processing data.

Furthermore, this solution entails a more severe problem, as demonstrated in the following example:

Assuming that the maximum detectable photon rate is *PhR_max* = 10 M photon/s and that the user launches an experiment of one-second duration (*T_exp* = 1 s), then according to

Equation 4.2, the maximum data size would be 10 M photon words (*D_max* = 10 M word = 40 MB). Following this solution, the mSGDMA should be configured with a data size of *DMA_size* = *D_max* = 40 MB. However, if the actual photon rate during the experiment were *PhR_exp* = 1 k photon/s, the user would have to wait for *T_acquisition* = 10000 seconds, which is the time required to detect 10 million photons in order to generate the 40 MB of data that would be transferred by the mSGDMA. Hence, this approach leads to a considerable delay before obtaining the measurement outcome. Moreover, only the first 4 kB of data would be relevant, as they correspond to the effective data size produced during the first second (*T_exp*), as shown in Figure 4.5-a. Another drawback of this approach is that the TDC system would continue measuring detected photons for the next 9999 seconds, which implies a high power consumption.

- **Proposed Solution:**

To address this issue, we modified the TDC system to write the measurement data to the FIFO at a constant and configurable rate (*Fix_rate*), slightly higher than the maximum possible photon rate (*PhR_max*), regardless of the actual photon rate during the measurement. We modified the data writing controller in the TDC system to write data at a fixed and configurable rate, determined by a frequency divider that generates the FIFO write request signal. As a result, the data writing controller continuously writes data words at this rate. When a photon/packet pulse is detected, it writes a photon/packet word to the FIFO and sets the last bit in this word to one. We refer to this bit as the data flag bit. Otherwise, it writes a zero word "0x0000" if no photon/packet pulse is detected. Later, during data processing, the algorithm discards these zero words and processes only the photon and packet words by checking the data bit flag.

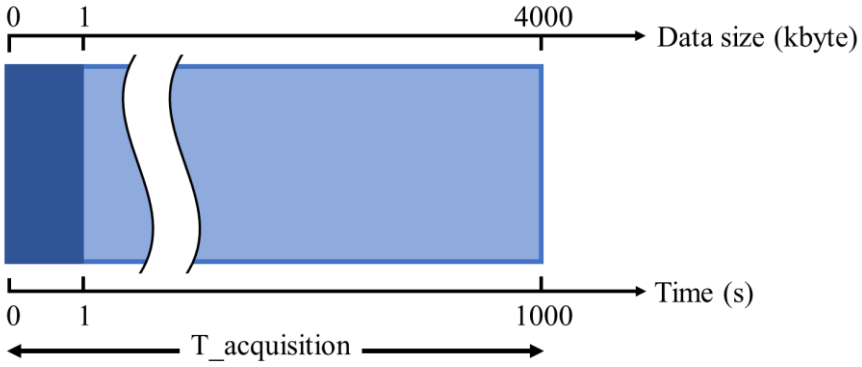
With this solution, the size of the experiment data (*D_exp*) can be estimated by Equation 4.4, since data are written at the constant rate *Fix_rate*. The mSGDMA can therefore be configured with a data size that corresponds exactly to the amount of data generated during the experiment.

$$D\_exp = T\_exp \times Fix\_rate \tag{4.4}$$

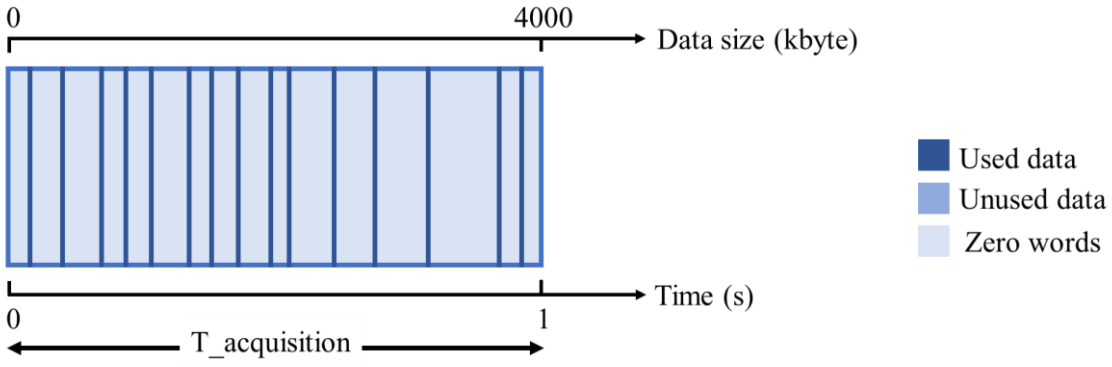
Hence, the time required to acquire the experiment data would be approximately equal to the experiment time (*T_acquisition* ≈ *T_exp*). Moreover, the transferred data would include the photon words corresponding to the photons detected during *T_exp*, i.e. these data would include only the relevant photon and packet words interspersed with zero words, and the TDC system would not measure the arrival time of irrelevant photons. The generated data would be directly transferred into the SDRAM by the mSGDMA to be ready for further processing.

Returning to the example discussed above, if the constant data writing rate (*Fix_rate*) is configured to be 10 M word/s, for *T_exp* = 1 s, the mSGDMA should be configured with *DMA_size = D_exp = 40 MB*. Regardless of the actual detected photon rate, the experiment data would be generated and transferred within about one second, and the user would receive the experiment outcome shortly after *T_exp* with a delay that depends on the data processing time. As shown in Figure 4.5-b, the transferred data includes only relevant photon and packet words interspersed with zero words.

Comparing the two approaches, it is clear that the second one is better in terms of time and power consumption since the acquisition time *T_aquisition* does not exceed *T_exp*, and no irrelevant photons would be detected and measured.



(a)



(b)

Figure 4.5 DMA data size in offline mode: (a) the maximum detectable photon rate approach, (b) the constant data writing rate approach.

2- Online Mode or Real-Time Mode:

In this mode, the system processes the data generated by the TDC in real time, and the measurement duration is not predetermined as in the Offline mode. The user launches an experiment, the TDC system starts measuring the arrival time of the detected photons, and the data generated by the TDC should be continuously processed and the outcome changes with time.

- **Data Buffering**

To enable continuous DMA-based data transfer for real-time data processing, data generated during the experiment will be buffered in the SDRAM in a pre-allocated space, referred to as the buffer, with a specific size. The buffer temporarily stores the data block transferred by the mSGDMA, until the processor processes these data, before receiving the next data block. Figure 4.6 depicts the buffering principle, illustrating the data movement and processing operations and the TDC FIFO status during the measurement. The buffering process can be divided into three stages:

Stage A: when the user starts a real-time measurement and the processor performs the following steps:

- First, it clears the TDC FIFO to discard any old data from previous measurements.
- Next, it loads the mSGDMA with a descriptor that specifies the address and the size of the buffer (*DMA_size = Buffer_size*), and triggers the mSGDMA to start transferring data from the TDC FIFO to the buffer.
- Then, it waits for the mSGDMA to finish the data transfer by checking its busy bit.

Meanwhile, the TDC system writes the detected photon and packet words into the FIFO. The mSGDMA reads the FIFO data and transfers it to the buffer allocated in the SDRAM. The data transfer rate to the SDRAM (the FIFO data reading rate) is typically much higher than the data writing rate to the FIFO, as discussed later, so the FIFO level remains close to 0.

Stage B: once the mSGDMA completes the data transfer, the processor accesses the buffer data and starts performing the data processing on these data according to the application requirement. Simultaneously, the TDC system continues measuring the arrival time of detected photons and writing the measurement data into the FIFO at the constant rate (*Fix_rate*), without any data being read by the mSGDMA. Consequently, the FIFO data level increases linearly with time during this stage.

Stage C: Upon completing the buffer data processing, the processor initiates a new data transfer operation by the mSGDMA. The mSGDMA first transfers the FIFO data that was written during stage B then proceeds to transfer the newly written data until the SDRAM buffer is refilled. Subsequently, the processor resumes the buffer data processing as in stage B, and so on. This cycle of stage B – stage C is repeated continuously until the user terminates the real-time measurement, as depicted in Figure 4.6.

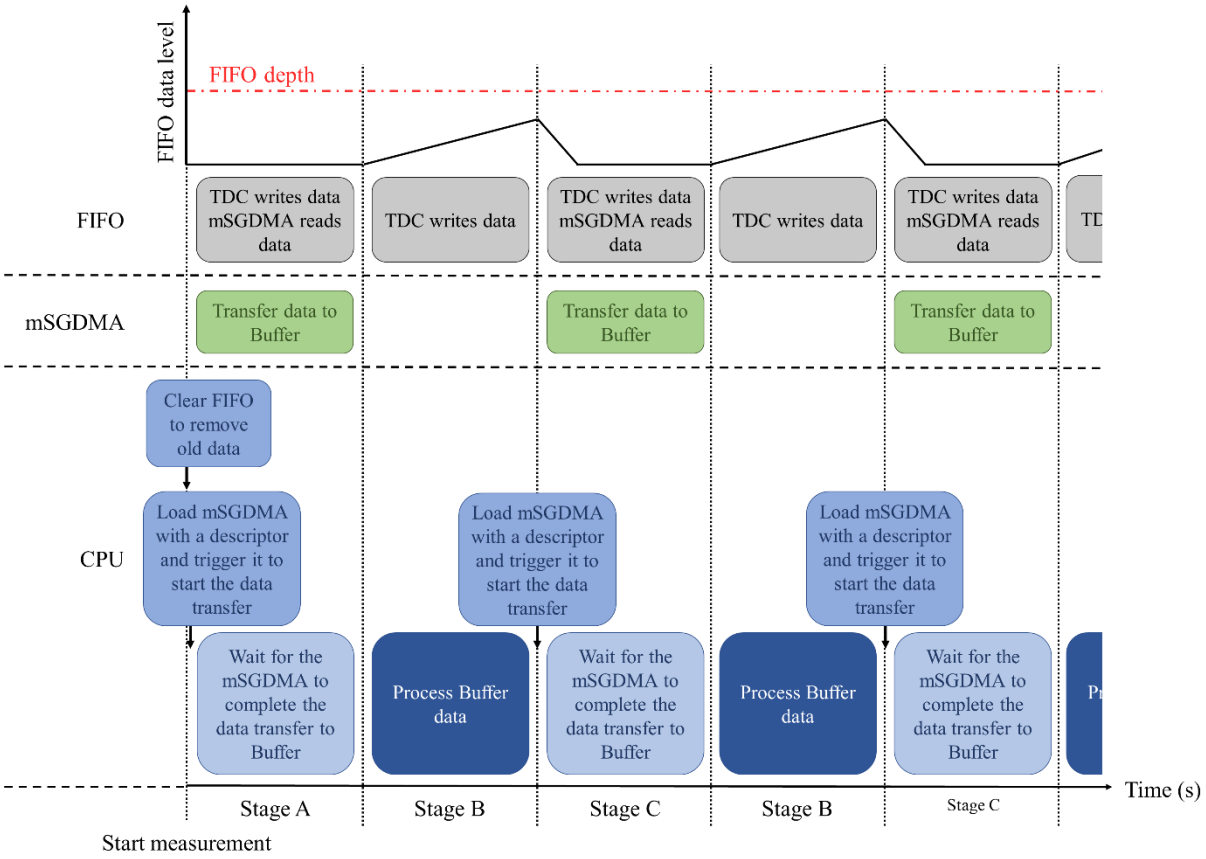


Figure 4.6 Data buffering and processing flow in real-time mode: data are buffered in the SDRAM before being processed. While buffered data are being processed, new data are being written into the FIFO.

A. Buffer Size and Latency

In DMA-based data transfer, the CPU cannot access the SDRAM while the mSGDMA performs a data transfer operation. Therefore, it must wait for the data transfer to finish before starting the data processing, which introduces a latency that can be calculated by Equation 4.5.

$$Latency = T\_buffer + T\_process \tag{4.5}$$

where *T_buffer* is the time needed for the mSGDMA to transfer data of a size equal to *Buffer_size* into the buffer (the duration of stage A or stage C), and *T_process* is the time needed for the CPU to process these data (the duration of stage B).

The system latency is a crucial factor in real-time applications. It can be reduced by minimizing *T_process* and *T_buffer*. The former term can be reduced by optimizing the data processing algorithms. However, the latter term depends largely on the buffer size and the data transfer rate. Although *T_buffer* is proportional to the data transfer rate to the SDRAM, it also depends on the data writing rate into the FIFO, since data must be generated and written to the FIFO before being transferred to the SDRAM. Nevertheless, the DMA data transfer rate using mSGDMA (*DMA_transfer_rate*) is significantly higher than the fixed rate at which data are written into the FIFO (*Data_writing_rate = Fix_rate*).

As illustrated in Figure 4.6, *T_buffer* in Stage C can be estimated by Equation 4.6.

$$T\_buffer = \frac{FIFO\_level}{DMA\_transfer\_rate} + \frac{Buffer\_size - FIFO\_level}{Data\_writing\_rate} \qquad (4.6)$$

where *FIFO_level* is the amount of data already written into the FIFO before the mSGDMA starts executing the descriptor, i.e. the FIFO data level at the end of stage B. These data are directly transferred into the SDRAM at the DMA transfer rate. Afterward, the new data words written to the FIFO will be immediately transferred into the SDRAM until the mSGDMA completes the transfer of *Buffer_size*. Therefore, we can consider that the data transfer rate of the new data words is equivalent to their writing rate in the FIFO, which is equal to *Fix_rate*.

Since we clear the FIFO before starting a new experiment, The FIFO level will be zero at the beginning of Stage A (*FIFO_level=0*). Hence, T_buffer in this stage is calculated by Equation 4.7.

$$T\_buffer = \frac{Buffer\_size}{Fix\_rate} \qquad (4.7)$$

If we configure the buffer size to be considerably larger than the TDC FIFO size (*Buffer_size >> FIFO_level*), since the data transfer rate is much higher than the data writing rate (*DMA_transfer_rate >> Data_writing_rate*), Equation 4.6 can be simplified to Equation 4.7. Thus, *T_buffer* is computed by Equation 4.7 in both Stages A and C.

Equation 4.7 demonstrates the importance of writing the TDC data at a constant rate, as the latency cannot be controlled if data are written into the FIFO at the variable detected photon

rate. However, by writing data into the FIFO at a constant, predefined rate (Fix_rate), the latency can be controlled and optimized by selecting an appropriate buffer size.

According to Equation 4.7, *T_buffer* and hence the latency can be minimized by reducing the buffer size. Nevertheless, there is an important limitation to take into account in the configuration of the buffer size; several studies have indicated that the mSGDMA transfer rate strongly depends, among other factors, on the data size parameter, which should be equal to the buffer size (*DMA_size* = *Buffer_size*). Thus, the buffer size should be carefully chosen to ensure a high data transfer rate that greatly exceeds the data writing rate into the FIFO.

Figure 4.7 depicts the relationship between the transfer rate and the data size when a DMA is used for the data transfer via the FPGA-to-SDRAM interface with bare metal application. This figure indicates that the maximum transfer rate is achieved for data size values larger than 250 kB. On the other hand, for data sizes below this threshold, the transfer rate decreases as the data size decreases and thus the latency increases. Based on these findings and our experimental observations, we set the *Buffer_size* and the *DMA_size* to 200 kB. With this configuration, we obtained a DMA transfer rate of approximately 1.1 GB/s. We also set the fixed data writing rate to 16 M data word/s which corresponds to 64 MB/s. With these settings, the system latency was about 3 ms, which is suitable for most of real-time applications.

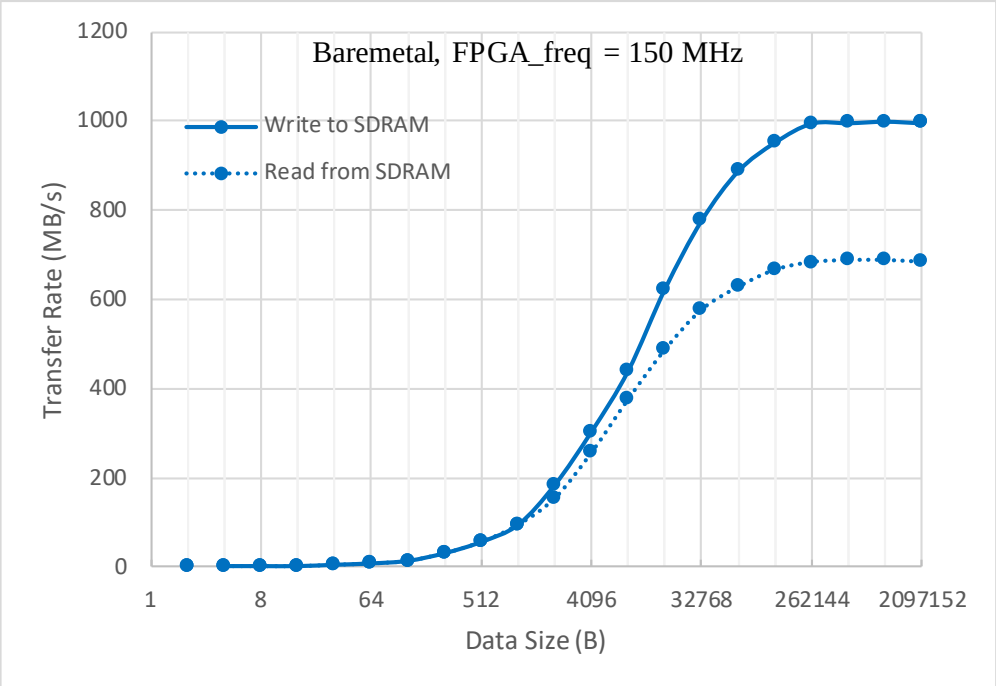


Figure 4.7 DMA transfer rate as a function of the data size (DMA_size) through the FPGA-to-SDRAM interface with a bare-metal application at an FPGA frequency of 150 MHz [152].

B. TDC FIFO Size and Data Processing Time

As shown above, while the processor performs online data processing on the buffer data (Stage B in Figure 4.6), the TDC system keeps writing the arrival time data of the detected photons into the FIFO. The FIFO data cannot be transferred to the SDRAM by the mSGDMA during this stage because the CPU reserves access to the SDRAM. To avoid FIFO overflow and hence loss of detectable photons, the size of the FIFO should be large enough to store all the measurement data generated during Stage B with a sufficient margin, as expressed by Equation 4.8.

$$FIFO\_size > T\_process\_max \times Data\_writing\_rate \quad (4.8)$$

where *T_process_max* is the maximum time required for the online data processing of buffer data, i.e. the time required for the slowest sequence of data processing steps performed in Stage B. This time varies depending on the type and complexity of the online processing tasks required for the application (e.g. histogram creation, photon counting, calibration, fluorescence lifetime (FLT) estimation, gravity center calculation, etc.). In our system, we experimentally found that a FIFO size of 128 K word (512 kB) ensure the functionality of the system in the different target applications without reporting any FIFO overflow.

3- Offline Mode with Data Buffering:

If the duration of a single measurement (*T_exp*) is so long that the amount of data generated during the measurement ($D\_exp = T\_exp \times Fix\_rate$) exceeds the available memory capacity, the offline mode is not applicable. Therefore, we also employed the data buffering approach in the offline mode, so that data is continuously transferred to the buffer during the measurement, and the buffer data are pre-processed online (e.g. by creating a histogram). After that, further processing can be applied to compute the final measurement value according to the application (e.g. extracting the FLT from the resulting histogram). The main difference between this mode and the real-time mode is that in the latter, the buffer data are fully processed in Stages B. Whereas, in the single measurement mode, the buffer data are pre-processed in Stages B and post-processed at the end of the measurement, as shown in Figure 4.8.

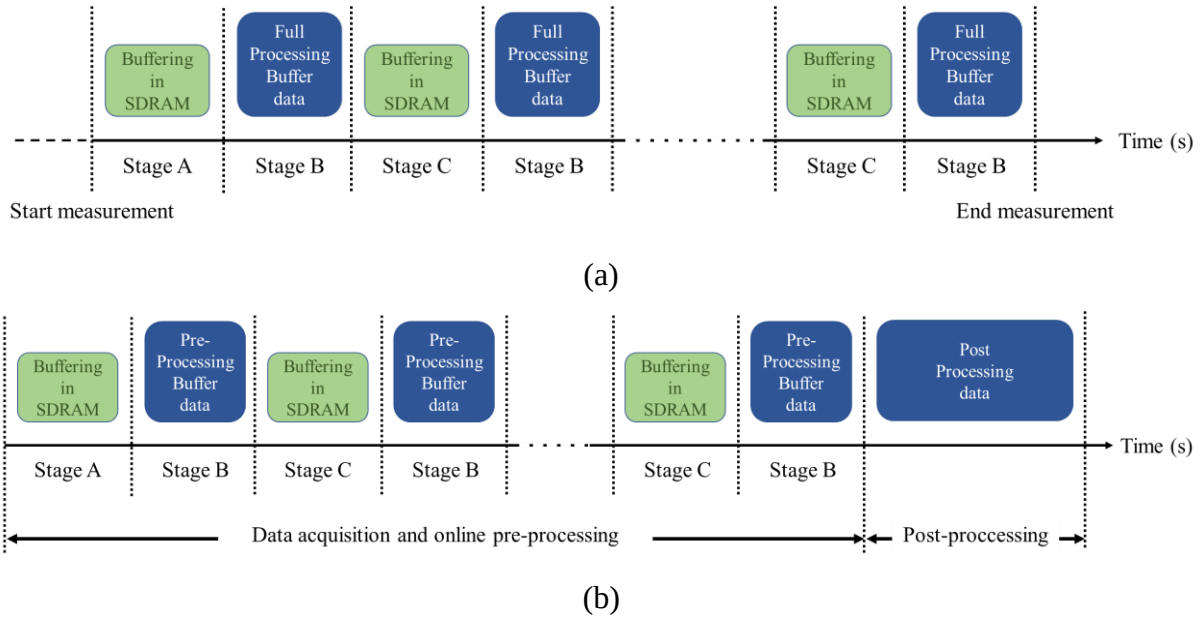


Figure 4.8 Data buffering and processing flow: (a) in real-time mode, data are fully processed in Stages B, (b) in single-measurement mode, data are pre-processed in Stages B and post-processed at the end of the measurement.

4.2 Data Processing

The data processing tasks in the TCSPC system differ according to the application, but there are some essential tasks that are common to all the applications. In our system, the entire data processing tasks are performed by the C program running on the HPS. This approach facilitates the customization of data processing and incorporating new tasks as per the application requirements, by integrating appropriate algorithms in the program. This section describes the various data processing tasks required for the four target applications.

4.2.1 Histogram Construction

The histogram construction is an essential step for most of the TCSPC applications. As our system operates in the individual arrival time (time-tag) mode, the first data processing step is to generate a histogram from the photon words, which contain the arrival time information. With the buffering data approach, this task is performed online in Stage B for both single measurement and real-time modes.

A. Synchronous System

In the synchronous system, the histogram is represented by a 1-D array, where each element corresponds to a bin. The number of bins in the histogram, and thus the size of the array, is

determined by the period of the excitation light, the clock period, and the number of bins in the TDL, as given in Equation 4.9.

$$Histogram\_size = TDL_{size} \times Ceil(\frac{T_{Laser}}{T_{clk}}) \quad (4.9)$$

Where TDL_{size} is the number of bins in the TDL, which is 256 in our system, T_{clk} is the TDC clock period, T_{Laser} is the period of the excitation light, which defines the measurement temporal window, and the term $Ceil(\frac{T_{Laser}}{T_{clk}})$ returns the number of clock period required to cover this window, which corresponds to the maximum coarse value (*Coarse_max*). The histogram creation algorithm reads the buffer data and discards the packet and zero words by checking the photon word flag bit. For each photon word, it increments by one the value of the array element, i.e. the histogram bin, that corresponds to the arrival time of that photon, as illustrated in Figure 4.9.

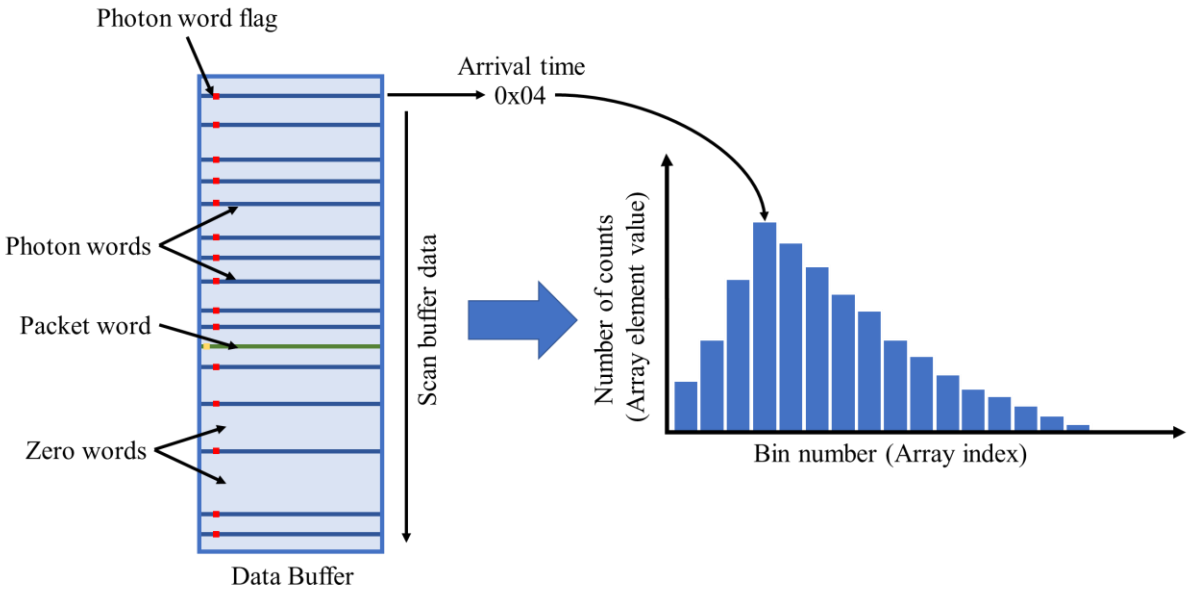


Figure 4.9 Histogram creation in synchronous system: the photon words are read from the buffer and mapped to the corresponding bins in the histogram array based on their arrival time.

B. Asynchronous System

As described in the previous chapter, the photon word in the asynchronous TDC includes three values: the START fine bin number (*fine1*), the STOP fine bin number (f*ine2*), and the coarse counter value (*Coarse*), as shown in Figure 3.28. These values indicate the arrival time of the photon. Hence, the measurement counts can be arranged in a 3D histogram. To construct this histogram in the SDRAM, we declare and allocate a 3D array, with a size of (256 × 256 ×

Coarse_max), since each of the START and STOP TDL has 256 bins. The algorithm scans the photon words in the buffer data and increments by one the value of the array cell that corresponds to the three values of each photon word, as depicted in Figure 4.10.

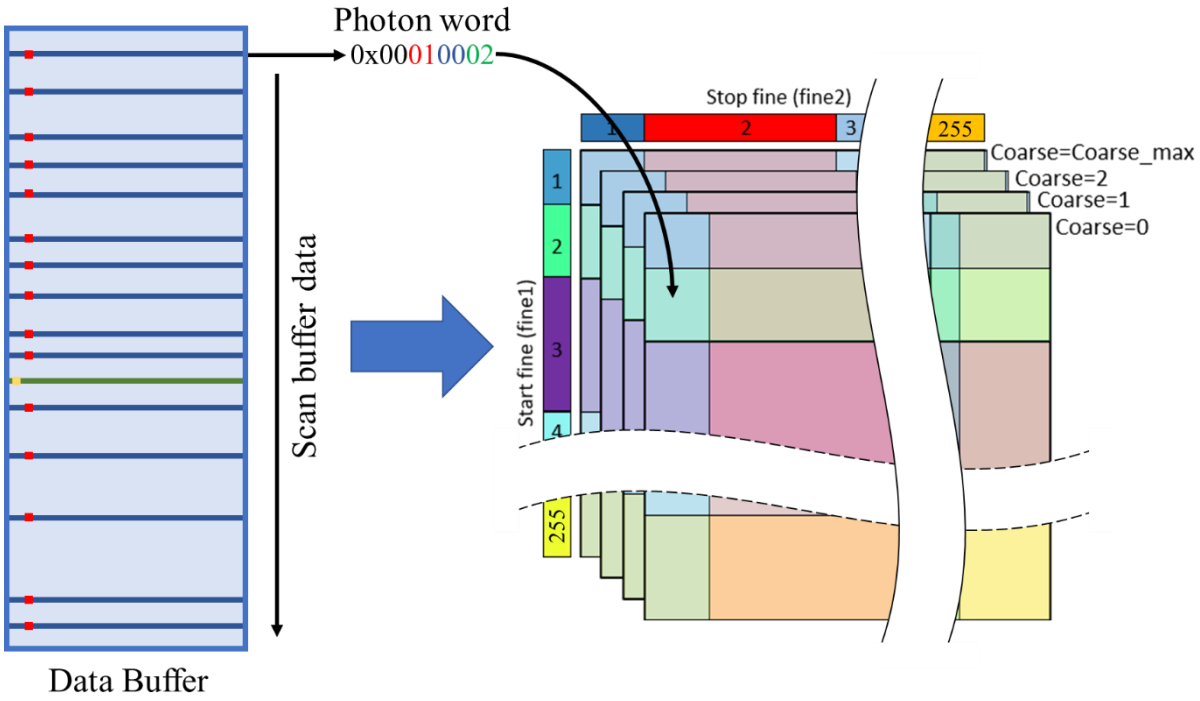


Figure 4.10 Histogram creation in asynchronous systems: the photon words are read from the buffer and mapped to the corresponding cell in the 3D array based on their fine1, fine2 and Coarse values.

4.2.2 Photon Counting (Intensity Measurement)

In TCSPC systems, the light intensity is quantified by the photon detection rate, which is the number of detected photons per unit of time. To capture fast fluctuations in intensity, the unit of time must be sufficiently short. For this purpose, the system employs the packet signal or the macro time clock. It measures the photon detection rate by counting the photons detected during each packet pulse, i.e. it scans the buffer data and counts the photon words between each pair of successive packet words. The photon counting is essential for different TCSPC applications and functions, such as:

- Real-time intensity measurement: the system displays the intensity variations in real time, as shown in Figure 4.11-a. This function is typically used during the alignment of the optical components of the TCSPC system. It is also used to adjust the photon count rate to the optimal level in the TCSPC technique, which is approximately 1% of the excitation light rate.

- Intensity mode: in this mode, the system constructs the intensity histogram, which represents the photon counts per unit time during the experiment time, as shown in Figure 4.11-b.
- Fluorescence intensity distribution analysis (FIDA): in this application, the system generates the histogram of the photon count distribution, which represents the distribution of the frequency of the obtained photon numbers [14], as illustrated in Figure 4.11-c. This application is usually used to quantify fluorescent particles in small volumes and the diffusion coefficient from the autocorrelation function of the fluorescence signal as described in detail in [159].
-

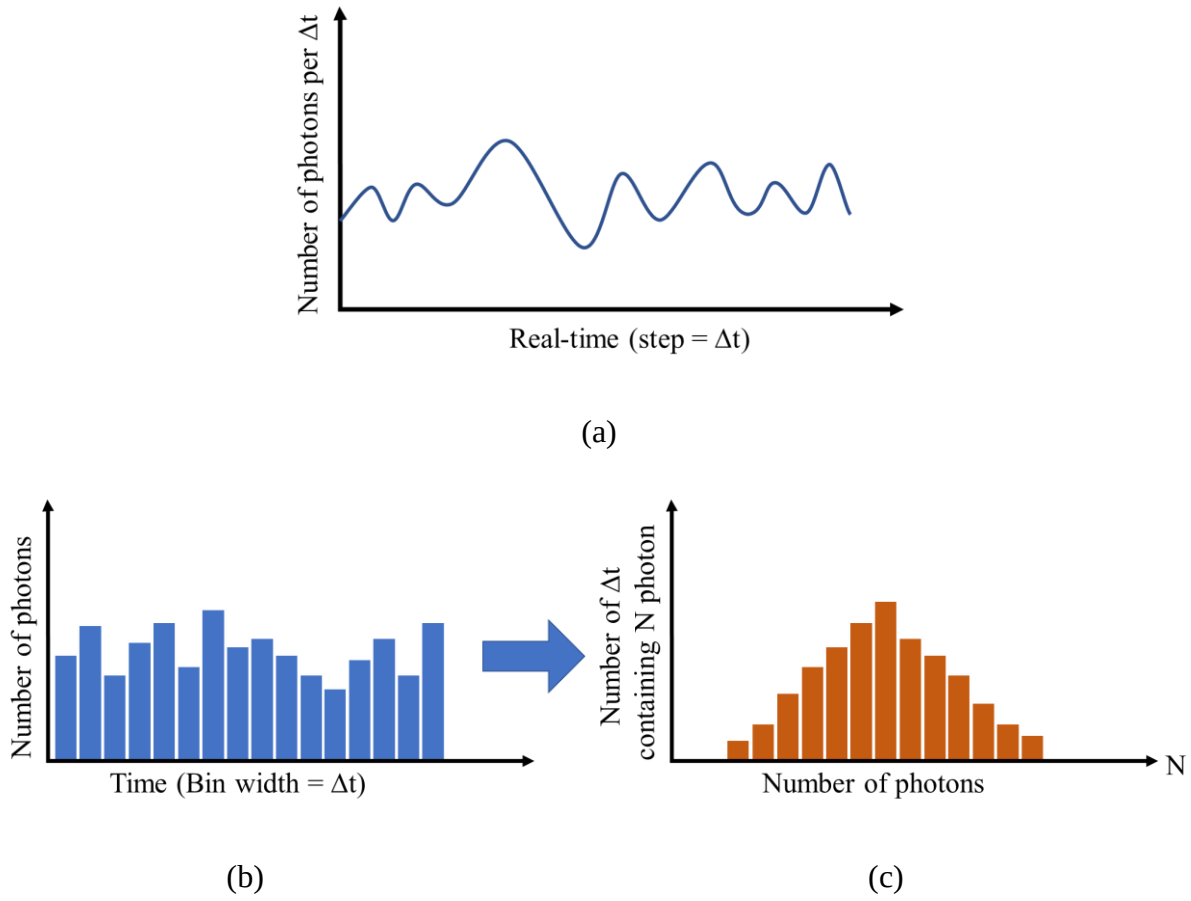


Figure 4.11 Photon counting: (a) Real-time photon counting mode, (b) Intensity histogram mode, (c) Fluorescence intensity distribution analysis (FIDA).

4.2.3 TDC Calibration

FPGA-based TDCs suffer from large variations in the bin width due to various factors, such as process variation and the routing topology. These variations induce the nonlinearity of the TDC,

which necessitates a calibration process. TDC calibration is a critical data processing step in TCSPC systems. To cover this important step in much detail, we dedicated the next chapter to discussing this step, where we describe and compare the most common calibration techniques and propose a novel and robust method for the calibration of asynchronous TDCs.

4.2.4 Background Signal Elimination

The elimination of the background signal from the measured signal is an important step for most TCSPC-based applications. This signal consists of two components: time-uncorrelated and time-correlated background. The former originates from either the ambient light or the dark count noise of the SPAD. The latter results from the excitation light that reaches the photon detector due to leakage and reflections in the optical setup, the fluorescence emission of the optical and microfluidic elements, and the fluorescence emission of trace impurities in the solvents or quenchers.

To eliminate the background signal, the system records this signal before recording the measured signal. Depending on the application, the background signal is recorded either with the excitation light Off or On. For example, in LiDAR applications, the background signal can only be recorded in the absence of the excitation light and is thus recorded while this light is Off. In this case, only the time-uncorrelated background, including the ambient light and the dark rate counts, is recorded. Conversely, in the measurement of the fluorescence signal of a solution, the background signal is recorded while the pulsed excitation light is On and no sample is present in the microfluidic cavity. This way, both the time-correlated and non-correlated background can be eliminated together. Furthermore, during the recording of the background signal, the cavity might be filled with the elution buffer in order to eliminate its fluorescence signal along with the background.

Usually, the photon rate of the background signal is much lower than that of the measured signal. Hence, the background signal should be acquired for a relatively long time (T_bg), much longer than the measurement time (T_exp), to increase the number of detected photons and thus improve the signal-to-noise ratio, since the latter is proportional to the square root of the photon number. Moreover, a brief acquisition time of the background signal results in a low photon count and thus a poor signal-to-noise ratio for the background. This, in turn, degrades the measurement's signal-to-noise ratio after subtracting the background signal from the measured one. To take into consideration the difference in acquisition time between the background and

the measurement signals, the background signal should be scaled by a factor of ($\frac{T_{Laser}}{T_{clk}}$) which reflects the ratio of acquisition times.

To eliminate the background contribution in the measurement signal, the background and measurement histograms are first constructed and calibrated, then the scaled background histogram is subtracted bin-by-bin from the measurement histogram. Figure 7.8 in Chapter 7 demonstrates the principle by a practical example of the background elimination from the fluorescence signal of the Benzo[a]pyrene in pure ethanol.

4.2.5 Afterpulsing Noise Suppression

Afterpulsing is a critical imperfection of most photon detectors. It refers to the probability that the detector generates a false pulse within some microseconds after the detection of a true photon. These false pulses are hence temporally correlated with the true ones. However, in our target applications, we operate at a high excitation repetition rate, which implies a relatively short temporal window of the measurement. For instance, at a repetition rate of 10 MHz, the temporal window is 100 ns. With such a relatively short temporal window, the afterpulsing counts are assumed to follow a uniform distribution over the measurement's temporal window [14, 160, 161], resulting in a constant offset. This offset can be considered as a signal-dependent noise since it is conditional on the detection of true photons of the signal under measurement. The magnitude of this noise is proportional to the signal intensity. In practice, the ratio between the afterpulsing noise level and the detected signal is determined by its afterpulsing probability, which is a key characteristic of photon detectors. The SPAD used in our system has an afterpulsing probability of about 0.5%, so to eliminate the afterpulsing noise, we reduce the total counts of the recorded histogram by a factor of 0.5%, and this reduction is uniformly distributed over the histogram bins.

4.2.6 Calculation of Additional Parameters

After constructing and calibrating the measurement histogram and removing the background and other noise signals, we can derive further parameters from the resulting histogram according to the application. These parameters may include the fluorescence intensity or lifetime (FLT), the center of gravity (COG), or the full width at half maximum (FWHM), among others.

As explained later in Chapter 7, in our target applications of water pollution sensing and real-time microfluidic droplet screening and sorting, we need to extract the fluorescence lifetime from the recorded fluorescence signal. In another target application, which is the stray light characterization, we compute the center of gravity of the recorded peaks to estimate the distances of the stray light sources.

- **Fluorescence Lifetime Estimation**

For the FLT estimation of monoexponential fluorescence decays, we improved and optimized a maximum likelihood estimator (MLE) algorithm, developed by our research team [162]. Compared to other estimations such as least square, the MLE provides superior performance in terms of speed, robustness, and precision, especially for signals with low photon counts and low signal-to-noise ratio. Furthermore, the optimized algorithm incorporates an iterative noise reduction procedure called auto-windowing, which improves the estimation accuracy and speed. Auto-windowing iteratively truncates the signal histogram to eliminate the noise caused by the high-weight late bins and recalculates the lifetime until a stable result is obtained. More details about this algorithm are available in [162].

- **Center of Gravity Calculation**

PGA-based TDCs suffer from various sources of errors that affect the precision of single-shot measurements, such as the TDC nonlinearity and jitter. Therefore, a common technique to enhance the accuracy is to use the center of gravity by averaging multiple measurements of same the time interval. The center of gravity is calculated by the sum of the count number of each bin weighted by the bin order divided by the total number of counts. Then, the result is scaled by the bin time width (*LSB*) to obtain the center of gravity in the time domain, as shown in Equation 4.10.

$$COG = LSB \times \frac{\sum_{i=0}^{B}(i \times N_i)}{\sum_{i=0}^{B} N_i} \tag{4.10}$$

where B is the number of bins in the histogram, i is the bin order and N_i is the number of counts in the i-th bin.

4.3 Conclusion

This chapter presented the data processing approach adopted in our TCSPC system. The data processing is performed by a bare-metal application running on the hard processor system (HPS) of the Cyclone V S0C-FPGA, rather than on a control PC, to enable the real-time operation of the system. To achieve high-rate data transfer from the FPGA to the HPS, we implemented a hardware-software co-design that utilizes the FPGA-to-SDRAM interface. Moreover, we employed a direct memory access (DMA) engine to offload the processor from the data movement operations and preserve its resources for data processing. In this chapter, we detailed the data transfer mechanism and the different components and interfaces involved in this co-design. We also discussed the different challenges and solutions related to the DMA-based data transfer, and proposed a data buffering approach that enables the DMA-based transfer of the TDC data at a constant rate, enabling online data processing. Moreover, we investigated the optimal configuration of the DMA data size, the data buffer size, and the FIFO size to minimize the system latency and avoid the FIFO overflow. This chapter also outlined the essential data processing tasks such as histogram construction and noise elimination, including background signal and afterpulsing noise. It also discussed further data processing steps required by our target applications, such as photon counting. FLT estimation and COG calculation.

Chapter 5: TDC Calibration

The nonlinearity of the TDC introduces errors in the raw histogram constructed from the TDC data. Therefore, the calibration of these histograms is an essential process before any further data processing. The average-bin-width and the bin-by-bin methods are the most commonly used calibration techniques [163-165]. In recent years, many studies have discussed these two methods, with a focus on the bin-by-bin calibration. However, most of these studies have employed this method for calibrating individual time interval measurements, such as single-shot and averaged measurements, and have not discussed the calibration of histogram measurements, which is an essential procedure for many TCSPC applications. Previous works [166-168], achieved time measurement standard deviations of around 18 ps (0.4 LSB), 12.2 ps (1.07 LSB), and 9 ps (0.34 LSB), respectively, by bin-by-bin calibration. Nevertheless, these works did not utilize this method for histogram measurements. Moreover, most of the studies about TDC calibration predominantly concentrated on synchronous TDCs [169, 170], and only a few of them have discussed asynchronous TDCs.

In this chapter, we present the conventional calibration techniques for synchronous and asynchronous TDCs. Next, we propose a novel robust technique, which we call the Matrix calibration, for the calibration of asynchronous TDCs and evaluate its performance against the conventional bin-by-bin technique.

5.1 Calibration of Synchronous TDCs

The most commonly used calibration methods for synchronous TDCs are the bin-by-bin and the average-bin-width methods. These methods involve performing a code density test to determine the widths of the raw bins. In the code density histogram, the number of counts in each bin is proportional to the width of that bin. To facilitate the explication, we consider a 5-bin TDC with a total delay of T. Figure 5.1 depicts a code density histogram of this simple TDC. The following subsections describe the two methods in detail.

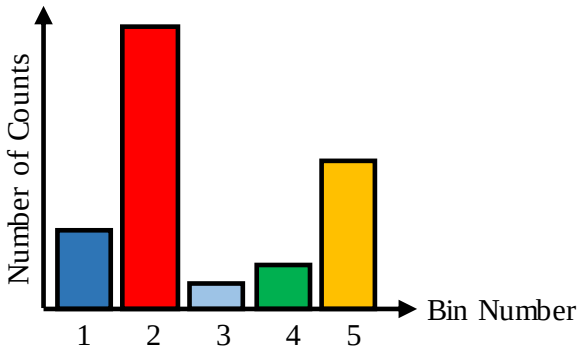


Figure 5.1 Code density histogram of a simple synchronous TDC with five bins and a total delay of T.

5.1.1 Bin-by-Bin Calibration for Synchronous TDCs

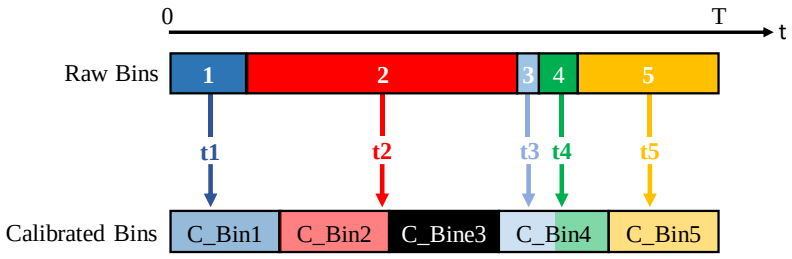
This method maps the TDC raw bins to calibrated times or calibrated bins using a lookup table (LUT). A code density test is conducted to determine the time distribution along the TDC bins. Then, the calibrated time corresponding to the center of each bin is computed from Equation 5.1.

$$ti = \left[\frac{N_i}{2} + \sum_{k=0}^{i-1} N_k\right] \times \frac{T}{N} \tag{5.1}$$

where *ti* is the calibrated time of the *i*-th bin, N_i is the number of counts in that bin, *N* is the total number of counts in the histogram and *T* is the total delay of the TDC.

The calibrated times of the raw bins are then stored in a bin-to-time lookup table (LUT) that maps each raw bin to its corresponding calibrated time. Moreover, the TDC measurement range is divided into calibrated bins of equal size, and the calibrated times of the raw bins are then projected onto the calibrated bins to determine which calibrated bin corresponds to each raw bin, as illustrated in Figure 5.2-a. Another LUT, termed the bin-to-calibrated_bin LUT, is then created to map each raw bin to its corresponding calibrated bin, as shown in Figure 5.2-b.

The bin-by-bin calibration method is appropriate for single-shot measurements where the bin-to-time LUT is used. However, this method may introduce dead bins when applied to histograms using the bin-to-calibrated_bin LUT. In this example, the third calibrated bin (C_Bin3) is a dead bin due to the excessive size of the second raw bin (Bin2), as illustrated in Figure 5.2-a.



(a)

Raw Bin	Calibrated Time (ps)	Calibrated Bin
1	t1	1
2	t2	2
3	t3	4
4	t4	4
5	t5	5

(b)

Figure 5.2 Calculating the calibrated times and calibrated bins of the raw bins; (b) Bin-by-bin calibration lookup table.

5.1.2 Average-Bin-Width Calibration

One version of this method corrects the bin width variation by applying scaling factors. The scaling factor of each bin is the ratio of the average bin width to the actual width of that bin. However, this approach amplifies the photonic noise for minuscule bins. We adopted a modified version of this method that involves redistributing the raw bins over calibrated bins with a uniform width according to their proportional overlap. From the code density histogram, the time width of the raw bins can be determined from Equation 5.2.

$$T_b = \frac{N_b}{N} \times T \tag{5.2}$$

Where T_b is the time width of the *b-th* raw bin, N_b is the number of counts in that bin, and T is the TDC measure range.

The concept of this calibration technique is to partition the measurement range of the TDC into M calibrated bins with equal time width (T_c), which is derived from Equation 5.3.

$$T_c = \frac{T}{M} \tag{5.3}$$

Since the calibrated bins have a uniform time width, the calibrated code density histogram should have calibrated bins with the same number of counts N_c, obtained from Equation 5.4.

$$N_c = \frac{N}{M} \tag{5.4}$$

Considering the code density histogram shown in Figure 5.1 that has five non-uniform raw bins. In order to obtain five calibrated bins identical in size, the counts of the raw bins are successively redistributed on the calibrated bins starting from the first bin. For each calibrated bin, the proportional shares of the raw bins are computed as illustrated in Figure 5.3 and stored in a specific table, called the calibration table, as presented in Figure 5.4. This table will be used later for the calibration of the measurement raw histogram.

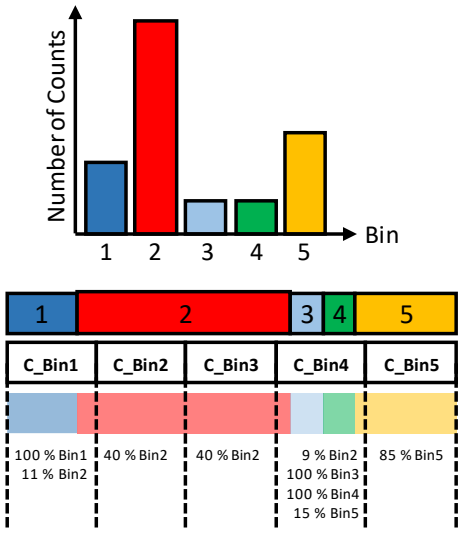


Figure 5.3 Average-bin-width calibration: redistributing the total counts on identical calibrated bins and creating the calibration table that defines the percentage share of the raw bins in each calibrated bin.

The average-bin-width method is suitable for the calibration of a histogram as it ensures a uniform size of the calibrated bins with no dead bins. In this example, the counts of the large raw bin (Bin2) are distributed over four calibrated bins (C_Bin1, C_Bin2, C_Bin3, and C_Bin4) with different percentages, while the fourth calibrated bin (C_Bin4) contains counts from four different raw bins (Bin2, Bin3, Bin4, and Bin5) also at different percentages, as illustrated in Figure 5.3.

C_ Bin	Raw Bins' Shares
1	Bin1 + 0,11 × Bin2
2	0,4 × Bin2
3	0,4 × Bin2
4	0,09 × Bin2 + Bin3 + Bin4 + 0,15 × Bin5
5	0,85 × Bin5

Figure 5.4 Calibration table: it describes the percentage share of the raw bins in the calibrated bins.

As the temporal resolution is determined by the bin size, this technique enables the downsampling and the adjustment of the temporal resolution by choosing a suitable value for the number of calibrated bins. Equation 5.5 gives the time resolution after calibration, which depends on the number of calibrated bins (*L)*.

$$T_c = \frac{T}{L} \tag{5.5}$$

5.2 Calibration of Asynchronous TDCs

In asynchronous TDCs, each count is represented by three values: Start fine (*fine1*), stop fine (*fine2*), and *Coarse*. Hence, the measurement counts are arranged in a 3D histogram, as explained in section 4.2.1. To simplify the explanation, we consider a 3D asynchronous coarse-fine TDC with 5 bins for both the start and stop TDL and a clock period *T*. Figure 5.5 depicts a 3D code density histogram of such a TDC. In the following, we discuss the calibration of this TDC using two methods.

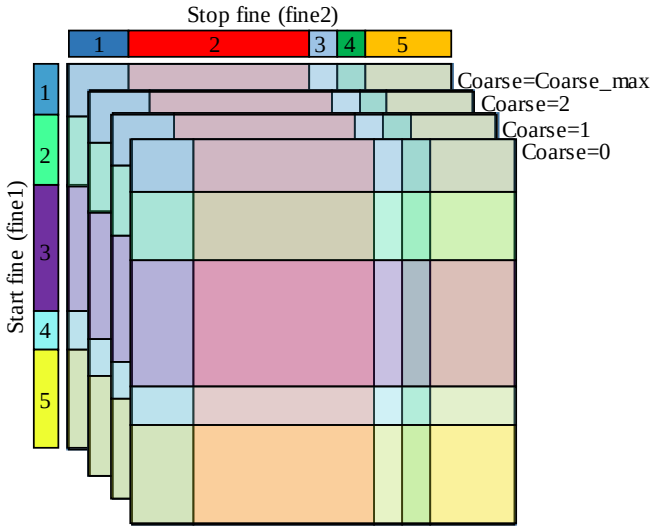


Figure 5.5 Three-dimension code density histogram of an asynchronous TDC: the Stop and Start fine bin numbers are respectively represented on x and y, whereas the Coarse value is represented on z.

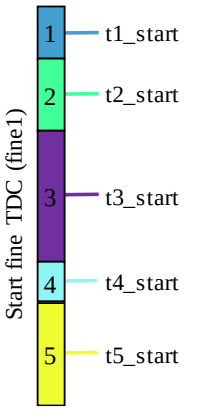
5.2.1 Bin-by-Bin Calibration for Asynchronous TDCs

Calibrating an asynchronous TDC using the bin-by-bin method requires performing a code density test to estimate the calibrated times of the raw bins and construct LUTs for the TDL of the *START* and *STOP* signals, as illustrated in Figure 5.6. These LUTs are then used to calculate the calibrated interval of each cell of the 3D histogram using Equation 5.6.

$$t\_cell = (Coarse \times T) + t\_stop - t\_start \tag{5.6}$$

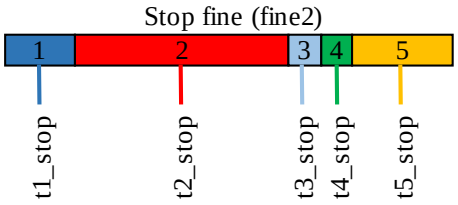
where *t_cell* is the calibrated time of the cell, *t_stop* are *t_start* are the calibrated times of the stop and start bins that represent the (x, y) coordinates of the cell, respectively, *T* is the system clock period and *Coarse* is the cell coarse value.

The calibrated intervals of all the cells are stored in a 3D LUT which is then used for the calibration of single-shot measurements. For the calibration of histograms, the TDC measurement range is partitioned into calibrated bins with uniform size, and a 3D cell-to-calibrated_bin LUT is created to map each cell of the 3D histogram to the corresponding calibrated bin in the 1D histogram based on its calibrated interval.



Start Bin	Calibrated Time (ps)
1	t1_start
2	t2_start
3	t3_start
4	t4_start
5	t5_start

(a)



Start Bin	Calibrated Time (ps)
1	t1_stop
2	t2_stop
3	t3_stop
4	t4_stop
5	t5_stop

(b)

Figure 5.6 Lookup tables of the two fine TDCs: (a) Start fine TDC LUT, (b) Stop fine TDL LUT.

5.2.2 Matrix Calibration

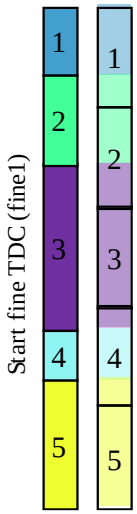
This method relies on the average-bin-width method to calibrate asynchronous coarse–fine TDCs. Considering the 3D code density histogram illustrated in Figure 5.5, each cell of this

histogram records the number of counts for which the *START* and *STOP* signals arrive in the bins corresponding to the x and y coordinates of that cell, respectively, with a coarse value equal to its z coordinate. For instance, the cell (4, 3, 1) records the number of counts where the *START* signal arrives in the fourth bin of the start TDL, the *STOP* signal arrives in the third bin, and the coarse value is equal to 1. Ideally, the start and stop TDL would have raw bins with a uniform width, and all the cells of the 3D code density histogram would have the same size. However, in a real TDC, since the TDC TDLs have non-uniform raw bins, resulting in different sizes for the cells of the 3D code density histogram. The size of each cell, indicated by its number of counts, depends on the width of its start and stop raw bins.

The concept of the Matrix calibration is to reallocate the raw cells onto calibrated cells with a uniform size. It redistributes the code density counts evenly on calibrated cells with equal size in four steps, as follows:

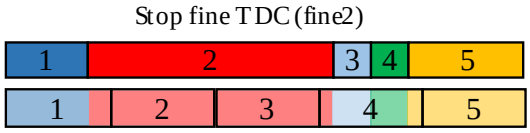
1- Individual calibration of the start and stop fine TDC:

By disregarding the stop fine and the coarse values of the 3D code density histogram cells, the columns are collapsed into one column. This column is a 1D histogram that represents a code density histogram of the start fine TDC. Similarly, by disregarding the start fine and the coarse values, the rows are collapsed into one row, which represents a 1D code density histogram of the stop fine TDC. These histograms can be used to construct the calibration tables of the start and stop fine TDCs, as illustrated in Figure 5.7.



C_Bin	Raw Bins' Shares %
1	100 % Bin1 + 40 % Bin2
2	60 % Bin2 + 30 % Bin3
3	55 % Bin3
4	15 % Bin3 + 100 % Bin4 + 25 % Bin5
5	75 % Bin5

(a)



C_Bin	Raw Bins' Shares %
1	100 % Bin1 + 14 % Bin2
2	40 % Bin2
3	40 % Bin2
4	6 % Bin2 + 100 % Bin3 + 100 % Bin4 + 25 % Bin5
5	75 % Bin5

(b)

Figure 5.7 Individual calibration tables: (a) Start fine TDC calibration table, (b) Stop fine TDC calibration table.

2- Column calibration:

In practice, each column of the 3D code density histogram can be considered as a 1D code density histogram of the start fine TDC, and thus can be calibrated using the start calibration table constructed in the previous step. The individual calibration of all the columns of the 3D histogram produces a semi-calibrated 3D histogram where all the cells have equal row height while the columns still have non-uniform widths, as illustrated in Figure 5.8.

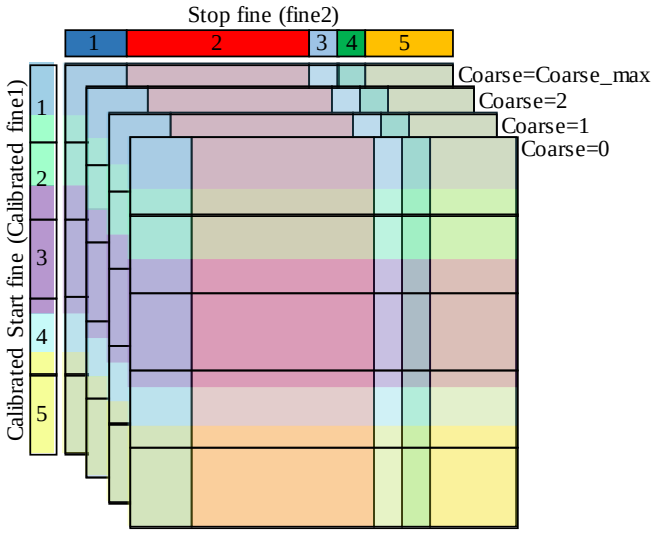


Figure 5.8 Column calibration: the individual calibration of the columns using the average-bin-width method gives a semi-calibrated histogram where all the rows have the same height.

3- Row calibration:

The rows of the semi-calibrated histogram obtained from the previous step are practically 1D code density histograms of the stop fine TDC. Hence, they can be calibrated using the calibration table of the stop TDC. The individual calibration of all the rows produces a calibrated 3D histogram where all the cells have identical sizes, as depicted in Figure 5.9.

4- Building 1D calibrated histogram:

The final step of the Matrix calibration is to convert the 3D histogram obtained from the preceding step into a 1D histogram. This entails adding the counts of each cell of the 3D calibrated histogram to its corresponding bin in the 1D calibrated histogram. The index of this bin is determined by Equation 5.7.

$$Bin\_index = (Coarse \times M) + C\_fine1 - C\_fine2 \qquad (5.7)$$

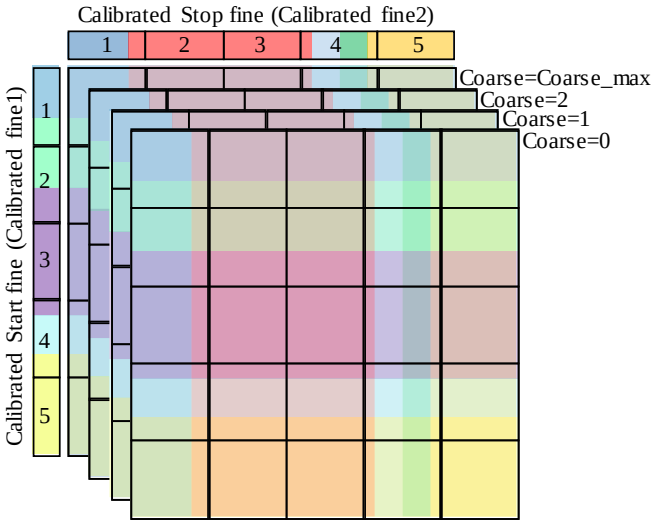


Figure 5.9 Row calibration; the average-bin-width calibration is applied to the rows, resulting in a calibrated 3D histogram with identical cell size.

where *C_fine1*, *C_fine2*, and *Coarse* denote, respectively, the x, y, and z coordinates of the calibrated cell, i.e., the numbers of its row, column, and slice, and *M* is the total number of calibrated bins in the stop TDL, which equals 5 in our example. For instance, the counts of the cell (3, 2, 4) should be added to the 21st bin of the calibrated 1D histogram (bin_index = (4×5) +3−2 = 21).

In this step, all the cells of the 3D calibrated histogram should be scanned and aggregated to their respective bins in the 1D calibrated histogram. Note that each bin of the 1D histogram encompasses multiple cells of the 3D calibrated histogram.

5.3 Simulation Results

We conducted a series of simulations using MATLAB to evaluate and compare the different calibration methods for synchronous and asynchronous TDCs. We modeled synchronous and asynchronous TDCs with the same temporal characteristics as our implemented coarse-fine TDCs. The TDC models comprised coarse counters operating at 190 MHz and two TDLs each with 256 DEs, a resolution of 20.56 ps (LSB = 20.56 ps), and the same time distribution as our implemented TDCs.

- **Synchronous TDCs**

In the first simulation, we generated 10 synchronous TDCs with different levels of Root Mean Square (RMS) differential non-linearity (DNL) varying from 0 to 1 LSB. For each simulated TDC, we generated 10^7 random photon counts to perform a code density test and create the bin-

by-bin LUTs and the average-bin-width calibration tables. Next, to evaluate and compare the two methods, we simulated another code density test with another 10^7 random counts and applied the two methods to calibrate the resulting code density histogram. After repeating these steps for each of the 10 simulated TDCs, we calculated the RMS DNL of the resulting calibrated histograms. Figure 5.10 shows the results and indicates that the bin-by-bin method did not improve the DNL of the TDC, while the average-bin-width calibration was independent of the noise of the TDC and significantly improved the DNL.

One may note that since the first TDC in this simulation was an ideal TDC with a DNL of 0 LSB, the DNL of the calibrated histogram should theoretically be 0 LSB. Nevertheless, the calibrated histogram exhibited a DNL of approximately 0.005 LSB. This is attributed to the shot noise, which is a fundamental limitation of code density tests and can be estimated by Equation 5.8.

$$Shot\ noise\ = \sqrt{\frac{Number\ of\ Bins}{Counts\ number}}\ (LSB) \tag{5.8}$$

For this simulation, the number of bins was 256 and the count number was 10^7, which gives a shot noise of around 0.005 LSB.

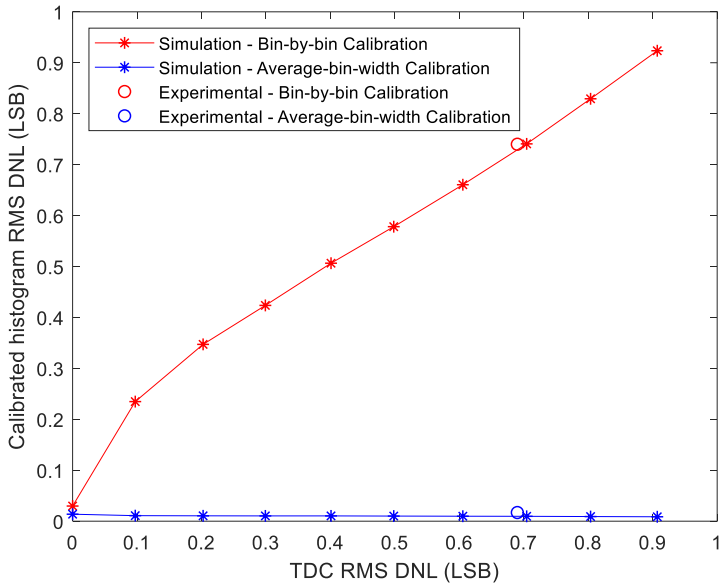
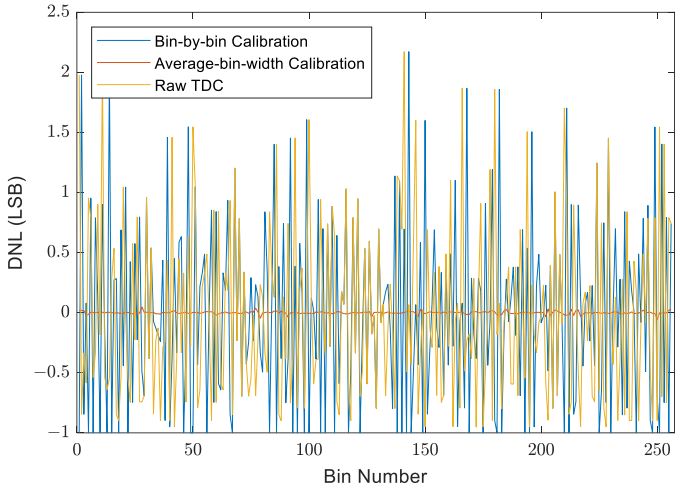


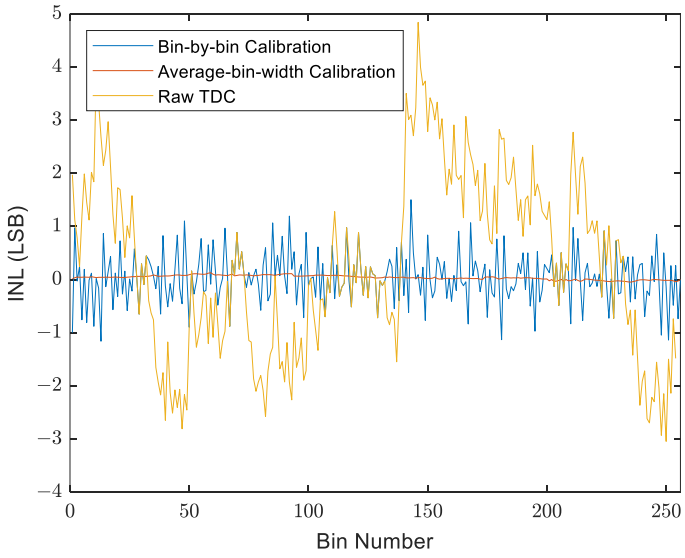
Figure 5.10 Simulation and experimental results for synchronous TDCs: the DNL after applying the calibration methods compared to the DNL of the raw TDC.

The second simulation aimed to evaluate the DNL and INL of the two methods when applied to our synchronous TDC model. We simulated two code density tests, with 10^7 events each,

one to create the bin-by-bin LUT and the average-bin-width calibration table and the other to apply and assess the two methods. Figure 5.11 presents the DNL and INL values of the raw code density histogram and those of the calibrated histograms, and Table 5.1 summarizes the data statistics of these values. The results show that the bin-by-bin calibration enhanced only the INL of the TDC without enhancing its DNL, whereas the average-bin-width calibration significantly improved both the DNL and the INL.



(a)



(b)

Figure 5.11 DNL and INL values for a synchronous TDC model, before and after applying the bin-by-bin and the average-bin-width calibration methods: (a) DNL values, (b) INL values.

Table 5.1 DNL and INL statistics for the synchronous TDC model, calculated for the raw histogram and the calibrated histograms of the bin-by-bin and average-bin-width calibration methods.

		Mean	**Median**	**STD**
DNL data statistics	Raw TDC	0	−0.1054	0.7362
	Bin-by-bin	0	−0.07453	0.7798
	Average-bin-width	0	−0.0003624	0.009665
INL data statistics	Raw TDC	0.2614	0.2038	1.676
	Bin-by-bin	0.003083	0.007194	0.4816
	Average-bin-width	0.0026	0.006688	0.03167

In the third simulation, we simulated a Gaussian signal and recorded it with our synchronous TDC model. We simulated 10^7 counts following a normal distribution with an arbitrary mean delay of 2.5 ns and standard deviation (sigma) of 0.3 ns. We then applied both methods to calibrate the recorded histogram. Figure 5.12 shows the calibrated histograms of the Gaussian signal and indicates that the average-bin-width calibration resulted in much lower noise than the bin-by-bin method. The results confirm the advantage of the average-bin-width over the bin-by-bin calibration for histograms.

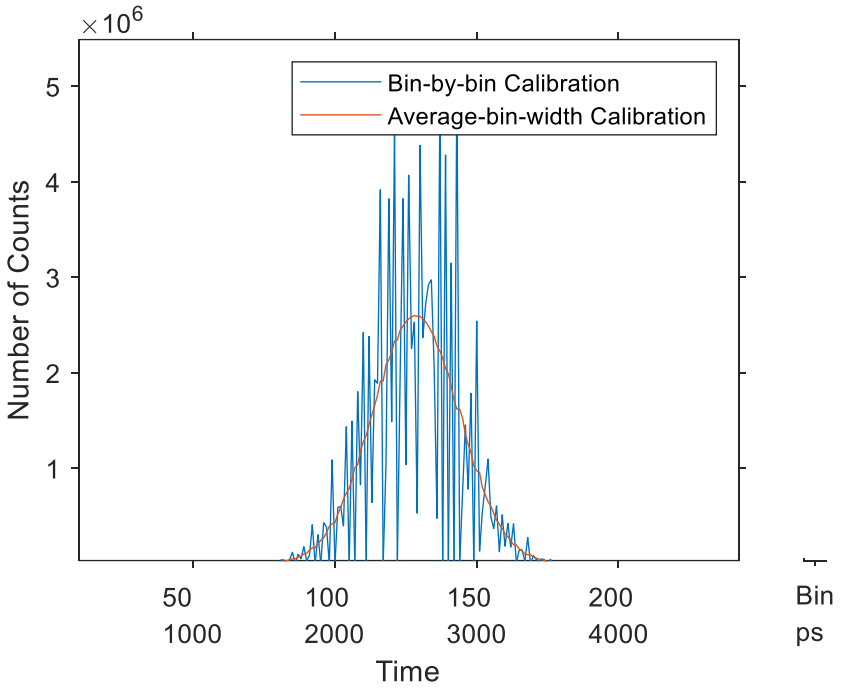


Figure 5.12 Calibrated histogram of a Gaussian signal using bin-by-bin and average-bin-width methods for synchronous TDCs.

- **Asynchronous TDCs**

In the first simulation, we simulated 10 asynchronous coarse-fine TDCs with RMS DNL values ranging from 0 LSB to 1 LSB. These DNL values were obtained after concatenating the raw bins of the two fine TDCs of each asynchronous TDC. Next, for each simulated TDC, we simulated a code density test with 10^7 random counts to construct the bin-by-bin LUTs and the calibration tables of the Matrix calibration. Then, to evaluate and compare the two methods, we simulated another code density test with 10^7 counts. For these counts, the *START* pulses arrived with random delays uniformly distributed over the range of the start fine TDC, and the *STOP* pulses arrived after their corresponding START pulses by time intervals that varied uniformly from 0 to 5 ns. We generated the 3D raw histogram from the arrival times of these counts by computing the coordinates of each count and incrementing the corresponding cell (start fine bin, stop fine bin, coarse value). Finally, we calibrated this histogram using both methods and measured the RMS DNL values of the calibrated 1D histograms. The results are presented in Figure 5.13.

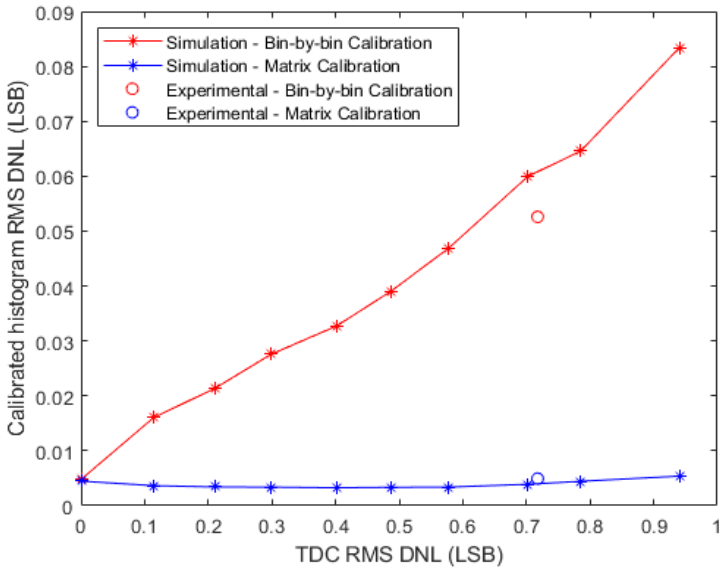
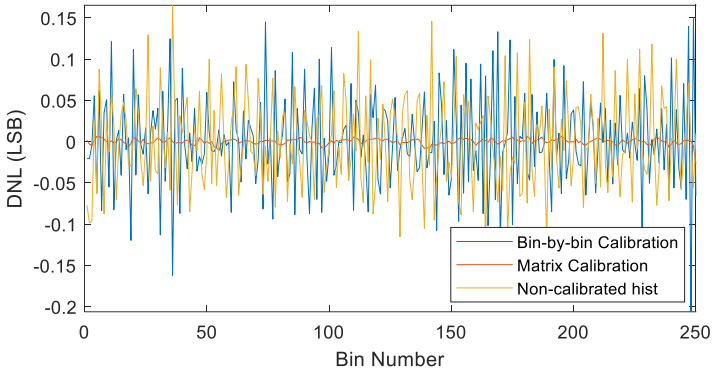


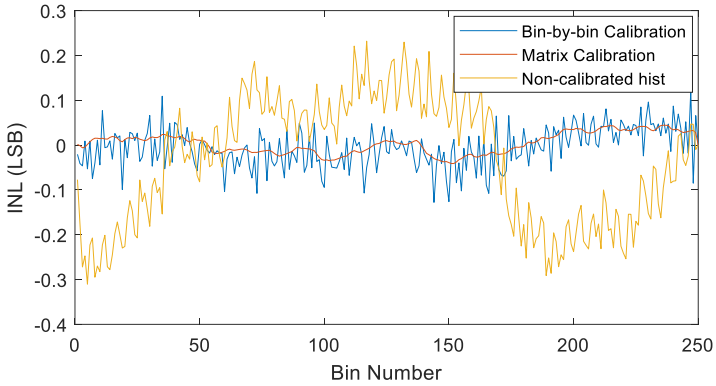
Figure 5.13 Simulation and experimental results for asynchronous TDCs: the DNL after applying the calibration methods compared to the DNL of the raw TDC.

The results demonstrate that the Matrix calibration method achieved better DNL than the bin-by-bin method and exhibited low sensitivity to the noise of the raw TDC with an almost flat response. The non-zero noise for the ideal TDC (DNL = 0 LSB) is also attributed to the shot noise, as discussed for the synchronous TDC, and could be estimated by Equation 5.8 as 0.005 LSB.

In the next simulation, we generated code density tests with 10^7 counts to construct the bin-by-bin LUTs and the Matrix calibration tables of our asynchronous TDC model. Then, we applied both methods to calibrate the raw histogram of another simulated code density test and evaluated the DNL and INL values. Figure 5.14 illustrates the DNL and INL values before and after applying the calibration methods and Table 5.2 summarizes the statistical data of these values.



(a)



(b)

Figure 5.14 DNL and INL values for the asynchronous TDC model, before and after applying the bin-by-bin and the Matrix calibration methods: (a) DNL values, (b) INL values.

The results indicate that the bin-by-bin method improved the INL, relative to the non-calibrated histogram, without improving the DNL. However, the Matrix calibration achieved significant improvements over the bin-by-bin method in both the DNL and the INL, with more than 10 times lower DNL and about 2 times lower INL.

A Comparison of these results with those obtained for the synchronous TDC, presented in Figure 5.11 and Table 5.1, reveals that the asynchronous TDC has an order of magnitude lower DNL and INL values of the non-calibrated histogram than the synchronous TDC. This indicates

that the asynchronous dual-chain TDC achieves a significantly higher accuracy than the synchronous TDC.

Table 5.2 DNL and INL statistics for the asynchronous TDC model, calculated for the non-calibrated histogram and the calibrated histograms using bin-by-bin and Matrix calibration methods.

		Mean	**Median**	**STD**
DNL data statistics	Non-calibrated histogram	~0	−0.0049	0.056
	Bin-by-bin calibration	~0	−0.0036	0.058
	Matrix calibration	~0	0.00033	0.0035
INL data statistics	Non-calibrated histogram	−0.034	−0.008	0.14
	Bin-by-bin calibration	−0.0064	−0.004	0.045
	Matrix calibration	0.0036	0.0048	0.023

In the last simulation, we measured a simulated Gaussian signal by our asynchronous TDC model and calibrated the resulting histogram using the bin-by-bin and Matrix calibration methods. We generated 10^7 counts with time intervals between the *START* and *STOP* signals following a normal distribution with an arbitrary mean interval of 2.5 ns and a standard deviation of 0.5 ns. Figure 5.15 depicts the obtained histograms after calibration. The Matrix calibration yielded lower noise than the bin-by-bin method. Moreover, the center of gravity of the calibrated histogram was closer to the true value for the matrix calibration (COG = 2499.1 ps, error = 0.9 ps) than for the bin-by-bin method (COG = 2489.5 ps, error = 10.5 ps).

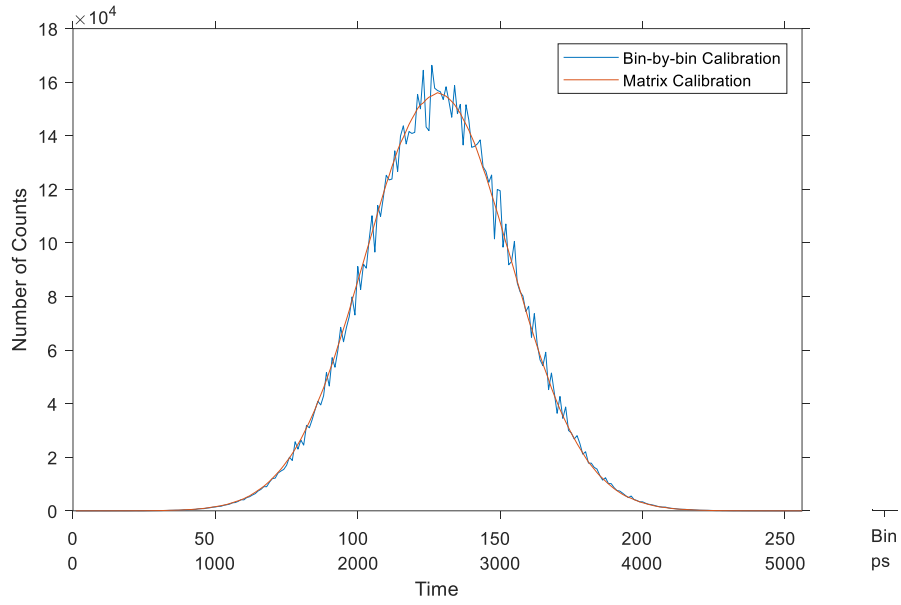


Figure 5.15 Calibrated histogram of a Gaussian signal using bin-by-bin and Matrix calibration methods for asynchronous TDCs.

5.4 Experimental Results

We performed a series of experiments to validate the simulation results using our synchronous and asynchronous TCSPC systems.

- **Synchronous TDC**

To construct the bin-by-bin LUT and the average-bin-width, we conducted a code density test with 10^7 counts by exposing the SPAD to ambient light at a low detected photon rate of about 1 M photon/s. At such a relatively low photon rate, the mean time between the arrival of two consecutive photons is 1 µs, which is 200 times larger than the TDL measurement range (5.26 ns). Thus, the *STOP* signal arrived with random delays that uniformly covered the measurement range of the TDC. From the code density histogram, we computed the RMS DNL of the raw TDC and it was about 0.69 LSB. We then performed another code density test with the same number of counts to evaluate and compare the two calibration methods. We calibrated the code density histogram using the two methods and computed the RMS DNL of the calibrated histograms. For the bin-by-bin method, the RMS DNL was 0.74 LSB, and only 0.017 LSB for the average-bin-width method. These experimental values confirm the simulation results, as plotted in Figure 5.10.

We also used our synchronous TDC system to record the fluorescence signal of a piece of paper excited by a 405 nm pulsed diode laser. Figure 5.16 shows the calibrated histograms of the recorded signal obtained by the two calibration methods. It confirms that the average-bin-width method introduces much lower noise than the bin-by-bin calibration.

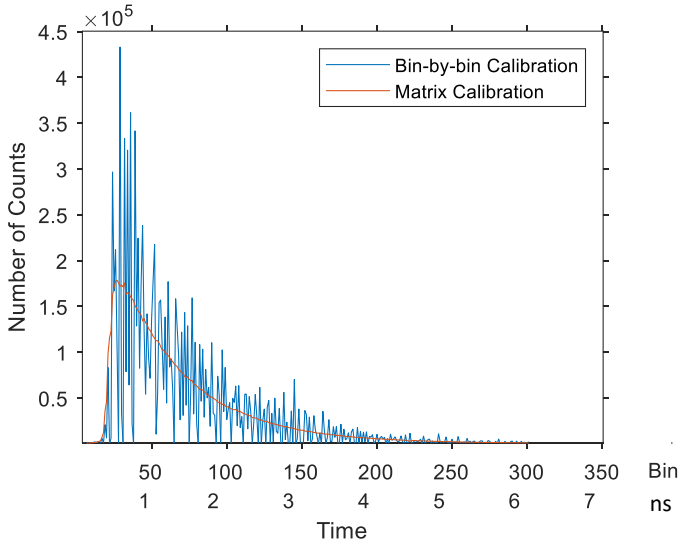


Figure 5.16 Calibrated histograms of the fluorescence signal of a piece of paper obtained with the bin-by-bin and the average-bin-width calibration methods for synchronous TDC.

- **Asynchronous TDC**

To validate our proposed Matrix calibration method experimentally and confirm the simulation results, we conducted a series of experiments using our asynchronous TCSPC system. We first performed a code density test with 10^7 counts by exposing the SPAD to ambient light at a photon rate of 1 M photon/s. From the resulting histogram, we created the bin-by-bin 3D LUT and the Matrix calibration tables and calculated the RMS DNL of the TDC TDLs which was 0.71 LSB. We then performed another code density test with the same number of counts to apply the two calibration methods and computed the RMS DNL of the calibrated histograms. The results showed an RMS DNL of 0.053 LSB for the bin-by-bin method and of only 0.005 LSB for the Matrix calibration, as plotted in Figure 5.13. To ensure comparability between the experimental and simulation results, the code density test should integrate the same number of counts in both cases, i.e., 10^7 counts.

In the final experiment, we used our asynchronous TCSPC system to record the fluorescence signal of a piece of paper. Figure 5.17 shows the calibrated histograms of the recorded signal obtained after applying the bin-by-bin and the Matrix calibration methods. This figure demonstrates that the Matrix calibration introduced lower noise than the bin-by-bin method.

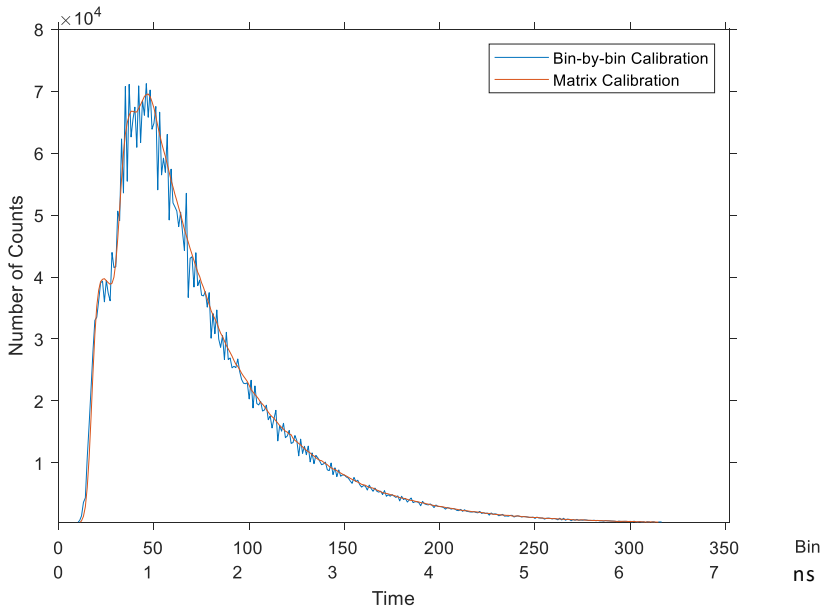


Figure 5.17 Calibrated histograms of the fluorescence signal of a piece of paper obtained with the bin-by-bin and the Matrix calibration methods for asynchronous TDC.

- **Comparison of Processing Speed**

The average-bin-width and the matrix calibration methods produce lower noise than the bin-by-bin method when applied to histograms. However, these methods require more complex signal processing involving more multiplications and data access. The average-bin-width calibration and the matrix calibration are post-processing techniques applied to the raw histogram. Similarly, the bin-by-bin method for asynchronous TDCs is very complicated to implement for online calibration, and it is more convenient to perform it as post-processing. Thus, all the studied calibration methods were implemented as post-processing on the HPS. To compare the different calibration methods in terms of speed, we conducted a set of experiments.

Table 5.3 compares the processing time between the bin-by-bin calibration for asynchronous TDCs and the Matrix calibration. It compares the time of applying the calibration processes without considering the time of creating the LUTs and the calibration tables. The processing time of the bin-by-bin calibration depends on the maximum value of Coarse (*Coarse_max*), since the 3D LUT size is always equal to the product of the number of raw bins in start and stop TDLs and *Coarse_max*, as explained in section 5.2.1. On the other hand, the processing time of the Matrix calibration depends on the total number of calibrated bins in the histogram. Furthermore, the ratio of the processing time of the two methods has an approximately linear relationship with the number of calibrated bins per clock period. This ratio reaches a maximum value of about 8 when the number of calibrated bins per clock period matches the number of raw bins of the TDL.

Table 5.3 Calibration processing speed comparison between the bin-by-bin and the Matrix calibration.

Total Number of Bins in Calibrated Histogram	Number of Calibrated Bins in Clock Period	Maximum Value of *COARSE*	Calibration Process Speed (CPU Tick)		Ratio Matrix/Bin-by-Bin
			Bin-by-Bin Calibration	Matrix Calibration	
4609	256	19	5432528	42229555	7.77
1537	256	7	2207046	18058416	8.18
513	256	3	1197789	9453249	7.89
385	256	2	952921	7493491	7.86
301	200	2	959247	5279525	5.5
226	150	2	936000	3681132	3.93
151	100	2	938161	2501956	2.67
76	50	2	930375	1467180	1.58

We conducted the same experiments on the synchronous TDC and observed that the speed ratio exhibited an approximately linear dependence on the number of calibrated bins per clock period. This ratio reached a maximum value of about 4 when the number of calibrated bins per clock period matched the number of raw bins of the TDL.

5.5 Conclusion

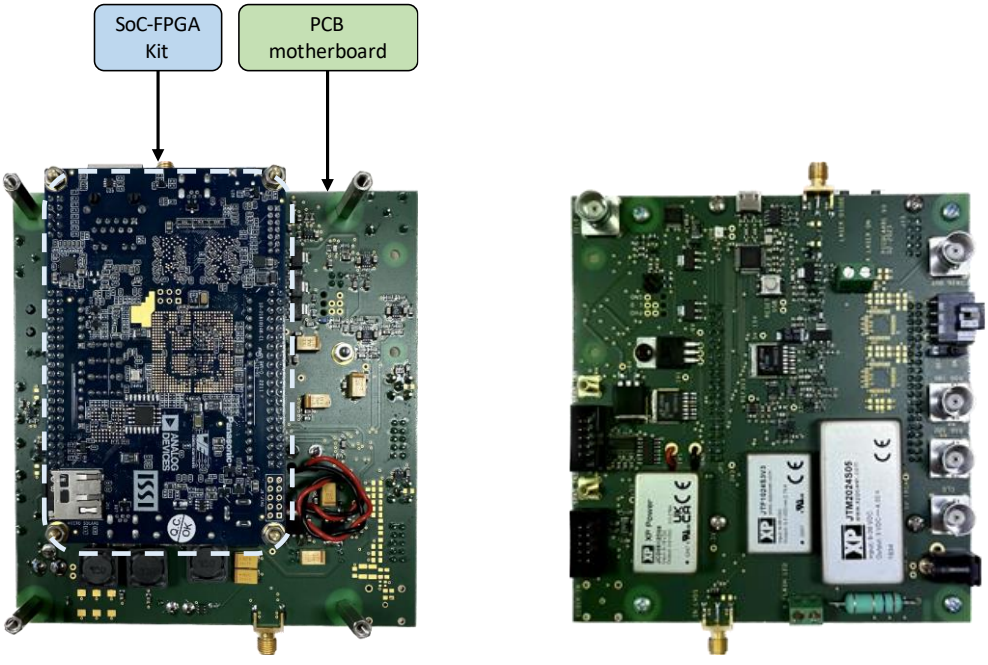
In this chapter, we presented and compared the prevalent calibration techniques for synchronous and asynchronous TDCs, namely the bin-by-bin and average-bin-width methods. We also proposed and evaluated a novel calibration technique for asynchronous TDCs, which we called “Matrix calibration”. We conducted simulations and experiments to assess the performance of these techniques. The results indicated that, for synchronous TDCs, the average-bin-width calibration outperformed the bin-by-bin method, and significantly reduced the DNL and INL of the raw TDC. The results also demonstrated that the proposed method for asynchronous TDCs exhibited high robustness to the DNL of the raw TDC and achieved up to 10 times better results than the bin-by-bin method applied to histograms. However, this improvement entailed a longer calibration time due to the complexity of the signal processing involving more multiplication and memory access instructions.

Given that all our target applications only involve histogram measurements and none of them require single-shot measurements, we opted for the average-bin-width and the Matrix methods as the calibration technique in our TCSPC systems.

Chapter 6: System Characterizations

Figure 6.1 presents a photograph of our TCSPC device, while Figure 6.2 shows the top and bottom views of the system electronics including the Cyclone V SoC-FPGA platform and the custom-designed PCB motherboard, that supplies power to the SoC-FPGA, enables USB communications with the PC, and incorporates a laser diode pulse generator, among other components.

Figure 6.1 Exterior view of the realized TCSPC device.



(a) (b)

Figure 6.2 Electronics of the realized TCSPC system with the Cyclone V SoC-FPGA platform and custom PCB motherboard: (a) bottom view, (b) top view.

In addition, we designed a user-friendly Labview interface that enables the user to configure the system with various parameters and visualize the results in different operating modes. Figure 6.3 illustrates this interface in the single-measurement mode.

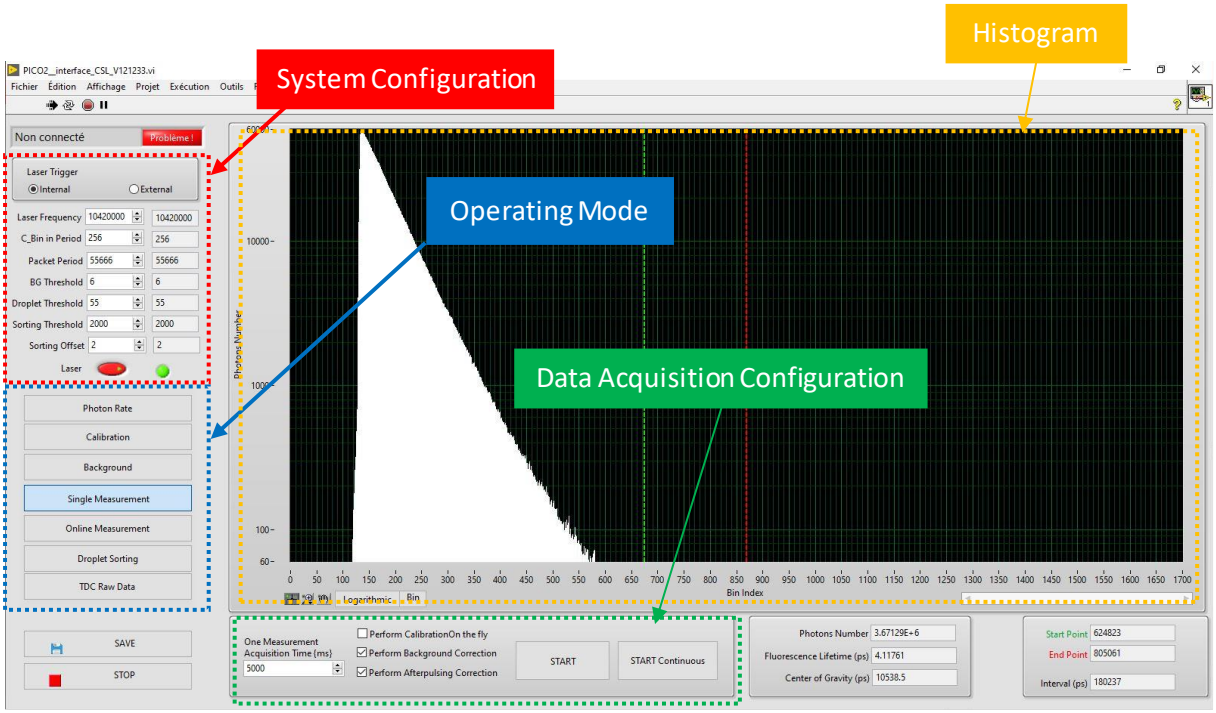


Figure 6.3 Labview interface of the TCSPC system in the single-measurement mode.

This chapter evaluates the temporal performance of the realized system in various aspects. Furthermore, it evaluates the performance of the TCSPC system by measuring its instrument response function (IRF) and recording the fluorescence signal of a reference fluorophore and estimating its fluorescence lifetime.

6.1 TDCs Characterization

There are several parameters that indicate the performance of a TDC. The DNL and INL of our synchronous and asynchronous TDCs were evaluated in the precedent chapter. This section evaluates these TDCs in terms of precision and accuracy by conducting cable delay tests.

6.1.1 RMS Precision

The TDC performance is affected by various error sources, such as the nonlinearity and the jitter, that introduce variations in the single-shot measurements of a constant time interval. Therefore, to mitigate these errors, we measure the time interval several times and compute the

average of all the measurement counts as the final results. Precision is an essential parameter that indicates the quality of the TDC. It is characterized by the standard deviation or the root mean square (RMS) of the results obtained when the TDC performs multiple measurements of a fixed time interval. To evaluate the RMS precision of our TDCs, we conducted cable delay tests using the synchronous and asynchronous TDCs.

- **Synchronous TDC**

In order to ensure a constant time interval (*TI*) between the *START* and *STOP* signals during the measurement, we output the *START* signal generated by the system via a BNC connector and connected it to the SMA connector of the *STOP* signal input using a coaxial cable. Therefore, the time interval between the *START* and *STOP* signals is determined by the length of the cable. We first performed a measurement with 2.5 M counts using a cable of a certain length and calibrated the resulting histogram with the average-bin-width method. Figure 6.4 shows the measurement calibrated histogram, where the mean value of 957.7 ps represents the measured time interval, and the standard deviation of 24.3 ps indicates the RMS precision.

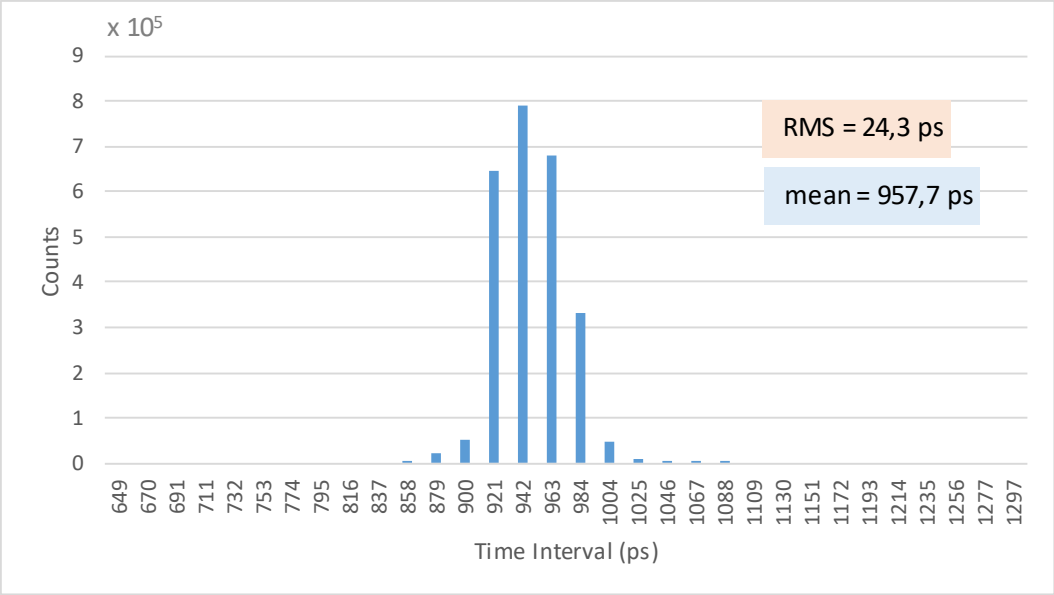


Figure 6.4 RMS precision of the synchronous TDC.

To evaluate the overall precision of the TDC, we used cables with different lengths to introduce different time intervals ranging from 0 to about 6 ns, which corresponds to the TDL measurement range, and measured the RMS precision at these time intervals. We only considered the TDL measurement range because the TDC precision is determined by the

precision of its fine time measurement. The results indicate that the precision across the studied range varies from 23 to 27 ps, depicted in Figure 6.5.

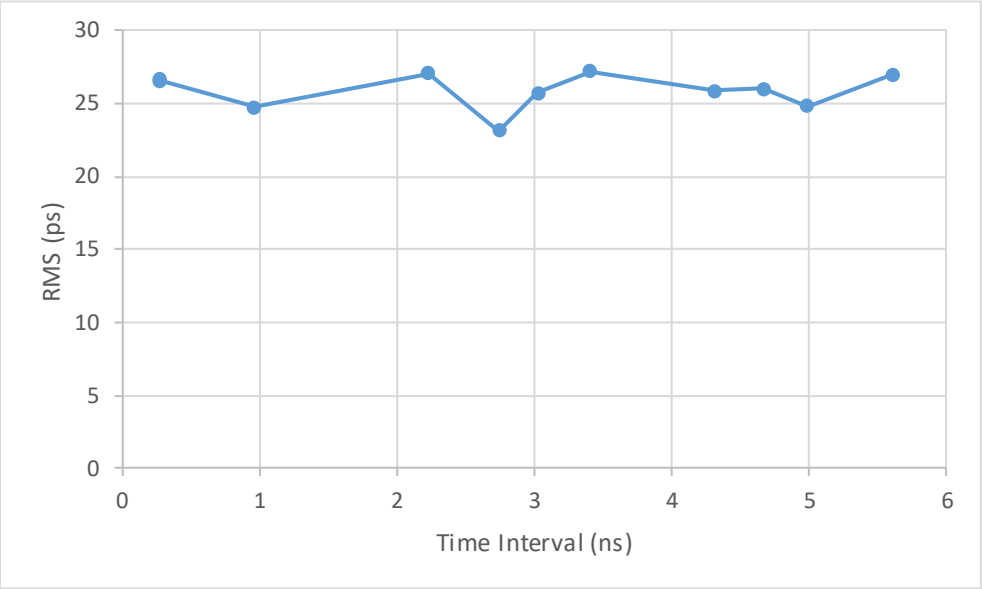
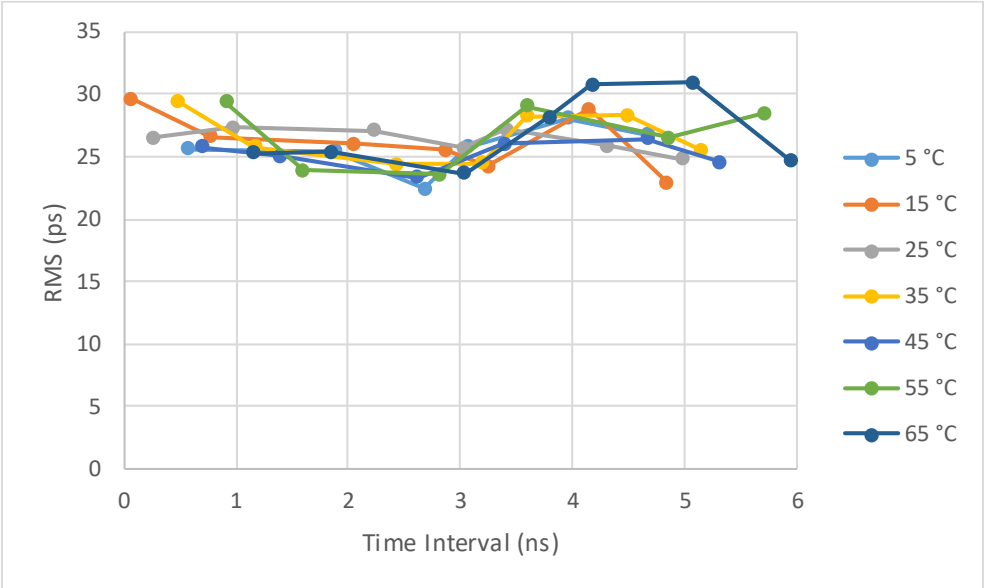


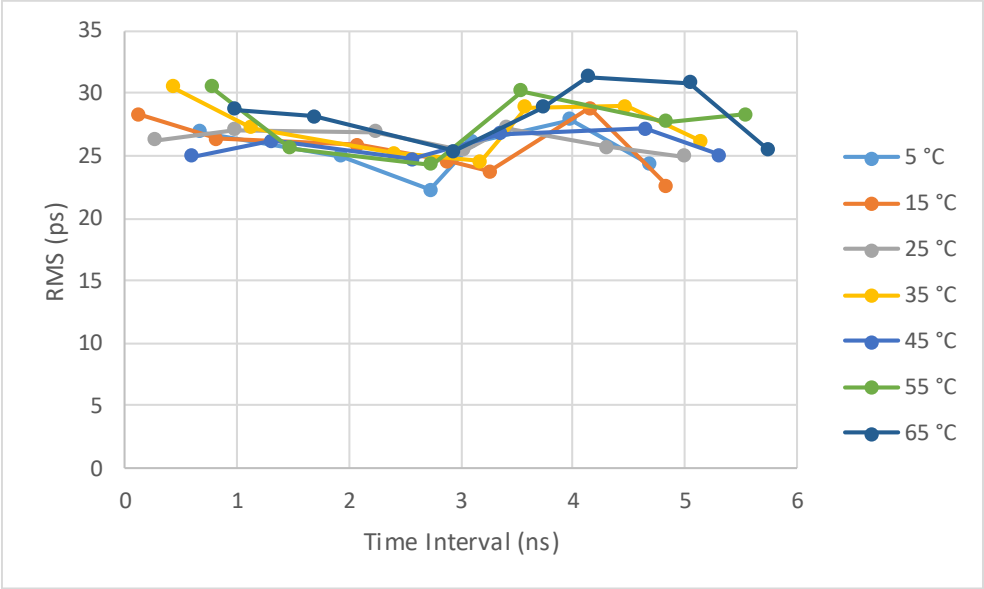
Figure 6.5 RMS precision variation across the TDL measurement range.

The cable delay tests reported above were conducted at room temperature. To investigate the effect of temperature on the TDC precision, we performed the same tests at different ambient temperatures ranging from 5 to 65 °C with a 5 °C increment. The TCSPC system was placed inside a thermal chamber and allowed to reach thermal equilibrium for about 15 minutes at each temperature. Then, we performed cable delay tests with different cable lengths and evaluated the RMS precision with different time intervals and at different temperatures. Moreover, to assess the efficacy of the used calibration method, we applied two types of calibration to the raw histogram of each measurement: one using a calibration table constructed at 25 °C, referred to as the base calibration, and another using a calibration table updated at the measurement temperature, referred to as the online calibration. Figure 6.6-a presents the RMS precision results obtained with the base calibration, and Figure 6.6-b shows the results obtained with the online calibration.

These results demonstrate that the TDC has almost the same overall RMS precision at the different temperatures. The results also show that the overall RMS precision with the base and online calibration is comparable, which confirms the robustness of the average-bin-width calibration and its low sensitivity to temperature variations in terms of RMS precision.



(a)



(b)

Figure 6.6 Overall RMS precision of synchronous TDC at different ambient temperatures: (a) with base calibration table, (b) with online calibration.

- **Asynchronous TDC**

In order to evaluate the RMS precision of the asynchronous TDC, we used an arbitrary function generator (AFG3252 from Tektronix) to generate the *START* and *STOP* signals at a frequency of 10.42 MHz, which corresponds to the pulsed laser frequency used in our target application

of stray light characterization. This frequency also avoids correlation with the TDC clock, ensuring the correct functioning of the asynchronous system. The square signals produced by the generator were fed to the *START* and *STOP* inputs via SMB connectors connected to the IO pins of the development board. The time interval between the rising edges of these signals was varied by adjusting the phase difference between the output channels. First, we set an arbitrary phase difference and measured the resulting time interval with approximately 2.6 million counts. Figure 6.7 shows the measurement histogram. The mean value was 9.044 ns and the RMS precision was about 34 ps.

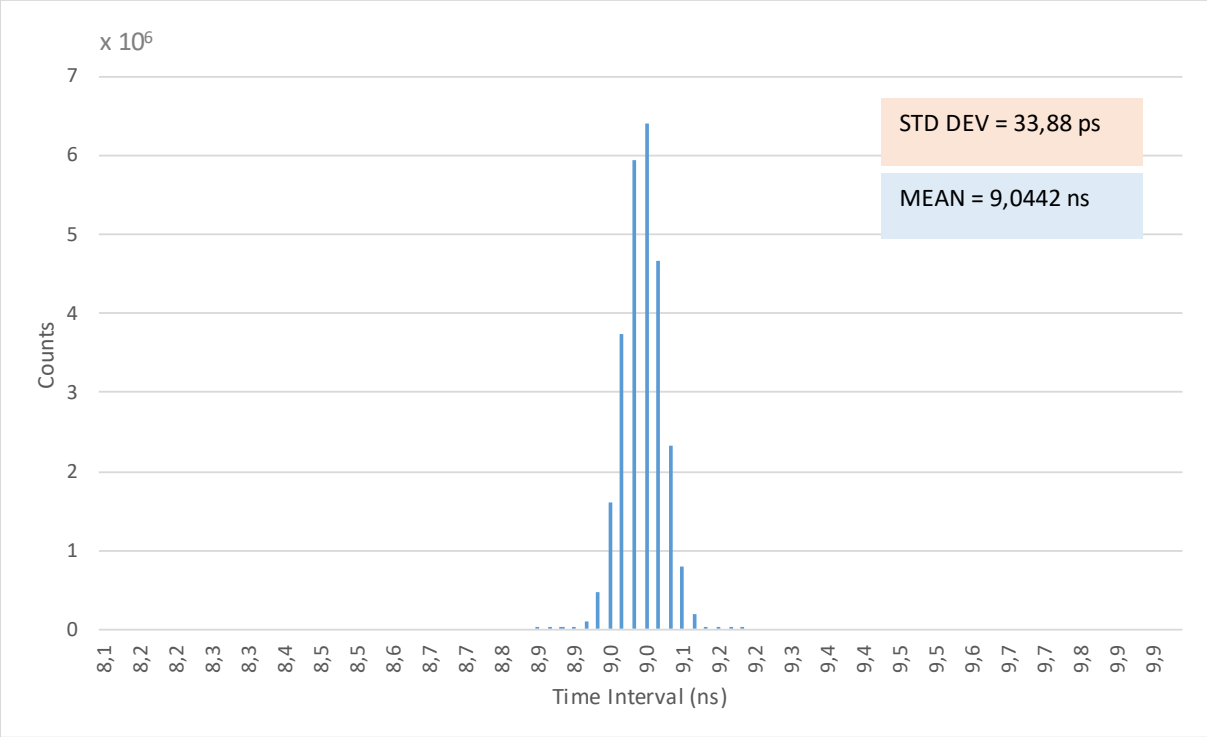


Figure 6.7 RMS precision of asynchronous TDC.

To assess the RMS precision across the entire range of the TDC, which corresponds to the *START* signal period, we varied the time interval between the *START* and STOP *signals* from 0 to 95.97 ns in steps of 533.16 ps, by changing their phase difference in increments of 2 degrees. For each time interval, we performed a measurement with approximately 2 million counts at room temperature and computed the RMS precision. Figure 6.8 depicts the RMS precision as a function of the time interval. It varies between 28 ps and 47 ps with an average value of about 38 ps. The results show a fluctuation in the RMS precision that might be due to the ground noise in the FPGA. Moreover, we observed a slight increasing trend of the RMS with the time interval. This trend may be attributed to the noise at the *START* and *STOP* signals input. In fact, when a signal pulse reaches the FPGA IO, it causes a current draw from the power supply

through parasitic inductances, which results in a drop in the VCCIO voltage proportional to ($L \times \frac{di}{dt}$). This noise affects the propagation time from the IO to the TDL, and the noise of each signal has a different influence on the propagation time of the other one. Assuming that the influence of the *STOP* signal is stronger than that of the *START* signal: For short time intervals, the *START* pulse arrives earlier than the *STOP* pulse and its propagation time is hence less influenced by the noise induced by the STOP pulse. Conversely, for long time intervals, the preceding *STOP* pulse arrives just before the *START* pulse and induces a significant noise that strongly affects the propagation time of the *START* signal, which causes an increase in the RMS precision. However, to consider the worst-case scenario regardless of the noise source causing the observed trend, we conducted the remaining characterization experiments within a range from about 80 ns to 95.97 ns, i.e. about three times the TDL range at the end of the total measurement range of the coarse-fine TDC.

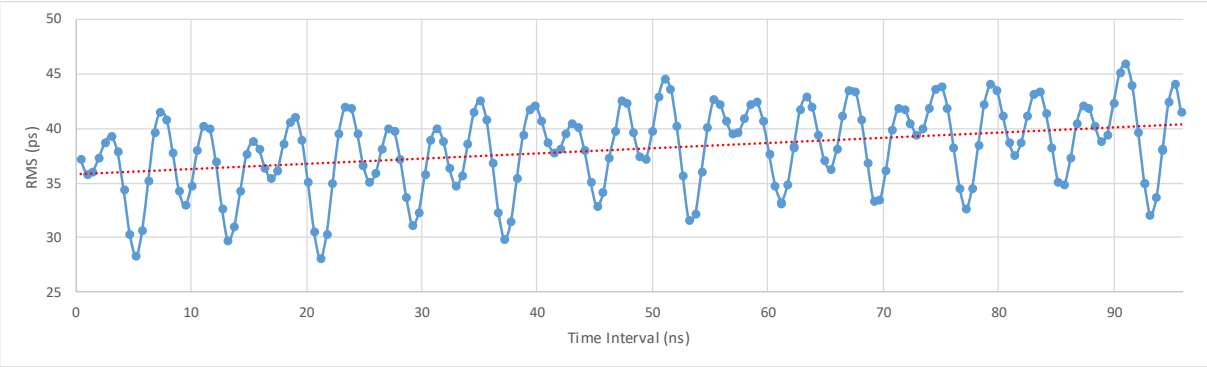


Figure 6.8 RMS precision of asynchronous TDC as a function of the time interval.

We also investigated the effect of temperature on the RMS precision by repeating the same previous measurement at different temperatures varying from 5 to 75 °C with a step of 5 °C. For the base calibration, the Matrix calibration tables were generated at 25 °C by exposing the SPAD to ambient light. In addition, new calibration tables were generated at each temperature for online calibration. The phase difference between the wave generator channels was changed to produce time intervals covering the range from 80 ns to 95.97 ns, and the RMS precision was evaluated for each time interval at the different temperatures. Figure 6.9 presents the results obtained using base calibration. The RMS precision degrades with the temperature variations. Moreover, the degradation is not uniform across the measurement range at a given temperature but rather exhibits an oscillating or periodic hyperbolic pattern with a period equal to the TDC clock period. This phenomenon can be explained by the influence of temperature variation on the propagation time through the TDL, and consequently on the DNL and INL, which introduce

errors in the fine values when calibration tables generated at a specific temperature are applied to calibrate a measurement histogram acquired at a different temperature. The resulting integral error, referred to as the calibration error, increases with the number of propagated delay elements (DEs) of the TDL. Thus, this error is inversely proportional to the bin order, which corresponds to the number of non-propagated DEs as illustrated in Figure 3.10. The hyperbolic pattern results from the differential calibration error of the *START* and *STOP* signals as a function of the measured time interval. When the time interval is close to zero or to the clock period (or its multiples), the *START* and *STOP* signals have close fine values and thus close calibration errors. These errors will be canceled out when the arrival time of the *START* signal is subtracted from that of the *STOP* signal resulting in a low degradation in the RMS precision. When the time interval is half the clock period (or a number of clock periods plus a half), two possible scenarios arise: 1) both the *START* and *STOP* signals have identical *Coarse* values and the fine value of the *STOP* signal is larger by a half period. In this case, the differential calibration error is negative. 2) The *STOP* signal has a higher *Coarse* value than the *START* signal. In this case, the fine value of the *STOP* signal is smaller than that of the *START* signal resulting in a positive differential calibration error. Consequently, the large disparity in calibration error obtained for the same time interval leads to a significant degradation of the RMS precision.

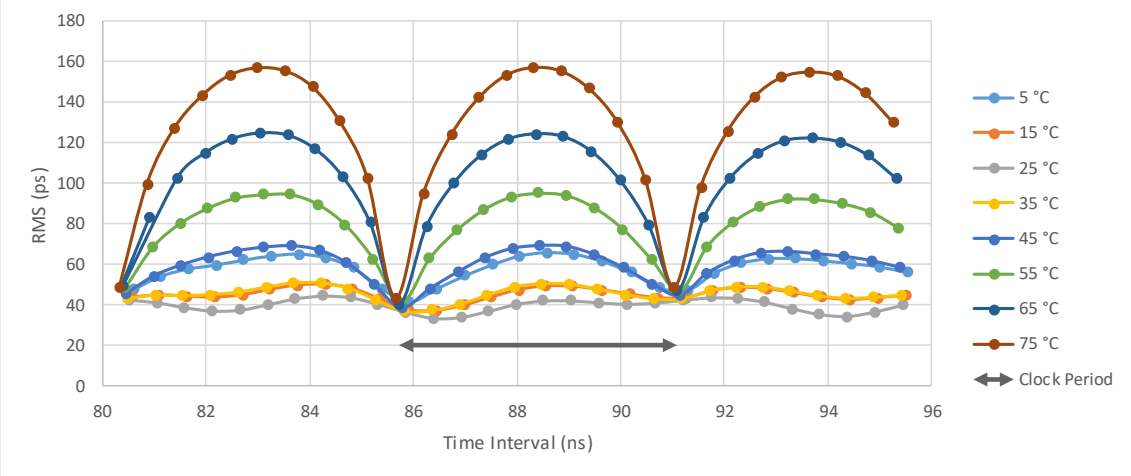


Figure 6.9 Overall RMS precision of asynchronous TDC at different ambient temperatures with base calibration.

Figure 6.10 illustrates the RMS precision obtained by applying online calibration, in which the calibration tables are generated at the measurement temperature. The results show that online calibration improves the RMS precision compared to the base calibration. Moreover, the RMS precision achieved with online calibration exhibits a consistent profile and almost stable values

at the different temperatures. Therefore, frequent updating of the calibration tables is highly recommended to minimize the calibration error and the loss of precision.

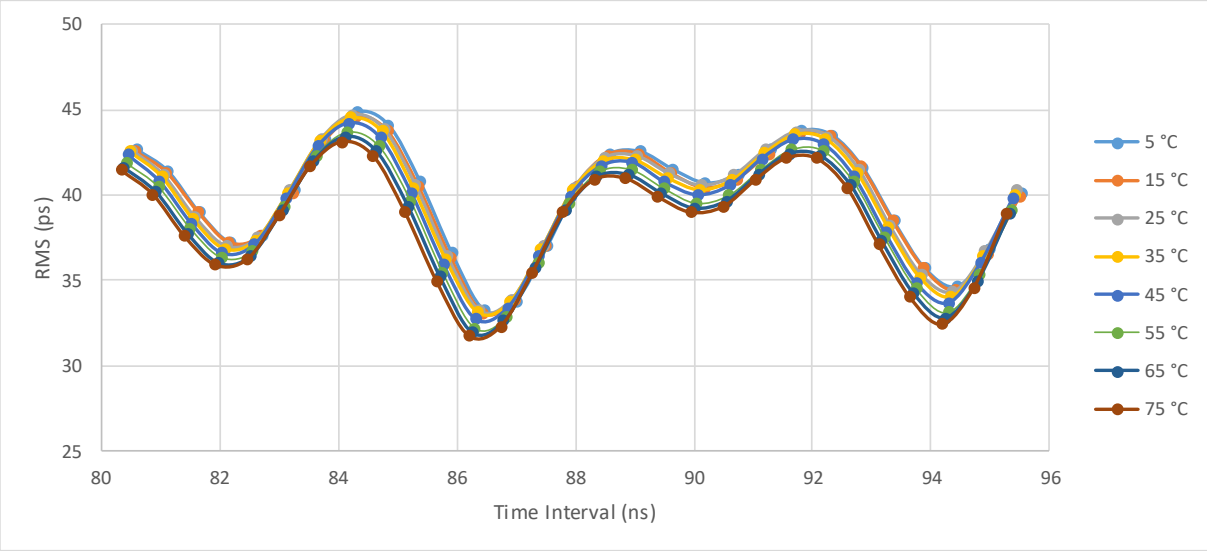


Figure 6.10 Overall RMS precision of asynchronous TDC at different ambient temperatures with online calibration.

6.1.2 Accuracy

Accuracy is another key parameter of TDCs that indicates the correctness of the measurement. It is determined by the deviation between the measured time interval and the actual interval. The accuracy of the proposed TDCs was assessed by conducting cable delay experiments.

- **Synchronous TDC**

As the exact delay introduced by a cable cannot be precisely determined, it is impossible to evaluate the absolute accuracy of the synchronous TDC. Nevertheless, we assessed the accuracy variation with temperature by comparing the mean time intervals measured at different temperatures with those measured at 25 °C. We used cables with various lengths to introduce different time intervals spanning the TDL measurement range of approximately 5.26 ns, and measured these intervals at different temperatures ranging from 5 to 65 °C in increments of 5 °C using both base and online calibration. For each cable length, we computed the deviation of the mean time interval obtained at different temperatures from the reference values obtained at 25 °C. Figure 6.11 depicts the obtained results.

The solid curves in Figure 6.11 show the deviations obtained with base calibration. These deviations can be attributed to offset and gain errors that arise from the temperature variation.

The offset error can be explained by the temperature-dependent delay of the measured signal (the STOP signal) routing path from the IO pin to the beginning of the TDL. When the temperature increases this delay increases causing an offset error in the measured time interval. The gain error, on the other hand, corresponds to the calibration error caused by the variation in the TDC response when the temperature changes. This error is due to the temperature-dependent propagation time along the TDL. It should be pointed out that the last values obtained at 55 and 65 °C deviate from the slope because their time intervals exceed the clock period, and thus they have a coarse value of 1 and small fine values corresponding to lower-order bins with larger calibration errors than the previous points. Moreover, the noise on the curves results from mechanical factors. It can be explained by the variation in the actual delay of the same cable when it was disconnected and reconnected during the experiments, as different cables were used to generate different time intervals at each temperature. These variations were also observed by repeatedly measuring the delay of the same cable using a benchtop oscilloscope (WaveMaster 813Zi from Teledyne LeCroy), after disconnecting and reconnecting it.

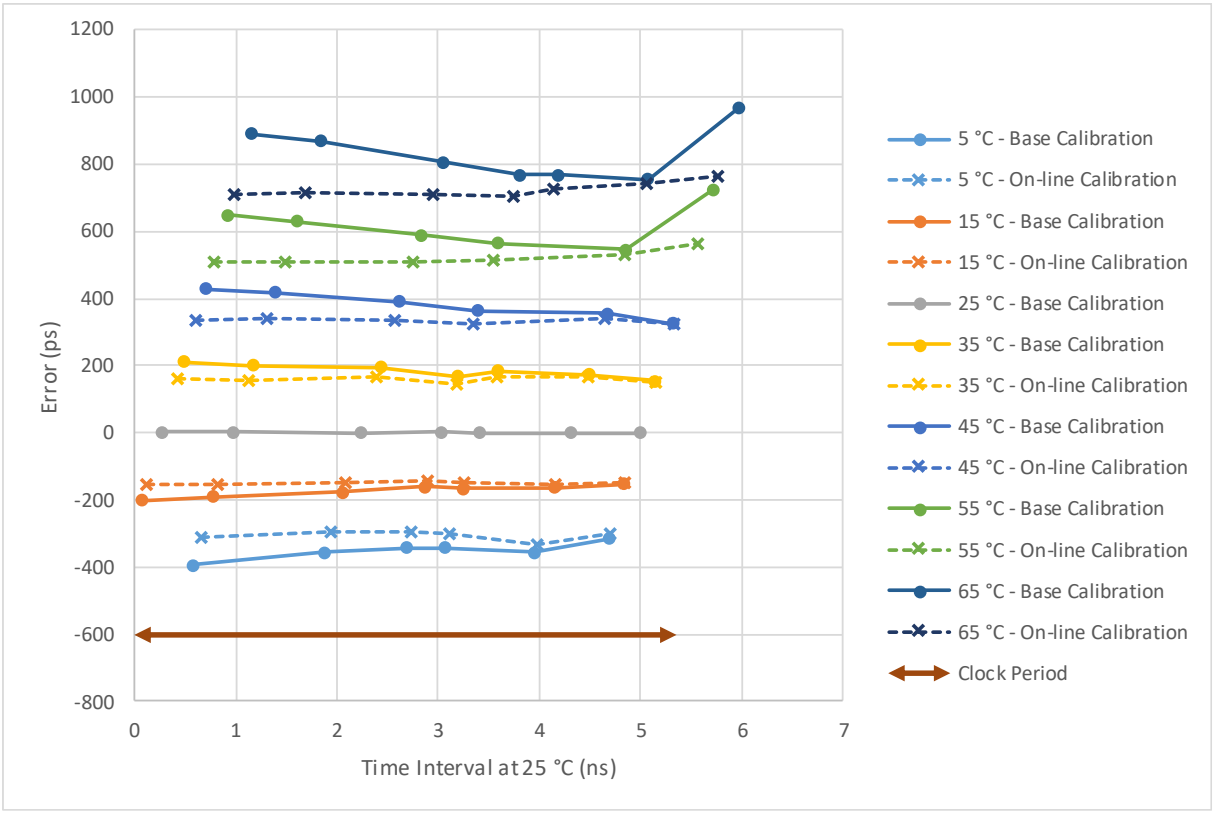


Figure 6.11 Deviations in the mean time interval measured by the synchronous TDC for different cable lengths and at different temperatures. The solid curves show the results obtained with base calibration, while the dotted curves show the results obtained with online calibration.

The dotted curves in Figure 6.11 show the deviations with online calibration, which have no gain error because the calibration error is eliminated by updating the calibration tables. However, online calibration does not reduce the offset error as the latter is independent of the calibration process.

The gain error, or the calibration error, can be effectively isolated by calculating the differences between the measurements obtained with base calibration and those obtained with online calibration, i.e. the error caused by base calibration. Figure 6.12 shows the deviations due to the gain error at different temperatures, which exhibit good linearity and indicate that the gain of this error is constant along the TDL at a constant temperature. The gain values, derived from the linear fit of the deviation values, vary from 1.76% to -3.18% as the temperature increases from 5 to 75 °C, as shown in Table 6.1 and plotted in Figure 6.13. This figure illustrates that the gain has a linear relationship with the temperature and can therefore be compensated with a simple correction coefficient using a temperature sensor that monitors the environmental temperature.

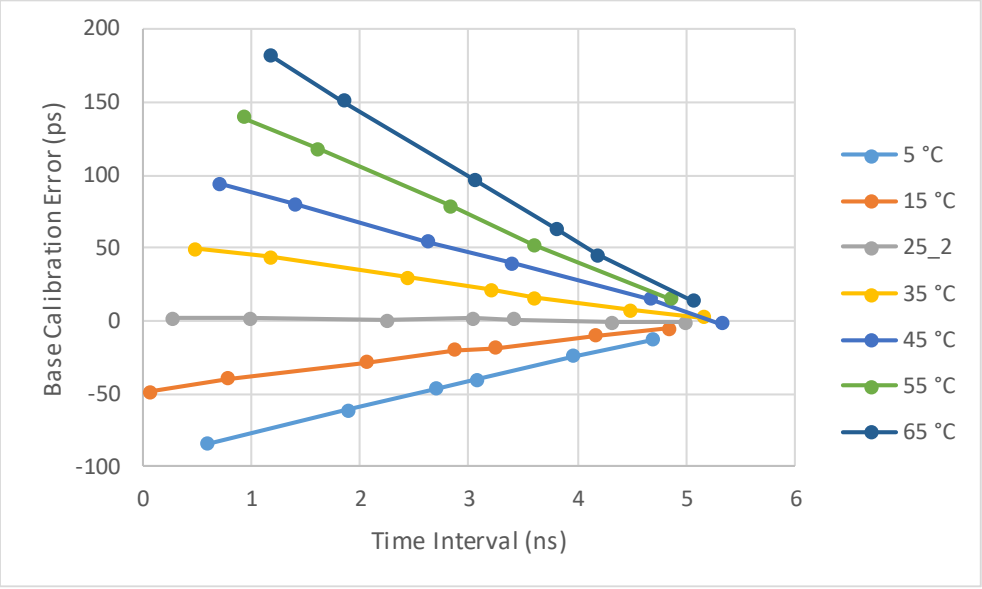


Figure 6.12 Deviations in the synchronous TDC's measurements due to the gain error at different temperatures.

On the other hand, the offset error can be estimated by the difference between results obtained with online calibration and the reference results obtained at 25 °C. Figure 6.14 shows a linear

correlation between the offset error and the temperature. Similar to the gain error, the offset error can be effectively compensated by a temperature-dependent correction coefficient.

Table 6.1 Gain error values in synchronous TDC at different temperatures.

Temperature °C	Slope
5	0,017579
15	0,009021
35	-0,000662
45	-0,010505
55	-0,020487
65	-0,031769

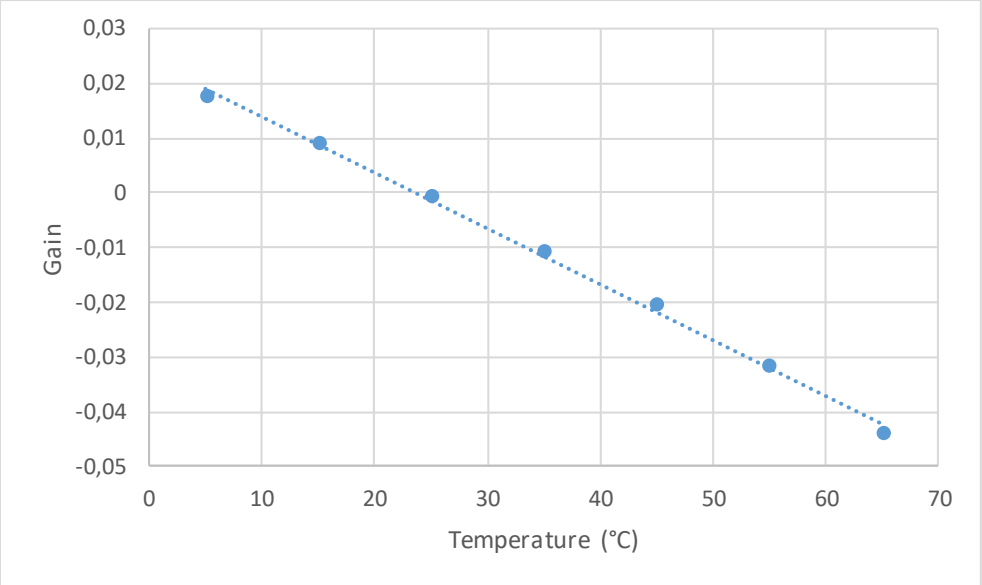


Figure 6.13 Gain values as a function of temperature in synchronous TDC.

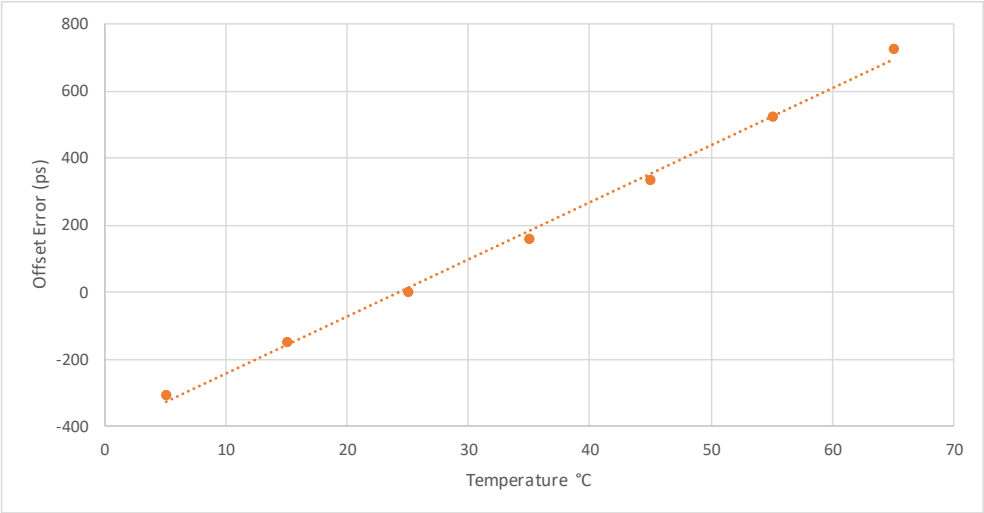


Figure 6.14 Offset error as a function of temperature in synchronous TDC.

- **Asynchronous TDC**

To evaluate the accuracy of the asynchronous TDC, we used the same experimental setup using the thermal chamber and the arbitrary function generator. The actual time interval between the *START* and *STOP* signals provided by the function generator was determined by the phase difference between these signals, with a precision limited by the generator's phase noise. We adjusted the frequency of the *START* and *STOP* signals to 10.42 MHz and varied their phase difference in steps of 4 degrees, corresponding to 1066.33-ps delay steps, to produce time intervals varying from 80 to 92 ns. This range covers approximately twice the clock period at the end of the measurement range, where the TDC exhibited the lowest precision, as explained above in section 6.1.1. For each phase difference, we measured the time interval between the *START* and *STOP* signals and computed the deviations between the obtained values and the expected values from the phase differences. The obtained results, presented in Figure 6.15, show that the deviations have an oscillatory pattern with a peak-to-peak amplitude of less than 6 ps and an offset of about 4.054 ns. The oscillatory pattern is likely due to the ground noise in the FPGA. Whereas the offset corresponds to the offset error of the TDC at 25 °C, which originates from the difference in propagation time of the two signals from the generator outputs to the TDC's tapped delay lines. This offset can be divided into two components: an external component, resulting from the difference in the length of the cables and connectors that link the generator outputs to the IO pins of the FPGA, and an internal component, corresponding to the difference in the routing path of the two signals from the IO pins to the TDLs.

After correcting the offset error, the peak-to-peak deviation along the considered measurement range is less than 6 ps at a constant temperature of 25 °C.

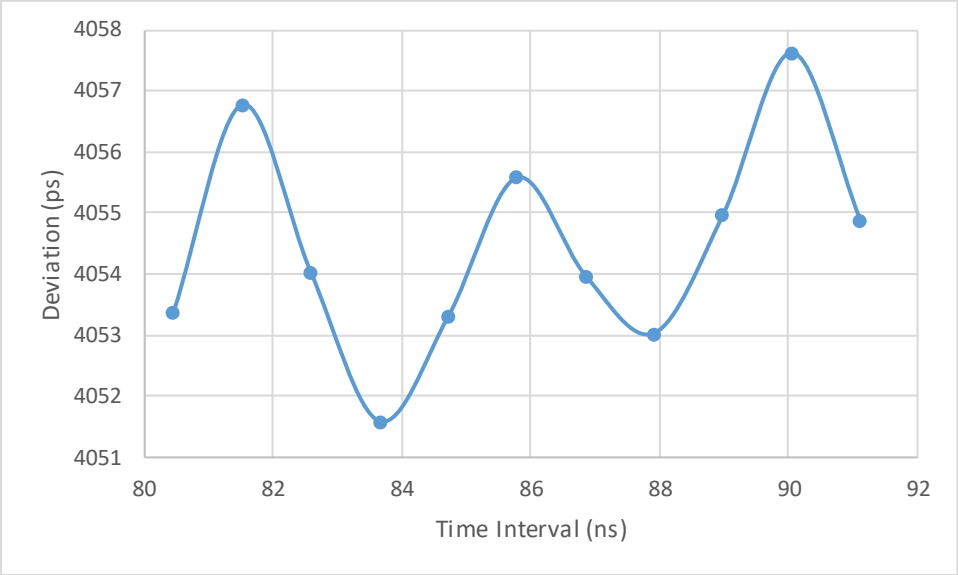


Figure 6.15 Deviation between the measured and expected time intervals using the asynchronous TDC.

To evaluate the influence of temperature on the TDC accuracy, we repeated the accuracy measurements at different temperatures ranging from 5 to 75 °C with a step of 5 °C. We computed the deviation between the measured and the expected value for each time interval at the different temperatures with base and online calibration. Figure 6.16 illustrates the results obtained after compensating the offset error at 25 °C to focus only on the offset variation with temperature. The solid curves represent the deviations obtained with base calibration and the dotted curves represent the deviations obtained with online calibration. The results demonstrate that the TDC exhibits almost the same accuracy profile across the studied range at the different temperatures, with an offset that increases with the temperature. A comparison of these results with the synchronous TDC results presented in Figure 6.11 shows that the asynchronous TDC is more stable and achieves higher accuracy than the synchronous TDC with about 5-fold lower offset error and negligible gain error.

As explained above, the synchronous TDC, which employs a single TDL, is susceptible to a large offset error caused by the temperature-dependent delay of the measured signal (the STOP signal) along the routing path from the IO pin to the TDL entry point. However, this error is substantially mitigated in the asynchronous TDC, which utilizes two TDLs and measures the difference between the arrival times of the two signals. These signals exhibit similar variations

in the delay from the IO pin to the TDL entry point which are eliminated when computing the difference between their arrival times. Therefore, only the variation corresponding to the initial difference in propagation time of the two signals induces the offset error.

The reduction in the gain error caused by the calibration error can also be explained by the same reasoning. In the asynchronous TDC, the calibration errors of the two TDLs are also canceled out by the subtraction of the arrival times of the two signals. When the time interval is measured repeatedly, since the *START* and *STOP* signals are both asynchronous to the TDC clock, the resulting calibration error may have positive or negative values depending on the fine values of the *START* and *STOP* signals, as discussed in section 6.1.1. Therefore, the final error will be considerably reduced when calculating the mean value of the measurement.

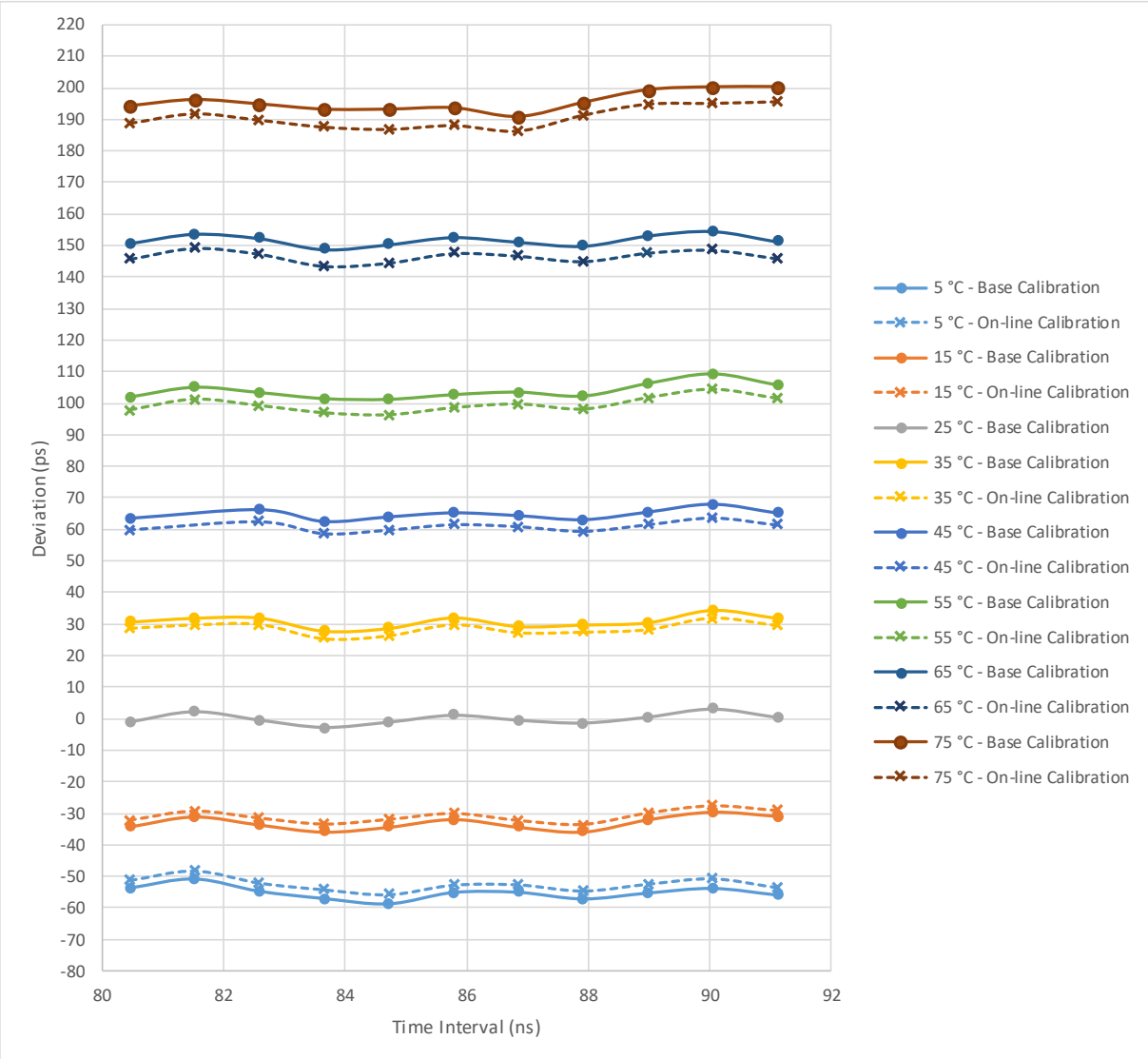


Figure 6.16 Deviation between measured and expected time intervals at different temperatures for the asynchronous TDC with base and online calibration.

Figure 6.17 illustrates the offset error variations with temperature, which are derived from the difference between the mean offset errors at the different temperatures and those obtained at 25 °C. This figure indicates that the offset error has a quadratic relationship with temperature.

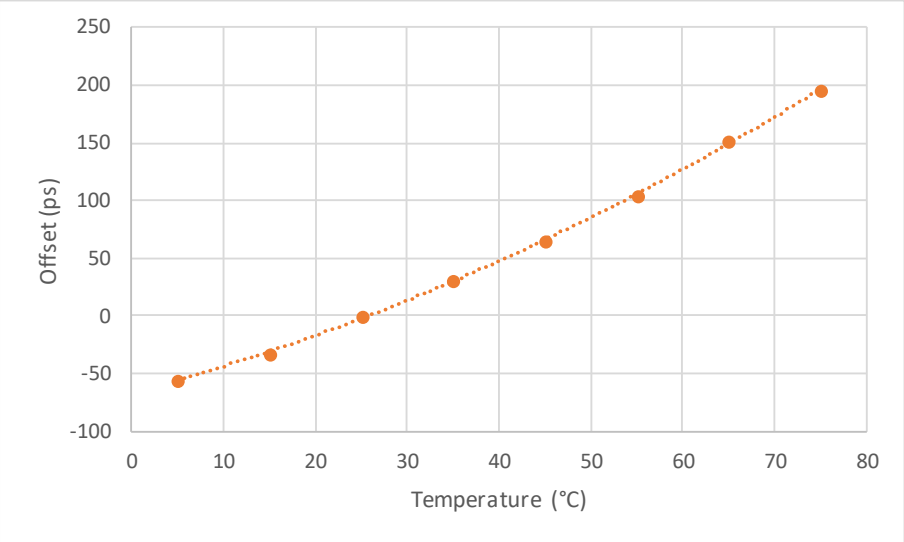


Figure 6.17 Offset error as a function of temperature in the asynchronous TDC.

- **On-the-Flight Calibration in Asynchronous TDCs**

Although the asynchronous TDC accuracy achieved with base calibration is comparable to that obtained with online calibration, it is highly recommended to frequently update the calibration tables to prevent precision degradation, as demonstrated in section 6.1.1, where using the base calibration resulted in a large decrease in the RMS precision.

One advantage of the asynchronous TDC over the synchronous one is that both the *START* and *STOP* signals are uncorrelated with the system clock, which allows the creation of the Matrix calibration tables from the measurement counts. The coarse values of the measurement counts are discarded and the fine values of the asynchronous *START* and *STOP* signals are used to generate the DNL histograms of the two TDLs and then to construct the calibration tables, similarly to the code density test. The calibration tables are then applied to calibrate the measurement histogram that is constructed from the coarse and fine values of the measurement counts. This technique, called on-the-flight calibration, ensures the precise calibration of the TDC and avoids the calibration error due to temperature variations since the calibration tables are updated in real time.

The TDC accuracy was evaluated at 25 °C with SPAD-based and on-the-flight calibration techniques. First, base calibration tables were generated from the arrival time of the ambient light photons by connecting the *STOP* signal to the SPAD that was exposed to ambient light. Next, the *STOP* signal was provided by the function generator and its phase relative to the *START* signal was varied to generate different time intervals. For each phase difference, the mean measured time interval was computed twice: once with the base calibration tables and once with the on-the-flight calibration tables that were generated from the measurement counts. As shown in Figure 6.18, the difference between the results obtained by the two calibration techniques is less than 1 ps across the studied measurement range.

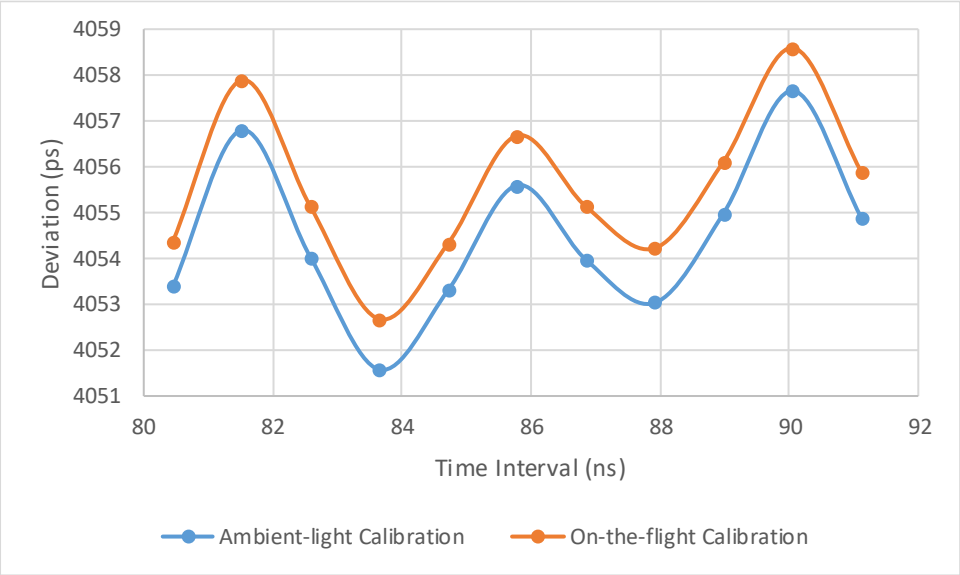


Figure 6.18 Comparison of the mean measured time intervals obtained by SPAD-based and on-the-flight calibration techniques.

6.1.3 Temperature Influence on the TDL Propagation Time

The previous experiments demonstrated that temperature variation significantly affects the precision, accuracy, and overall performance of the TDC. This effect arises from the change in signal propagation speed in the FPGA due to the temperature variation. To better understand the temperature influence on the TDL, we conducted code density tests at various temperatures and analyzed the variations in the propagation time through the delay elements (DEs) with temperature. In practice, the propagation time of the DE increases with temperature. Hence, the number of effective DEs required to cover the clock period decreases as the temperature rises, and the number of non-effective DEs increases. The non-effective DEs results result in dead bins at the beginning of the DNL histogram, as shown in Figure 6.19.. Figure 6.20 depicts the average bin width of the TDL as a function of temperature.

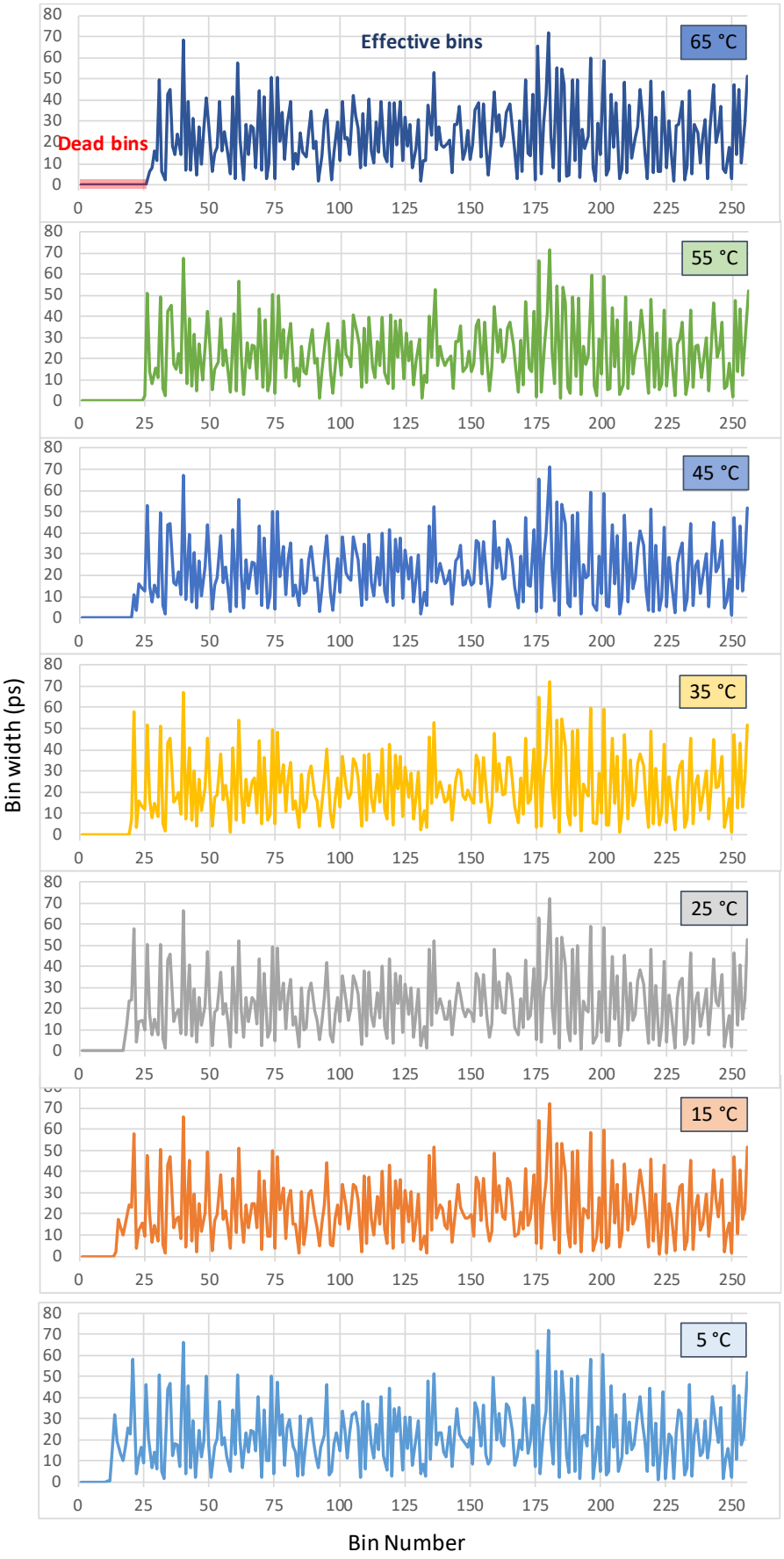


Figure 6.19 TDL bins' widths at different temperatures.

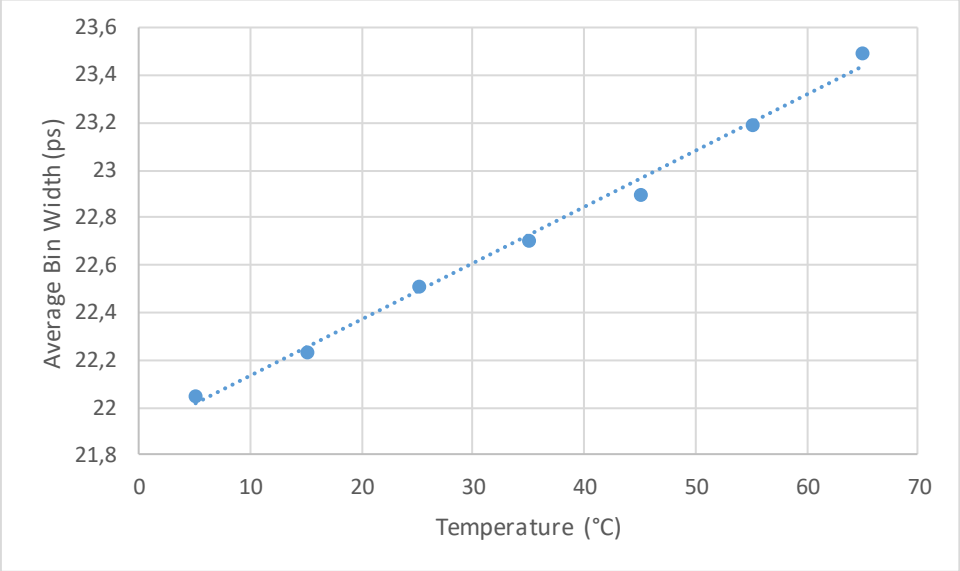


Figure 6.20 Average bin width of the TDL as a function of temperature.

6.2 TCSPC System Characterization

To evaluate the performance of the proposed TCSPC system, we measured its instrument response function system (IRF) by directly recording the excitation light pulse. We also used the system to record the fluorescence signal of a reference fluorophore and estimate its fluorescence lifetime.

6.2.1 IRF Measurement

In this experiment, we used a laser diode with a wavelength of 405 nm, coupled with a high-repetition-rate pulse generator [50], as the excitation light source. We recorded the shape of the generated light pulses by focusing the laser beam on the SPAD. To prevent the SPAD saturation and the distortion of the measured signal, we oriented the laser beam obliquely on the detection area of the SPAD such that the detected photon rate was about 2% of the laser repetition rate. Figure 6.21 shows the recorded signal. The laser pulse had a full width at half maximum (FWHM) of about 130 ps and a full width at 1% of maximum of 630 ps. The noise around the peak contributes only about 0.6% of the signal intensity. Figure 6.22 shows the signal after eliminating this noise by considering only the bins with photon counts greater than 0.1% of the maximum.

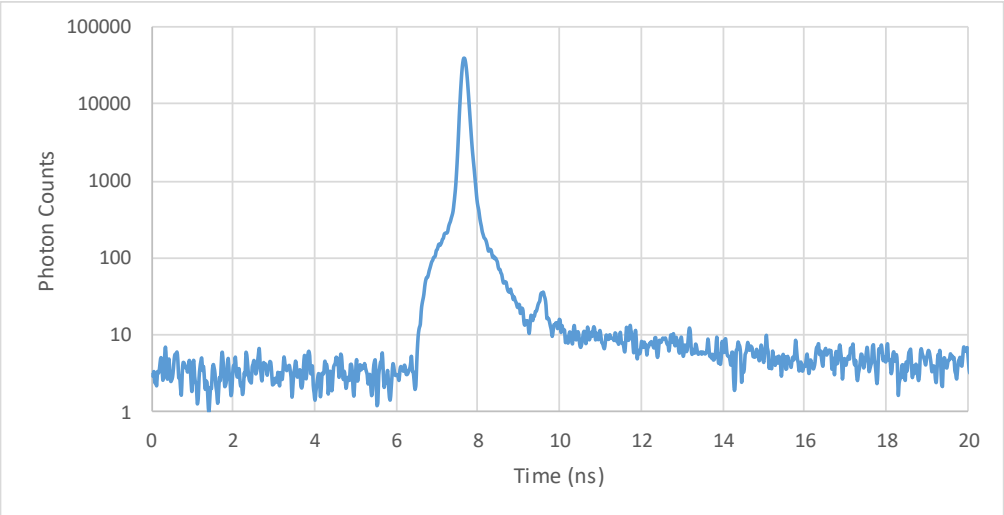


Figure 6.21 Instrument response function of the synchronous TCSPC system using a 405-nm pulsed laser diode. The FWHM is about 130 ps.

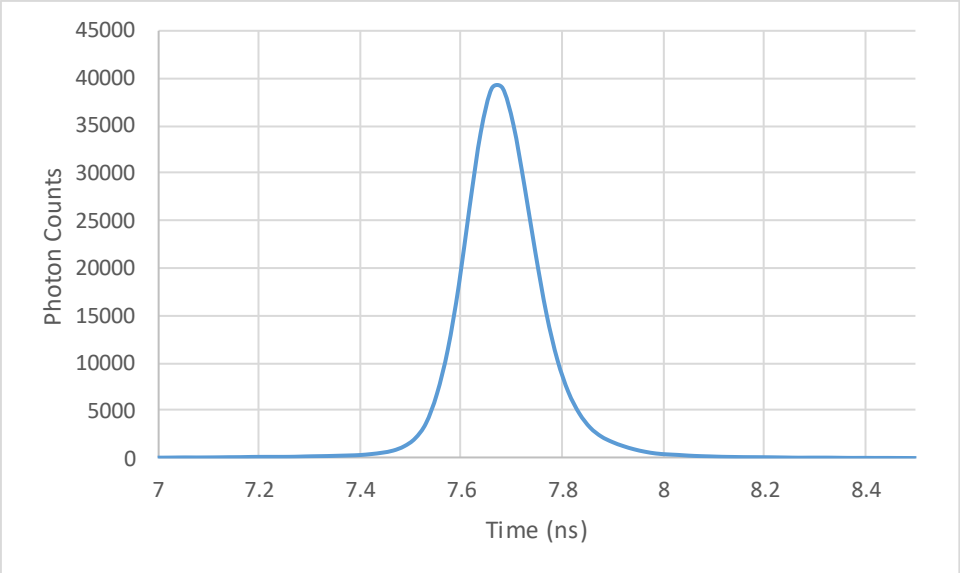


Figure 6.22 Laser pulse signal after excluding the noise caused by ambient light.

6.2.2 Florescence Lifetime Measurement

We used our synchronous TCSPC system to measure the fluorescence signal of a reference fluorophore, which is the sodium fluorescein (NaFl), which is a standard fluorophore with good photostability and a monoexponential decay. Its fluorescence lifetime depends on the pH and its value is about 4090 ± 50 ps at a pH of 7.4 [171, 172]. In this experiment, we measured the fluorescence signal of NaFl in phosphate-buffered saline (PBS)whose pH value is 7.4. For excitation light, we used a 405-nm laser diode as the excitation light source, despite its suboptimal overlap with the absorption spectrum of fluorescein. We adjusted the detected

photon rate to be about 14% of the laser repetition rate and performed a measurement with 1.18 M photon counts. Figure 6.23 shows the recorded signal, which demonstrates a clear monoexponential decay. We applied the maximum Likelihood-based algorithm to estimate the fluorescence lifetime from the recorded signal. The estimated lifetime was 4.11 ns, which is consistent with the literature value.

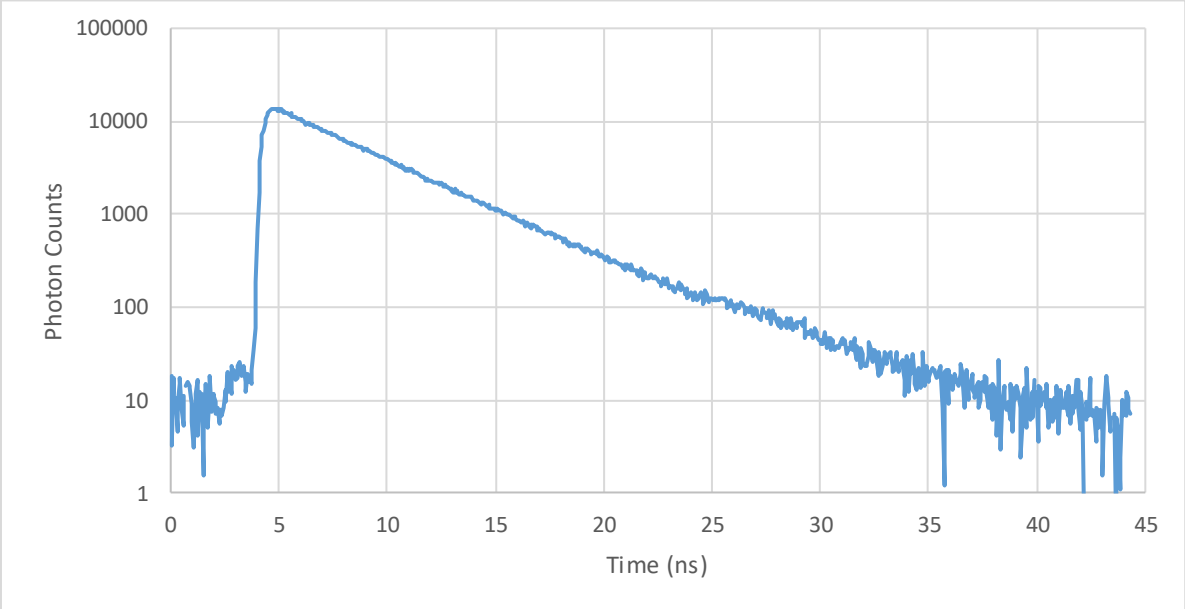


Figure 6.23 Fluorescence signal of Sodium Fluorescein in PBS measured by the synchronous TCSPC system. The estimated FLT is 4.11 ns.

6.3 Conclusion

This chapter presented the evaluation of the temporal performance of our TDC and TCSPC systems. We conducted different experiments to evaluate the precision and accuracy of the synchronous and asynchronous TDCs. We also investigated the effect of temperature on the TDC performance and discussed the different sources of errors caused by the temperature variations and proposed solutions to mitigate these errors, such as the on-the-flight calibration method for asynchronous TDC and the different correction coefficients. Moreover, we measured the IRF of our TCSPC signal and evaluated its performance by recording the fluorescence signal of a reference fluorophore and estimating its fluorescence lifetime.

The results showed that our synchronous TDC achieved an RMS precision ranging between 23 ps to 27 ps across its TDL measurement range of 5.26 ns at different temperatures, while the accuracy had large variations with temperature. The asynchronous TDC achieved an RMS precision ranging from 28 ps to 47 ps and an accuracy deviation of about 6 ps at a constant

temperature across the measurement range. The asynchronous TDC exhibited higher stability and robustness than the synchronous TDC with respect to temperature variations, thanks to its dual-TDL architecture that cancels out most of the temperature-dependent errors. Furthermore, we demonstrated the advantage of updating the calibration tables on the TDC performance and proposed on-the-flight calibration of asynchronous TDC, which generates the calibration tables in real time from the measurement counts.

The TCSPC system exhibited a high temporal resolution acceptable for our target applications with an IRF FWHM of about 130 ps. It also demonstrated a high accuracy in measuring the fluorescence lifetime of fluorescein, which was estimated to be 4.11 ns, in agreement with the literature value.

Chapter 7: Applications

We applied our TCSPC system in four different applications in environmental, biological, and astronomical fields. These applications are: water pollution sensing, time-correlated optical turbidity measurement, real-time high-throughput microfluidic droplet screening and sorting, and stray light characterization. The diversity of target applications enabled the comprehensive verification and debugging of the proposed TCSPC system. Moreover, the specific optimizations and requirements of each application enabled incorporating many enhancements and modifications to improve the system's versatility. For instance, in the optical turbidity measurement, we identified and addressed a laser stability problem, where the laser intensity fluctuated with time likely due to temperature variations. We added electronic and logic circuits to stabilize the laser intensity during the measurement, which improved the system performance. Moreover, in the stray light characterization, a long continuous measurement of two weeks was performed without any problem which proved the system's reliability.

This chapter provides a detailed description of the four target applications.

7.1 Water Pollution Sensor (WPS)

With the growth in industrial and agricultural activities throughout Europe in the past decades, many chemical products end up in the wastewater at some point. However, current wastewater treatment processes cannot remove all harmful chemicals, which then contaminate various water bodies. Therefore, monitoring the quality of drinking water is a major concern for local authorities to prevent adverse effects on human health. Current methods for pollutant detection rely on manual sampling of drinking water and laboratory analysis. Moreover, the sampling frequency depends on the water consumption rate and the population density in the supplied area. For instance, in areas with low population density, the drinking water quality is only monitored sporadically. Therefore, increasing the sampling frequency and reducing the decision time are urgent needs for public health. Several technologies have already been investigated to detect micropollutants in drinking water, but none of them meet all the requirements of autonomous in-situ monitoring. The Water Pollution Sensor (WPS) is an interregional project that aims to develop an autonomous device for the continuous

measurement of the concentration of chemical micropollutants at very low concentrations (of the order of some µg/L or ng/L) in drinking water, which enables the quick alert in case of excessive presence of traces of certain pesticides at certain points in the network. The targeted pollutants are glyphosate, atrazine, and polycyclic aromatic hydrocarbons which are the most common pollutants in the Upper Rhine Region. However, the proposed device is designed to be easily adapted to other pollutants. This device combines three detection methods: pulsed amperometry (PAD), nuclear magnetic resonance (µNMR), and time-resolved fluorescence.

Figure 7.1 depicts a schematic diagram of the proposed device. It consists of three main parts:

1. A Replaceable Concentrator Unit (RCU): the role of this part is to perform selective adsorption and desorption processes to prepare solutions that only contain the targeted pollutants at concentration levels that meet the detection limits of the respective sensing technique.
2. A Measurement Unit (MU): this part includes the three detectors, namely the pulsed amperometry (PAD) detector, nuclear magnetic resonance (µNMR) detector, and the optical analysis system (OAS). These detectors perform measurements on the solutions provided by the RCU to identify and quantify the existing pollutant.
3. A Control and Data Processing Unit (CDPU): this unit drives the entire device. It is responsible for the control of the three detectors, the data processing, and the external communications.

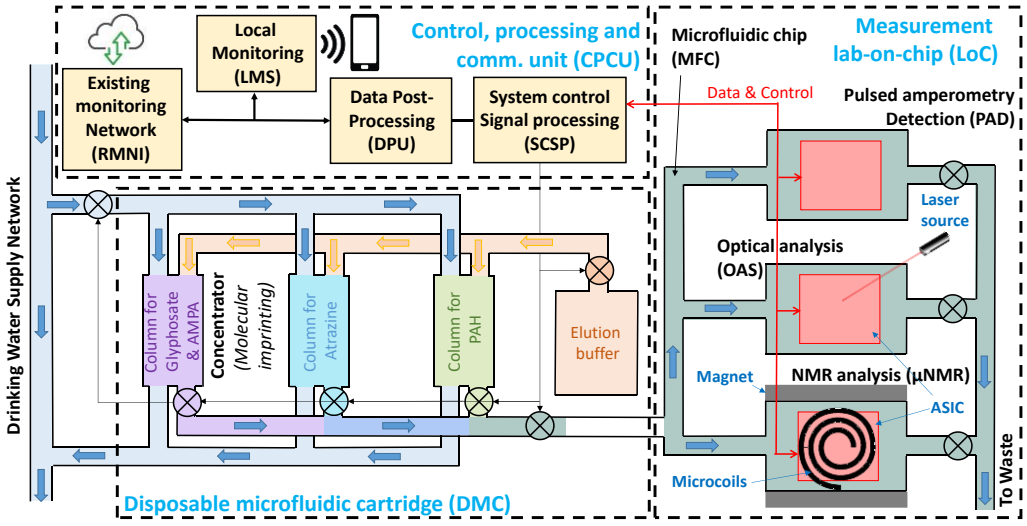


Figure 7.1 Schematic diagram of the water pollution sensor (WPS) device.

Our contribution to this project is the design and implementation of the optical analysis system (OAS).

7.1.1 Optical Analysis System (OAS)

The optical analysis system aims to detect Polycyclic Aromatic Hydrocarbons (PAHs), such as Benzo[a]pyrene (B[a]P), Benzo[a]anthracen, Benzo[a]fluoranthene, and Chrysen. The Council Directive 98/83/EC (CEU, 1998) of the European Union specifies the maximum acceptable concentrations of these PAHs in drinking water: 0.01 µg/L for benzo(a)pyrene and 0.1 µg/L for the Sum of the other PAHs (benzo(b)fluoranthene, benzo(k)fluoranthene, benzo(ghi)perylene, indeno(1,2,3-cd)pyrene).

We implemented the OAS sensor based on our asynchronous TCSPC system. It detects the presence and concentration of PAHs in the water sample by measuring its fluorescence signal. The Fluorescence lifetime value determines the type of PAH, while the fluorescence intensity quantifies its concentration. Figure 7.2 illustrates the schematic diagram of the optical analysis system with the related components of the RCU and CDPU parts.

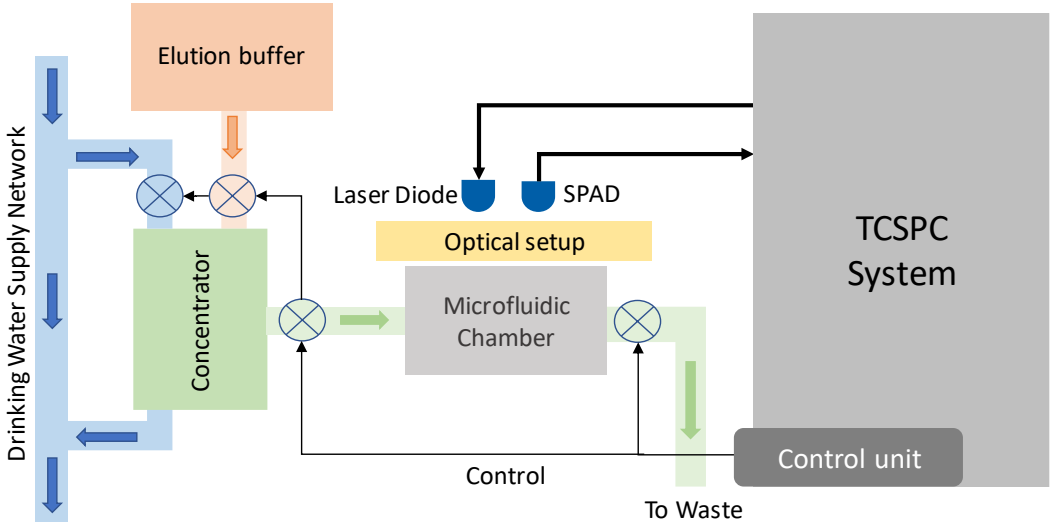


Figure 7.2 Global architecture of the optical analysis system.

7.1.2 Experimental Study

To validate the concept, we carried out preliminary experiments to measure the fluorescence signal of the B[a]P in ethanol. We selected ethanol as a solvent because of the low solubility of the B[a]P in water. Ethanol is also a potential candidate for the elution buffer in the RCU.

We conducted these experiments with B[a]P at a concentration of 750 µM, which is the saturation limit in ethanol. The aim was to start with a high concentration and then progressively decrease it to determine the limit of detection.

A. Experimental Setup

We coupled the TCSPC with an optical setup and a microfluidic circuit. The optical setup comprises optical fibers, mirrors, filters, and a microscope objective. It delivers and focuses the excitation light beam on the water sample, and isolates and focuses the fluorescence signal on the photon detector. The microfluidic circuit includes a microfluidic chamber and a valve that enables the measurement of the background signal to eliminate it from the sample signal.

- **Optical Setup**

Figure 7.3 shows the optical setup. For the excitation light, we used a 405-nm laser diode driven by a high-repetition-rate pulse generator [50]. A pass-band filter (FB405-10, Thorlabs) was used to block unwanted wavelengths from the excitation beam. The beam was then transmitted by a dichroic mirror (DMSP425R, Thorlabs) and focused onto the sample by an objective microscope (N20X-PF, NIKON). The fluorescence emission of the sample was collected by the same objective and reflected by the dichroic mirror orthogonally to the excitation path. The emission signal was then directed to the SPAD by an optical fiber after passing through a notch filter (NF405-13, Thorlabs) and a colored glass filter (C3 3-72) to reject any residual excitation light.

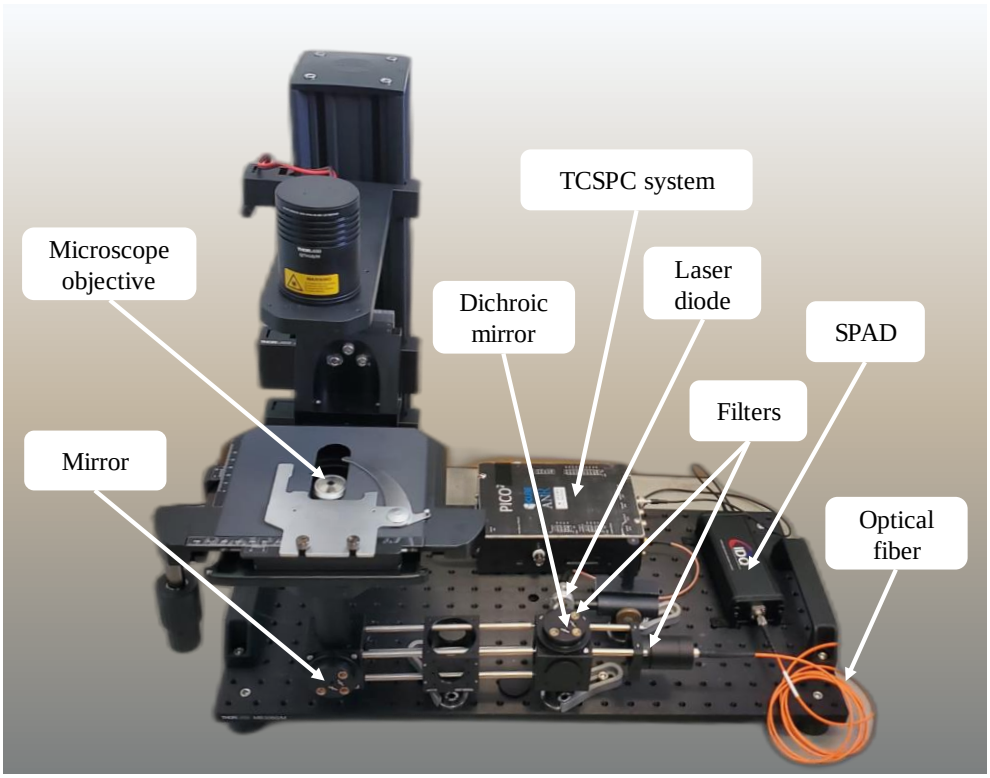


Figure 7.3 Optical setup.

- **Microfluidic Circuit**

To ensure the safe manipulation of the hazardous pollutants under investigation, we incorporated a microfluidic circuit into the OAS. The circuit comprises a microfluidic chamber, syringe pumps, and a waste reservoir, that prevents any contact with the dangerous fluids. Moreover, the circuit facilitates adjusting the pollutant concentration by diluting the base solution with solvent, which can be done by controlling the flow rates of the base solution and the solvent using the syringe pumps. Furthermore, we added a valve to enable the accurate measurement and subtraction of the background signal. The valve allows filling the chamber with the solvent to measure the background signal, and then with the sample containing the solvent and the pollutant to measure the fluorescence signal of the sample. The background signal is then subtracted from the sample signal to obtain a histogram that represents the fluorescence signal of the pollutant. Figure 7.4 depicts a schematic diagram of the microfluidic circuit.

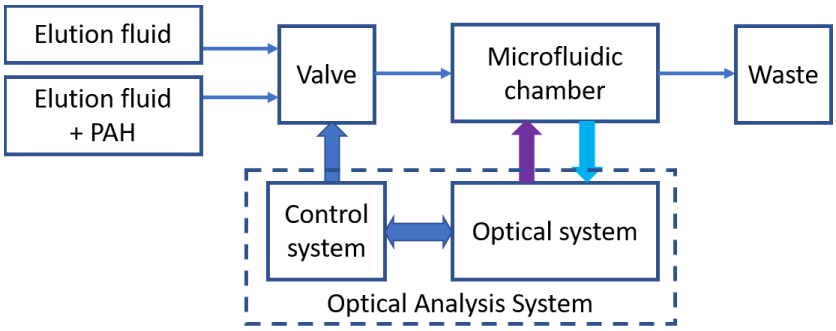


Figure 7.4 Schematic diagram of microfluidic circuit.

B. Experimental Results

We filled the microfluidic chamber with the B[a]P dissolved in pure ethanol at a concentration of 750 µM and measured its fluorescence signal for 100 s, collecting 5.5 million photon counts. Figure 7.5 shows the semi-logarithmic plot of the recorded signal, which exhibited a multiexponential decay with an offset. However, this signal is composed of the fluorescence signal of the B[a]P and the background signal detected during the measurement.

To eliminate the background signal, we filled the microfluidic chamber with pure ethanol using the valve and recorded the background signal. It is important to note that the background signal should be acquired for a considerably longer time than the sample measurement time. This reduces the photonic noise and prevents introducing additional noise when the background signal is subtracted, as explained in 4.2.4. The background signal is then scaled and subtracted

from the measurement signal. In this experiment, we acquired the background signal for 200 s, collecting about 7 M photon counts. Thus, the photon counts of the scaled background, i.e. the background detected during the B[a]P signal acquisition time of 100 s, was around 3.5 M photons. Figure 7.6 shows the semi-logarithmic plot of the scaled background signal, which consists of two components: a multiexponential decay and an offset. The multiexponential decay can be attributed to the autofluorescence of the optical and microfluidic parts and the fluorescence emission of the residual deposits from previous measurements. Whereas the offset can be explained by the ambient light and the SPAD noise due to the dark count rate and the afterpulsing effect.

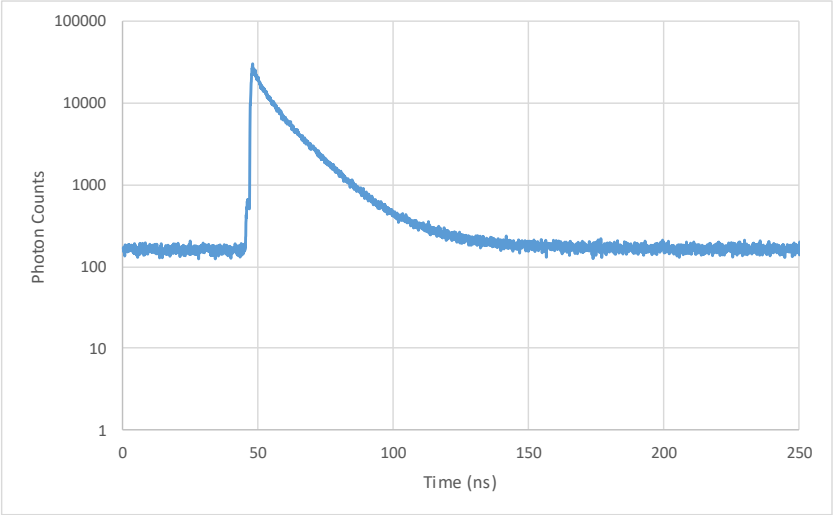


Figure 7.5 Fluorescence signal of B[a]P in pure ethanol at a concentration of 750 µM without background suppression.

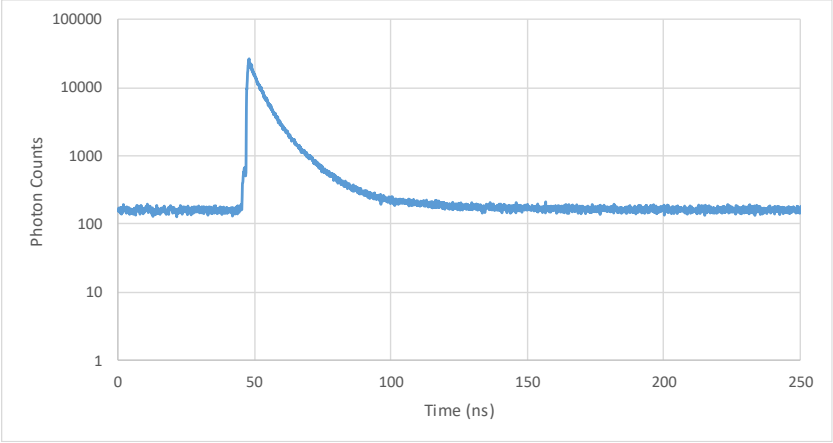


Figure 7.6 Scaled background signal.

To eliminate the background signal, we scaled it by a factor of (100 / 200 = 0.5) and then subtracted the scaled background signal from the B[a]P sample signal. This method effectively

removed the background signal and resulted in a monoexponential decay, as depicted in Figure 7.7, with a lifetime of 14.16 ns. The principle of background elimination is illustrated in Figure 7.8.

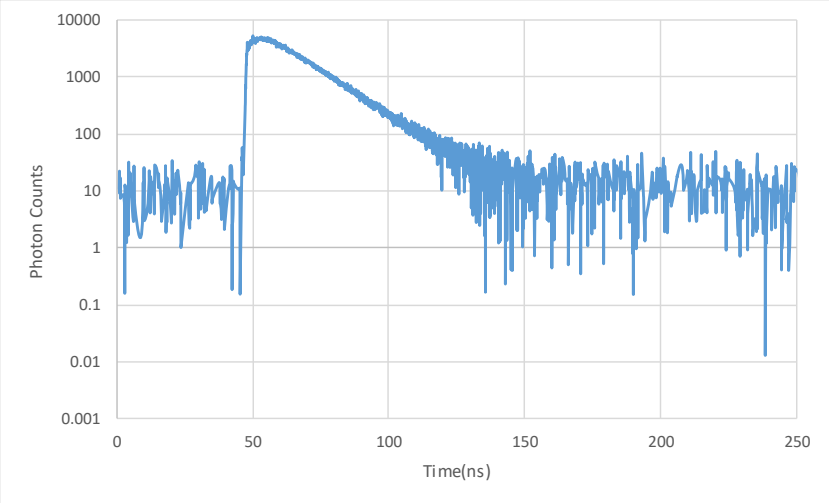


Figure 7.7 Fluorescence signal of B[a]P in pure ethanol at a concentration of 750 µM after background suppression.

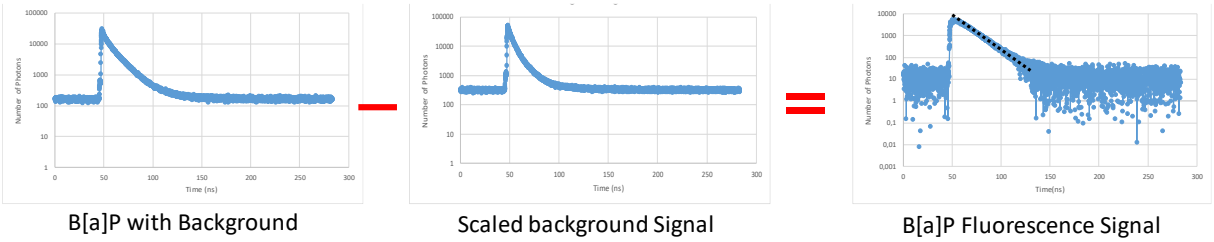


Figure 7.8 Principle of background signal suppression.

To estimate the fluorescence lifetime (FLT) and intensity of the recorded signal, we selected a temporal window that covers the monoexponential segment of the signal, as shown by the orange region in Figure 7.9. We only considered the photon counts within this window for the fluorescence analysis. We used the exponential fitter of Microsoft Excel to fit this segment of signal and obtained an FLT of about 14.16 ns. We also applied our maximum likelihood-based algorithm to estimate the FLT of the recorded signal. The result obtained by our algorithm was approximately 14.18 ns, which is close to the value obtained by the Microsoft Excel fitter. This demonstrates the reliability and accuracy of our algorithm.

After estimating the fluorescence lifetime, the measured value is compared with the expected values for the target pollutants. The presence of a pollutant is confirmed if the measured FLT

matches the expected FLT for that pollutant and then the concentration is estimated from the fluorescence intensity. Otherwise, the measurement is discarded as no pollutant is detected.

In the next experiment, we reduced the concentration of the B[a]P in pure ethanol to 40 µM. We first recorded the background signal for 500 seconds, collecting 23.5 M photon counts. Next, we recorded the fluorescence signal of the B[a]P sample for 250 seconds, obtaining a total photon count of approximately 16 million. We then subtracted the background signal, after scaling it by a factor of 0.5, from the sample signal. Figure 7.9 shows the semi-logarithmic plot of the original, the background, and the corrected signals. The fluorescence lifetime of the corrected signal was estimated to be 13.8 ns.

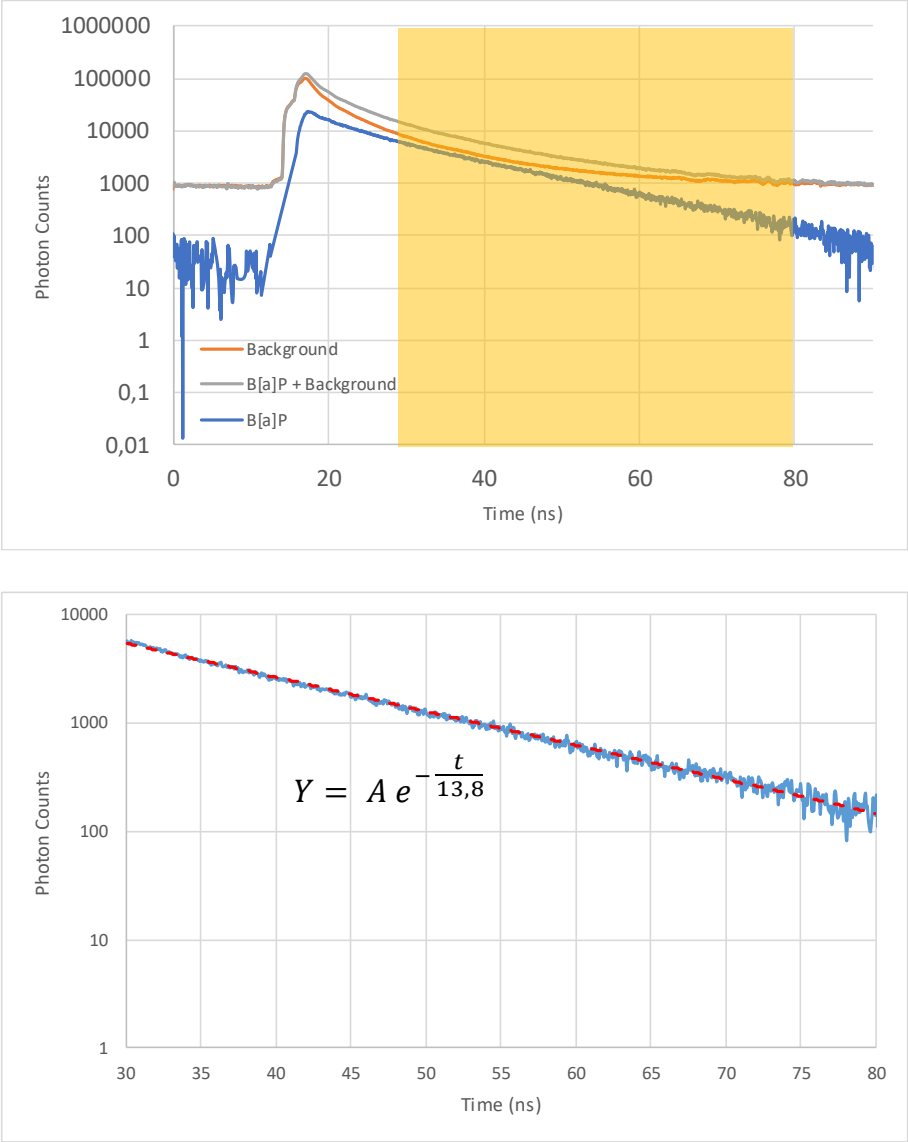


Figure 7.9 Fluorescence signal of B[a]P in pure ethanol at a concentration of 40 µM after background signal suppression.

7.2 Microfluidic Droplet Screening and Sorting

High-throughput screening (HTS) of biomolecules is a technique that allows the identification of molecules with novel properties or biologically active molecules from a chemical library. It is widely used in pharmacology and biochemistry. A microfluidic circuit coupled with fluorescence detection is an excellent tool for HTS [173] and fluorescence-activated cell sorting (FACS) [174]. The fluorescence quantum yield (QY) of fluorescently labeled molecules can be significantly reduced upon interaction with other molecules. For example, in Förster resonance energy transfer (FRET)-based assays, the fluorescence yield of the excited "donor" molecule is quenched if its excitation energy is efficiently transferred to a FRET "acceptor" in close proximity. Consequently, the measurement of the fluorescence QY enables indicating the biomolecular interactions. However, the fluorescence QY is usually detected by measuring the fluorescence intensity which can be affected by many other factors, such as the fluorophore concentration and the excitation light intensity. In contrast, fluorescence lifetime (FLT) detection is potentially more accurate for assaying biomolecular interactions because it is an intrinsic property of the fluorescence QY that is independent of these factors. This is why fluorescence lifetime imaging microscopy (FLIM) has been extensively developed for cell biology analysis [175]. Similarly, FLT detection is expected to enhance the reliability of biomolecular interaction assays [176-178]. It has been implemented in well-plate readers for HTS [179, 180]. However, current commercial plate readers take about 0.5 seconds to read one microwell in FLT mode (they can read a plate with up to 1536 microwells in a few minutes) [181], while a flow rate of a few thousand droplets per second can be achieved with a microfluidic channel. Furthermore, microfluidic chips allow for reducing the quantity of reactants and offer an easy solution to sort the targeted samples based on the fluorescence measurement results. Thus, the use of microfluidics can potentially accelerate drug discovery by a factor of 100. Nevertheless, combining FLT-based assays with microfluidics is challenging, especially when a sorting action is needed, because this implies performing both data acquisition and processing in real time. For example, a droplet rate of 1 kHz requires both the acquisition and the lifetime assessment of the droplet to be performed in less than 1 ms. Moreover, the thin thickness of about 50 µm of the droplets in the microchannel results in low absorption and weak fluorescence signal.

In [182], the authors present a real-time droplet sorting system to enrich the population of fluorescent proteins expressed in bacteria, by extracting the FLT value from the phase between the emission and the excitation signals. their system achieves a high droplet rate of about 2.5 K

droplet/second. However, their system relies on costly components, such as the electro-optic modulator (EOM).

It has been demonstrated that TCSPC can be applied to measure the FLT of microfluidic droplets under high-throughput conditions [183]. The setup proposed by [184] can screen up to several hundred independent reactions per second based on the TCSPC technique. However, this setup is not compatible with droplet sorting, as the FLT measurement is performed offline rather than in real time. So far, using the TCSPC-based FLT measurement, the highest reported rate of droplet sorting in real time is only a few droplets per second [185].

Thanks to the online data processing feature, our TCSPC system can perform FLT-based droplet sorting in real time at high droplet rates, comparable to those achieved by the system proposed by [182], but with a much lower cost.

7.2.1 System Modifications for Droplet Detection and Sorting

To enable the droplets' data processing in real time, we chose to employ our synchronous TCSPC system in this application as it offers a faster calibration process than the asynchronous system. Indeed, the drawbacks of the synchronous system, such as its higher sensitivity to temperature variations, have a negligible impact in this application because the experiment time is short enough to assume temperature stability. Furthermore, the fluorescence decays of the droplets are implicitly aligned since the bins preceding the decay peak are discarded by the FLT estimation algorithm.

For this application, our system requires some modifications to enable the detection and sorting of microfluidic droplet.

- **Droplet Detection**

The droplets flow through the microfluidic channel separated by the continuous phase, i.e. the medium in which they flow. The SPAD detects the fluorescence emission of each droplet as it passes through the detection spot. The detected signal is termed the droplet signal. On the other hand, the signal detected when the continuous phase passes through the detection spot is termed the background signal.

As described in section 3.1.5 and illustrated in Figure 3.22, the *PACKET* signal is a periodic signal that indicates the elapsed from the beginning of the experiment, and the photon words corresponding to the photons detected within a packet period are grouped between two packet

words. In this application, the packet period is configured to be much shorter than the droplet period, which is the time required for a droplet to cross the detection spot that can be estimated by multiplying the droplet duty cycle by the inverse of the droplet flow rate. Therefore, the photon words of each droplet will be organized into multiple packets, termed the droplet packets. Likewise, the photon words of the background signal will also be organized in multiple packets, termed the background packets. Since the droplet signal has a higher intensity than the background signal, the droplet packets have more photons than the background packets. Thus, a threshold value can be defined to discriminate droplets from background packets by comparing the photon count of each packet with this value. However, applying this method as soft data processing by the C program requires two readings of the buffer data: one to determine the number of photons in each packet and classify it as background or droplet packet, and another to add its photon words to the corresponding histogram. This decreases the data processing speed and consequently the maximum detectable droplet rate.

We significantly accelerated the droplets' data processing by preprocessing the TDC data in the FPGA to discriminate between droplet and background packets. We count the photons of each packet and classify it as a droplet or a background packet based on the threshold value. We used a specific bit in the packet and assigned it to '1' for droplet packets and to '0' in the background ones. To implement this, we modified the data writing controller by adding a counter that is reset to zero at each *PACKET* pulse and increments with every new detected photon. At the end of each packet, at the following *PACKET* pulse, the data writing controller generates a packet word that includes the packet type flag bit. This packet word comes after its photon words in the buffer data as a packet footer. Consequently, the C program can process the buffer data with a single reading in reverse order. It checks the packet type bit at each packet word to classify the packet as a droplet or a background packet. Then it continues reading the buffer data and directly adds the photon words of the packet to its corresponding histogram until it reaches the next packet word.

However, discriminating droplet packets from background packets based on a single threshold may result in erroneous classification due to noisy packets, such as packets F1, F2, and F3 in Figure 7.10. To improve the robustness of the droplet classification, we propose a dual-threshold method with two flag bits. We use a background threshold and a droplet threshold to classify the packets based on their photon count, as illustrated in Figure 7.10. We define three types of packets: 'background packet' for packets with photon count below the background threshold, 'droplet packet' for packets with photon count above the droplet threshold, and

'unclassified packet' for packets with photon count between the two thresholds. The droplet classification is done as follows: the droplet flag is set to '1' after three successive droplet packets (packets A′ and C′), indicating the detection of a droplet. This flag remains '1' until the end of the droplet, which is identified by the first background packet that succeeds at least two successive unclassified packets (packets B′ and D′). Starting from this packet until the beginning of the next droplet, the background flag is set to '1' for every background packet, while both flags are set to '0' for the unclassified and droplet packets. This approach requires that the packet period be much shorter than the droplet period to improve the resolution of the droplet detection.

To label the droplets, we implemented a 6-bit counter that increments with each new detected droplet to give it a unique number. This number is then included in the droplet packet words to enable the C program to distinguish the droplets in the case the data buffer contains data of multiple droplets. However, the total number of droplets in an experiment is not limited by this counter to 64 since the C program assigns new identification numbers to the processed droplets.

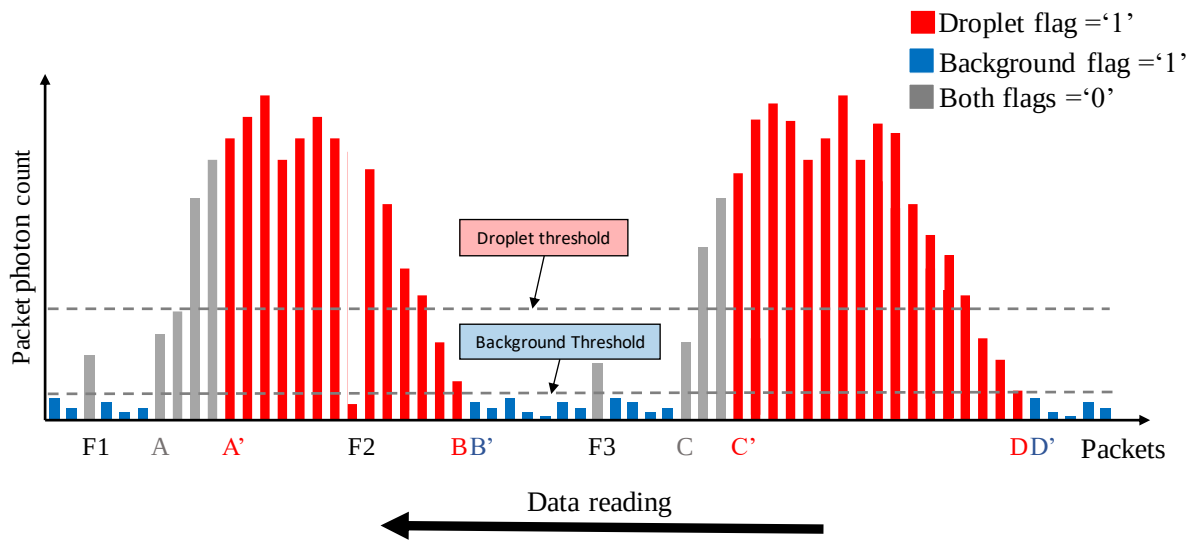


Figure 7.10 Droplet detection: two thresholds are defined to determine the beginning and the end of the droplet and avoid noise spikes.

- **Droplet Data Processing**

The TDC data preprocessing in the FPGA simplifies the subsequent data processing by the C program. To generate the raw histograms of the droplets and the background, the C program reads the data buffer in reverse order and checks the droplet and background flag bits of each

packet word. If the packet has a background flag bit of ‘1’ (blue packets in Figure 7.10), the program updates the background histogram by incrementing the bin corresponding to the arrival time of each photon of this packet by one. Once it encounters a packet with a droplet flag bit of ‘1’ (packets B or D), it starts constructing the droplet histogram by adding the photon words to this histogram until it reaches a packet footer with a background flag bit of ‘1’ indicating the end of the droplet data (packets A-1 or C-1). Moreover, using the 6-bit droplet identifier included in the packet words, the C program discriminates the packets belonging to different droplets in the data buffer and creates a separate histogram for each droplet.

The background photons detected among the droplets are accumulated in a single background histogram throughout the experiment. This reduces the photonic noise, which is inversely proportional to the square root of the photon number. The background histogram will be subsequently used to eliminate the background signal from the droplet fluorescence signal.

After creating a droplet histogram and updating the background histogram, the C program applies the calibration routine to these histograms to compensate for the nonlinearity of the TDC. It then subtracts the background histogram, scaled by the ratio of the packet numbers of the droplet and the background histograms, from the droplet histogram to remove the background signal. The resulting histogram represents the pure fluorescence decay of the droplet. From this histogram, the C program estimates the droplet fluorescence intensity from its photon count and its FLT using the maximum likelihood-based algorithm.

- **Droplet Sorting**

The proposed system enables sorting droplets according to their fluorescence properties, either intensity or lifetime. The user sets a sorting threshold value for the desired fluorescence property, and the system measures the value of this property for each droplet and compares it with the threshold. Then, the system sends the sorting information to a sorting block implemented in the FPGA. The sorting information includes the packet number of the first packet of the droplet (indicating the arrival moment of the droplet at the detection spot), the number of packets in the droplet (indicating the droplet duration), and a binary action bit that determines the sorting decision based on the predefined criteria. Moreover, the system estimates the delay between the droplet arrival time at the detection spot and the sorting point from the flow rate and the distance between these points. Using this information, the sorting block generates a *SORTING* signal that activates an actuator to sort the droplet, as depicted in Figure 7.11.

For optimal performance in applications requiring sorting action, it is recommended to configure the data buffer size to include data of one sequence droplet-background, in order to satisfy the sorting delay and ensure the real-time processing and sorting of droplets.

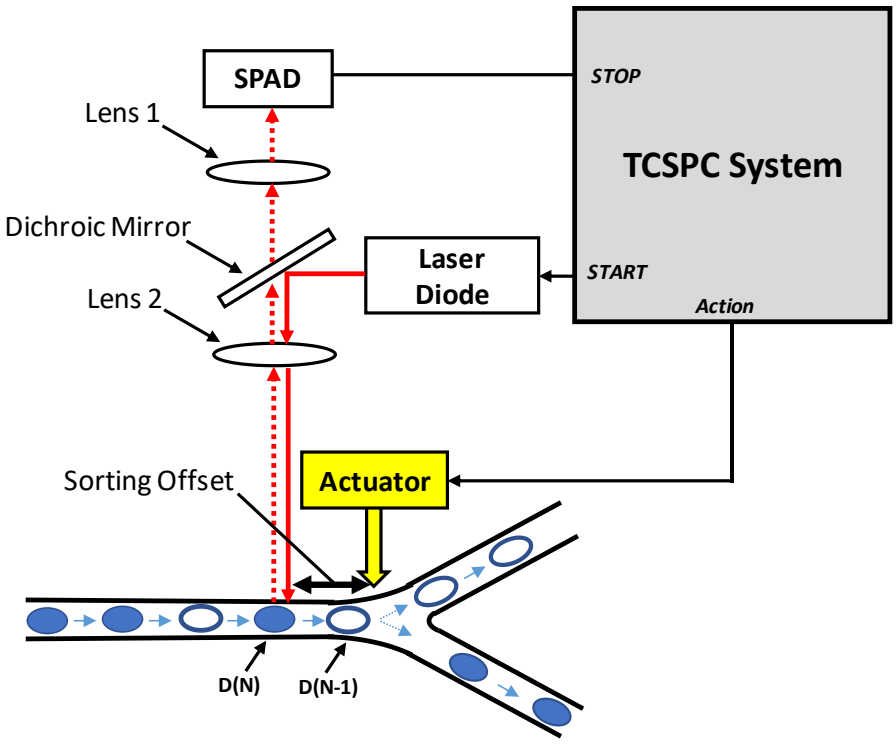


Figure 7.11 Schematic diagram of the droplet sorting system with the optical and microfluidic parts.

7.2.2 Experimental Study

We integrated the modified TCSPC system into a comprehensive system that comprised the optical setup used in the WPS application and a microfluidic chip that generates fluorescent droplets. To validate the real-time droplet detection and screening and evaluate the achievable droplet rate, we conducted a set of experiments on microfluidic droplets of fluorescein in PBS at a concentration of 1 mM.

A. Experimental Setup

We employed the same optical setup as in the WPS application, using a 405-nm pulsed laser diode as the excitation light source. This wavelength is not optimal for the excitation of the studied fluorescent dye. However, the aim of these experiments was to demonstrate the maximum detectable droplet rate. To generate the microfluidic droplets, we used a microfluidic chip fabricated with soft lithography of PDMS on glass. This chip was placed on the XYZ microscopy stage of the optical setup, as shown in Figure 7.12. The flow focusing junction

produced aqueous droplets containing fluorescein in perfluorinated HFE-7500 oil with surfactant from Raindance company. To enhance the weak fluorescence emission of fluorescein resulting from the suboptimal excitation wavelength, we used a high concentration of 1 mM for the fluorophore. The droplet-oil sequence flows through a squared channel of 25 µm, which is sufficiently small to generate droplet rates of up to 10 thousand droplets per second.

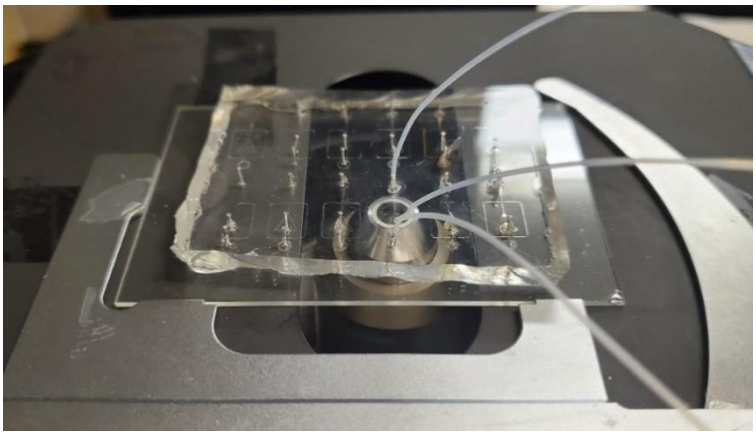


Figure 7.12 Microfluidic chip.

B. Experimental Results

- **Maximal Achieved Droplet Rate**

We adjusted the flow rate of the oil and the fluorescein solution to generate 3573 droplets per second with a duty cycle of about 50%. At this rate, the droplet period was around 140 µs. To achieve a high resolution in the droplet detection, we configured the packet period to about 16.7 µs, as discussed above. We also configured the background and droplet thresholds to 3 and 7 photons, respectively. These threshold values were experimentally chosen. With these configurations, our system detected the droplets based on the packets' intensity with a good resolution, as shown in Figure 7.13, and estimated their FLT in real time. At such a high droplet rate, the average number of detected photons per droplet was only about 43 photons. Nevertheless, the system could detect all the droplets and estimate their FLT with an average value of 4.02 ns, which agrees with the literature value [171, 172]. Figure 7.14 shows the distribution of the FLT values of 10,000 droplets. The standard deviation of the estimated FLTs was 0.72 ns, while the theoretical photonic noise contributed a standard deviation of 0.6 ns. Thus, the system introduced an increase in the standard deviation of only 20% of that caused by the photonic noise, indicating that the total noise of the system is very close to the physical limits.

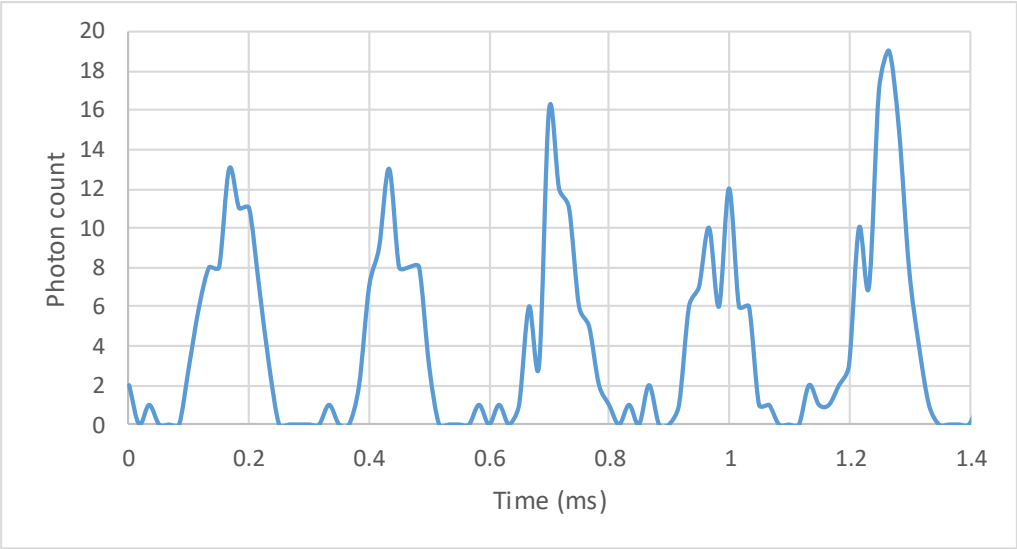


Figure 7.13 Droplet detection based on the packets' photon counts at a droplet rate of 3573 droplets/second.

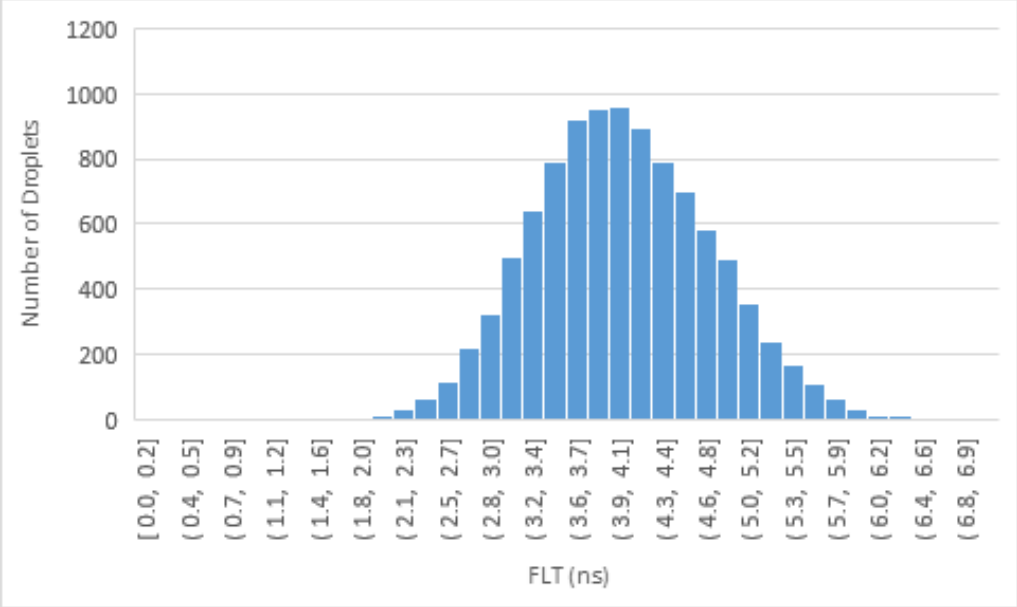


Figure 7.14 FLT distribution of 10000 droplets of fluorescein in PBS at 3573 droplets/second.

The photonic noise can be reduced by increasing the photon detection rate. In this experiment, we operated the laser diode at a frequency of 8.5 MHz. The mean photon rate during the passage of a droplet was approximately 428 K photon/s, which corresponds to about 5 % of the laser rate. If this ratio is accepted to be 10%, we can increase the photon rate by a factor of 2 by selecting an optimal wavelength for the excitation light. This would entail a loss of precision on the FLT due to the pileup effect of only 3%. Moreover, at a laser frequency of 8.5 MHz, the measurement temporal window was about 117 ns. However, the FLT estimation algorithm

requires a temporal window of only five times larger than the FLT value, which is 20.5 ns in the case of fluorescein. Therefore, we can increase the laser frequency to 42.5 MHz which boosts the photon rate by factor 5. By combining these two optimizations, we can increase the photon rate by factor 10 and consequently reduce the standard deviation by approximately factor 3.

- **Practical Droplet Rate**

In the second set of experiments, we varied the oil and the fluorescein solution flow rates to generate a droplet rate of 1000 droplets per second. We set the packet duration to about 55.7 µs and the background and droplet thresholds to 6 and 55 photons, respectively. The system detected the droplets in real time and estimated their FLTs with an average value of 3.92 ns. This value is slightly lower than the theoretical value due to the pileup effect. The laser pulse rate was about 8.5 MHz and the average number of detected photons per droplet was about 718 photons, which corresponds to about 16% of the laser rate. At such a high ratio the pileup effect distorts the recorded decay and leads to an underestimation of the FLT measurement. Figure 7.15 shows the distribution of the FLT values of 2000 detected droplets. The standard deviation of the FLTs values was 0.18 ns, while the photonic noise contributed to a standard deviation of 0.15 ns. Thus, the system noise was only 20% higher than the theoretical noise limit, indicating a low noise level and a high precision. However, the error in the average FLT value can be reduced by increasing the laser pulse rate to 42.5 MHz while maintaining the photon detection rate to have a photon rate of 3.2% of the laser rate.

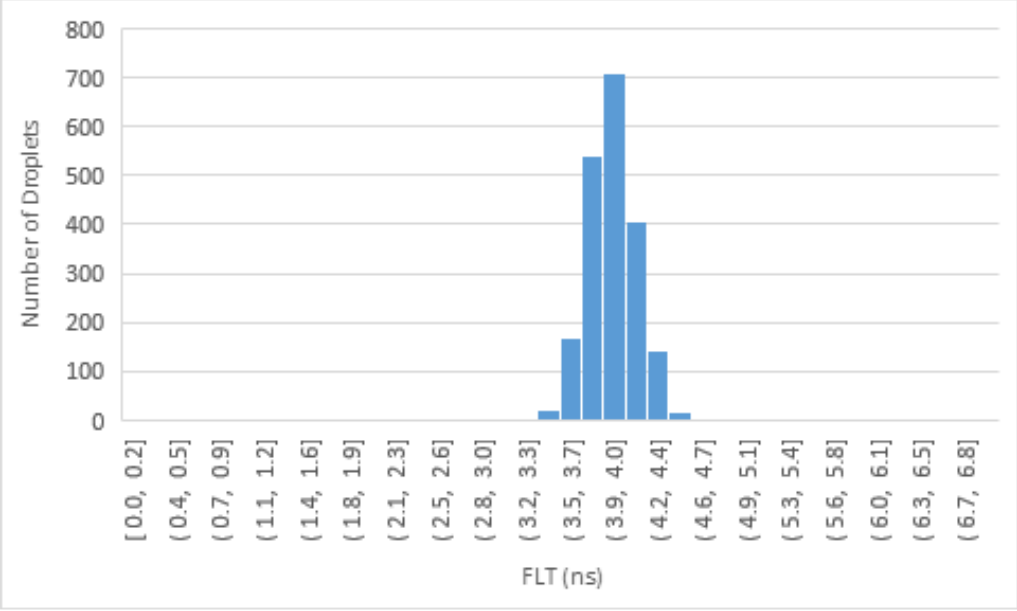


Figure 7.15 FLT distribution of 2000 droplets of fluorescein in PBS at 1000 droplets/second.

7.2.3 Discrimination of Two Populations of Samples

In screening and sorting applications, the measured FLT is compared with a predefined threshold to make a decision. Therefore, it is crucial to achieve a clear separation between two populations of samples with distinct FLT values, as depicted in Figure 7.16-a. where the histograms of the two populations do not overlap. However, due to large standard deviation values, the distribution of the two populations may overlap, leading to erroneous decisions of false positive and false negative decision, as illustrated in Figure 7.16-b. A quantitative measure of the separation quality between two populations is given by the Z' factor [186], which is defined by Equation 7.1.

$$Z' = 1 - \frac{3\sigma^+ + 3\sigma^-}{\tau^+ - \tau^-} \approx 1 - \frac{3}{\sqrt{N^+}(1-\sqrt{r})} \qquad (7.1)$$

where σ^+ and σ^- are the standard deviation of the higher FLT population and the lower FLT population, respectively, τ^+ and τ^- are the FLT of these populations, r is the ratio between τ^- and τ^+, and N^+ is the average number of photons in the droplet of the higher FLT population.

A Z factor of 0.5 ensures that the means of the two populations are separated by 3 σ of the first population (σ^+) plus 3 σ of the other one (σ^-).

Assuming that the population mentioned in the first set of experiments in section 7.2.2 is the higher-FLT population, we can determine the maximum value of FLT for the lower-FLT population (τ^-) that can be discriminated with a Z factor of at least 0.5. In this case, N^+=43 photon, τ^+ = 4.02 ns. Using Equation 7.1, we obtain that τ^- must be less than or equal to 0.029 ns. This value is impractical due to the high photonic noise. However, if we apply the optimization mentioned above where N^+ is increased to 430 photons, then the maximum value of τ^- can be increased to 2.03 ns.

For the case of the second set of experiments in section 7.2.2, where N^+ =719 photons and τ^+ =3.922 ns, the maximum accepted value of τ^- is 2.36 ns. This value can be further increased up to 2.69 ns by applying the same optimization as above.

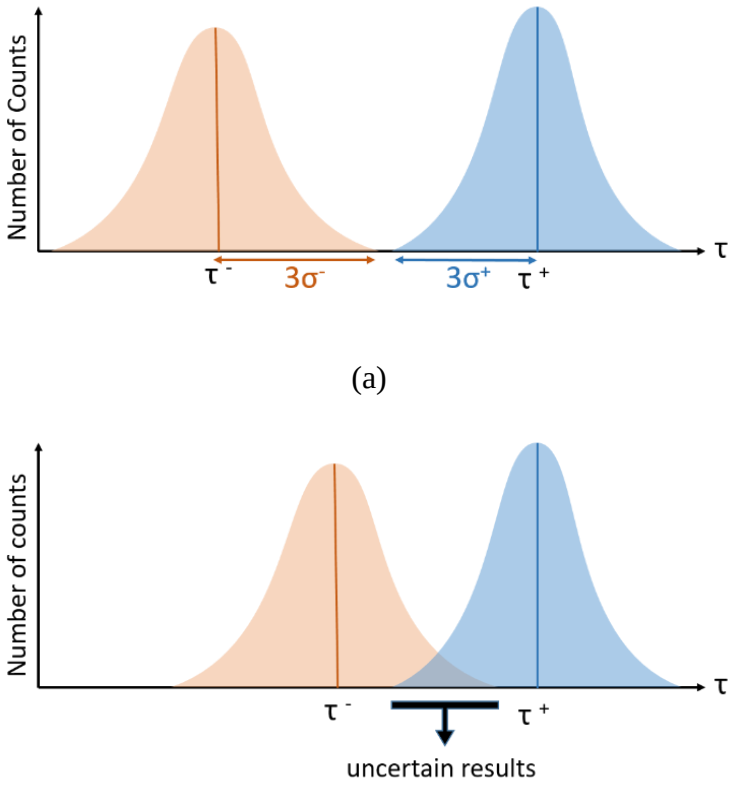


Figure 7.16 Discrimination of two populations of samples: (a) the distributions of the FLT values for the two populations are clearly separated, (b) the distributions of the FLT values for the two populations overlap, resulting in uncertain results.

7.3 Time-Resolved Optical Turbidity

We employed our TCSPC system in collaboration with the "Mécanique des Fluides" research time of our laboratory to investigate the time-resolved optical turbidity technique. Turbidity is a key parameter to evaluate water quality. It indicates the amount and size of suspended particles in water. The Water Framework Directive (WFD) [187] in Europe requires regular monitoring of water quality. One of the main water quality parameters is the total suspended solids (TSS) load, which should be quantified in both riverine and wastewater systems. The objective is to monitor sediment flux that may transport pollutants and cause occlusions.

Similar to the detection of water pollutants, periodic laboratory analysis of samples is the most common method for measuring TSS concentration. However, this method is costly and time-consuming. Furthermore, a major drawback of this method is the long delay between sampling and analysis; the sample is susceptible to chemical and physical modifications during the storage and transport operations, which may affect the turbidity.

Real-time measurement of TSS concentration provides continuous information that enables potential real-time management of wastewater systems or treatment facilities. Optical turbidity and acoustic backscattering are the most frequently used techniques for real-time TSS monitoring. These methods have some limitations. In fact, the relationship between optical turbidity and particle concentration is considered to be linear [188, 189]. Nevertheless, this relationship depends on the consistency of the particles' characteristics such as their shape and size [190]. Moreover, it suffers from ambiguity and saturation at high concentrations [191]. On the other hand, the acoustic backscattering technique has a linear relationship with particle concentration over a larger range but is less sensitive to low particle concentrations (< 0.1 g/L) [192]. Furthermore, this technique is less sensitive to small particles (<10 micrometers).

To overcome the limitations of conventional methods, time-resolved optical turbidity (TROT) is a new promising technique for turbidity measurement based on the TCSPC technique. Figure 7.17 illustrates the TROT principle: A short periodic pulse of light with an FWHM of typically less than 100 ps is emitted into the turbid medium via an optical fiber, referred to as the emission fiber. A second optical fiber, the reception fiber, is placed at a distance r of a few tens of mm from the emission fiber and collects the photons that are backscattered from the suspended particles in the medium. These photons are then detected by a photon detector. The time of flight of these photons depends on the optical paths they follow in the medium (red and green paths In Figure 7.17). The arrival time of the scattered photons at the photon detector is used to construct a histogram. Due to the stochastic nature of the scattering process, it introduces a temporal broadening of the initial light pulse resulting in a temporal point spread function (TPSF) that spans a few nanoseconds. The number of detected photons and the shape of the TPSF are expected to have a relationship with the particle concentration and thus the medium turbidity. In practice, the number of detected photons represents the intensity, which is the property on which classical optical turbidimeters are based. However, TROT also utilizes other characteristics of the TPSF, such as the mean arrival time of the detected photons, which is demonstrated to have a strong correlation with the turbidity level.

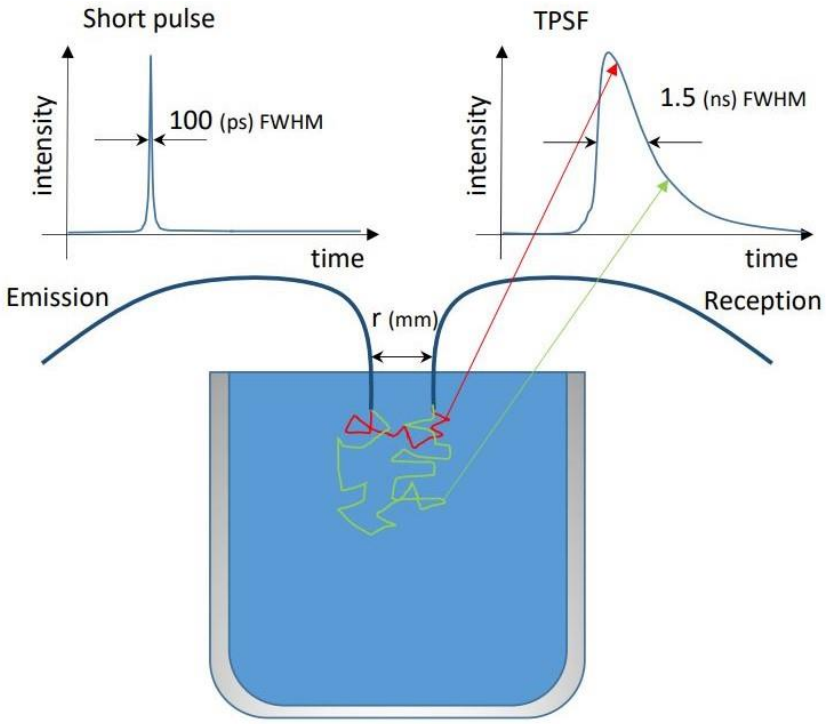


Figure 7.17 Principle of Time-Resolved Optical Turbidity (TROT).

7.3.1 System Modifications for TROT

We employed our TCSPC system in the TROT application to record the temporal distribution of the detected photons and construct the TPSF histogram. We initially used our synchronous TCSPC system for this purpose. However, we encountered measurement instability issues due to temperature variations, which caused shifting of the TPSF peak and resulted in inaccurate measurements. We solved this problem by switching to our asynchronous TCSPC system, which exhibits better stability with temperature variation than the synchronous one.

Subsequent experiments revealed another source of measurement instability which is the fluctuations of the laser power. To investigate this issue, we conducted periodic measurements with the laser diode operating continuously in pulsed mode for a long duration. In these measurements, we observed variations in the photon detection rate due to the laser power fluctuation. To address this issue, we implemented a feedback control loop that utilized the photodiode integrated in the laser diode, analog-to-digital (ADC), and digital-to-analog (DAC) converters, and a control logic implemented on the FPGA. This control loop works as follows: The DAC, which is part of the laser pulse generator, modulates the emitted intensity. The photodiode monitors the mean intensity and provides an analog value to the ADC. The control logic reads the ADC value every second and compares it with a reference value that corresponds

to the desired intensity. According to the difference between the current and the reference values, the control algorithm configures the DAC with a new value to regulate the emitted power and converge it to the reference level. Figure 7.18 illustrates the schematic of the feedback control loop.

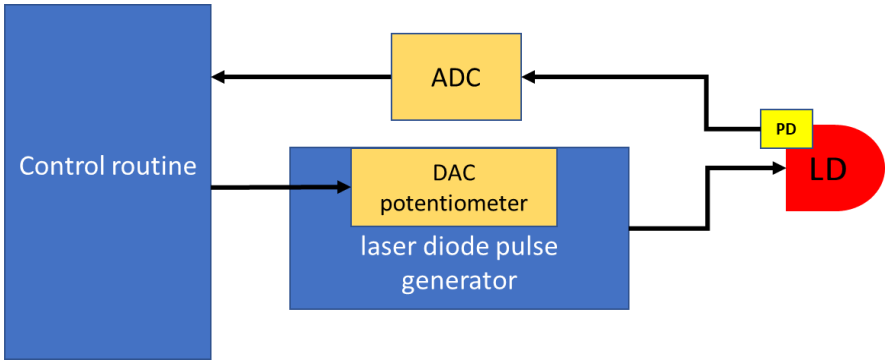


Figure 7.18 Feedback control loop for stabilizing the laser power.

7.3.2 Experimental Results

We conducted a series of experiments to demonstrate the TROT technique and to assess its robustness against biofouling, a significant obstacle for optical sensors in contact with water.

- **TROT Demonstration**

We carried out a series of experiments to investigate the TROT technique. We prepared different levels of turbidity using a beaker filled with 4L of water to which we gradually added varying amounts of Formazine. We used a 630-nm laser diode as the light source and positioned the emission and reception fibers at an angle of 120 degrees. We recorded the TPSF histogram for 10 seconds at each turbidity level. We analyzed the characteristics of the recorded TPSF and examined several parameters that varied with the turbidity. The most relevant parameters were the number of detected photons and the mean arrival time of these photons. Figure 7.19 depicts the number of detected photons as a function of the concentration (the turbidity level). The number of photons increases with turbidity until a certain level where it starts to decrease.

Figure 7.20 shows the mean arrival time of detected photons as a function of the concentration (the turbidity level). This parameter is independent of the detected intensity and exhibits a close-to-linear trend with the turbidity level over a wide range.

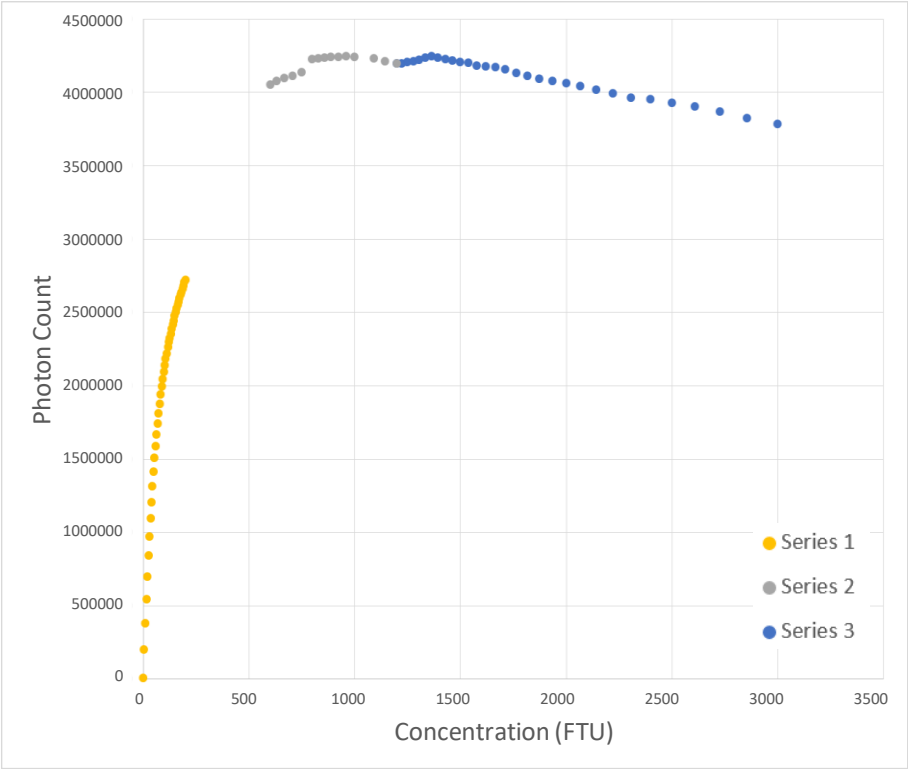


Figure 7.19 Detected photon count as a function of concentration.

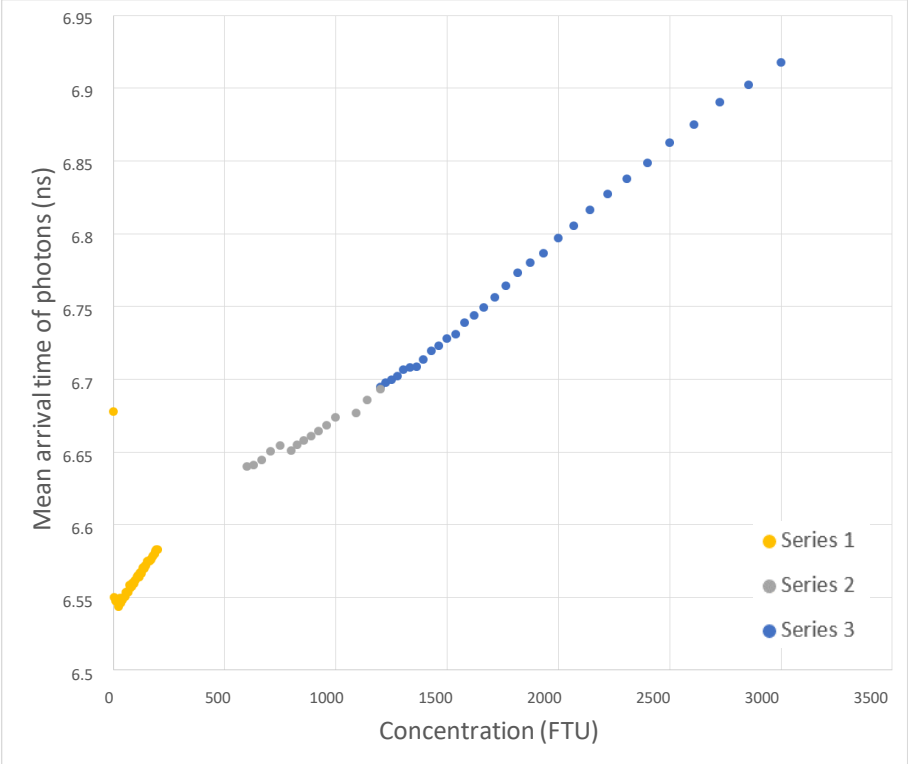


Figure 7.20 Mean arrival time of detected photons as a function of concentration.

- **Robustness against Fouling**

One of the major challenges of optical turbidimeters is biofouling. The deposition of undesirable materials such as microorganisms, organic matter, and minerals on the parts in contact with water can impair the performance and accuracy of the device. Biofouling affects the transduction interfaces of the sensors within a few days and introduces unacceptable errors in the measurements of the in-situ monitoring system. Figure 7.21 shows the biofouling accumulated on an optical turbidimeter after being placed in a wastewater treatment plant for 30 days (figure 7.21-a) and then after undergoing ultrasonic cleaning (figure 7.21-b).

(a) (b)

Figure 7.21 Biofouling effect in optical turbidimeter: (a) after 30 days in a wastewater treatment plant, (b) after ultrasonic cleaning.

To investigate the influence of biofouling on the performance of the time-resolved optical turbidimeter, we conducted a series of experiments using optical attenuators with different optical densities to simulate different levels of biofouling by inserting them between the laser diode and the reception fiber, as depicted in Figure 7.22. We used nine optical attenuators with optical densities from 0.01 to 4, representing 9 levels of biofouling. For each attenuator, we recorded the transmitted signal while maintaining the laser intensity constant.

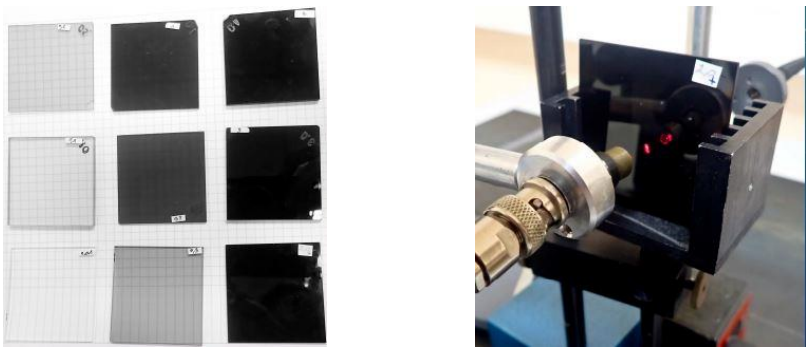


Figure 7.22 Simulating different levels of biofouling by inserting optical attenuators with different optical densities between the laser diode and the SPAD.

Figure 7.23 shows the signals recorded at the different levels of attenuation. From each signal, we extracted the photon count and the mean arrival time of photons. Figure 7.24 plots the photon count and the mean arrival time of photons against the attenuator optical density.

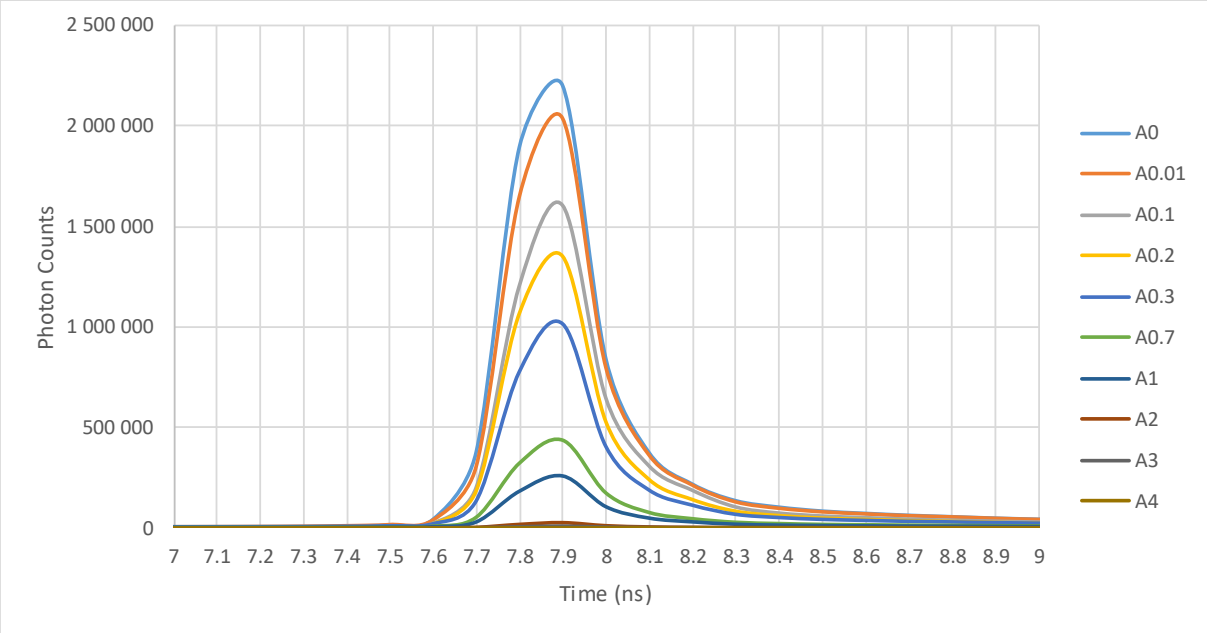


Figure 7.23 Laser pulse recorded at different levels of attenuation that simulate different levels of biofouling.

In practice, the photon count indicates the optical intensity, which is the parameter employed by conventional optical turbidimeters to estimate the turbidity level. Whereas the mean arrival time is the parameter used by TROT, Figure 7.24 shows that the photon count significantly decreases as the optical density of the attenuator increases, while the mean arrival time is unaffected by the attenuation introduced by the different attenuators. This demonstrates that TROT is much more robust than conventional optical turbidimeters against the optical intensity reduction caused by biofouling.

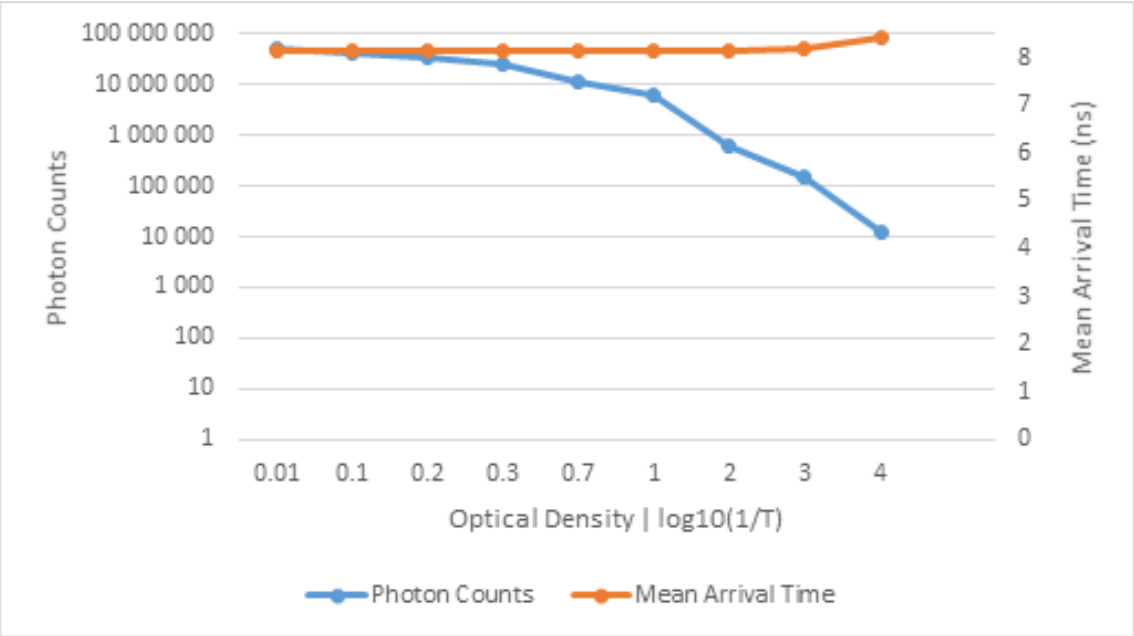


Figure 7.24 Photon count and mean arrival time of photons of the signals recorded at different levels of attenuation.

In the next experiment, we performed measurements at different turbidity levels with different attenuators simulating different degrees of biofouling. We computed the mean arrival times of detected photons for each measurement. Figure 7.25 shows the results and demonstrates that the mean arrival time of photons is consistent across a wide range of attenuation for each turbidity level.

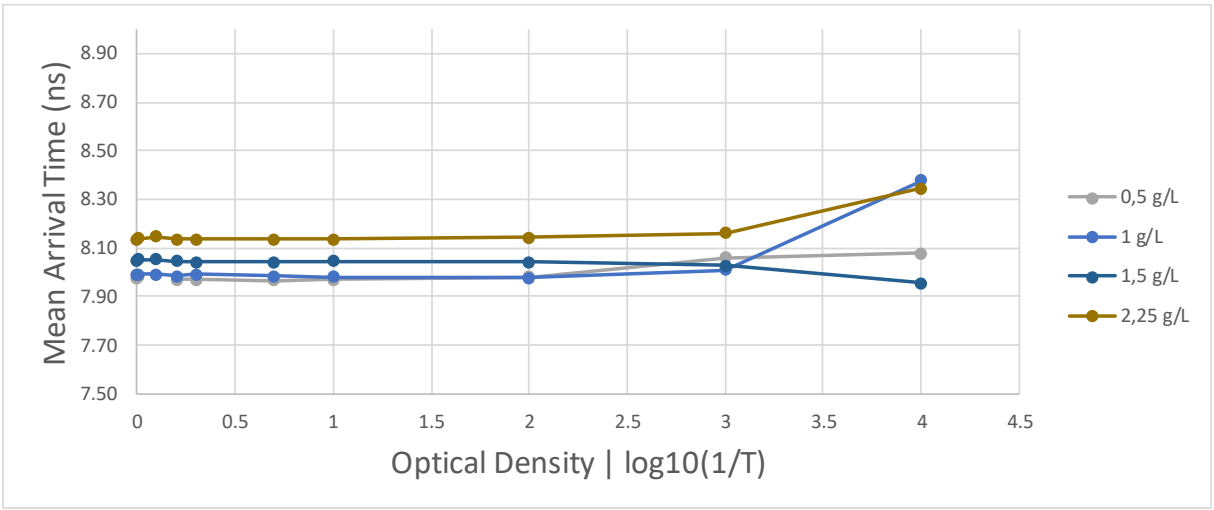


Figure 7.25 Mean arrival times of detected photons at different turbidity and attenuation levels.

7.4 Stray Light Characterization

We employed the developed TCSPC system in a novel approach for the stray light characterization, in collaboration with Liège Space Center "Centre Spatial de Liège (CSL)" in Belgium.

Stray light is a major challenge for optical instruments, especially in space applications. It refers to unwanted light that deviates from the intended image-forming beam and reaches the detector. It can originate from different sources, such as ghost reflections, scattering on optical or non-optical surfaces, or straight shots. Figure 7.26 depicts these different phenomena [193].

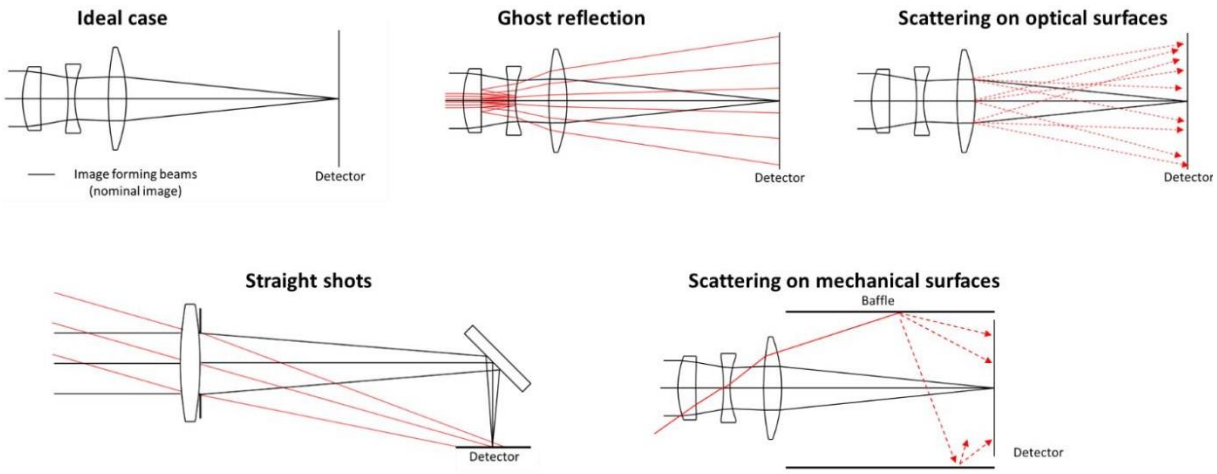


Figure 7.26 Illustration of different types of stray light effects.

Stray light degrades the image quality and thus the performance of the optical system by reducing the signal-to-noise ratio, the resolution, and introducing artefact features. Figure 7.27 illustrates an example of observing a point-like source object, such as a star [193]: the left figure shows the expected image, while the right figure shows the stray light features superimposed on the nominal image.

Stray light poses a more severe challenge for space instruments as it can degrade the quality of measurements and compromise the scientific objectives. Therefore, it is imperative to conduct experimental characterization of stray light before launching the system into the space, where any error would be irreversible. Experimental characterization can help identify and resolve any stray light issue on Earth. Moreover, it provides the necessary information for the development of stray light correction algorithms [194, 195].

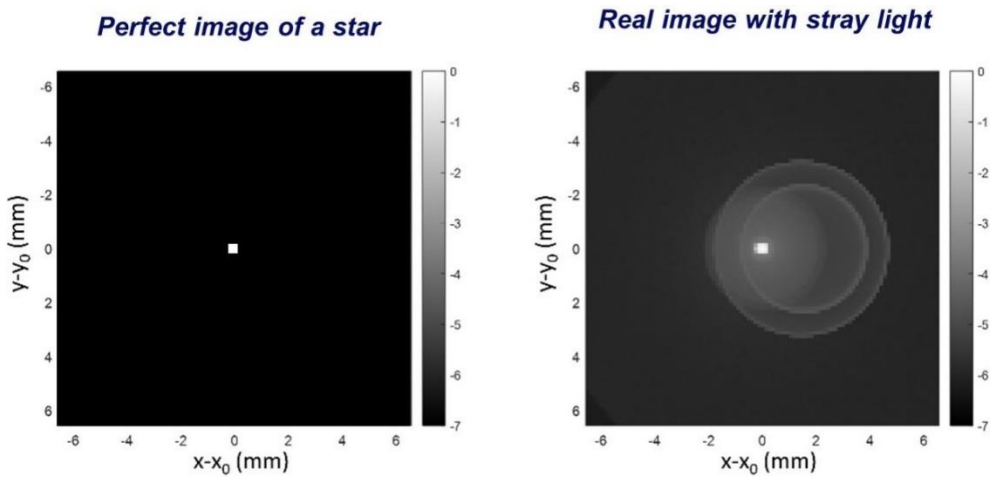


Figure 7.27 An example of the stray light effect: (a) perfect image of a star, (b) realistic image of that star with stray light effects.

The conventional method for the characterization of stray light is based on exposing the instrument to a point-like source and observing the signal at the detector. Any signal that deviates from the expected pixel is considered stray light (as long as it exceeds the dark current signal). This pattern is often referred to as the spatial point source transmittance (SPST). When integrated over a specific zone, either a pixel or the entire detector area, it is simply called the point source transmittance (PST). This approach has many limitations [196, 197]. First, the interpretation of stray light measurements is difficult. For instance, Figure 7.27 shows the superposition of multiple ghosts. Only three are clearly visible, but more are actually present. In more complex situations, several hundreds of stray light features can be superimposed. Second, the source of stray light is not identified, which makes it difficult to troubleshoot any issue and find a solution. The only information available is the level of stray light and whether it meets the specification or not. If the performance is satisfactory, the instrument can be launched into space. If there is any deviation from the expected performance, the cause of the deviation needs to be determined and addressed. However, without knowing the cause of a problem, it is impossible to devise a solution. This is a complex, time-consuming, and costly process. Third, stray light could originate from the measurement device itself. This means that the predicted performance may not reflect the actual performance of the instrument under test, but the combination of the instrument and the measurement device. For example, there could be a situation where the instrument has acceptable performance, but the measurement device inflates the stray light level beyond the specification. In that case, a spurious problem would be investigated.

The CSL has proposed and demonstrated a novel approach for the characterization of stray light by ultrafast time of flight imaging, which overcomes the limitations of conventional techniques. This method entails exposing the optical instrument under test to a pulsed laser beam and measuring the signal at the focal plane with an ultrafast sensor, as depicted in Figure 7.28. In practice, the image-forming beam and the various stray light components would traverse optical paths with different lengths, and consequently different times of flight. Hence, each component would arrive at the focal plane at a distinct time. Unlike a regular detector, an ultrafast sensor can discern the differences in time of flight and measure each path separately. This enables the decomposition and identification of the different stray light components based on their time of flight. This approach overcomes the three fundamental limitations of conventional methods, as it facilitates the interpretation, allows locating the potential sources of error, and distinguishes the stray light of the instrument under study from that of the measurement device.

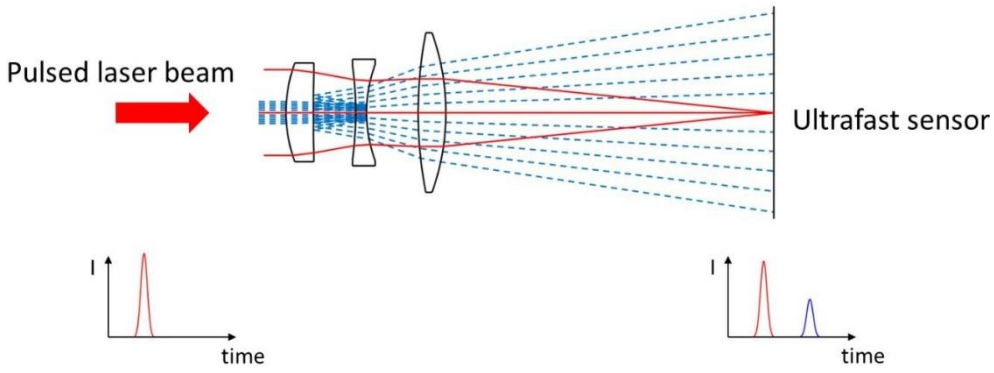


Figure 7.28 Illustration of TOF-based stray light characterization: when a pulsed beam illuminates the optical system, the nominal beam and the different ghost paths reach the detector at different times due to their different optical path lengths.

In the case of short instruments, such as refractive telescopes, the optical path lengths of stray light sources vary by very small amounts, requiring a sensor with sub-pico-second temporal resolution, such as a streak camera, to discriminate paths with sub-millimeter precision. However, for large facilities with optical path lengths ranging from several centimeters to tens of meters, the differences in time of flight are larger, allowing the use of a less expensive sensor with lower temporal resolution typically of a few picoseconds.

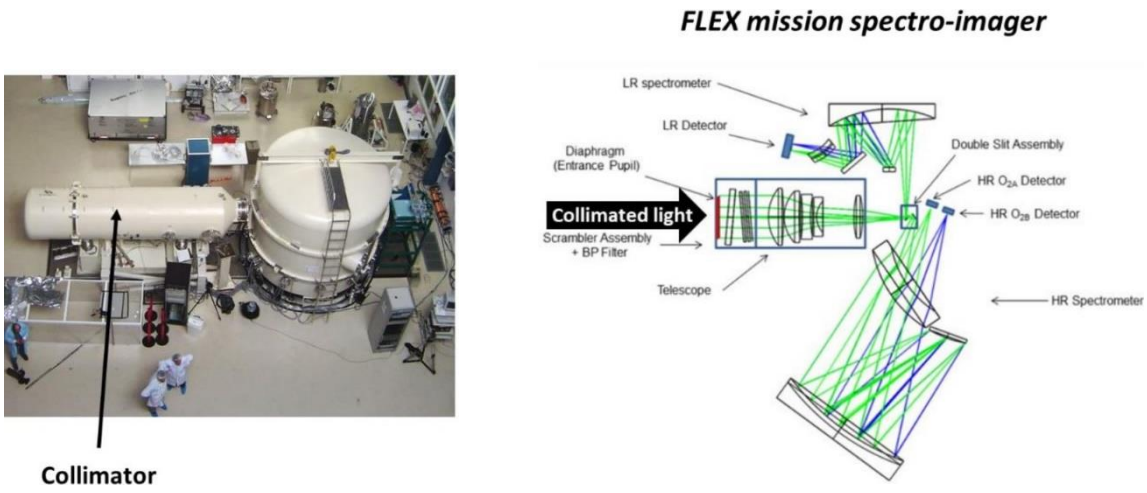
Our TCSPC system was used by the CSL for the validation and improvement of the stray light facility for the FLEX Earth observation instrument. In fact, this application was the direct motivation for the development of our asynchronous TCSPC system as it used an external pico-second pulsed laser that could not be driven by our synchronous system.

7.4.1 Experimental Results

The time of flight-based stray light characterization method was applied to measure the stray light and validate the performance of the facility.

A. Experimental Setup

Figure 7.29-a shows the optical characterization facility at CSL, which operates under vacuum conditions to simulate the space environment. The facility has large dimensions, with approximately 5 meters between the collimator and the sensor. Consequently, the stray light components to be characterized may have significant variations in the optical path lengths of up to several meters. Hence, our TCSPC system which has a temporal resolution of about 20 ps and a large measurement range of 1.3 µs, corresponding to a spatial resolution of less than 1 cm and a spatial measurement range of about 390 m, is appropriate for detecting the different stray light components with much lower cost than a streak camera.



(a) (b)

Figure 7.29 (a) Optical calibration facility in a vacuum chamber, (b) Sketch of the FLEX mission spectro-imager.

A picosecond pulsed laser operating at a frequency of 10.42 MHz was utilized for the light emission. A large collimator reproduced a quasi-point light source that illuminates the instrument. A commercial SPAD from ID Quantique was utilized as a photon detector. The FLEX Earth observation instrument, shown in Figure 7.29-b, was a representative example of the instrument tested in this facility. Figure 7.30 presents a 3D schematic of the facility with the instrument under test.

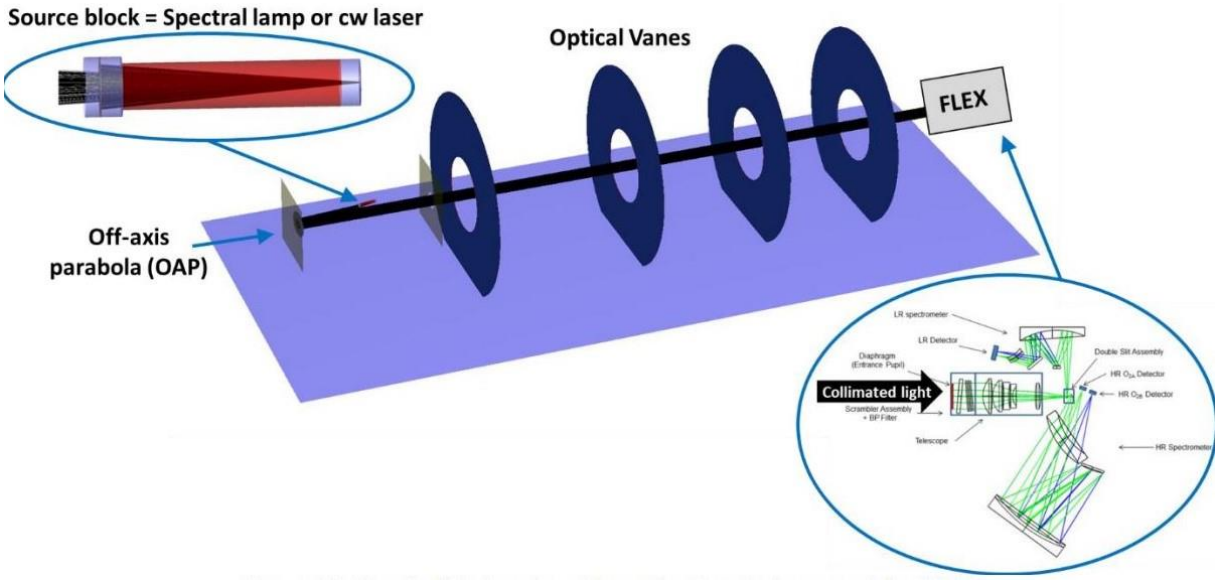


Figure 7.30 Three-dimensional sketch of the optical calibration facility with the FLEX instrument.

Figure 7.31 illustrates the setup, where the instrument under test was replaced by the TCSPC system with the SPAD placed at the position of the instrument entrance. This allowed the measurement of the stray light produced by the facility and incident on the instrument, and the temporal and spatial discrimination of its components. Here, the spatial dimension refers to the angular distribution of the stray light.

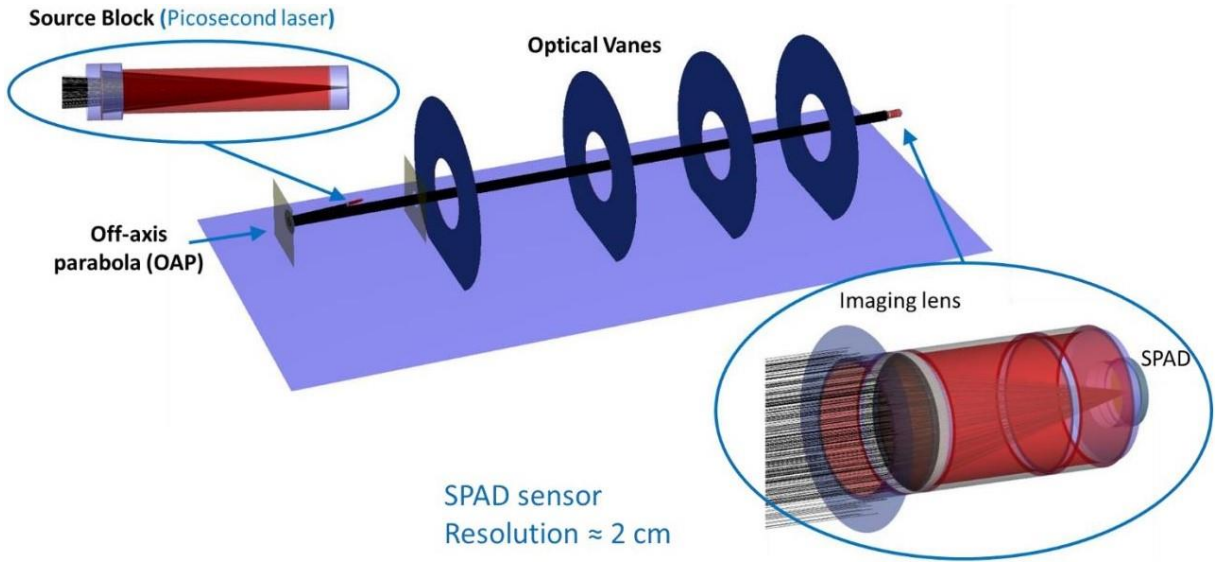


Figure 7.31 Experimental setup: the FLEX instrument is replaced by the SPAD of the TCSPC system to validate the characterization facility.

An imaging lens was placed in front of the SPAD to reproduce the effect of the telescope. The SPAD, having a single-pixel sensing area, only detects the light that arrives directly towards the optical axis of the imaging lens and within a narrow angular range, termed the detector field of view (FOV). For the spatial discrimination of the stray light components, the SPAD was mounted on an angular scan mechanism that enabled performing 2D angular scanning. The angular mapping of stray light, decomposed by the time of flight, was obtained by combining the images together,

B. Experimental Results

The 2D angular scan results, shown in Figure 7.32, demonstrate the discrimination and the angular distribution of the different stray light components that take different optical path lengths. For instance, the left image depicts the stray light arriving with an optical path length of 1.1 meters shorter than the nominal one. This component originates from light that directly illuminates the optical fiber before the collimator. In addition, the stray light from the individual vanes or from the wall behind the collimator is also discriminated, as illustrated in the middle and right images. These images, along with others corresponding to different times of flight, provide the angular distribution of the different stray light paths and enable the identification of their origins based on their time of flight.

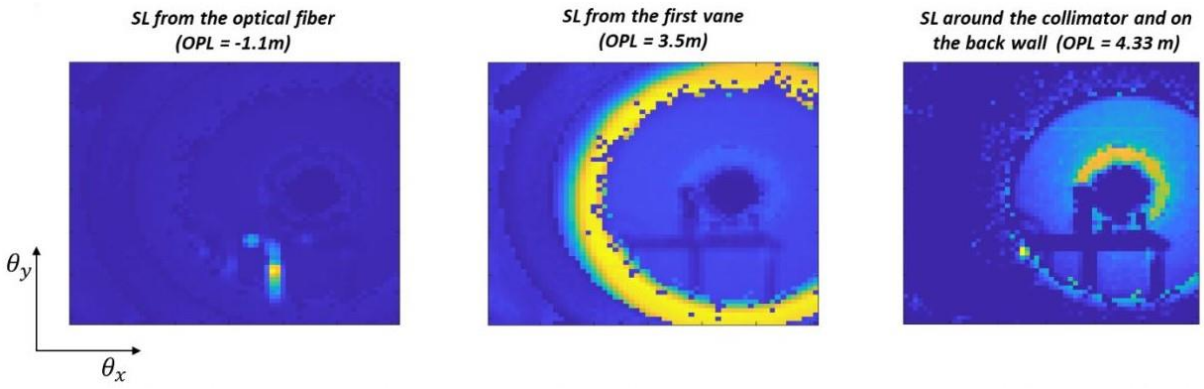


Figure 7.32 Illustration of stray light pattern arriving at the SPAD detector at different moments, showing different stray light contributors.

Figure 7.33 presents the temporal integration of the near field stray light mapping, which corresponds to the stray light in the vicinity of the image-forming beam. The left image exhibits multiple features, such as tilted squares and spider-like patterns. The nominal image signal appears in the central region as a bright spot. A zoomed-in view of this region (in linear scale)

is displayed on the right image, which confirms the nominal signal with an angular extent consistent with that of the collimator, and some broadening due to the finite extent of the detector's field of view (FOV).

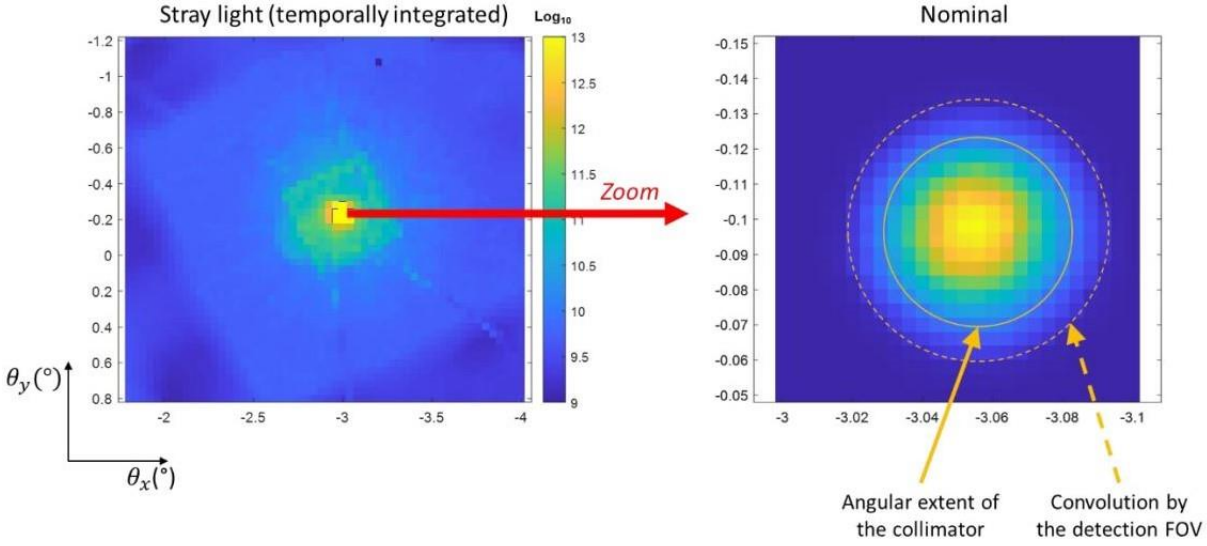


Figure 7.33 Stray light from the facility arriving at the same moment as the nominal.

The origin of the square and spider features was determined based on the time of flight information. They originate from the ghost reflection in the SPAD itself. Figure 7.34 depicts the front side of the SPAD, which clearly exhibits these features. This phenomenon occurs when direct light that illuminates the side of the SPAD is reflected back and creates a ghost in a window near the sensitive area. This means that the SPAD sees itself and it could blur the signal and decrease the detection sensitivity. A possible solution for this problem is to place a mask before the SPAD to ensure that only the sensitive area is exposed to light.

Figure 7.34 Picture of the SPAD front side, showing the sensitive area and the surrounding window.

7.5 Conclusion

In this chapter, we demonstrated the employment of our TCSPC system in four applications in environmental, biological, and astronomical fields, namely water pollution sensing, real-time microfluidic droplet screening and detection, time-resolved optical turbidity, and stray light characterization. The diversity of these applications enabled us to validate the performance of our system in various aspects and to implement different modifications and enhancements to improve the performance. For example, we integrated a droplet detection feature for the microfluidic droplet screening and sorting application, which enabled real-time data processing at high droplet rates. We also implemented a feedback control loop to stabilize the laser intensity, which enhanced the accuracy and consistency of the measurements.

Some of the remarkable results that we achieved in these applications include:

- In the water pollution sensor application, we demonstrated that our system could detect and quantify the presence of B[a]P by measuring the fluorescence signal of the sample, and obtained a clear monoexponential decay with a FLT of about 14 ns.
- In the droplet screening and sorting application, we showed that our droplet detection approach enabled the detection and the FLT estimation of fluorescein droplets at a high throughput of more than 3500 droplets per second, which is the state-of-the-art rate using the TCSPC technique, with an average FLT of the detected droplets in agreement with the theoretical value and a standard deviation close to the theoretical value.
- In the TROT application, we demonstrated the TROT principle and its advantages over the conventional turbidity measurement techniques, such as its robustness against biofouling by using the mean arrival time of photons as a turbidity indicator.
- In the stray light characterization, we demonstrated the TOF-based technique and conducted a long continuous measurement of stray light for two weeks, which proved the reliability of our system.

These results validated the versatility of our TCSPC system and its potential for different applications in various scientific and industrial domains.

Chapter 8: Conclusion and Future Work

8.1 Conclusion

This thesis presented the design, implementation, and evaluation of a low-cost, high-performance, and versatile TCSPC system based on a SoC-FPGA platform. It also discussed the employment of this system in four different applications in various domains.

The thesis started with a comprehensive overview of the TCSPC technique and its main components, such as excitation light sources, photon detectors, and TDCs, in Chapter 2. It evaluated and compared the various choices for each component, highlighting their advantages and drawbacks.

Chapter 3 introduced the design and implementation of a synchronous coarse-fine TDC. This design employed the TDL structure to achieve high temporal resolution and integrated a robust coarse structure that avoids the metastability problem caused by the asynchronous sampling of the counter value. This chapter also addressed several timing issues that were encountered due to the limitations of low-cost FPGA devices and proposed solutions to ensure the correct functionality and reliability of the TDC system. Moreover, the TDC design was extended to support asynchronous excitation light sources, by adding another time measurement channel for the excitation light reference signal.

Chapter 4 presented the data processing approach that utilized the hard processor system (HPS) of the Cyclone V S0C-FPGA to perform on-board data processing. It described the data transfer mechanism and the different components and interfaces involved in the data transfer co-design, which employed a direct memory access (DMA) engine to offload the processor from the data movement operations and preserve its resources for data processing. This chapter also proposed a data buffering technique that enables the DMA-based transfer of the TDC data at a constant rate, enabling online data processing. Furthermore, it outlined the essential data processing tasks such as histogram construction, noise elimination, photon counting, FLT estimation, and COG calculation, which were implemented by an optimized bare-metal C application running on the embedded processor.

Chapter 5 presented and compared the prevalent calibration techniques for synchronous and asynchronous TDCs, namely the bin-by-bin and average-bin-width methods. It also proposed and evaluated a novel calibration technique for asynchronous TDCs, called the Matrix calibration, which achieved up to 10 times better results than the bin-by-bin method applied to histograms.

Chapter 6 evaluated the temporal performance of our TDCs and TCSPC system. The evaluation included assessing the precision and accuracy of the synchronous and asynchronous TDCs, investigating the effect of temperature on the TDC performance, and proposing solutions to mitigate the temperature-induced errors, such as the on-the-flight calibration method for asynchronous TDCs and different correction coefficients. It also evaluated the performance of our TCSPC system by measuring its IRF and recording the fluorescence signal of a reference fluorophore and estimating its fluorescence lifetime.

Chapter 7 demonstrated the employment of our TCSPC system in four distinct applications in the environmental, biological, and astronomical fields, namely water pollution sensing, real-time microfluidic droplet screening and detection, time-resolved optical turbidity, and stray light characterization. These applications allowed us to evaluate the capabilities of our system and to introduce different adaptations and enhancements to optimize its performance.

Our TCSPC system demonstrated high temporal resolution and precision acceptable for our target applications. Furthermore, it yielded fluorescence lifetime estimations that are consistent with the expected values from the literature for various fluorophores. Tables 8.1 and 8.2 summarize the different specifications of our TCSPC system operating in synchronous and asynchronous mode, respectively.

Table 8.1 Performance parameters of the TCSPC system in synchronous mode.

Parameter	Min	Typical	Max	Unit
Laser repetition rate (f_L)	1.10-7		95	MHz
Raw Bin size	0	20.56	65	ps
Calibrated Bin size	20	20.56	22	ps
Dead time (TDC)		15.79		ns
Photon detection rate	0		63.3	MHz
RMS precision	23	24.3	27	ps
Jitter	50	130	200	ps FWHM

Table 8.2 Performance parameters of the TCSPC system in asynchronous mode.

Parameter	Min	Typical	Max	Unit
Laser Repetition rate (f_L)	1.10-7		100	MHz
Raw Histogram Bin size	14.3	21,9	29.7	ps
Calibrated bin size	19.35	20.57	22.15	ps
Dead time (TDC)		15.79		ns
Photon detection rate (TDC)	0		Min(63.3, f_L/2)	MHz
RMS precision	28	37.5	47	ps
Jitter	50	170	300	ps FWHM

The system also demonstrated high versatility and potential for different applications in various scientific and industrial domains. It was successfully employed in four distinct applications, and it can be easily adapted to other applications with different data processing requirements, by integrating appropriate data processing algorithms in the C program and updating the execution file on the microSD card of the SoC-FPGA kit.

8.2 Future Work

The proposed TCSPC system offers a versatile and high-performance solution for various applications that require time-resolved measurements. However, there are still several aspects that can be improved and extended in the future. Some possible directions for future work are:

- **Enhancing the System Performance and Functionality**

Improving the temporal resolution and precision of the TDC by employing more advanced FPGA devices or by implementing other TDC architectures, such as the Wave-Union (WU) [122] method and dual-sampling architectures [127] that can enhance the resolution beyond the cell delay limit.

Reducing the dead time of the TDC by implementing a pipeline encoding architecture, which can increase the throughput and avoid data loss.

Improving the stability of the TDC with temperature variations by applying correction coefficients or by regulating the voltage using temperature sensors and feedback loops to compensate for the temperature-introduced errors. These improvements can enable the TDC to operate reliably and accurately in various environments and conditions.

- **Expanding the System Capabilities and Applications**

Extending the system to support multi-channel or multi-dimensional measurements by adding multiple TDC channels. This can widen the range of applications that can benefit from the system, such as biomedical imaging.

Incorporating additional algorithms in the software to perform more sophisticated data processing tasks, such as deconvolution, multiexponential curve fitting, and machine learning. This can facilitate supporting more applications and improve the quality and accuracy of the extracted parameters.

Exploring other applications that require high-resolution time measurements, such as quantum cryptography, radar systems, or medical imaging. The proposed TCSPC system can be adapted and customized to meet the specific requirements and challenges of these applications.

- **Future Work on the Studied Applications**

For WPS application: optimizing the detected photon rate by using a more powerful light source with an optimal excitation wavelength, which enhances the signal-to-noise ratio and thus lowers the limit of detection. Replacing the current optical setup with a filterless compact design based on the time gating technique, that enables simplifying the system and making it suitable for in-situ measurements.

For the microfluidic droplet sorting application: conducting further experiments to examine the sorting function and evaluate its efficiency and accuracy.

Bibliography

[1] L. M. Bollinger and G. Neil Thomas, “Measurement of the Time Dependence of Scintillation Intensity by a Delayed-Coincidence Method,” *Review of Scientific Instruments*, vol. 32, no. 9, pp. 1044–1050, Sep. 1961.

[2] Dr. Wolfgang Becker, *The bh TCSPC handbook*. Becker & Hickl GmbH, 2021.

[3] R. Pike, “Lasers, photon statistics, photon-correlation spectroscopy and subsequent applications,” *Journal of the European Optical Society: Rapid Publications*, vol. 5, Sep. 2010.

[4] L. M. Hirvonen and K. Suhling, “Wide-field TCSPC: methods and applications,” *Measurement Science and Technology*, vol. 28, no. 1, p. 012003, Dec. 2016.

[5] Schuyler R and I. Isenberg, “A Monophoton Fluorometer with Energy Discrimination,” *Review of Scientific Instruments*, vol. 42, no. 6, pp. 813–817, Jun. 1971.

[6] J. Yguerabide, “Nanosecond fluorescence spectroscopy of macromolecules,” *Methods Enzymol*, vol. 26, pp. 498–578, Jan. 1972.

[7] R. Z. Bachrach, “A Photon Counting Apparatus for Kinetic and Spectral Measurements,” *Review of Scientific Instruments*, vol. 43, no. 5, pp. 734–737, May 1972.

[8] Th. Binkert, H. P. Tschanz, and P. E. Zinsli, “The measurement of fluorescence decay curves with the single-photon counting method and the evaluation of rate parameters,” *Journal of Luminescence*, vol. 5, no. 3, pp. 187–217, Jul. 1972.

[9] C. M. Lewis, W. R. Ware, L. J. Doemeny, and T. L. Nemzek, “The Measurement of Short-Lived Fluorescence Decay Using the Single Photon Counting Method,” *Review of Scientific Instruments*, vol. 44, no. 2, pp. 107–114, Feb. 1973.

[10] S. Cova, M. Bertolaccini, and C. Bussolati, “The measurement of luminescence waveforms by single-photon techniques,” *physica status solidi (a)*, vol. 18, no. 1, pp. 11–62, Jul. 1973.

[11] B. Leskovar, Chung Mau Lo, P. R. Hartig, and K. Sauer, “Photon counting system for sub-nanosecond fluorescence lifetime measurements,” *Review of Scientific Instruments*, vol. 47, no. 9, pp. 1113–1121, Sep. 1976.

[12] D. V. O'Connor and D. Phillips, *Time-correlated Single Photon Counting*. London: Academic Press, 1984.

[13] B. Leskovar, "Nanosecond Fluorescence Spectroscopy," *IEEE Transactions on Nuclear Science*, vol. 32, no. 3, pp. 1232–1241, Mar. 1985.

[14] W. Becker, *Advanced Time-Correlated Single Photon Counting Techniques*. Springer Science & Business Media, 2005.

[15] I. Bugiel, K. Konig, and H. Wabnitz, "Investigation of cells by Fluorescence Laser Scanning Microscopy with Sub-nanosecond Time Resolution," *Lasers in the Life Sciences*, vol. 3, pp. 47–53, 1989.

[16] E. Gratton, "Fluorescence lifetime imaging for the two-photon microscope: time-domain and frequency-domain methods," *Journal of Biomedical Optics*, vol. 8, no. 3, p. 381, Jul. 2003.

[17] S. Kinoshita, H. Ohta, and Takashi Kushida, "Sub-nanosecond fluorescence-lifetime measuring system using single photon counting method with mode-locked laser excitation," *Review of Scientific Instruments*, vol. 52, no. 4, pp. 572–575, Apr. 1981.

[18] M. Y. Berezin and S. Achilefu, "Fluorescence Lifetime Measurements and Biological Imaging," *Chemical reviews*, vol. 110, no. 5, pp. 2641–2684, May 2010.

[19] S. M. Hickey *et al.*, "Fluorescence Microscopy—An Outline of Hardware, Biological Handling, and Fluorophore Considerations," *Cells*, vol. 11, no. 1, p. 35, Dec. 2021.

[20] W. BECKER, "Fluorescence lifetime imaging - techniques and applications," *Journal of Microscopy*, vol. 247, no. 2, pp. 119–136, May 2012.

[21] L. Marcu, "Fluorescence Lifetime Techniques in Medical Applications," *Annals of Biomedical Engineering*, vol. 40, no. 2, pp. 304–331, Jan. 2012.

[22] J. L. Lagarto *et al.*, "Multispectral Depth-Resolved Fluorescence Lifetime Spectroscopy Using SPAD Array Detectors and Fiber Probes," *Sensors*, vol. 19, no. 12, p. 2678, Jan. 2019.

[23] K. Suhling, P. M. W. French, and D. Phillips, "Time-resolved fluorescence microscopy," *Photochemical & Photobiological Sciences*, vol. 4, no. 1, pp. 13–22, 2005.

[24] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*. New York, Ny: Springer Us, 2006.

[25] G. O. Fruhwirth, S. Ameer-Beg, R. Cook, T. Watson, T. Ng, and F. Festy, "Fluorescence lifetime endoscopy using TCSPC for the measurement of FRET in live cells," *Optics Express*, vol. 18, no. 11, pp. 11148–11158, May 2010.

[26] J. Rapp, A. Dawson, and V. K. Goyal, “Improving Lidar Depth Resolution with Dither,” in *25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 1553–1557.

[27] M.-C. Amann, T. M. Bosch, M. Lescure, R. A. Myllylae, and M. Rioux, “Laser ranging: a critical review of unusual techniques for distance measurement,” *Optical Engineering*, vol. 40, no. 1, pp. 10–19, Jan. 2001.

[28] B. Behroozpour, P. A. M. Sandborn, M. C. Wu, and B. E. Boser, “Lidar System Architectures and Circuits,” *IEEE Communications Magazine*, vol. 55, no. 10, pp. 135–142, Oct. 2017.

[29] W. Liu *et al.*, “Learning to Match 2D Images and 3D LiDAR Point Clouds for Outdoor Augmented Reality,” in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 2020, pp. 654–655.

[30] D. M. Winker, R. H. Couch, and M. P. McCormick, “An overview of LITE: NASA’s Lidar In-space Technology Experiment,” *Proceedings of the IEEE*, vol. 84, no. 2, pp. 164–180, Feb. 1996.

[31] J. S. Massa, G. S. Buller, A. C. Walker, S. Cova, M. Umasuthan, and A. M. Wallace, “Time-of-flight optical ranging system based on time-correlated single-photon counting,” *Applied Optics*, vol. 37, no. 31, p. 7298, Nov. 1998.

[32] T. Wu and T.-C. Hsueh, “A High-Resolution Single-Photon Arrival-Time Measurement With Self-Antithetic Variance Reduction in Quantum Applications: Theoretical Analysis and Performance Estimation,” *IEEE Transactions on Quantum Engineering*, vol. 3, pp. 1–15, 2022.

[33] A. W. Sloman and M. D. Swords, “A fast and economical gated discriminator,” *Journal of Physics E: Scientific Instruments*, vol. 11, no. 6, p. 521, 1978.

[34] J. Arlt, D. Tyndall, B. R. Rae, D. D.-U. . Li, J. A. Richardson, and R. K. Henderson, “A study of pile-up in integrated time-correlated single photon counting systems,” *Review of Scientific Instruments*, vol. 84, no. 10, p. 103105, Oct. 2013.

[35] N. Franch, O. Alonso, J. Canals, A. Vilà, and A. Dieguez, “A low cost fluorescence lifetime measurement system based on SPAD detectors and FPGA processing,” *Journal of Instrumentation*, vol. 12, no. 02, p. C02070, 2017.

[36] J. Bouchard *et al.*, “A Low-Cost Time-Correlated Single Photon Counting System for Multiview Time-Domain Diffuse Optical Tomography,” *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 10, pp. 2505–2515, 2017.

[37] X. T. Nguyen *et al.*, “An efficient FPGA-based database processor for fast database

analytics," in *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1758–1761.

[38] M. Wahl, "Modern TCSPC Electronics: Principles and Acquisition Modes," in *Advanced Photon Counting: Applications, Methods, Instrumentation*, P. Kapusta, M. Wahl, and R. Erdmann, Eds., Cham: Springer International Publishing, 2015, pp. 1–21.

[39] W. Becker, "Optical Signal Recording," in *Advanced Time-Correlated Single Photon Counting Techniques*, C. A. W, J. P. Toennies, and W. Zinth, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 1–9.

[40] P. R. Hartig, K. Sauer, Chung Mau Lo, and B. Leskovar, "Measurement of very short fluorescence lifetimes by single-photon counting," *Review of scientific instruments*, vol. 47, no. 9, pp. 1122–1129, Sep. 1976.

[41] J. Yguerabide, "[24] Nanosecond fluorescence spectroscopy of macromolecules," *Methods in Enzymology*, vol. 26, pp. 498–578, Jan. 1972.

[42] B. Wellegehausen, H. Welling, and R. Beigang, "A narrowband jet stream dye laser," *Applied physics*, vol. 3, no. 5, pp. 387–391, 1974.

[43] M. Kress *et al.*, "Time-resolved microspectrofluorometry and fluorescence lifetime imaging of photosensitizers using picosecond pulsed diode lasers in laser scanning microscopes," *Journal of Biomedical Optics*, vol. 8, no. 1, p. 26, Jan. 2003.

[44] M. Maus *et al.*, "New picosecond laser system for easy tunability over the whole ultraviolet/visible/near infrared wavelength range based on flexible harmonic generation and optical parametric oscillation," *Review of Scientific Instruments*, vol. 72, no. 1, pp. 36–40, Jan. 2001.

[45] A. Maccarone *et al.*, "Custom-Technology Single-Photon Avalanche Diode Linear Detector Array for Underwater Depth Imaging," *Sensors*, vol. 21, no. 14, p. 4850, Jan. 2021.

[46] Y. Zhang, S. A. Soper, L. R. Middendorf, J. A. Wurm, R. Erdmann, and M. Wahl, "Simple Near-Infrared Time-Correlated Single Photon Counting Instrument with a Pulsed Diode Laser and Avalanche Photodiode for Time-Resolved Measurements in Scanning Applications," *Appl. Spectrosc*, vol. 53, no. 5, pp. 497–504, 1999.

[47] D. Wu *et al.*, "Multibeam single-photon LiDAR with hybrid multiplexing in wavelength and time," *Optics & Laser Technology*, vol. 145, p. 107477, 2022.

[48] K. Bantounos, T. M. Smeeton, and I. Underwood, "Towards a solid-state light detection and ranging system using holographic illumination and time-of-flight image sensing," *J Soc Inf Display*, vol. 30, no. 5, pp. 363–372, 2022.

[49] W. Becker, “Practice of TCSPC Experiments,” in *Advanced Time-Correlated Single Photon Counting Techniques*, C. A. W, J. P. Toennies, and W. Zinth, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 263–346.

[50] Wilfried Uhring, C. Zint, and J. Bartringer, “A low-cost high-repetition-rate picosecond laser diode pulse generator,” in *Proc. SPIE*, Sep. 2004, pp. 583–590.

[51] J. A. Lau, V. B. Verma, D. Schwarzer, and A. M. Wodtke, “Superconducting single-photon detectors in the mid-infrared for physical chemistry and spectroscopy,” *Chemical Society Reviews*, vol. 52, no. 3, pp. 921–941, Jan. 2023.

[52] R. Cheng, Y. Zhou, S. Wang, M. Shen, T. Taher, and H. X. Tang, “A 100pixel photon-number-resolving detector unveiling photon statistics,” *Nature Photonics*, vol. 17, no. 1, pp. 112–119, 2023.

[53] H. Gan *et al.*, “A flat spectral photon flux source for single photon detector quantum efficiency calibration,” in *2015 11th Conference on Lasers and Electro-Optics Pacific Rim (CLEOPR)*, 2015, pp. 1–2.

[54] Y. Xu, P. Xiang, and X. Xie, “Comprehensive understanding of dark count mechanisms of single-photon avalanche diodes fabricated in deep submicron CMOS technologies,” *Solid-State Electronics*, vol. 129, pp. 168–174, 2017.

[55]P. K. Dubey, S. L. Jain, B. C. Arya, and P. S. Kulkarni, “Discriminator threshold selection logic to improve signal to noise ratio in photon counting,” *MAPAN*, vol. 25, no. 1, pp. 63–70, Mar. 2010.

[56] “PHOTOMULTIPLIER TUBES - Basics and Applications, FOURTH EDITION, Hamamatsu.” Accessed: Dec. 16, 2022. [Online]. Available:

[57] A. Bülter, “Single-Photon Counting Detectors for the Visible Range Between 300 and 1,000 nm,” in *Advanced Photon Counting: Applications, Methods, Instrumentation*, P. Kapusta, M. Wahl, and R. Erdmann, Eds., Cham: Springer International Publishing, 2015, pp. 23–42.

[58] R. Mansmann, T. A. Sipkens, J. Menser, K. J. Daun, T. Dreier, and C. Schulz, “Detector calibration and measurement issues in multi-color time-resolved laser-induced incandescence,” *Applied Physics B*, vol. 125, no. 7, p. 126, 2019.

[59] W. Becker, “Detectors for Photon Counting,” in *Advanced Time-Correlated Single Photon Counting Techniques*, C. A. W, J. P. Toennies, and W. Zinth, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 213–261.

[60] P. Windischhofer and W. Riegler, “Passive quenching, signal shapes, and space charge effects in SPADs and SiPMs,” *Nuclear Instruments and Methods in Physics Research*

Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, vol. 1045, p. 167627, 2023.

[61] L. Liu *et al.*, "Performance of Active-Quenching SPAD Array Based on the Tri-State Gates of FPGA and Packaged with Bare Chip Stacking," *Sensors*, vol. 23, no. 9, 2023.

[62] ID QUANTIQUE, "ID100 Visible Single Photon Detector," Jan. 2023. Accessed: Jun. 03, 2023. [Online]. Available: https://marketing.idquantique.com/acton/attachment/11868/f-0236/1/-/-/-/-/ID100_Brochure.pdf

[63] W. Becker, "Overview of Photon Counting Techniques," in *Advanced Time-Correlated Single Photon Counting Techniques*, C. A. W, J. P. Toennies, and W. Zinth, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 11–25.

[64] J. Kalisz, "Review of methods for time interval measurements with picosecond resolution," *Metrologia*, vol. 41, no. 1, p. 17, 2004.

[65] Z. Cheng, X. Zheng, M. J. Deen, and H. Peng, "Recent Developments and Design Challenges of High-Performance Ring Oscillator CMOS Time-to-Digital Converters," *IEEE Transactions on Electron Devices*, vol. 63, pp. 1–17, Dec. 2015.

[66] J. Doernberg, H. .S. Lee, and D. A. Hodges, "Full-speed testing of A/D converters," *IEEE Journal of Solid-State Circuits*, vol. 19, no. 6, pp. 820–827.

[67] P. Napolitano, A. Moschitta, and P. Carbone, "A survey on time interval measurement techniques and testing methods," in *2010 IEEE Instrumentation & Measurement Technology Conference Proceedings*, pp. 181–186.

[68] J. Szyduczyński, D. Kościelnik, and M. Miśkowicz, "Time-to-digital conversion techniques: a survey of recent developments," *Measurement*, vol. 214, p. 112762, 2023.

[69] V. Sesta, A. Incoronato, F. Madonini, and F. Villa, "Time-to-digital converters and histogram builders in SPAD arrays for pulsed-LiDAR," *Measurement*, vol. 212, p. 112705, 2023.

[70] S. Henzler and S. Henzler, "Time-to-Digital Converter Basics," in *Time-to-Digital Converters*, Dordrecht: Springer Netherlands, 2010, pp. 5–18.

[71] K. Maatta and J. Kostamovaara, "A high-precision time-to-digital converter for pulsed time-of-flight laser radar applications," *IEEE Transactions on Instrumentation and Measurement*, vol. 47, no. 2, pp. 521–536.

[72] E. RaisanenRuotsalainen, T. Rahkonen, and J. Kostamovaara, "A high resolution time-to-digital converter based on time-to-voltage interpolation," in *Proceedings of the 23rd European Solid-State Circuits Conference*, pp. 332–335.

[73] Mattada, Mahantesh P and H. Guhilot, “Time-to-digital converters—A comprehensive review,” *Int J Circ Theor Appl*, vol. 49, no. 3, pp. 778–800, 2021.

[74] J. Szyduczyński, D. Kościelnik, and M. Miśkowicz, “Time-to-digital conversion techniques: a survey of recent developments,” *Measurement*, vol. 214, p. 112762, 2023.

[75] Mattada, Mahantesh P and H. Guhilot, “Time-to-digital converters—A comprehensive review,” *Int J Circ Theor Appl*, vol. 49, no. 3, pp. 778–800, 2021.

[76] Y. Chen, “Time Resolution Improvement Using Dual Delay Lines for Field-Programmable-Gate-Array-Based Time-to-Digital Converters with Real-Time Calibration,” *Applied Sciences*, vol. 9, no. 1, 2019.

[77] M. P. Mattada and H. Guhilot, “Area efficient vernier Time to Digital Converter(TDC) with improved resolution using identical ring oscillators on FPGA,” in *INTERNATIONAL CONFERENCE ON SMART STRUCTURES AND SYSTEMSICSSS’13*, pp. 125–130.

[78] K. Cui, Z. Ren, X. Li, Z. Liu, and R. Zhu, “A High-Linearity, Ring-Oscillator-Based, Vernier Time-to-Digital Converter Utilizing Carry Chains in FPGAs,” *IEEE Transactions on Nuclear Science*, vol. 64, no. 1, pp. 697–704.

[79] F. Nolet *et al.*, “22 μW, 5.1 ps LSB, 5.5 ps RMS jitter Vernier time-to-digital converter in CMOS 65 nm for single photon avalanche diode array,” *Electron. Lett.*, vol. 56, no. 9, pp. 424–426, 2020.

[80] L. Vercesi, A. Liscidini, and R. Castello, “Two-Dimensions Vernier Time-to-Digital Converter,” *IEEE Journal of Solid-State Circuits*, vol. 45, no. 8, pp. 1504–1512.

[81] T. .i. Otsuji, “A picosecond-accuracy, 700MHz range, Si bipolar time interval counter LSI,” *IEEE Journal of Solid-State Circuits*, vol. 28, no. 9, pp. 941–947.

[82] L. Vercesi, A. Liscidini, and R. Castello, “Two-Dimensions Vernier Time-to-Digital Converter,” *IEEE Journal of Solid-State Circuits*, vol. 45, no. 8, pp. 1504–1512.

[83] K. Cui, Z. Ren, X. Li, Z. Liu, and R. Zhu, “A High-Linearity, Ring-Oscillator-Based, Vernier Time-to-Digital Converter Utilizing Carry Chains in FPGAs,” *IEEE Transactions on Nuclear Science*, vol. 64, no. 1, pp. 697–704.

[84] P. Dudek, S. Szczepanski, and J. V. Hatfield, “A high-resolution CMOS time-to-digital converter utilizing a Vernier delay line,” *IEEE Journal of Solid-State Circuits*, vol. 35, no. 2, pp. 240–247.

[85] N. Xing, H. Song, D. K. Jeong, and S. Kim, “A PVT-insensitive time-to-digital converter using fractional difference Vernier delay lines,” in *2009 IEEE International SOC Conference (SOCC)*, pp. 43–46.

[86] L. Vercesi, A. Liscidini, and R. Castello, “Two-Dimensions Vernier Time-to-Digital Converter,” *IEEE Journal of Solid-State Circuits*, vol. 45, no. 8, pp. 1504–1512.

[87] P. Lu, Y. Wu, and P. Andreani, “A 2.2ps Two-Dimensional Gated-Vernier Time-to-Digital Converter With Digital Calibration,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 63, no. 11, pp. 1019–1023.

[88] R. Enomoto, T. Iizuka, T. Koga, T. Nakura, and K. Asada, “A 16bit 2.0ps Resolution Two-Step TDC in 0.18μ m CMOS Utilizing Pulse-Shrinking Fine Stage With Built-In Coarse Gain Calibration,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 27, no. 1, pp. 11–19.

[89] L. Xiang, P. Yang, T. Wu, and M. Zhou, “Ultra compact pulse shrinking TDC on FPGA,” *Measurement*, vol. 203, p. 111874, 2022.

[90] P. Chen, S. Liu, and J. Wu, “A CMOS pulse-shrinking delay element for time interval measurement,” *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 47, no. 9, pp. 954–958.

[91] L. Xiang, P. Yang, T. Wu, and M. Zhou, “Ultra compact pulse shrinking TDC on FPGA,” *Measurement*, vol. 203, p. 111874, 2022.

[92] R. Szplet and K. Klepacki, “An FPGA-Integrated Time-to-Digital Converter Based on TwoStage Pulse Shrinking,” *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 6, pp. 1663–1670.

[93] Y. Sano, Y. Horii, M. Ikeno, O. Sasaki, M. Tomoto, and T. Uchida, “Sub-nanosecond time-to-digital converter implemented in a Kintex7 FPGA,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 874, pp. 50–56, 2017

[94] F. S. Deng, H. Liang, J. Y. Tang, and B. J. Ye, “Design of a 32channel NIM TDC module based on single FPGA for the μSR spectrometer prototype at the China Spallation Neutron Source,” *Journal of Instrumentation*, vol. 14, no. 06, p. T06002, 2019

[95] L. N. Cojocariu, D. FouldsHolt, F. Keizer, V. M. Placinta, and S. Wotton, “A multichannel TDC-in-FPGA with 150 ps bins for time-resolved readout of Cherenkov photons,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, p. 168483, 2023.

[96] J. Qin, D. Guo, L. Zhao, S. Lan, Y. Wang, and Q. An, “Design and performance of a 16channel coarse-fine TDC prototype ASIC,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated*

Equipment, vol. 1050, p. 168167, 2023.

[97] K. C. Choi, S. W. Lee, B. C. Lee, and W. Y. Choi, "A Time-to-Digital Converter Based on a Multiphase Reference Clock and a Binary Counter With a Novel Sampling Error Corrector," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 59, no. 3, pp. 143–147.

[98] Z. Cheng, X. Zheng, M. J. Deen, and H. Peng, "Recent Developments and Design Challenges of High-Performance Ring Oscillator CMOS Time-to-Digital Converters," *IEEE Transactions on Electron Devices*, vol. 63, no. 1, pp. 235–251

[99] M. Z. Straayer and M. H. Perrott, "A Multi-Path Gated Ring Oscillator TDC With First-Order Noise Shaping," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 4, pp. 1089–1098.

[100] A. Elshazly, S. Rao, B. Young, and P. K. Hanumolu, "A Noise-Shaping Time-to-Digital Converter Using Switched-Ring Oscillators—Analysis, Design, and Measurement Techniques," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 5, pp. 1184–1197.

[101] Z. Cheng, M. J. Deen, and H. Peng, "A Low-Power Gateable Vernier Ring Oscillator Time-to-Digital Converter for Biomedical Imaging Applications," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 10, no. 2, pp. 445–454

[102] P. Lu, Y. Wu, and P. Andreani, "A 2.2ps Two-Dimensional Gated-Vernier Time-to-Digital Converter With Digital Calibration," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 63, no. 11, pp. 1019–1023.

[103] M. Z. Straayer and M. H. Perrott, "An efficient high-resolution 11bit noise-shaping multipath gated ring oscillator TDC," in *2008 IEEE Symposium on VLSI Circuits*, pp. 82–83.

[104] R. Nutt, "Digital Time Intervalometer," *Review of Scientific Instruments*, vol. 39, no. 9, pp. 1342–1345, Sep. 1968.

[105] S. Huang *et al.*, "Design of a multichannel high-resolution TDC based on FPGA," *Journal of Instrumentation*, vol. 17, no. 07, p. P07037, 2022.

[106] S. Tancock, E. Arabul, and N. Dahnoun, "A Review of New Time-to-Digital Conversion Techniques," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 10, pp. 3406–3417.

[107] R. Machado, J. Cabral, and F. S. Alves, "Recent Developments and Challenges in FPGA-Based Time-to-Digital Converters," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 11, pp. 4205–4221.

[108] Z. Cheng, X. Zheng, M. J. Deen, and H. Peng, "Recent Developments and Design

Challenges of High-Performance Ring Oscillator CMOS Time-to-Digital Converters," *IEEE Transactions on Electron Devices*, vol. 63, no. 1, pp. 235–251.

[109] J. Qin, D. Guo, L. Zhao, S. Lan, Y. Wang, and Q. An, "Design and performance of a 16channel coarsefine TDC prototype ASIC," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 1050, p. 168167, 2023.

[110] R. Machado, J. Cabral, and F. S. Alves, "Recent Developments and Challenges in FPGABased TimetoDigital Converters," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 11, pp. 4205–4221.

[111] J. Szyduczyński, D. Kościelnik, and M. Miśkowicz, "Time-to-digital conversion techniques: a survey of recent developments," *Measurement*, vol. 214, p. 112762, 2023.

[112] A. I. Hussein, S. Vasadi, and J. Paramesh, "A 450 fs 65nm CMOS Millimeter-Wave Time-to-Digital Converter Using Statistical Element Selection for All-Digital PLLs," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 2, pp. 357–374.

[113] M. Liu, X. Lai, and Y. Wang, "A Novel 550fs Time-Resolution 7.5bit Stochastic Time-to-Digital Converter Based on Two Arbiter Groups," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–11, 2023.

[114] R. K. Henderson *et al.*, "A 192×128 Time Correlated Single Photon Counting Imager in 40nm CMOS Technology," in *ESSCIRC 2018IEEE 44th European Solid State Circuits Conference (ESSCIRC)*, pp. 54–57.

[115] A. Boutros and V. Betz, "FPGA Architecture: Principles and Progression," *IEEE Circuits and Systems Magazine*, vol. 21, no. 2, pp. 4–29.

[116] Xilinx, "AMD Adaptive Computing Documentation Portal," *docs.xilinx.com*, Nov. 07, 2022. https://docs.xilinx.com/v/u/en-US/ds891-zynq-ultrascale-plus-overview (accessed Jul. 01, 2023).

[117] Intel, "Intel® Stratix® 10 GX/SX Device Overview," *Intel.com*, Aug. 18, 2022. https://cdrdv2.intel.com/v1/dl/getContent/670537?fileName=s10-overview-683729-670537.pdf (accessed Jul. 01, 2023).

[118] Terasic Inc., "DE10-Nano Cyclone V SoC with Dual-core ARM Cortex-A9 User Manual," *Terasic.com.tw*, 2017. https://www.terasic.com.tw/cgi-bin/page/archive_download.pl?Language=Taiwan&No=1046&FID=1c19d1d50e0ee9b21678e881004f6d81 (accessed Jul. 01, 2023).

[119] J. Kalisz, R. Szplet, J. Pasierbinski, and A. Poniecki, "Field-programmable-gate-array-based time-to-digital converter with 200ps resolution," *IEEE Transactions on*

Instrumentation and Measurement, vol. 46, no. 1, pp. 51–55.

[120] C. Liu, Y. Wang, P. Kuang, D. Li, and X. Cheng, "A 3.9 ps RMS resolution time-to-digital converter using dual-sampling method on Kintex UltraScale FPGA," in *2016 IEEENPSS Real Time Conference (RT)*, pp. 1–3.

[121] M. A. Daigneault and J. P. David, "A novel 10 ps resolution TDC architecture implemented in a 130nm process FPGA," in *Proceedings of the 8th IEEE International NEWCAS Conference 2010*, pp. 281–284.

[122] J. Wu and Z. Shi, "The 10ps wave union TDC: Improving FPGA TDC resolution beyond its cell delay," in *2008 IEEE Nuclear Science Symposium Conference Record*, pp. 3440–3446.

[123] R. Szplet and K. Klepacki, "An FPGA-Integrated Time-to-Digital Converter Based on Two-Stage Pulse Shrinking," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 6, pp. 1663–1670.

[124] B. Markovic *et al.*, "A compact Time-to-Digital Converter (TDC) module with 10 ps resolution and less than 1.5% LSB DNL," in *IEEE Photonics Conference 2012*, pp. 26–27.

[125] O. Petura, "True random number generators for cryptography : Design, securing and evaluation," Université de Lyon, 2019. Accessed: Mar. 03, 2023. [Online]. Available: https://theses.hal.science/tel-02895861/preview/These-Oto-Petura-2019.pdf

[126] M. Fishburn, L. H. Menninga, C. Favi, and E. Charbon, "A 19.6 ps, FPGA-Based TDC With Multiple Channels for Open Source Applications," *IEEE Transactions on Nuclear Science*, vol. 60, no. 3, pp. 2203–2208.

[127] Y. Wang, J. Kuang, C. Liu, and Q. Cao, "A 3.9ps RMS Precision Time-to-Digital Converter Using Ones-Counter Encoding Scheme in a Kintex7 FPGA," *IEEE Transactions on Nuclear Science*, vol. 64, no. 10, pp. 2713–2718.

[128] J. Wang, S. Liu, Q. Shen, H. Li, and Q. An, "A Fully Fledged TDC Implemented in Field-Programmable Gate Arrays," *IEEE Transactions on Nuclear Science*, vol. 57, no. 2, pp. 446–450.

[129] F. Dadouche, T. Turko, W. Uhring, I. Malass, J. Bartringer, and J. Le, "Design Methodology of TDC on Low Cost FPGA Targets Case Study: Implementation of a 42 ps Resolution TDC on a Cyclone IV FPGA Target," in *SENSORCOMM*, Venise, Italy, Aug. 2015, pp. 29–34.

[130] J. Wu, Z. Shi, and I. Y. Wang, "Firmware-only implementation of time-to-digital converter (TDC) in field-programmable gate array (FPGA)," in *2003 IEEE Nuclear*

Science Symposium. Conference Record (IEEE Cat. No.03CH37515), pp. 177–181 Vol.1.

[131] Altera, "Cyclone V Device Handbook Volume 1: Device Interfaces and Integration Subscribe Send Feedback," 2016. Accessed: Jul. 03, 2023. [Online]. Available: http://www.ee.ic.ac.uk/pcheung/teaching/E2_experiment/C5%20handbook%20v1.pdf

[132] F. Dadouche, T. Turko, A. Skilitsi, I. Malass, W. Uhring, and J. Leonard, "Design, Implementation and Characterization of Time-to-Digital Converter on Low-Cost FPGA," in *Advances in Sensors: Reviews*, Sensors and Applications in Measuring and Automation Control Systems, 2016, pp. 205–229.

[133] J. Kim, J. H. Jung, Y. Choi, J. Jung, and S. Lee, "Linearity improvement of Ultra-Scale+ FPG-Abased time-to-digital converter," *Nuclear Engineering and Technology*, vol. 55, no. 2, pp. 484–492, 2023.

[134] R. Szplet and A. Czuba, "Two-Stage Clock-Free Time-to-Digital Converter Based on Vernier and Tapped Delay Lines in FPGA Device," *Electronics*, vol. 10, no. 18, 2021.

[135] H. Xia, G. Cao, and N. Dong, "A 6.6 ps RMS resolution time-to-digital converter using interleaved sampling method in a 28 nm FPGA," *Review of Scientific Instruments*, vol. 90, no. 4, p. 044706, Apr. 2019.

[136] E. Dikopoulos, M. Birbas, and A. Birbas, "An Adaptive Downsampling FPGA-Based TDC Implementation for Time Measurement Improvement," *Chips*, vol. 1, no. 3, pp. 175–190, 2022.

[137] A. Tontini, L. Gasparini, L. Pancheri, and R. Passerone, "Design and Characterization of a Low-Cost FPGA-Based TDC," *IEEE Transactions on Nuclear Science*, vol. 65, no. 2, pp. 680–690.

[138] F. Ceccarelli, G. Acconcia, A. Gulinatti, M. Ghioni, and I. Rech, "83ps Timing Jitter With a Red-Enhanced SPAD and a Fully Integrated Front End Circuit," *IEEE Photonics Technology Letters*, vol. 30, no. 19, pp. 1727–1730.

[139] E. Sall and M. Vesterbacka, "Thermometer-to-binary decoders for flash analog-to-digital converters," in *2007 18th European Conference on Circuit Theory and Design*, pp. 240–243.

[140] M. P. Ajanya and G. T. Varghese, "Thermometer code to Binary Code Converter for Flash ADCA Review," in *2018 International Conference on Control, Power, Communication and Computing Technologies (ICCPCCT)*, pp. 502–505.

[141] A. Chunn and R. K. Sarin, "Comparison of thermometer to binary encoders for flash ADCs," in *2013 Annual IEEE India Conference (INDICON)*, pp. 1–4.

[142] A. V. Kale, P. Palsodkar, and P. K. Dakhole, “Comparative Analysis of 6 Bit Thermometer-to-Binary Decoders for Flash Analog-to-Digital Converter,” in *2012 International Conference on Communication Systems and Network Technologies*, pp. 543–546.

[143] S. Kumar, M. K. Suman, and K. L. Baishnab, “A novel approach to thermometer-to-binary encoder of flash ADCs-bubble error correction circuit,” in *2014 2nd International Conference on Devices, Circuits and Systems (ICDCS)*, pp. 1–6.

[144] C. Ljuslin, J. Christiansen, A. Marchioro, and O. Klingsheim, “An integrated 16channel CMOS time to digital converter,” *IEEE Transactions on Nuclear Science*, vol. 41, no. 4, pp. 1104–1108.

[145] O. Bourrion and Laurent Gallin-Martel, “Development of a TDC to equip a liquid xenon PET prototype,” 2005.

[146] Intel Corporation, “Intel ® Quartus ® Prime Pro Edition User Guide Design Compilation,” Nov. 2021. Accessed: Jul. 05, 2023. [Online]. Available: https://cdrdv2.intel.com/v1/dl/getContent/666883?fileName=ug-qpp-compiler-683236-666883.pdf

[147] Intel Corporation, *Intel® Quartus® Prime Standard Edition User Guide: Design Compilation*. Intel Corporation, 2018. Accessed: Mar. 05, 2023. [Online]. Available: https://cdrdv2-public.intel.com/666615/ug-qps-compiler-683283-666615.pdf

[148] R. W. Nock, X. Ai, Y. Lu, Naim Dahnoun, and J. G. Rarity, “FPGA based time-to-digital converters,” in *Proc. SPIE*, Mar. 2020, p. 1134719.

[149] J. Wu, “Several Key Issues on Implementing Delay Line Based TDCs Using FPGAs,” *IEEE Transactions on Nuclear Science*, vol. 57, no. 3, pp. 1543–1548.

[150] S. Isbaner *et al.*, “Dead-time correction of fluorescence lifetime measurements and fluorescence lifetime imaging,” *Opt. Express*, vol. 24, Art. no. 9, 2016.

[151] Intel Corporation, “AN 796: Cyclone® V and Arria® V SoC Device Design Guidelines,” *Intel.com*, Mar. 30, 2022. https://cdrdv2.intel.com/v1/dl/getContent/666598?fileName=an-cv-av-soc-ddg-683360-666598.pdf (accessed Apr. 16, 2023).

[152] R. F. Molanes, J. J. RodríguezAndina, and J. Fariña, “Performance Characterization and Design Guidelines for Efficient Processor–FPGA Communication in Cyclone V FPSoCs,” *IEEE Transactions on Industrial Electronics*, vol. 65, no. 5, pp. 4368–4377.

[153] R. Novickis and M. Greitāns, “FPGA Master based on chip communications architecture for Cyclone V SoC running Linux,” in *2018 5th International Conference*

on Control, Decision and Information Technologies (CoDIT), pp. 403–408.

[154] R. F. Molanes, F. Salgado, J. Fariña, and J. J. RodríguezAndina, "Characterization of FPGA-master ARM communication delays in Cyclone V devices," in *IECON 201541st Annual Conference of the IEEE Industrial Electronics Society*, pp. 004229–004234.

[155] Intel Corporation, "Embedded Peripherals IP User Guide," *Intel.com*, Mar. 28, 2022. https://cdrdv2.intel.com/v1/dl/getContent/723705?fileName=ug_embedded_ip-683130-723705.pdf (accessed Jul. 16, 2023).

[156] M. Ruiz and A. Carpeno, *Using Quartus and Buildroot for building Embedded Linux Systems with DE1-SOC*. 2017.

[157] Intel Corporation, "Intel ® SoC FPGA Embedded Development Suite User Guide," Sep. 2018. Accessed: Mar. 16, 2023. [Online]. Available: https://cdrdv2.intel.com/v1/dl/getContent/705461?fileName=ug_soc_eds-18-1-683187-705461.pdf

[158] Intel Corporation, "FPGA-to-HPS Bridges Design Example," *Intel*. https://www.intel.com/content/www/us/en/support/programmable/support-resources/design-examples/horizontal/fpga-to-hps-bridges-design-example.html?_ga=2.217702806.1081597139.1559811462-417086916.1558618295&erpm_id=6634842 (accessed Jul. 16, 2022).

[159] Y. Chen, Müller, Joachim D, P. T. C. So, and E. Gratton, "The Photon Counting Histogram in Fluorescence Fluctuation Spectroscopy," *Biophysical Journal*, vol. 77, no. 1, pp. 553–567, 1999.

[160] M. Wahl, "The Principle of Time-Correlated Single Photon Counting," 2014. Accessed: Feb. 17, 2022. [Online]. Available: https://www.picoquant.com/images/uploads/page/files/7253/technote_tcspc.pdf

[161] M. Wahl, A. Benda, M. Hof, and J. Enderlein, "Fluorescence Lifetime Correlation Spectroscopy," *Journal of Fluorescence*, vol. 17, pp. 43–48, Feb. 2007.

[162] T. Lieske, W. Uhring, N. Dumas, J. Leonard, and D. Fey, "Embedded fluorescence lifetime determination for high throughput real-time droplet sorting with microfluidics," in *2017 Conference on Design and Architectures for Signal and Image Processing (DASIP)*, pp. 1–6.

[163] Y. Xu *et al.*, "A Bin-by-Bin Calibration with Neural Network for FPGA-Based Tapped-Delay-Line Time-to-Digital Converter," in *2022 IEEE International Conference on Real-time Computing and Robotics (RCAR)*, pp. 681–686.

[164] Z. Song *et al.*, "An 8.8 ps RMS resolution time-to-digital converter implemented in a

60 nm FPGA with real-time temperature correction," *Sensors*, vol. 20, Art. no. 8, 2020.

[165] K. Choi and D. Jee, "Design and Calibration Techniques for a Multichannel FPGABased Time-to-Digital Converter in an Object Positioning System," *IEEE Transactions on Instrumentation and Measurement*, vol. PP, pp. 1–1, Jul. 2020.

[166] G. Cao, H. Xia, and N. Dong, "An 18ps TDC using timing adjustment and bin realignment methods in a Cyclone-IV FPGA," *Review of Scientific Instruments*, vol. 89, p. 054707, May 2018.

[167] X. Mao, F. Yang, F. Wei, J. Shi, J. Cai, and H. Cai, "A low temperature coefficient time-to-digital converter with 1.3 ps resolution implemented in a 28 nm FPGA," *Sensors*, vol. 22, Art. no. 6, 2022.

[168] W. Zhang *et al.*, "TDC with uncontrolled delay lines: calibration approaches and precision improvement methods," *Journal of Instrumentation*, vol. 18, no. 01, p. C01011, 2023.

[169] S. Ito *et al.*, "Stochastic TDC architecture with self-calibration," in *2010 IEEE Asia Pacific Conference on Circuits and Systems*, pp. 1027–1030.

[170] C. W. Yao *et al.*, "A 14nm 0.14psrms Fractional-N Digital PLL With a 0.2-ps Resolution ADC-Assisted Coarse/Fine-Conversion Chopping TDC and TDC Nonlinearity Calibration," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 12, pp. 3446–3457.

[171] M. Hammer, D. Schweitzer, S. Richter, and E. Königsdörffer, "Sodium fluorescein as a retinal pH indicator?," *Physiological measurement*, vol. 26, pp. N9-12, Sep. 2005.

[172] M. Alhibah et al., "Penetration Depth of Propylene Glycol, Sodium Fluorescein and Nile Red into the Skin Using Non-Invasive Two-Photon Excited FLIM," *Pharmaceutics*, vol. 14, p. 1790, Aug. 2022.

[173] O. J. Miller *et al.*, "High-resolution dose–response screening using droplet-based microfluidics," *Proceedings of the National Academy of Sciences*, vol. 109, no. 2, pp. 378–383, Dec. 2011.

[174] J.-C. Baret *et al.*, "Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity," *Lab Chip*, vol. 9, no. 13, pp. 1850–1858, 2009.

[175] Bastiaens, Philippe I.H and A. Squire, "Fluorescence lifetime imaging microscopy: spatial resolution of biochemical processes in the cell," *Trends in Cell Biology*, vol. 9, no. 2, pp. 48–52, 1999.

[176] Gakamsky, Dmitry M, R. B. Dennis, and S. S. Desmond, "Use of fluorescence lifetime technology to provide efficient protection from false hits in screening applications,"

Analytical Biochemistry, vol. 409, no. 1, pp. 89–97, 2011.

[177] S. Pritz *et al.*, "A Fluorescence Lifetime-Based Assay for Abelson Kinase," *SLAS Discovery*, vol. 16, no. 1, pp. 65–72, 2011.

[178] R. L. Cornea *et al.*, "High-throughput FRET assay yields allosteric SERCA activators," *Journal of Biomolecular Screening*, vol. 18, Art. no. 1, 2013.

[179] K. Petersen, K. Peterson, J. Muretta, S. Higgins, G. Gillispie, and D. Thomas, "Fluorescence lifetime plate reader: Resolution and precision meet high-throughput," *The Review of scientific instruments*, vol. 85, p. 113101, Nov. 2014.

[180] D. Alibhai *et al.*, "Automated fluorescence lifetime imaging plate reader and its application to Förster resonant energy transfer readout of Gag protein aggregation," *Journal of biophotonics*, vol. 6, May 2013.

[181] P. Hansson *et al.*, "A comparative study of fluorescence assays in screening for BRD4," *ASSAY and Drug Development Technologies*, vol. 16, Art. no. 7, 2018

[182] S.-T. Hung, S. Mukherjee, and R. Jimenez, "Enrichment of rare events using a multi-parameter high throughput microfluidic droplet sorter," *Lab Chip*, vol. 20, no. 4, pp. 834–843, 2020.

[183] J. Léonard *et al.*, "High-throughput time-correlated single photon counting," *Lab Chip*, vol. 14, no. 22, pp. 4338–4343, 2014.

[184] A. I. Skilitsi *et al.*, "Towards sensitive, high-throughput, biomolecular assays based on fluorescence lifetime," *Methods and Applications in Fluorescence*, vol. 5, no. 3, p. 034002, 2017.

[185] S. Hasan, D. Geissler, K. Wink, A. Hagen, J. J. Heiland, and D. Belder, "Fluorescence lifetime-activated droplet sorting in microfluidic chip systems," *Lab Chip*, vol. 19, no. 3, pp. 403–409, 2019.

[186] J. Léonard *et al.*, "High-throughput time-correlated single photon counting," *Lab Chip*, vol. 14, no. 22, pp. 4338–4343, 2014.

[187] European Commission, "Water Framework Directive," *environment.ec.europa.eu*, 2023. https://environment.ec.europa.eu/topics/water/water-framework-directive_en (accessed Apr. 06, 2023).

[188] C. J. Gippel, "The Use of Turbidity Instruments to Measure Stream Water Suspended Sediment Concentration," Ph.D thesis, The University of New South Wales, Australia, 1988.

[189] I. Foster, R. Millington, and R. GREW, "The impact of particle size controls on stream turbidity measurement; some implications for suspended sediment yield estimation,"

Erosion and sediment monitoring programmes in river basins. Proc. international symposium, Oslo, 1992, vol. 210, Jan. 1992.

[190] J. Downing, "Twenty-five years with OBS sensors: The good, the bad, and the ugly," *Special Issue in Honor of Richard W. Sternberg's Contributions to Marine Sedimentology*, vol. 26, no. 17, pp. 2299–2318, 2006.

[191] N. Voichick, D. Topping, and R. Griffiths, "Technical Note: False low turbidity readings during high suspended-sediment concentrations," *Hydrology and Earth System Sciences Discussions*, pp. 1–11, Oct. 2017.

[192] A. Pallarès, P. Schmitt, and W. Uhring, "Comparison of Time Resolved Optical Turbidity Measurements for Water Monitoring to Standard Real-Time Techniques," *Sensors*, vol. 21, no. 9, 2021.

[193] L. Clermont, "Stray light control in space instruments: overcoming the conventional limits," Ph.D thesis, University of Liége, 2021.

[194] L. Clermont *et al.*, *Straylight calibration and correction for the MetOpSG 3MI mission*. 2018, p. 7.

[195] L. Clermont, C. Michel, and Y. Stockman, "Stray Light Correction Algorithm for High Performance Optical Instruments: The Case of Metop3MI," *Remote Sensing*, vol. 14, no. 6, 2022.

[196] L. Clermont, Wilfried Uhring, and M. P. Georges, "Time resolved characterization of stray light," in *Proc.SPIE*, Jun. 2021, p. 1178211.

[197] L. Clermont, W. Uhring, and M. Georges, "Stray Light Characterization With Ultrafast Time-of-flight Imaging," *Scientific Reports*, vol. 11, May 2021.

List of Publications

[1] W. Khaddour, W. Uhring, F. Dadouche, N. Dumas, and M. Madec, “Calibration Methods for Time-to-Digital Converters,” *Sensors*, vol. 23, no. 5, pp. 2791–2791, Mar. 2023.

[2] W. Khaddour, N. Dumas, F. Dadouche, M. Madec, and W. Uhring, “A low-cost high t hroughput microfluidics system for multi-thousand droplets per second sorting based on time resolved fluorescence measurement,” in *2022 Symposium on Design, Test, Integration and Packaging of MEMS/MOEMS (DTIP)*, pp. 1–4.

[3] W. Khaddour, W. Uhring, F. Dadouche, N. Dumas, and M. Madec, “High precision calibration method for asynchronous time-to-digital converters,” in *2022 20th IEEE Interregional NEWCAS Conference (NEWCAS)*, pp. 421–425.

*** Best Student Paper Award THIRD PLACE.**

[4] L. Clermont, W. Uhring, M. Georges, W. Khaddour, P. Blain, and E. Mazy, “Advances in stray light characterization by ultrafast time-of-flight imaging,” in *ICSO 2022 - International Conference on Space Optics*, Dubrovnik, Croatia, Oct. 2022, pp. 1–11.

[5] L. Clermont, W. Uhring, M. Georges, W. Khaddour, P. Blain, and E. Mazy, “A new paradigm in the field of stray light control and characterization enabled by ultrafast time-of-flight imaging,” in *Proc.SPIE*, Aug. 2022, p. 121881F.

[6] L. Clermont, Wilfried Uhring, Wassim Khaddour, P. Blain, E. Mazy, and M. P. Georges, “Ultrafast time-of-flight imaging with SPAD and picosecond laser for validation of the stray light rejection in an optical calibration facility,” in *Proc.SPIE*, May 2022, p. PC1213606.

[7] E. Aguénounon *et al.*, “Design and Characterization of an Asynchronous Fixed Priority Tree Arbiter for SPAD Array Readout,” *Sensors*, vol. 21, no. 12, p. 3949, Jun. 2021.

[8] W. Khaddour, F. Dadouche, W. Uhring, V. Frick, and M. Madec, “Design Methodology and Timing Considerations for implementing a TDC on a Cyclone V FPGA Target,” in *2020 18th IEEE International New Circuits and Systems Conference (NEWCAS)*, pp. 126–129.

[9] W. Khaddour, W. Uhring, F. Dadouche, V. Frick, and M. Madec, “Time-Resolved

fluorescence measurement system for real-time high-throughput microfluidic droplet sorting," in *2020 18th IEEE International New Circuits and Systems Conference (NEWCAS)*, pp. 246–249.

UNIVERSITE DE STRASBOURG

ECOLE DOCTORALE MATHEMATIQUES SCIENCES DE L'INFORMATION ET DE L'INGENIEUR

RESUME DE LA THESE DE DOCTORAT

Discipline : Électronique, microélectronique, optique et lasers, optoélectronique, micro-ondes, robotique

Spécialité (facultative) : Micro et Nanoélectronique

Présentée par : KHADDOUR Wassim

Titre : Système polyvalent de comptage de photons uniques corrélé en temps et Applications

Unité de Recherche : UMR 7357 - Laboratoire des sciences de l'Ingénieur, de l'Informatique et de l'Imagerie (ICube)

Directeur de Thèse : MADEC Morgan - MCF HDR

Co-Directeur de Thèse : UHRING Wilfried - Professeur

Localisation : 23, rue du Loess 67037 Strasbourg Cedex 02

Thèse confidentielle : ☑ NON ☐ OUI

1. Introduction

Le comptage de photons uniques corrélés en temps (Time-Correlated Single Photon Counting en anglais, TCSPC) est une technique puissante pour enregistrer des signaux optiques périodiques avec une haute sensibilité, résolution et précision. Elle a un large éventail d'applications en physique, biologie, ingénierie et technologies quantiques telles que la mesure de durée de vie de fluorescence résolue dans le temps, les applications de détection et télémétrie par ondes lumineuses (Light detection and ranging en anglais, LIDAR) et les mesures quantiques. Le principe général du TCSPC est illustré sur la Figure 1. Il repose sur la détection de photons uniques d'un signal optique périodique, la mesure du temps d'arrivée relatif des photons détectés dans la période du signal et la création d'un histogramme à partir des temps d'arrivée. En accumulant un grand nombre de comptages de photons, l'histogramme construit représente la forme d'onde du signal périodique.

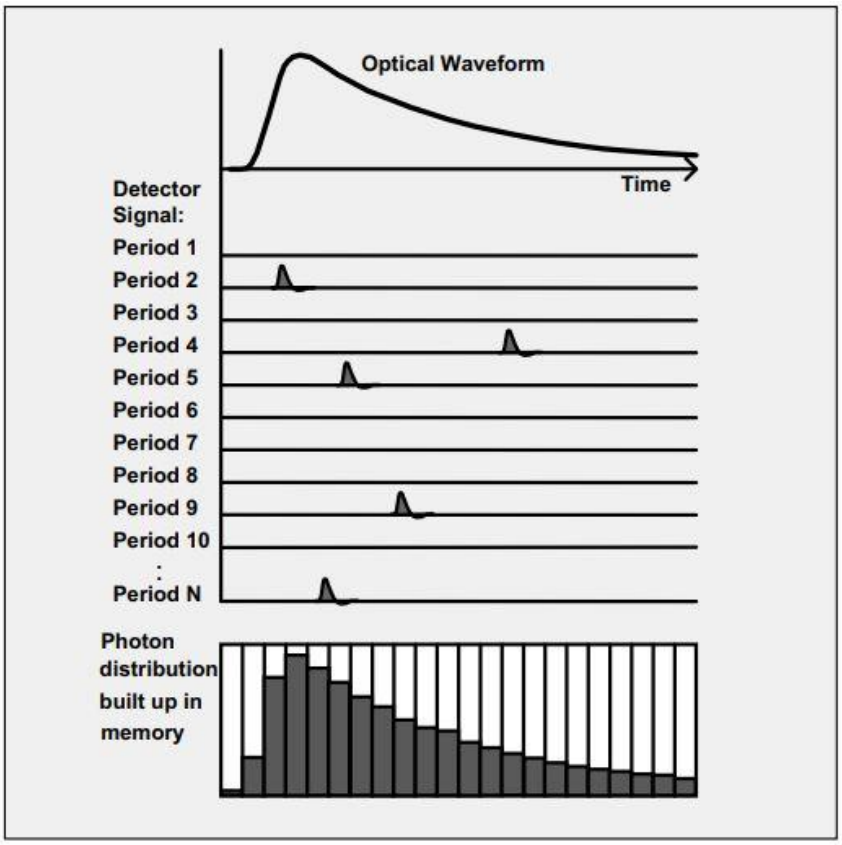


Figure 1. Principe général du comptage de photons uniques corrélés en temps [14].

Un système TCSPC traditionnel comprend généralement un détecteur de photons qui génère une impulsion pour chaque photon détecté, une photodiode qui génère un signal de référence

synchronisé avec la source lumineuse, et un canal de mesure du temps qui mesure le temps d'arrivée des photons détectés comme suit : un convertisseur temps-amplitude (Time-to-Amplitude converter en anglais, TAC) reçoit les signaux photon et de référence et génère un signal analogique dont l'amplitude est proportionnelle à l'intervalle de temps entre ces impulsions. La sortie du TAC est ensuite amplifiée et envoyée à un convertisseur analogique-numérique (CAN) pour être convertie en une valeur numérique représentant le temps d'arrivée du photon. La valeur numérique est utilisée pour adresser une unité mémoire qui stocke le nombre de photons détectés dans chaque intervalle de temps. Ainsi, un histogramme de la distribution des photons en fonction du temps est construit dans la mémoire. Cet histogramme est ensuite envoyé à un PC pour analyse. Les systèmes TCSPC traditionnels ont été largement utilisés dans les recherches scientifiques et industrielles. Cependant, ces systèmes présentent plusieurs inconvénients tels qu'un coût élevé, une taille encombrante et des performances faibles. Pour surmonter ces limitations, des systèmes TCSPC intégrés ont été développés dans des circuits intégrés spécifiques à une application (application specific integrated circuits en anglais, ASIC). Ces systèmes intègrent les différents composants du système TCSPC ainsi que l'électronique de traitement des données sur une seule puce de silicium, permettant la conception de systèmes TCSPC compacts et performants. Cependant, les systèmes ASIC présentent également de nombreux inconvénients tels qu'un temps de développement long, un coût élevé pour les petites séries et un manque de flexibilité. D'autre part, les progrès récents de la technologie des réseaux logiques programmables (field programmable gate array en anglais, FPGA) ont permis la mise en œuvre de systèmes TCSPC performants avec de nombreuses caractéristiques telles qu'un faible coût, un cycle de développement court, une flexibilité et une reconfigurabilité. De plus, le système sur puce FPGA (SoC-FPGA) intègre un système processeur dur (hard processor system en anglais, HPS) puissant avec le tissu FPGA sur la même puce, offrant un excellent outil pour la réalisation d'applications de coprocesseur matériel/logiciel performantes. Cette thèse présente la conception et la mise en œuvre d'un système TCSPC polyvalent sur une plateforme SoC-FPGA à faible coût, à savoir le Cyclone V DE10-Nano d'Altera, qui peut être utilisé pour différentes applications. La thèse propose deux conceptions de convertisseur temps-numérique (Time-to-Digital Converter en anglais, TDC), l'une pour les sources lumineuses synchrones et l'autre pour les sources asynchrones. Elle propose également une conception efficace pour le transfert de données à haute vitesse entre la partie FPGA et le HPS basée sur l'accès direct à la mémoire, et un programme C bare-metal qui s'exécute sur le HPS et effectue tout le traitement des données TCSPC à bord. La Figure 2 montre un schéma du co-design matériel/logiciel proposé du système TCSPC. Le TDC sera

implémentée sur la logique FPGA, tandis que le traitement des données sera exécuté sur le système processeur dur de la plateforme SoC-FPGA. Le système intégrera également une carte mère PCB personnalisée qui fournit une alimentation stable au SoC-FPGA, permet les communications USB avec le PC, et intègre un générateur d'impulsions à diode laser, entre autres composants.

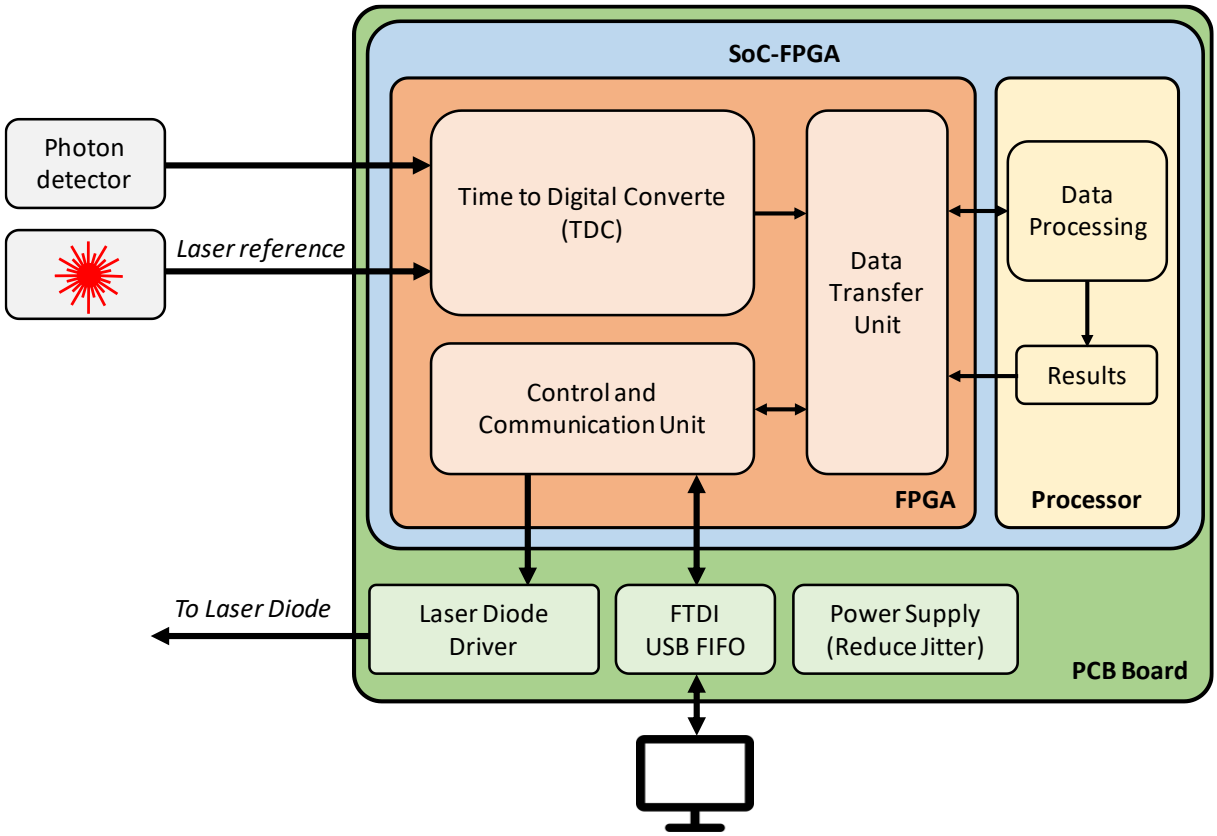


Figure 2. Schéma du co-design matériel/logiciel proposé du système TCSPC.

2. Convertisseur temps-numérique

Le convertisseur temps-numérique (TDC) est le cœur du système TCSPC. Son rôle est de mesurer le temps d'arrivée des photons détectés avec une haute résolution en mesurant l'intervalle de temps entre deux signaux : le signal START qui représente l'émission de la lumière d'excitation et le signal STOP qui est l'impulsion du détecteur de photons. Différentes structures de TDC ont été proposées dans la littérature. Nous avons adopté la structure à ligne à retard (tapped-delay line en anglais, TDL), qui offre la meilleure résolution pour les TDC basés sur FPGA. Dans cette structure, le signal START se propage à travers des éléments de retard (DE) cascadés, généralement des tampons, provoquant le passage du niveau logique de la sortie des éléments propagés d'un niveau à un autre. Les sorties des DEs sont échantillonnées

par des bascules à l'arrivée du signal STOP, produisant un code thermomètre qui est ensuite codé en un code binaire représentant l'intervalle de temps entre les deux signaux. Pour améliorer la résolution, nous avons construit le TDL en utilisant les chemins de chaîne de report disponibles dans Cyclone V. Dans des conceptions complexes et de haute précision comme les TDC, la topologie du routage a un impact critique sur le comportement et les performances du système en raison du réseau dense de connexion et de configuration du FPGA. Cependant, les dispositifs FPGA à faible coût comme Cyclone V ne supportent pas le routage et le placement manuels. Par conséquent, nous avons rencontré de nombreux problèmes et défis lors de la mise en œuvre du TDC, qui ont nécessité d'imposer de nombreuses contraintes pour contrôler le routage et le placement des différentes parties du système, pour gérer les délais de propagation qui assurent le bon fonctionnement du système, et pour améliorer sa performance.

2.1 TDC synchrone

Dans la conception du TDC, nous avons combiné la structure TDL avec la méthode de Nutt pour étendre la plage de mesure (MR) sans augmenter la consommation de ressources. La Figure 3 représente l'architecture globale de notre TDC, qui comprend un bloc contrôleur de signal, deux blocs de mesure (un bloc fin et un bloc grossier), et un bloc d'écriture de données. Le bloc grossier compte le nombre de périodes d'horloge entre les signaux START et STOP, tandis que le bloc fin évalue l'intervalle entre le signal STOP et le front montant d'horloge suivant. Ainsi, le temps d'arrivée peut être calculé par l'Equation 1, comme le montre la Figure 4.

$$TI = T_{coarse} - T_{fine} \quad (01)$$

Le bloc fin est un TDL de 256 DE composé d'additionneurs et de DFFs. Pour éviter le placement des DFFs dans d'autres modules logiques adaptatifs (ALM) que ceux de leurs additionneurs correspondants, nous avons créé une région de verrouillage logique pour le bloc fin TDL et ajusté sa largeur à 1. Nous avons également proposé une technique d'échantillonnage réduit basée sur les caractéristiques du temps de propagation le long de la chaîne de report de Cyclone V, qui a amélioré la linéarité du TDC avec une résolution temporelle satisfaisante pour les applications cibles. Pour coder le code thermomètre résultant à la sortie du TDL, nous avons implémenté un encodeur anti-bubbles basé sur l'architecture du compteur d'unités.

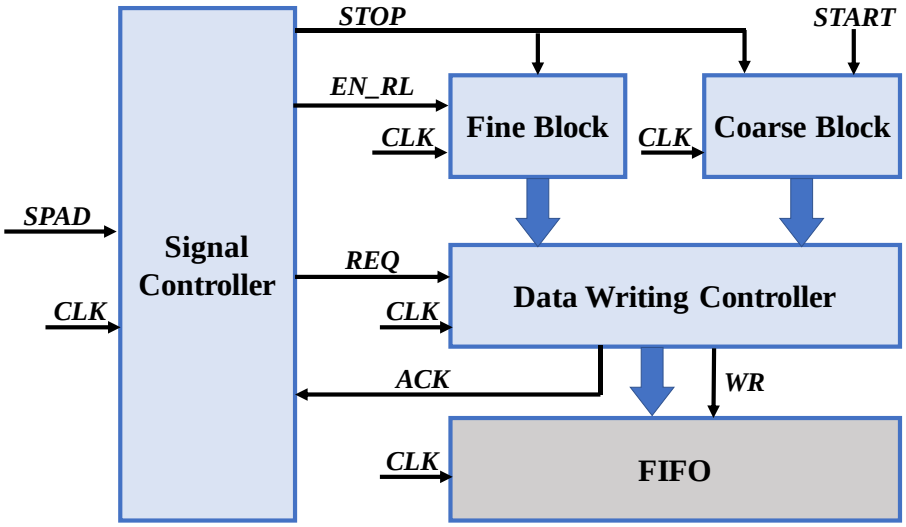


Figure 3. Architecture globale du TDC synchrone.

De plus, nous avons proposé une structure grossière robuste qui évite les problèmes de métastabilité résultant de l'échantillonnage asynchrone de la valeur du compteur. Cette structure emploie deux compteurs fonctionnant en phase l'un avec l'autre et une conception logique pour sélectionner la valeur stable en fonction de la valeur fine.

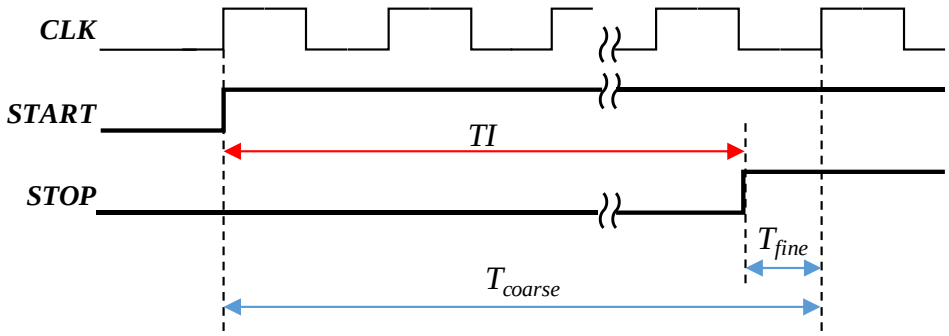


Figure 4. Diagramme de synchronisation du TDC synchrone. TI est décomposé en deux composantes T_{coarse} et T_{fine}.

2.2 TDC asynchrone

Nous avons étendu la conception du TDC synchrone pour supporter des sources de lumière d'excitation asynchrones, en ajoutant un deuxième canal TDL et un deuxième bloc grossier pour la mesure du temps d'arrivée du signal START. Les systèmes TDC synchrones et asynchrones ont été implémentés avec succès sur le SoC-FPGA Cyclone V, atteignant une résolution temporelle moyenne d'environ 20 ps et un taux de détection maximal d'environ

66,7 M photon/s pour le système synchrone. En revanche, le taux de détection du système asynchrone est proportionnel à la fréquence de la lumière d'excitation.

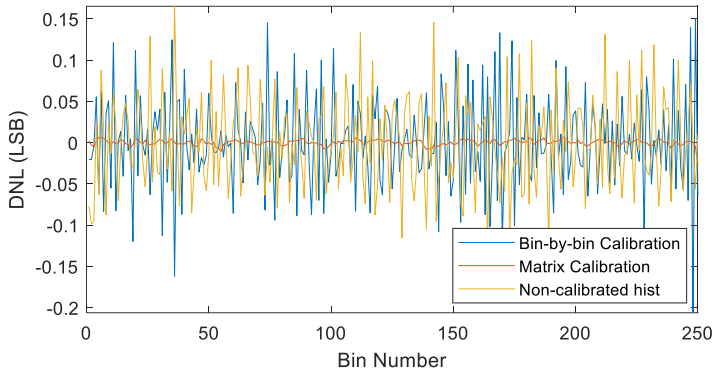
3. Source de lumière d'excitation et détecteur de photons

Pour assurer le faible coût du système, nous avons utilisé une diode laser couplée à un générateur d'impulsions à haut taux de répétition développé par notre équipe, comme source de lumière d'excitation, au lieu de lasers pulsés coûteux. Cette source génère des impulsions laser courtes avec une largeur à mi-hauteur (full width at half maximum en anglais, FWHM) d'environ 100 ps. Pour la détection des photons, nous avons utilisé une diode avalanche à photon unique (single-photon avalanche diode en anglais, SPAD) commerciale, qui peut être ensuite remplacée par un réseau de SPAD conçu par notre équipe pour réduire encore plus le coût.

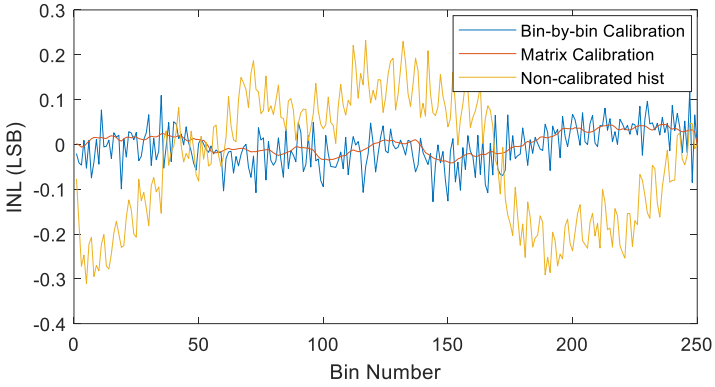
4. Traitement des données

Contrairement aux systèmes TCSPC conventionnels dans lesquels les données sont envoyées à un PC de contrôle pour le traitement, notre système effectue tout le traitement des données localement à bord en utilisant le HPS de la carte SoC-FPGA. Pour transférer les données de la partie FPGA vers le HPS à haut débit, nous avons mis en œuvre une conception matérielle-logicielle qui utilise l'interface FPGA-vers-SDRAM. De plus, nous avons employé un moteur d'accès direct à la mémoire (direct memory access en anglais, DMA) pour décharger le processeur des opérations de mouvement des données et préserver ses ressources pour le traitement des données. Nous avons également proposé une approche de mise en mémoire tampon des données qui permet le transfert basé sur le DMA des données du TDC à un taux constant, permettant un traitement en ligne des données. De plus, nous avons étudié la configuration optimale de la taille des données DMA, de la taille du tampon de données et de la taille du FIFO TDC pour minimiser la latence du système et éviter le débordement du FIFO. Cette approche permet au système de fonctionner en temps réel. Pour le traitement des données, nous avons développé un programme C bare metal qui s'exécute sur le HPS et effectue toutes les tâches de traitement des données, telles que l'étalonnage du TDC, la suppression du bruit de fond et d'autres corrections, ainsi que le traitement final des données selon l'application (par exemple, mesure de l'intensité ou de la durée de vie de la fluorescence). Pour l'étalonnage du TDC synchrone, nous avons adopté la méthode d'étalonnage par largeur moyenne de bin. De plus, nous avons introduit une nouvelle méthode d'étalonnage "Calibration matricielle" pour les TDC asynchrones qui a permis d'obtenir des améliorations significatives par rapport à la

méthode classique bin par bin, tant au niveau du DNL que de l'INL, avec un DNL plus de 10 fois inférieur et un INL environ 2 fois plus faible, comme le montre la Figure 5. De plus, nous avons développé et optimisé un algorithme d'estimation du maximum de vraisemblance (maximum likelihood estimation en anglais, MLE). Cet algorithme intègre une procédure itérative de réduction du bruit appelée auto-fenêtrage, qui améliore la précision et la vitesse d'estimation. L'auto-fenêtrage consiste à tronquer itérativement l'histogramme du signal pour éliminer le bruit causé par les bacs tardifs à fort poids et à recalculer la durée de vie jusqu'à convergence vers un résultat stable.



(a)



(b)

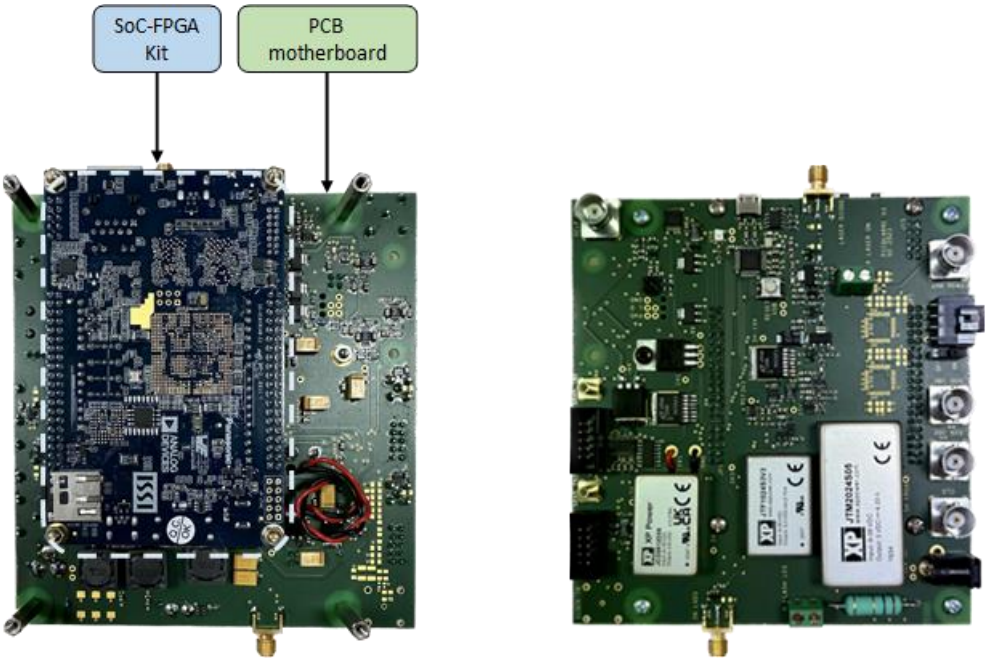
Figure 5. Valeurs de DNL et de INL du TDC asynchrone avant et après appliquer la calibration bin-par-bin et la calibration matricielle : (a) valeurs de DNL, (b) valeurs de INL.

5. Caractérisation du système

La Figure 6 présente une photographie de notre dispositif TCSPC, tandis que la Figure 7 montre les vues de dessus et de dessous de l'électronique du système, comprenant la plateforme Cyclone V SoC-FPGA et la carte mère PCB conçue sur mesure, qui fournit l'alimentation au

SoC-FPGA, permet les communications USB avec le PC, et intègre un générateur d'impulsions de diode laser, entre autres composants.

Figure 6. Vue extérieure du dispositif TCSPC réalisé.



(a) (b)

Figure 7. Électroniques du système TCSPC réalisé incluant la plateforme Cyclone V SoC-FPGA et la carte mère PCB personnalisée : (a) vue de dessous, (b) vue de dessus.

Nous avons évalué les performances des TDC en termes de résolution, de précision et d'exactitude, et étudié l'effet de la température sur leurs performances. Nous avons également proposé des solutions pour atténuer les erreurs causées par les variations de température. De plus, nous avons évalué les performances du système TCSPC en mesurant sa fonction de

réponse instrumentale (instrument response function en anglais, IRF) et en enregistrant le signal de fluorescence d'un fluorophore de référence pour estimer sa durée de vie de fluorescence. La résolution du TDC, évaluée par la largeur moyenne des bacs, variait de 22 à 23,5 ps lorsque la température variait de 5 à 65 °C, comme le montre la Figure 8.

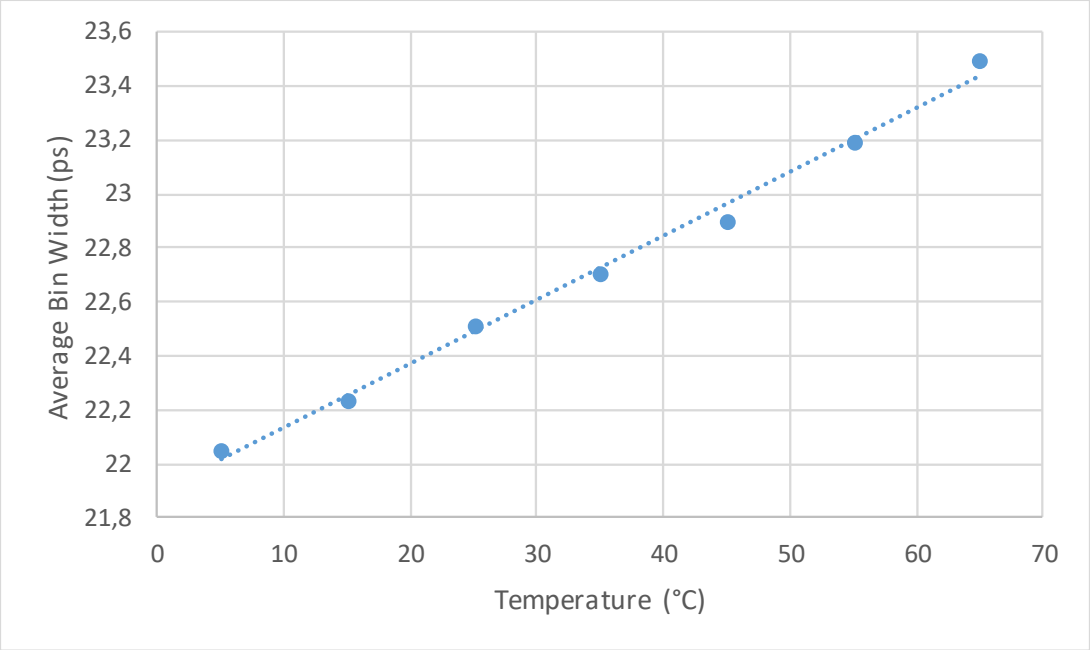


Figure 8. Largeur moyenne des bins du TDL en fonction de la température. La largeur des bins augmente linéairement avec la température.

Nous avons évalué la précision et l'exactitude du TDC synchrone en effectuant des tests de délai de câble en utilisant des câbles de différentes longueurs pour couvrir la plage de mesure du TDL de 5,26 ns. Le TDC synchrone a atteint une précision RMS variant de 23 ps à 27 ps sur sa plage de mesure à différentes températures, tandis que l'exactitude présentait de grandes variations allant de -400 ps à +900 ps lorsque la température variait de 5 à 65 °C. Ces variations peuvent être attribuées à deux types d'erreurs : une erreur d'offset causée par le retard dépendant de la température du chemin de routage depuis la broche d'entrée jusqu'au point d'entrée du TDL, et une erreur de gain (également appelée erreur d'étalonnage) due à la variation de la réponse du TDC avec la température. Les Figures 9 et 10 montrent les variations de précision et d'exactitude RMS avec la température le long de la plage de mesure du TDL.

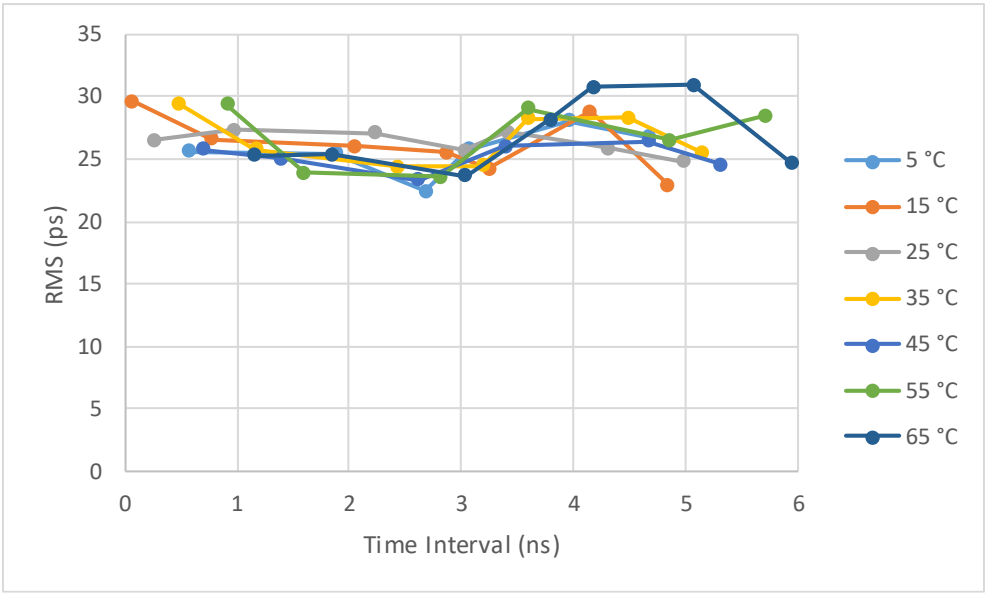


Figure 9. Précision RMS globale du TDC synchrone à différentes températures ambiantes.

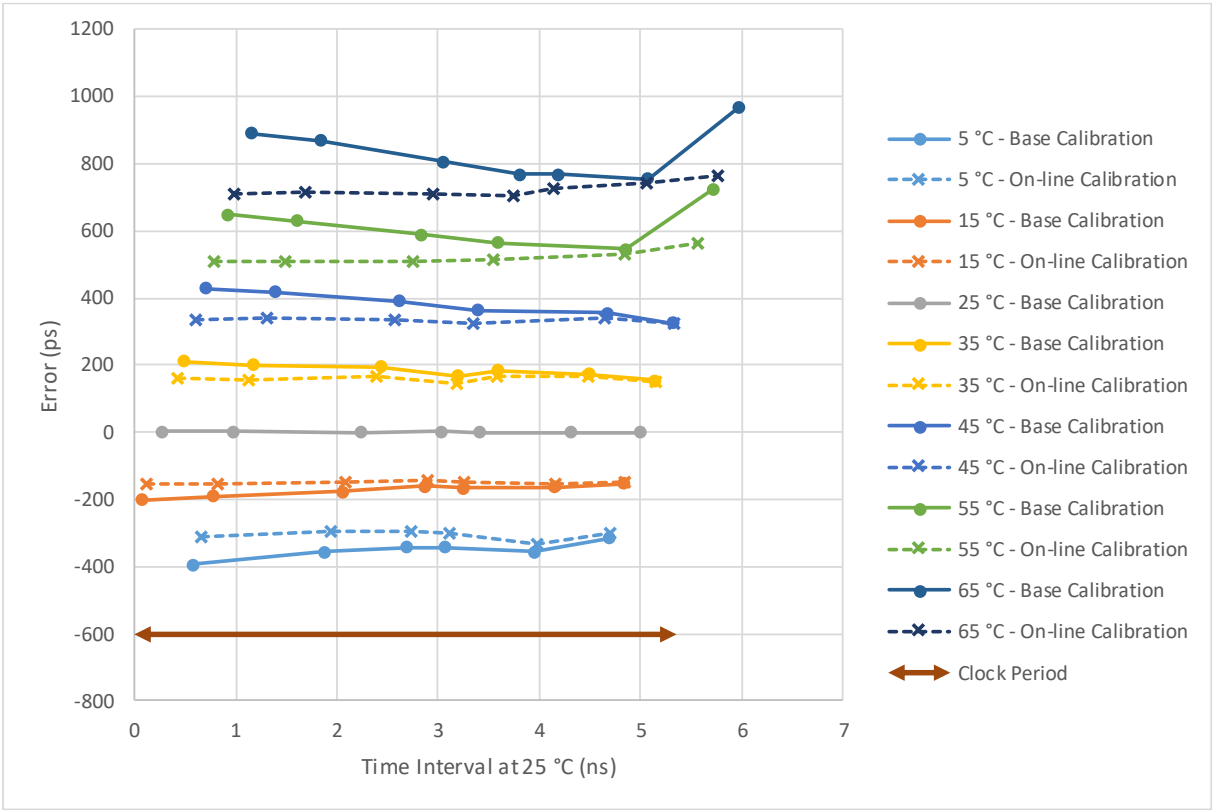


Figure 10. Déviations des mesures obtenus par le TDC synchrone en fonction de la température. Les courbes pleines montrent les résultats avec la calibration de base, tandis que les courbes pointillées montrent les résultats avec la calibration en ligne.

L'erreur d'offset présente une relation linéaire avec la température et peut être efficacement compensée par un coefficient de correction dépendant de la température en utilisant un capteur de température qui surveille la température ambiante. D'autre part, l'erreur de gain présente également une relation linéaire avec la température et peut également être corrigée par un simple coefficient de correction ou en effectuant un étalonnage en ligne pour compenser les variations de la réponse du TDC. Les Figures 11 et 12 représentent les valeurs d'offset et de gain en fonction de la température.

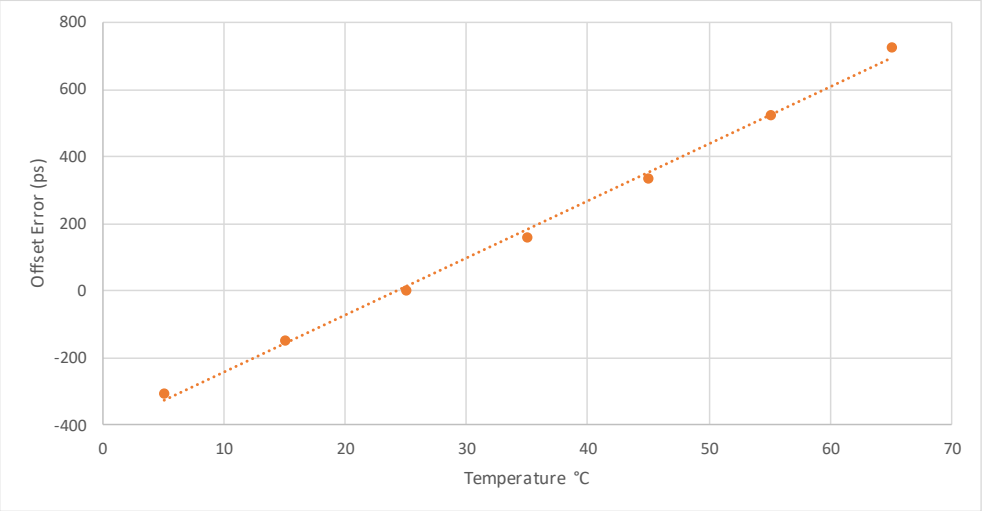


Figure 11. Erreur d'offset en fonction de la température dans le TDC synchrone.

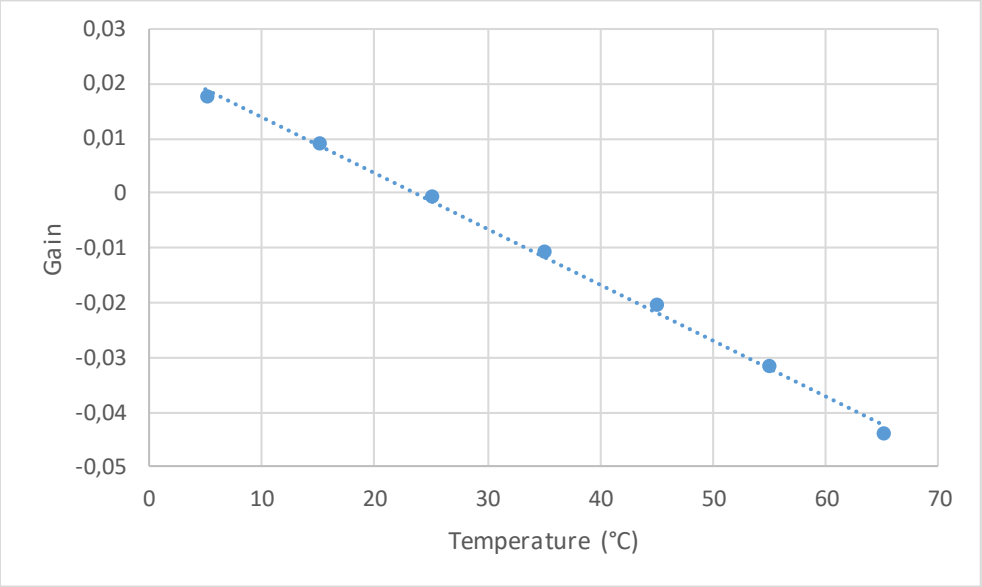
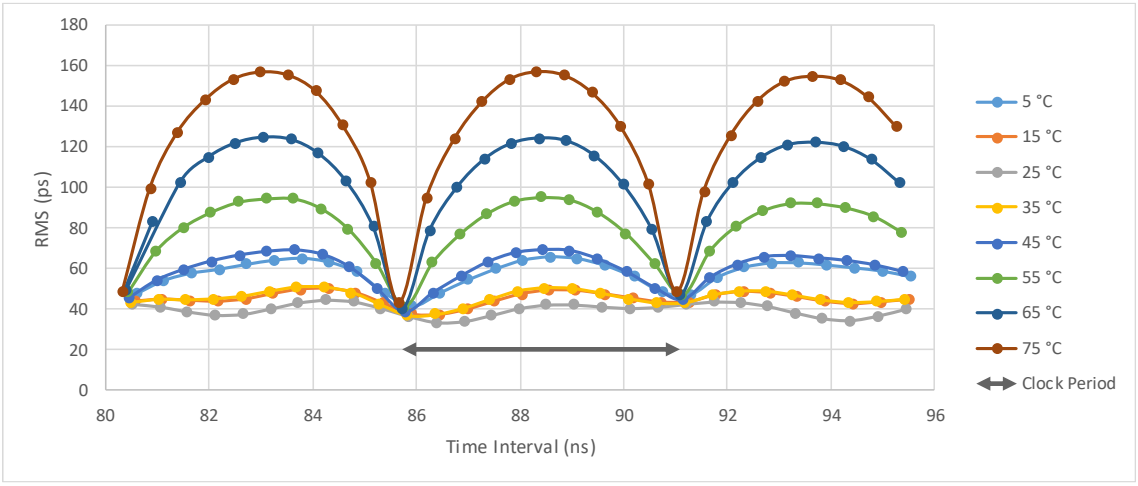
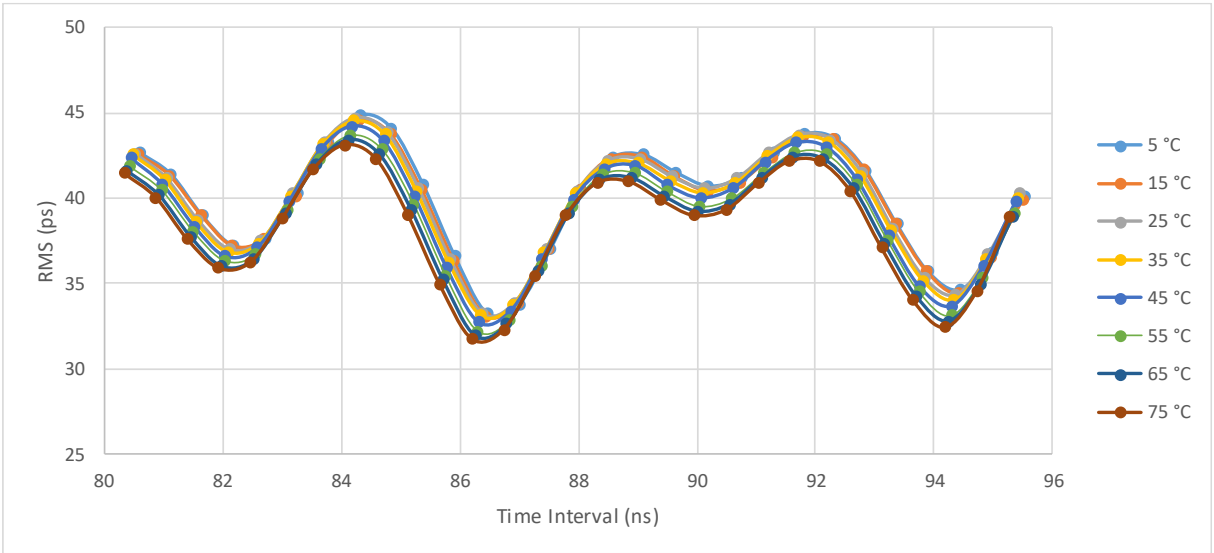


Figure 12. Valeurs de gain en fonction de la température dans le TDC synchrone.

Pour évaluer la précision et l'exactitude du TDC asynchrone, nous avons utilisé un générateur de fonctions arbitraires pour produire différents intervalles de temps prédéfinis entre deux signaux qui étaient appliqués aux entrées START et STOP. Lorsque la température variait de 5 à 65 °C, la précision RMS du TDC asynchrone variait de 37 ps à 160 ps avec un étalonnage hors ligne, et de 31 ps à 45 ps avec un étalonnage en ligne, comme le montre la Figure 13. La déviation de l'exactitude variait quant à elle de -50 ps à +150 ps dans la même plage de température, comme le montre la Figure 14.



(a)



(b)

Figure 13. Précision RMS globale du TDC asynchrone à différentes températures ambiantes : (a) avec la calibration de base, (b) avec la calibration en ligne.

Le TDC asynchrone a présenté une meilleure stabilité et robustesse que le TDC synchrone grâce à son architecture à double TDL qui annule la plupart des erreurs dépendantes de la température. Ces résultats démontrent l'avantage d'actualiser les tables d'étalonnage sur les performances du TDC. Par conséquent, nous avons proposé un étalonnage en vol du TDC asynchrone, qui génère les tables d'étalonnage en temps réel à partir des données de mesure.

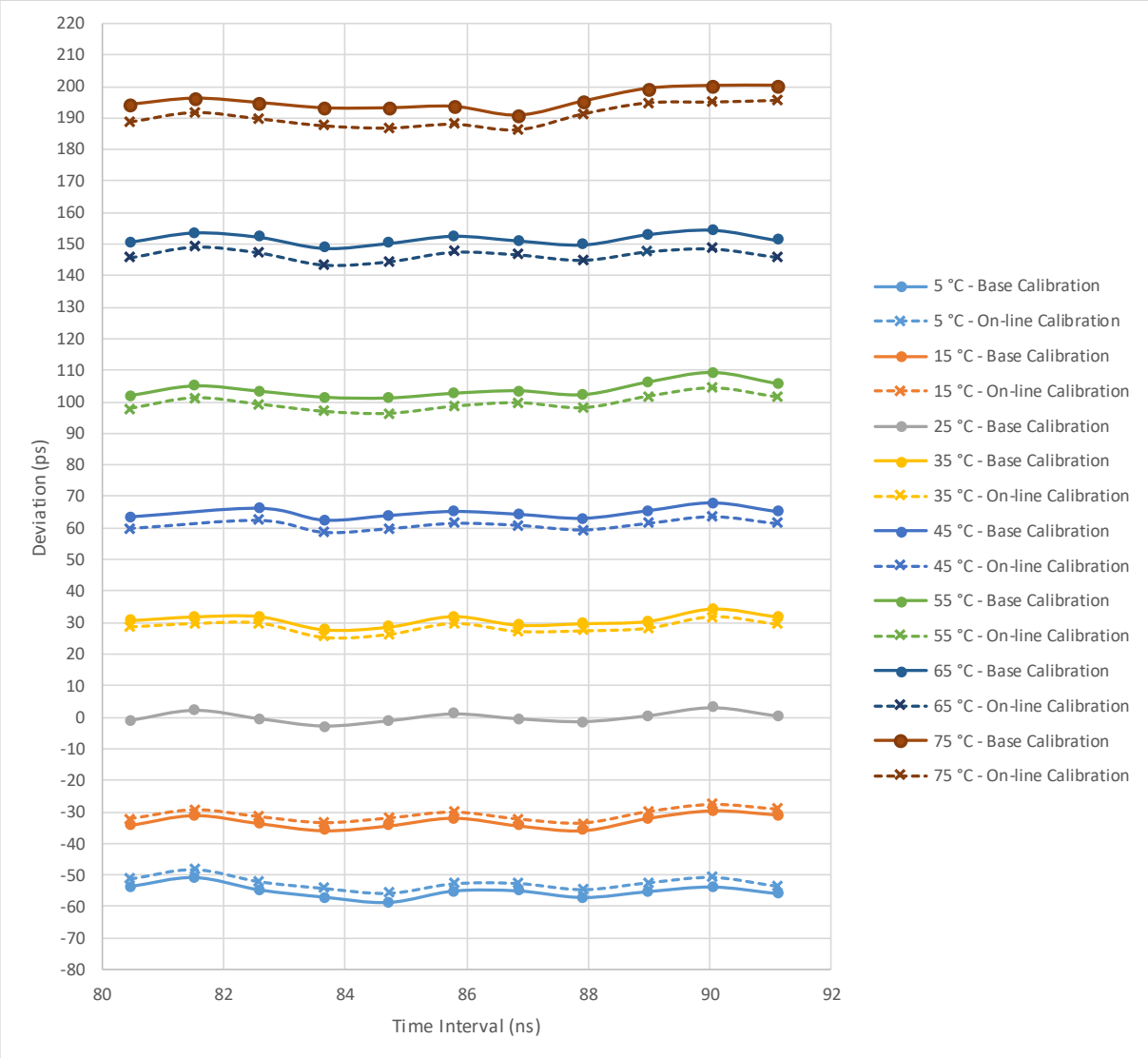


Figure 14. Déviations des mesures obtenus par le TDC asynchrone en fonction de la température. Les courbes pleines montrent les résultats avec la calibration de base, tandis que les courbes pointillées montrent les résultats avec la calibration en ligne.

Pour évaluer les performances du système TCSPC, nous avons mesuré sa fonction de réponse instrumentale (IRF) en enregistrant la forme de l'impulsion de lumière d'excitation. Les résultats ont montré que l'impulsion générée par une diode laser à 405 nm avait une largeur à

mi-hauteur (FWHM) d'environ 130 ps et une largeur à 1 % du maximum de 630 ps, comme le montre la Figure 15.

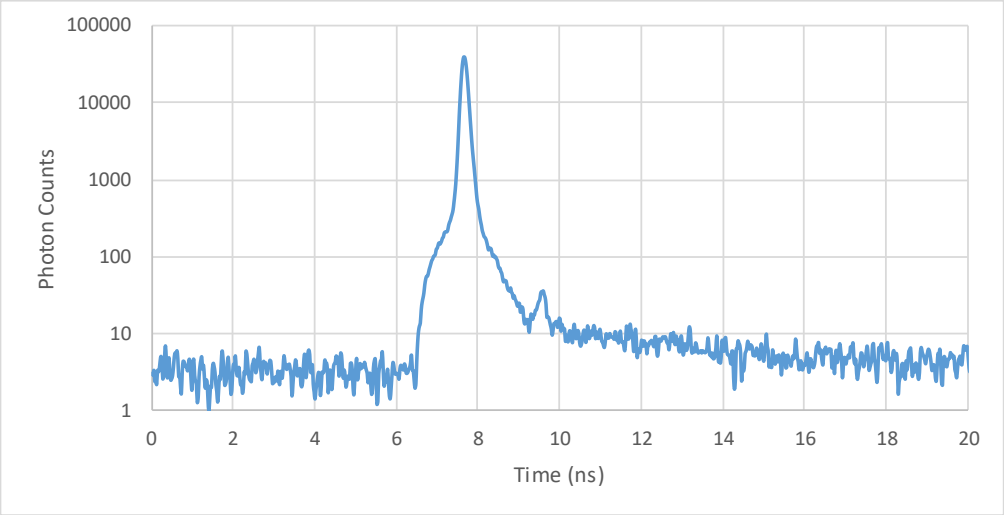


Figure 15. IRF du système TCSPC synchrone, en utilisant une diode laser pulsée à 405 nm. La FWHM est d'environ 130 ps.

Nous avons étalonné le système TCSPC en enregistrant le signal de fluorescence de la fluorescéine dans le PBS à une concentration de 1 mM, qui est un fluorophore standard avec une bonne photostabilité et une décroissance monoexponentielle. Nous avons mesuré la durée de vie de fluorescence de ce fluorophore. La FLT obtenue était de 4,11 ns, ce qui est conforme à la valeur de la littérature. La Figure 16 montre le signal enregistré, qui démontre une décroissance monoexponentielle nette.

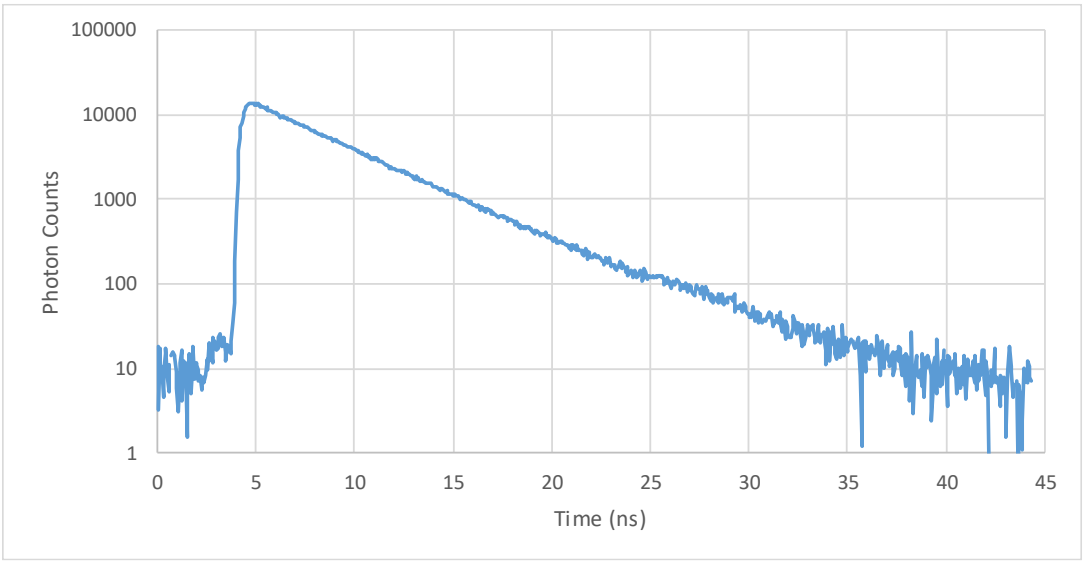


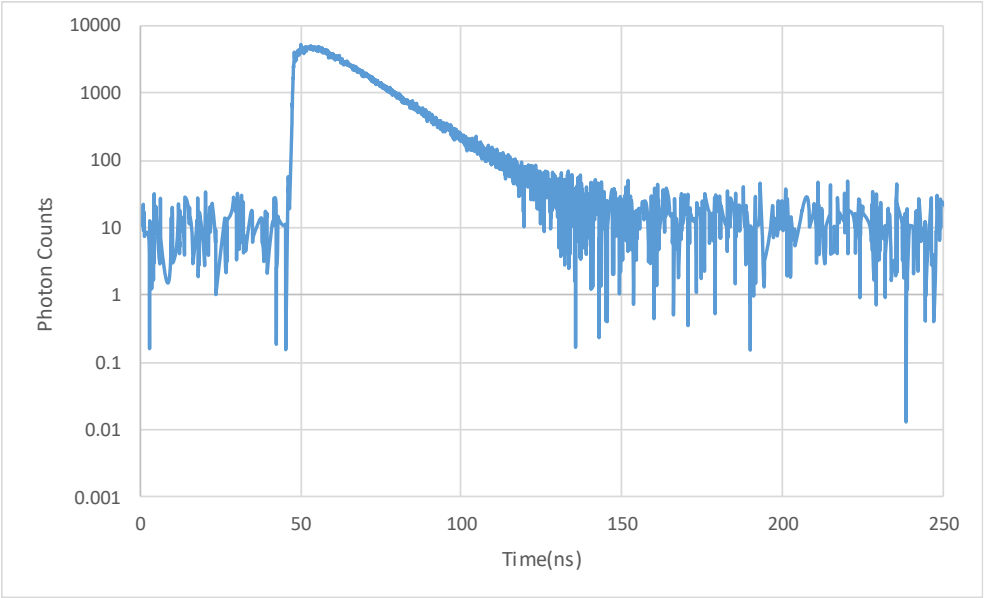
Figure 16. Signal de fluorescence de la fluorescéine dans le PBS mesuré par le système TCSPC synchrone. Le FLT estimé est de 4,11 ns.

6. Applications

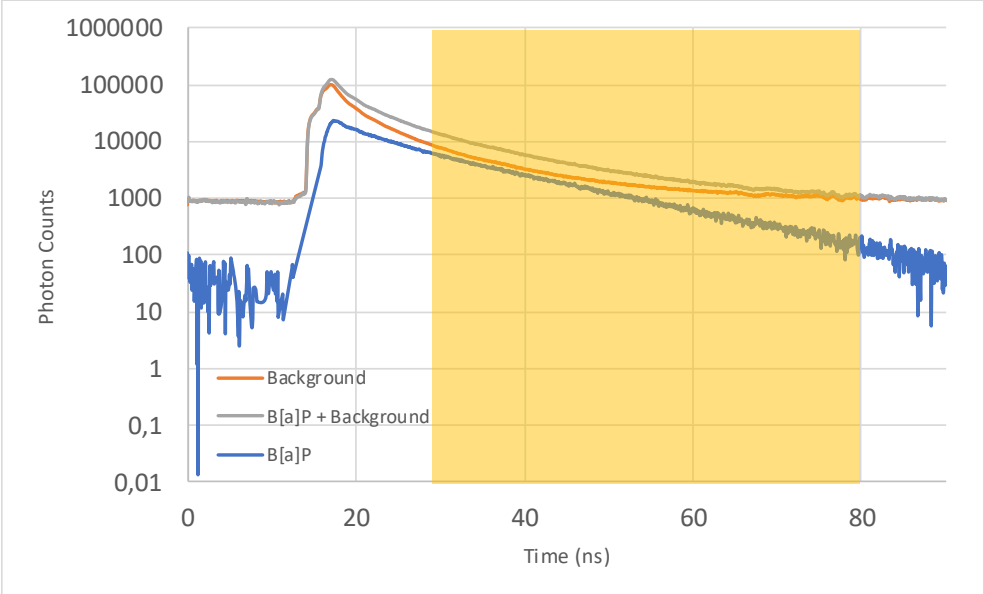
Pour démontrer la polyvalence du système proposé, nous l'avons employé dans quatre applications différentes dans les domaines environnemental, biologique et astronomique, qui sont : un capteur de pollution de l'eau (water pollution sensor en anglais, WPS), turbidité optique résolue en temps (time-resolved optical turbidity en anglais, TROT), criblage et tri de gouttelettes microfluidiques, et caractérisation de la lumière parasite. La diversité des applications cibles a permis la vérification et le débogage complets du système. De plus, les optimisations et les exigences spécifiques de chaque application individuelle nous ont permis d'incorporer de nombreuses améliorations et modifications pour rendre le système plus polyvalent.

6.1 Capteur de pollution de l'eau (WPS)

Le projet Interreg WPS vise à développer un dispositif autonome pour la mesure continue de la concentration de micropolluants chimiques dans l'eau potable, en utilisant trois méthodes de détection. L'une de ces méthodes est l'analyse optique basée sur la mesure de la fluorescence. Le dispositif doit permettre une alerte rapide en cas de contamination excessive par certains pesticides, ce qui n'est pas possible avec les approches de détection actuelles qui nécessitent un échantillonnage manuel et une analyse en laboratoire. La méthode d'analyse optique cible les hydrocarbures aromatiques polycycliques (HAP) (polycyclic aromatic hydrocarbons en anglais, PAHs), tels que le benzo[a]pyrène. Le système mesure le signal de fluorescence de l'échantillon d'eau et détermine la présence et la concentration du HAP. La valeur de la durée de vie de la fluorescence permet d'identifier le HAP présent, tandis que l'intensité de la fluorescence permet d'estimer la concentration. Nous avons utilisé notre système TCSPC dans ce projet pour la méthode d'analyse optique. Nous avons couplé le système TCSPC avec un dispositif optique pour délivrer et focaliser la lumière d'excitation sur l'échantillon d'eau et pour collecter et focaliser le signal de fluorescence sur le détecteur de photons, et un circuit microfluidique comprenant une chambre microfluidique et une vanne qui assure la manipulation sécurisée des polluants dangereux. Nous avons effectué une série d'expériences pour mesurer le signal de fluorescence du B[a]P dans l'éthanol pur à deux concentrations différentes de 750 µM et 40 µM. Les signaux enregistrés étaient monoexponentiel avec une durée de vie d'environ 14 ns. La figure 17 montre les signaux enregistrés après avoir éliminé le signal de fond pour les deux concentrations.



(a)



(b)

Figure 17. Signal de fluorescence du B[a]P dans l'éthanol pur après la suppression du signal de fond : (a) à une concentration de 750 µM, (b) à une concentration de 40 µM.

6.2 Criblage et tri de gouttelettes microfluidiques

Le criblage à haut débit (high-throughput screening en anglais, HTS) de biomolécules est une technique puissante qui permet d'identifier des molécules biologiquement actives. L'interaction de molécules marquées avec d'autres molécules entraîne une diminution de leur rendement quantique de fluorescence (fluorescence quantum yield en anglais, FQY), qui peut être utilisé

comme indicateur d'interactions biomoléculaires. Cependant, le FQY est généralement mesuré par l'intensité, qui est influencée par de nombreux facteurs, tels que la concentration et l'intensité de la lumière d'excitation. La durée de vie de la fluorescence (FLT) est une propriété intrinsèque du FQY plus précise qui est indépendante de ces facteurs. Par conséquent, la microscopie à fluorescence par imagerie de durée de vie (FLIM) a été largement utilisée pour l'analyse de la biologie cellulaire. De même, la détection de la FLT devrait améliorer la fiabilité des tests d'interaction biomoléculaire. Elle a été mise en œuvre dans des lecteurs de plaques pour le HTS. Cependant, les lecteurs de plaques commerciaux actuels prennent environ 0,5 seconde pour lire un micropuits en mode FLT. D'autre part, les circuits microfluidiques couplés à la détection de fluorescence sont d'excellents outils pour le HTS et le tri cellulaire activé par fluorescence (FACS) avec un débit de quelques milliers de gouttelettes par seconde. De plus, les puces microfluidiques permettent de réduire la quantité de réactifs et de trier les échantillons ciblés en fonction des résultats de mesure de la fluorescence. Ainsi, la microfluidique peut potentiellement accélérer la découverte de médicaments d'environ 100 fois. Néanmoins, combiner les tests basés sur la FLT avec la microfluidique est un défi, surtout lorsqu'une action de tri est nécessaire, car elle nécessite une acquisition et un traitement des données en temps réel. Grâce à la fonctionnalité de traitement des données en ligne, notre système TCSPC peut effectuer un tri de gouttelettes basé sur la FLT en temps réel à des débits élevés et à faible coût en même temps. Pour cette application, nous avons développé une méthode robuste pour la détection des gouttelettes microfluidiques en temps réel basée sur un seuil prédéfini d'intensité instantanée détectée, nous avons également développé un algorithme à haute vitesse pour extraire les FLTs des gouttelettes. Nous avons intégré notre système TCSPC avec un dispositif optique et une puce microfluidique qui génère des gouttelettes à haut débit. Nous avons effectué des expériences pour évaluer le débit atteignable en utilisant des gouttelettes de fluorescéine en PBS à une concentration de 1 mM. Les résultats ont montré que le système pouvait détecter et cribler les gouttelettes en fonction de leurs FLTs à un débit supérieur à 3500 gouttelettes/seconde. Le nombre moyen de photons par gouttelette n'était que de 43 photons. Cependant, le système pouvait détecter toutes les gouttelettes et estimer leur FLT avec une valeur moyenne de 4,02 ns, qui correspond à la valeur trouvée dans la littérature. La Figure 18 montre la distribution des valeurs de FLT de 10000 gouttelettes. L'écart type des FLTs estimées était de 0,72 ns, tandis que le bruit photonique théorique contribuait à un écart type de 0,6 ns.

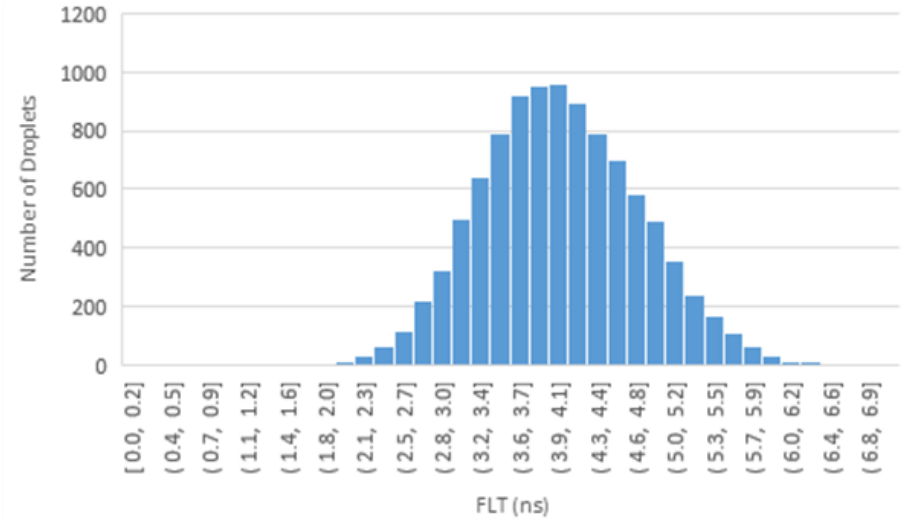


Figure 18. Distribution des valeurs de FLT mesurés pour 10000 gouttelettes de fluorescéine dans le PBS à un débit de 3573 gouttelettes/seconde.

Nous avons effectué un autre ensemble d'expériences à un taux de gouttelettes de 1000 gouttelettes par seconde. Le nombre moyen de photons détectés par gouttelette était d'environ 718. Le système a détecté les gouttelettes en temps réel et estimé leurs FLTs avec une valeur moyenne de 3,92 ns. Cette valeur est légèrement inférieure à la valeur théorique en raison de l'effet de pileup, car le rapport entre le taux de photons détectés et le taux de répétition du laser était relativement élevé (beaucoup plus que 1%). Le taux d'impulsions laser était d'environ 8,5 MHz et le nombre moyen de photons détectés par gouttelette était d'environ 718 photons, ce qui correspond à environ 16% du taux de laser. La Figure 19 montre la distribution des valeurs de FLT de 2000 gouttelettes détectées. L'écart type des valeurs de FLTs était de 0,18 ns, tandis que le bruit photonique contribuait à un écart type de 0,15 ns.

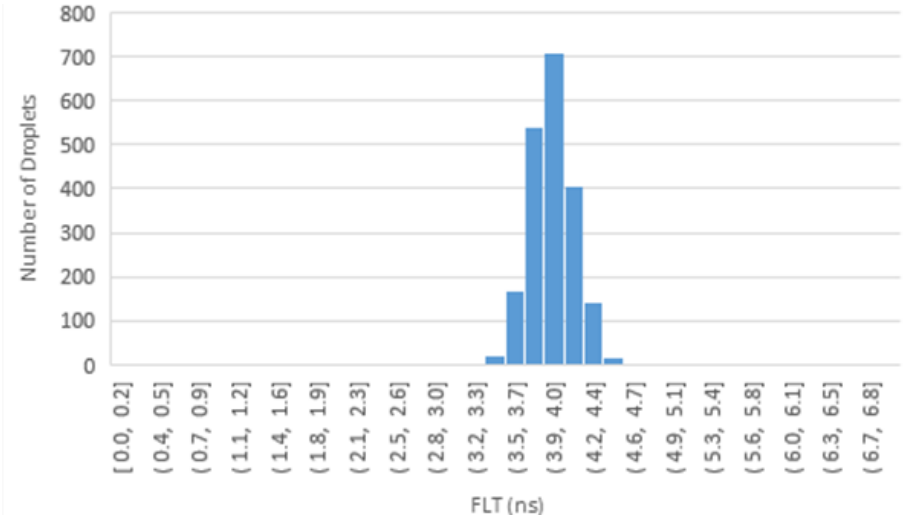


Figure 19. Distribution des valeurs de FLT mesurés pour 10000 gouttelettes de fluorescéine dans le PBS à un débit de 1000 gouttelettes/seconde.

6.3 Turbidité optique résolue en temps (TROT)

Les méthodes conventionnelles d'évaluation de la turbidité reposent sur des analyses périodiques en laboratoire d'échantillons qui sont coûteuses et longues. D'autre part, la turbidité optique et la rétrodiffusion acoustique, qui sont les techniques les plus utilisées pour la mesure de la turbidité en temps réel, présentent plusieurs inconvénients dus aux limitations de saturation et à leur dépendance vis-à-vis de la forme et de la taille des particules en suspension. La turbidité optique résolue en temps (TROT) est une nouvelle technique prometteuse pour la mesure de la turbidité basée sur le TCSPC. Dans cette technique, des impulsions laser courtes et périodiques sont émises dans le milieu trouble à travers une fibre optique. Une deuxième fibre optique, placée à une courte distance de la première, collecte les photons qui sont rétrodiffusés par les particules en suspension dans le milieu. Ces photons sont ensuite détectés par un détecteur de photons et le système TCSPC mesure leurs temps d'arrivée. Le temps de vol de ces photons dépend des chemins optiques qu'ils suivent dans le milieu. La distribution temporelle des photons atteignant le détecteur est liée au niveau de turbidité du milieu.

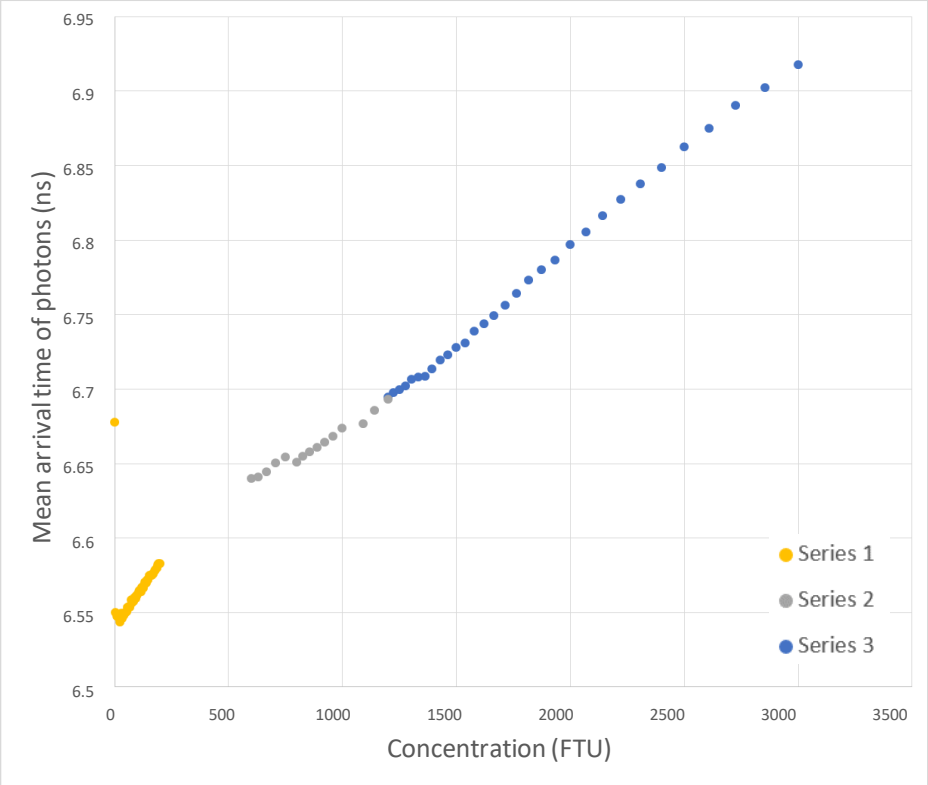


Figure 20. Temps moyen d'arrivée des photons détectés en fonction de la concentration.

Notre système a été utilisé par l'équipe de recherche "Mécanique des Fluides" de notre laboratoire pour étudier la technique TROT. Les expériences ont montré que plusieurs

caractéristiques, telles que le nombre de photons détectés et le temps d'arrivée moyen de ces photons, ont une relation directe avec le niveau de turbidité.

Nous avons réalisé une série d'expériences pour étudier la technique TROT. Nous avons préparé différents niveaux de turbidité. Nous avons enregistré l'histogramme de la fonction de diffusion ponctuelle temporelle (temporal point spread function en anglais, TPSF) pendant 10 secondes à chaque niveau de turbidité. Nous avons analysé les caractéristiques des TPSFs enregistrés et examiné plusieurs paramètres qui variaient avec la turbidité. Le paramètre le plus pertinent était le temps d'arrivée moyen des photons qui présente une tendance proche de la linéarité avec le niveau de turbidité sur une large gamme, comme le montre la Figure 20.

Nous avons également étudié l'influence de l'encrassement biologique sur les performances du turbidimètre optique résolu en temps, nous avons réalisé une série d'expériences en utilisant des atténuateurs optiques avec différentes densités optiques pour simuler différents niveaux d'encrassement biologique. Les résultats ont montré que le nombre de photons, qui est le paramètre utilisé par les turbidimètres optiques conventionnels, diminue significativement lorsque la densité optique de l'atténuateur augmente, tandis que le temps d'arrivée moyen n'est pas affecté par l'atténuation introduite par les différents atténuateurs, comme le montre la Figure 21. Ce qui démontre que le TROT est beaucoup plus robuste que les turbidimètres optiques conventionnels face à la réduction d'intensité optique causée par l'encrassement optique.

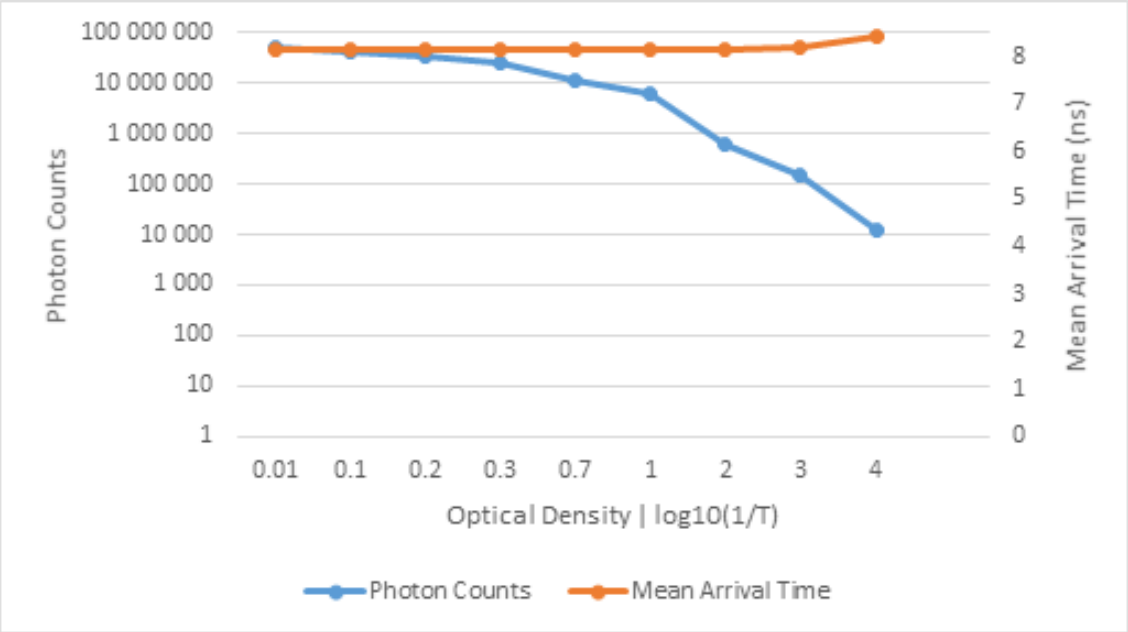


Figure 21. Comparaison entre le nombre de photons et le temps moyen d'arrivée des photons des signaux enregistrés à différents niveaux d'atténuation.

Dans cette application, nous avons identifié un problème de stabilité du laser, où l'intensité du laser variait avec le temps probablement en raison des variations de température. Nous avons résolu ce problème en mettant en œuvre une boucle de régulation à rétroaction qui maintient la stabilité du laser tout au long de la mesure, ce qui a amélioré les performances du système.

6.4 Caractérisation de la lumière parasite

Notre système TCSPC a été utilisé en collaboration avec le Centre Spatial de Liège (CSL) en Belgique dans une nouvelle méthode de caractérisation de la lumière parasite dans les instruments optiques. La lumière parasite est une lumière indésirable qui atteint le photodétecteur provenant de diverses sources, telles que les réflexions, la diffusion ou les tirs directs. Elle dégrade la qualité d'image et les performances des instruments optiques, surtout pour les applications spatiales. Par conséquent, la lumière parasite doit être caractérisée expérimentalement avant de lancer l'instrument dans l'espace, afin de valider ses performances et de corriger les éventuels problèmes. La méthode conventionnelle pour caractériser la lumière parasite est basée sur la mesure de la transmittance ponctuelle spatiale (spatial point source transmittance en français, SPST) de l'instrument. Cette méthode présente de nombreuses limitations, telles que la difficulté à interpréter et à identifier les sources de lumière parasite, et l'interférence de la lumière parasite provenant du dispositif de mesure lui-même. Pour surmonter ces limitations, le CSL a proposé et démontré une méthode de caractérisation basée sur la technique TCSPC. Cette méthode consiste à utiliser un faisceau laser pulsé et un capteur ultra-rapide pour mesurer le signal au plan focal. Cela permet de séparer le faisceau formant l'image et les composantes de la lumière parasite, qui suivent des chemins optiques différents, en fonction des différences dans leurs temps d'arrivée. Cette approche simplifie l'interprétation, identifie l'origine des composantes de la lumière parasite et discrimine la lumière parasite de l'instrument de celle du dispositif de mesure.

Notre système TCSPC a été utilisé par le CSL dans la validation de cette méthode et dans l'amélioration du dispositif de mesure de la lumière parasite pour l'instrument d'observation terrestre FLEX. Pour la discrimination spatiale des composantes de la lumière parasite, le SPAD a été monté sur un mécanisme de balayage angulaire qui a permis d'effectuer un balayage angulaire 2D. La Figure 22 montre la discrimination et la distribution angulaire des différentes composantes de la lumière parasite qui prennent des longueurs de chemin optique différentes.

Dans ces expériences, des mesures continues longues de trois semaines ont été effectuées sans aucun problème, ce qui a prouvé la fiabilité de notre système TCSPC.

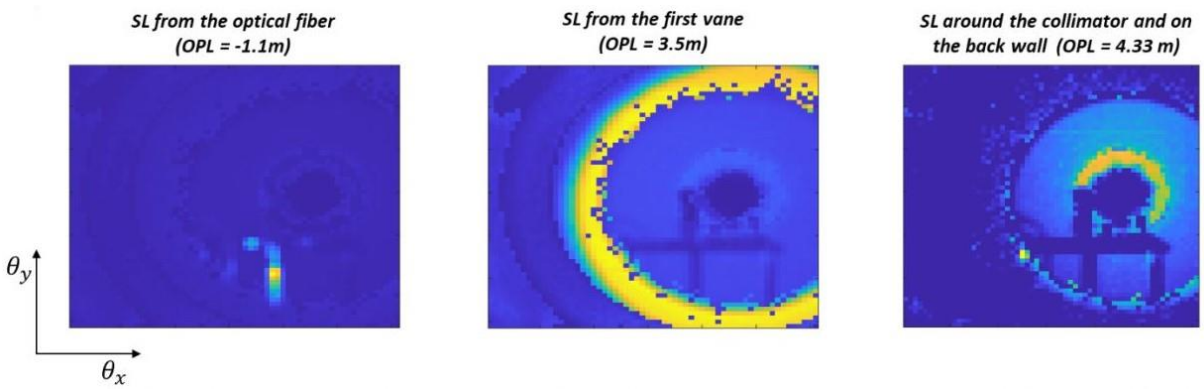


Figure 22. Illustration du motif de lumière parasite arrivant au détecteur SPAD à différents moments, montrant différents contributeurs de lumière parasite.

7. Conclusion

Cette thèse présente un système TCSPC à faible coût, flexible et performant pour diverses applications qui nécessitent une mesure du temps de haute précision. Cette thèse introduit également une nouvelle méthode de calibration pour les TDC asynchrones qui améliore la linéarité d'un facteur 10. Elle introduit aussi une méthode robuste pour la détection des gouttelettes microfluidiques qui permet un criblage à bas coût des gouttelettes basé sur la FLT à un débit de plus de 3500 gouttelettes par seconde en utilisant la technique TCSPC.

Le système réalisé a démontré une grande polyvalence et un potentiel pour diverses applications dans différents domaines scientifiques et industriels. Il a été utilisé dans quatre applications distinctes dans les domaines environnemental, biologique et astronomique, et il peut être facilement adapté à d'autres applications avec des besoins spécifiques en matière de traitement de données, en intégrant des algorithmes de traitement de données appropriés dans le programme C. Les contributions de cette thèse font avancer l'état de l'art des systèmes TCSPC basés sur FPGA et du criblage à haut débit de gouttelettes.