UNIVERSITÀ DEGLI STUDI DI TRIESTE

XXXV Ciclo del Dottorato di Ricerca in Ingegneria Industriale e dell'Informazione

Real-time Features Extraction for Trigger-less Data Acquisition Systems in Particle Physics Experiments

Settore scientifico-disciplinare: ING-INF/01 Electronics

Dottorando:	Bruno VALINOTI
Coordinatore:	Alberto TESSAROLO
Supervisore di tesi UNITS:	Sergio CARRATO
Supervisora di tesi ICTP:	María Liz CRESPO

Anno accademico 2021/2022

Abstract

The COmmon Muon and Proton Apparatus for Structure and Spectroscopy (COM-PASS) spectrometer at CERN is upgrading part of its detectors and data acquisition system to handle larger event sizes, owing to a higher channel quantity and higher event rates. Starting in 2023, the experiment will be formally called AMBER, and its spectrometer will contain approximately 100 million data channels. Owing to the progress in micro-electronics and SoC-FPGA technology, it has become possible to read and process data coming from detectors without the classical data reduction of a Level-1 trigger schema. For AMBER data acquisition, migration to a Level-1 trigger free architecture, known as trigger-less DAQ, was proposed. This architecture upgrade allows for the generation of more statistics for further analysis.

The trigger-less operation requires a new type of frontend electronics capable of working in a free-running mode for detectors readout. Due to the intensive processing requirements and the need for high-speed serial data transmission, some of the devices that are currently installed for the readout of the detectors are not capable of working in a trigger-less DAQ, so they need to be replaced.

The main purpose of this thesis is to propose, design and develop a new trigger-less readout system for the 3068-channel ECAL2 detector of the COMPASS/AMBER experiment. For this purpose, a hardware platform for front-end electronics was developed using the current digitizer board (MSADC, FPGA-based) and as core component a high-performance MPSoC Ultrascale+ SoM device for processing the data coming from 16

channels at 960 Mbit/s each. The modular hardware design allows its use in different front-end readouts and SoM families. The firmware for FPGAs and processors was developed and a digital pulse processor (DPP) for real-time data features extraction in a trigger-less system was implemented. The DPP is the core component for pulse processing, including methods for extracting the pulse amplitude value and pulse shape discrimination. These methods are based on a mathematical model of typical pulses acquired from a detector prototype. Besides features extraction, the DPP is also prepared to send traces containing whole pulses for further studies. With this new frontend electronics, the ECAL2 can work in a free-running and trigger-less mode and with the addition of a DPP the amount of transmitted data is drastically reduced, saving further processing time and storage resources.

Acknowledgements

From the first time I arrived at Trieste, I felt it was home, maybe because of the warm people I met or perhaps because it was what I was looking for. I would like to thank Dr. Maria Liz Crespo and Dr. Sergio Carrato for their support and scientific guidance. In particular, Maria Liz is present in daily detail, not only from a professional sphere but also from a human perspective. Thanks to Stefano Levorato from INFN-TS for all the support, for pushing the project and believing in me, and for always being a step ahead with the things I need for working. Special thanks go to my friends and colleagues from the ICTP-Mlab for sharing with me the most precious and valuable piece we have, the time. Thanks to Andres Cicuttin for sharing his knowledge, experience, friendship, and support. I would like to thank Andrea for the daily good mood, and for being the best flatmate I ever had, and to Werner for the same things, but two years before. I am eternally grateful to Manu, Fede, Uriel, Nico, Alessandro and Matilde for making my days and nights after working the best moments for music and aperitivos. Thanks to Liliana Fraigi for encouraging me to start with a Ph.D. and to my colleagues from INTI for their support during my stay in Trieste. Thanks to Fiorella for being there, for the love and support of every day, and for helping me in this endeavour from the very beginning. I would then like to thank my family, especially my mother Alejandra, father Mario, and sisters and brothers, Marina, Martin, Julian, Guido, and Tiziana. With you, everything is easier. Finally, I thank the ICTP Multidisciplinary Laboratory, the University of Trieste, and INFN-TS for giving me the opportunity to complete the Ph.D. program.

Contents

List of Figures

List of Tables

List of Acronyms

μC Microcontroller.

μP Microprocessor.

ACF Autocorrelation Function.

ADC Analog to Digital Converter.

AMBER Apparatus for Meson and Baryon Experimental Research.

AXI Advanced eXtensible Interface.

BOS Beginning of Spill.

CDC Clock Domain Crossing.

CIAA-ACC Computadora Industrial Abierta Argentina - Alta Capacidad de Computo.

CLI Command Line Interface.

COMBLOCK Communication Block.

COMPASS Common Muon and Proton Apparatus for Structure and Spectroscopy.

CSA Charge Sensitive Amplifier.

DAC Digital to Analog Converter.

DAQ Data Acquisition System.

DCM Digital Clock Manager.

DDR Double Data Rate.

DMA Direct Memory Access.

DPLMS Digital Penalized Least Mean Squares.

DSP Digital Signal Processing.

ECAL2 Electromagnetic Calorimeter two.

EDA Electronic Design Automation.

EOI Event of Interest.

EOS EOS Open Storage.

FFeCCa FPGA-SoC Frontend Carrier Card.

FIFO First In First Out.

FIR Finite Impulse Response.

FMC FPGA Mezzanine Connector.

FPGA Field Programmable Gate Array.

FWHM Full Width at Half Maximum.

GD Geometrically Derived.

GPIO General Purpose Input Output.

HEP High Energy Physics.

HPC High Pin Count.

HV High Voltage.

HVPSCS High Voltage Power Supply Control System.

HVPSS High Voltage Power Supply System.

I2C Inter-Integrated Circuit.

IBERT IP Integrated Bit Error Ratio Tester.

IC Integrated Circuit.

IDDR Input DDR.

IIR Infinite Impulse Response.

IP Intellectual Property.

JTAG Joint Test Action Group.

L1 Level one.

LAN Local Area Network.

LAS Large angle scatterings.

LED Light Emitting Diode.

LHC Large Hadron Collider.

LPC Low Pin Count.

LUT Lookup Table.

LVCMOS Low Voltage Complementary Metal Oxide Semiconductor.

LVDS Low Voltage Differential Signaling.

LVTTL Low Voltage Transistor Transistor Logic.

lwIP Light Weight IP.

MGT Multi-Gigabit Transceiver.

MMCM Mixed-Mode Clock Manager.

MUX Multiplexer.

ODDR Output DDR.

OFSODA Open Framework for SoC-FPGA DAQ based platforms.

OPC Open Platform Communications.

PCB Printed Circuit Board.

PCI Pearson Coefficient Index.

PDN Power Distribution Network.

PL Programmable Logic.

PLD Programmable Logic Device.

PMT Photomultiplier Tube.

PS Processing System.

PSA Pulse Shape Amplifier.

SAS Small Angle Scatterings.

SATA Serial Advanced Technology Attachment.

SDD Silicon Drift Detector.

SDRAM Synchronous Dynamic Random-Access Memory.

SERDES Serializer-Deserializer.

SFP+ Small Form-factor Pluggable.

SNR Signal to Noise Ratio.

SoC System on Chip.

SoM System on Module.

SPCI Simplified Pearson Coefficient Index.

SPI Serial Peripheral Interface.

TCS Timing Control System.

TPC Time Projection Chamber.

TTL Transistor Transistor Logic.

UART Universal Asynchronous Receiver-Transmitter.

UDMA Universal Direct Memory Access.

USB Universal Serial Bus.

Chapter 1

Introduction

High-energy physics (HEP) aims at understanding the elementary particle constituents of matter and their interactions. The Standard Model explains the constituents and their interactions. Nevertheless, there are many questions that the Model leaves unanswered, or partially answered. The most direct way to verify the particles predicted by the Standard Model and to understand those that are outside the description is through high-energy physics experiments. These experiments rely on the collision of high-energy particles to study their interactions. Research on HEP and its evolution depends in a symbiotic manner on its ability to extract information from these interactions. Several detectors and techniques have been developed and evolved since the first HEP experiments, from stereoscopic photographs in bubble chambers to state-of-the-art neutrino detectors. Recent advances in electronics, photonics, communications, signal processing, and, in particular, field programmable gate arrays (FPGA) and integration with microprocessors and high-speed communication transceivers in System on Chip (SoC) have allowed the development of edge instrumentation. The new readout solutions using these technological advances offer the possibility of extracting more information from the detector's measurement at higher rates, increasing the quality and amount of statistics for the measurements. This increase in data rate is accompanied by

more complex electronic systems for readout and data curation.

This thesis is framed in the electronics upgrade and real-time data features extraction techniques for data acquisition systems in HEP and for multichannel detectors. More precisely, the COMPASS experiment at the European Organization for Nuclear Research (CERN) is going through an update and upgrade on part of the detectors and data acquisition system. HEP experiments are characterized by a large number of detectors with a large number of channels ($> 10^4$) that produce an extremely large amount of data per unit of time (for example, TB/sec). All these data must be acquired and processed online to apply complex algorithms for data reduction and filtering for subsequent offline data analysis, reducing the amount of information to be stored in tapes and the communication traffic inside the whole system. These requirements make SoC-FPGAs optimal devices for processing and transmission because of the large amount of input/outputs, parallel processing capabilities, and high-speed transceivers they have. Another advantage of these devices over processor-based systems is their high reconfigurability, allowing reusability and dynamic updates to adapt to different hardware setups and experimental conditions.

The COMPASS data acquisition system (DAQ) migrates from a triggered to trigger-less operation and incorporates an online trigger processor, implementing an event reconstruction at the FPGA level and in real-time, before storing the data in hard disks. Migration requires a new type of frontend electronics for digitizing all detector channels without the intervention of an external trigger, analyzing, and processing the data online and in real-time.

This work focuses on the design of a DAQ platform prepared for a trigger-less operation and on the implementation of algorithms and methods for real-time feature extraction and high-speed multichannel data transmission for the COMPASS ECAL2 calorimeter frontend. The modular hardware for the frontend readout system comprises the FFeCCa carrier board harnessed with a mezzanine sampling analog to digial converter (MSADC) digitizer board, based on a Virtex-4 FPGA and a 12-bit, 80 Msps, 16-channel

ADC, and an MPSoC Ultracasle+ [1] System on Module (SoM) for data processing and communication.

During the Ph.D. program, a full-stack development of the different components of the DAQ platform was carried out, including the design and production of different printed circuit boards (PCB). Besides the FFeCCa carrier, an adapter board was designed to connect the digitizer with any commercial SoC-FPGA platform having a standard FPGA mezzanine connector (FMC).

The adapter board allowed to start with the development of the firmware for the FPGA of the digitizer, the FPGA and microprocessor firmware of the SoC-FPGA, and the remote control software to interact with a PC. With this last part of the firmware, an open-source framework for controlling SoC-FPGA-based systems from a PC called UDMA, with a GNU license, was developed and released for the community. Then a DAQ framework for use with Xilinx SoC-FPGA was also developed and released under BSD-3 license, designed in order to be easy-readaptable as modular way, with the possibility of using multichannel ADCs with minimum porting efforts.

The FFeCCa data acquisition platform was tested and stressed under beam time during the first AMBER pilot run, reading 16 channels at 80 Msps each, from an ECAL2 prototype. AMBER is the next-generation successor of the COMPASS experiment starting its data-taking in 2023. The data obtained during that period were then used to adjust the algorithms used for lossless data compression and digital pulse processing for features extraction.

This research work has been carried out as part of a collaboration between the Multidisciplinary Laboratory of the Abdus Salam International Centre for Theoretical Physics (MLAB-ICTP) and the COMPASS Group of the Italian National Institute of Nuclear Energy, Trieste Section (INFN-TS).

Objectives of the Thesis

The objectives and main contributions of this thesis can be summarised as follows:

1. Design and implementation of the hardware for a data acquisition platform able to work in trigger-less mode

    (a) Develop a modular carrier board (FFeCCa) to allow its use with different frontend readouts and MPSoC Ultrascale+ SoMs.

    (b) Develop a hardware adapter board for testing the digitizer board (MSADC) and measuring the data transfer rates using different commercial SoC-FPGA evaluation boards.

    (c) Validate the designed platform under beam conditions at CERN.

    (d) Design a modular open-source framework for SoC-FPGA based DAQ.

2. Development and implementation of firmware for free-running operation and the control software for remote access to the data acquisition platform

    (a) Implement the MSADC FPGA firmware for multichannel acquisition and lossless data compression for a free-running operation.

    (b) Implement the firmware of the SoC-FPGA for reading multiple channels in parallel, decompressing, detecting pulses, extracting their main features, and sending data using high-speed optical transceivers.

    (c) Develop the microprocessor firmware for slow control services such as system configuration and status information.

    (d) Design a general-purpose user interface for remote control from a PC through Ethernet.

3. Design and implementation of a Digital Pulse Processor for real-time features extraction in a trigger-less system

(a) Acquire experimental data from an ECAL2 prototype using the developed platform for off-line pulse study and characterization.

(b) Analyze the acquired data to extract the mathematical model of typical ECAL2 pulses and characterize the noise.

(c) Study and develop methods for features extraction based on the parameters provided by a mathematical model of typical pulses and pulse-shape discrimination algorithms.

(d) Implement a Digital Pulse Processor for two operation modes, the first extracting the main features of the signals, and a second sending a trace containing full pulse capture.

4. Release the developed designs to the scientific community

(a) Release the developed hardware with an open hardware license.

(b) Release the framework for working with SoC-FPGA multichannel data acquisition platforms with an open software/hardware license.

Scientific Publications

During this research, a total of 14 publications in journals and conferences have been produced, as follows:

Journals

[1] K.S. Mannatunga, B. Valinoti, W. Florian Samayoa, M.L. Crespo, A. Cicuttin, J. Folla Kandem, L.G. Garcia, S. Carrato, (2022). Data Analysis and Filter Optimization for Pulse-Amplitude Measurement: A Case Study on High-Resolution X-ray Spectroscopy. Sensors 22(13), 4776; https://doi.org/10.3390/s22134776

[2] A. Cicuttin, I.R. Morales, M.L. Crespo, S. Carrato, L.G. García, R.S. Molina, B. Valinoti, J. Folla Kamdem. (2022). A Simplified Correlation

Index for Fast Real-Time Pulse Shape Recognition. Sensors, 22, 7697. https://doi.org/10.3390/s22207697

[3] W. Florian Samayoa, B. Valinoti, R. Molina, L. G. Garcia, M.L. Crespo, S. Carrato, A. Cicuttin, S. Levorato (2023). Diagnostic Analytics for Pixelated Particle Detectors: A Case Study. In R. Berta & A. De Gloria (Eds.), Applications in Electronics Pervading Industry, Environment and Society, pp. 216-221, Springer Nature Switzerland, https://doi.org/10.1007/978-3-031-30333-3_28.

[4] M.L. Crespo, F. Foulon, A. Cicuttin, M. Bogovac, C. Onime, C. Sisterna, ... , B. Valinoti. (2021). Remote Laboratory for E-Learning of Systems on Chip and Their Applications to Nuclear and Scientific Instrumentation. Electronics, 10(18), 2191, DOI: 10.3390/electronics10182191

[5] L.G. García, M.L. Crespo , S. Carrato, A. Cicuttin, W. Florian, R. Molina, B. Valinoti, S. Levorato (2021). High Voltage Isolated Bidirectional Network Interface for SoC-FPGA Based Devices. A Case Study: Application to Micro-pattern Gaseous Detectors. In S. Saponara & A. De Gloria (Eds.), Applications in Electronics Pervading Industry, Environment and Society, Lecture Notes in Electrical Engineering, Vol 738, pp. 280–285, Springer, Cham. DOI: 10.1007/978-3-030-66729-0_34

[6] W.O. Florian Samayoa, B. Valinoti, L.G. Garcia Ordoñez, M. Cervetto, E. Marchi, M.L. Crespo, S. Carrato and A. Cicuttin (2022). An Open-Source Hardware/Software Architecture for Remote Control of SoC-FPGA based Systems. In: Saponara, S., De Gloria, A. (eds) Applications in Electronics Pervading Industry, Environment and Society. Lecture Notes in Electrical Engineering, vol 866. Springer, Cham. https://doi.org/10.1007/978-3-030-95498-7_10

[7] S. Carrato, C. Chatterjee, A. Cicuttin, P Ciliberti, M.L. Crespo, S. Dalla-Torre, S. Dasgupta, W. Florian, L. García Ordóñez, M. Gregori, A. Kosoveu, S. Levorato,

M. Mannatunga, G. Menon, F. Tessarotto, Triloki, B. Valinoti, Y. X. Zhao. A scalable High Voltage power supply system with SoC control for Micro Pattern Gaseous Detectors. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 963 (2020) 163763. DOI: 10.1016/j.nima.2020.163763

[8] G. D. Alexeev et al. [COMPASS collaboration] "Exotic meson $\pi 1(1600)$ with JP C = 1 $\pm$ and its decay into $\rho(770)\pi$". In: Phys. Rev. D 105 (2022), p. 012005. doi: 10.1103/PhysRevD

[9] G. D. Alexeev et al. [COMPASS collaboration] "Double J/ψ production in pion-nucleon scattering at COMPASS". In: Physics Letters B 838 (2023), p. 137702. doi: 10.1016/j.physletb.2023.137702

Conferences

[1] "A Hardware/Software Architecture for Remote Control of SoC-FPGA Based Reconfigurable Virtual Instrumentation", Second International Conference on Advances in Electrical, Electronic and System Engineering (ICAEESE 2019), Gauhati University, Guwahati, Assam, India, November 2-3, 2019

[2] "Real-time quantum-accurate voltage waveform synthesis". In: 2020 Conference on Precision Electromagnetic Measurements (CPEM). 2020, pp. 1–2. doi: 10.1109/CPEM49742.2020.9191715

[3] "SoC-based trigger-less data acquisition for multichannel detectors", Società Italiana di Fisica, 106° CONGRESSO NAZIONALE, 14-18 SETTEMBRE 2020, atticon13007

[4] "New Electronics for ECAL2", DAQFEET-2021, Monday, 8 February 2021 - Wednesday, 10 February 2021, Online

[5] "The high voltage system the novel MPGD-based photon detectors of COMPASS RICH-1 and its developement towards a scalable High Voltage Power Supply System with system on chip control for Micro Pattern Gaseous Detectors". RICH2022, University of Edinburgh, Edinburgh, Scotland, 12–16 Sept 2022.

Thesis Outline

The remainder of this thesis is structured as follows: Chapter 2 presents the basic concepts of data acquisition systems, the trigger, and the manner in which the main high-energy particle physics experiments implement DAQ according to their requirements. Chapter 3 describes the COMPASS experiment and its main constituents. Chapter 4 presents the physics and concepts of electromagnetic calorimetry, ECAL2 detector, readout system, and upgrading of the free-running proposal. Chapter 5 describes all the hardware developed for evaluating and testing the digitizer board as an intermediate step before the final design. Chapter 6 presents the development of the FPGA-SoC Frontend Carrier Card, justifying all main component selections and how the most critical parts were tested. Chapter 7 describes the details of the firmware developed for testing the hardware, an open software/hardware framework as well as the main firmware implementations. In Chapter 8 the features extraction method developed for pulse amplitude measurement is described and the data analysis for pulse model estimation and noise characterization is presented. Chapter 9 presents the results of the experiments conducted during the research program related to data feature extraction and trigger-less operation. Finally, the conclusions, remarks and future work are presented in Chapter 10.

Chapter 2

Data Acquisition Systems for High Energy Particle Physics

Data acquisition systems (DAQ) are used to acquire the signals from sensors, digitize them, and record the data from the measurements for further processing and analysis. In most cases, these measurements need to also be plotted or shown on a graphical user interface as they occur over time. By this way, a DAQ is an essential component of simple instruments, such as a pocket-size thermometer, and of complex apparatus, like those used in high-energy physics (HEP) experiments, with hundreds of thousands channels that need to be digitized at high-speed and high-resolution.

In a general way a DAQ can be described as a system composed by a series of building blocks with very specific and defined tasks. Figure 2.1 shows a block diagram of a typical DAQ in a wide sense. Even if some of the components can be neglected for simple systems, a DAQ is typically composed by [2]:

- A detector for sensing and transforming a physical measurable quantity into a proportional voltage or current signal [3].

- A frontend electronics that includes a signal conditioning stage [4] for adapting the
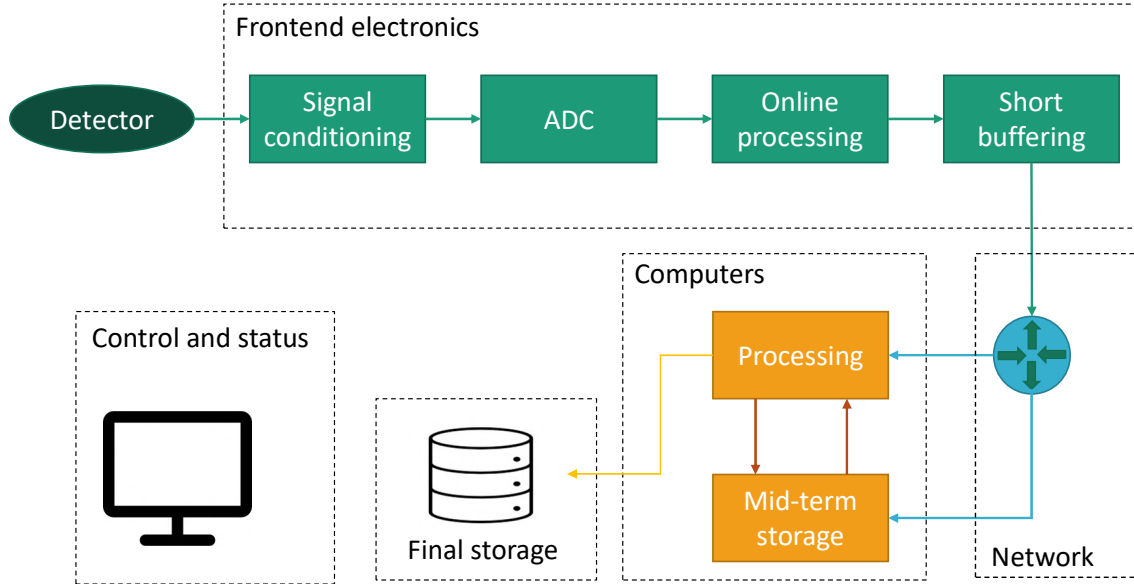
Figure 2.1: Typical DAQ schema.

measured signal to a proper level and bandwidth for analog-to-digital conversion (ADC). An online data processing stage for filtering the data or re-adapting some characteristic to the digital system. The online processing at this stage of the DAQ is made with FPGAs [5] or microcontrollers (μC) [6], depending mostly on the processing requirements. The processed data are then temporary stored in short data buffers, until the system asks for them when possible. These short buffers are important owing to the different latencies involved in the readout chain and they are usually implemented as First-In First-Out (FIFO) memories [7].

- The frontend electronics [8] is connected to some type of network link or communication channel for transferring the data to a computer or data server. As one of the most important parameters on a DAQ is the ability to save relevant data in the final storage, network resources play a central role in the system.

- Before being stored in permanent resources, the data are usually saved in middle term data servers, in such a way they can be accessed, processed, and filtered by

processing servers. The main purpose of this is to reduce and optimize the amount of information to be saved on the final storage.

- The DAQ may also incorporate a software for giving support to all services of the data acquisition, including the evaluation of data integrity, and monitoring and control of the hardware status. This software should be easily configurable and provide a friendly interface for the user.

A DAQ must work predictable, safe, and preferably with the ability to recover from hardware failures without losing any data.

2.1 High Energy Physics Experiments

HEP experiments seek to identify the fundamental constituents of matter (quarks and leptons) and reveal the laws governing their interactions (electromagnetic, strong, and weak). To study and uncover these constituents and interactions, the experiments are built with different types of detectors, each of which is sensitive to specific characteristics. These characteristics are then tied to the type of particle or to a property of interaction between particles [9]. In addition, extremely high energies are required to reveal these characteristics. The interactions between two very energetic particles spread a series of subparticles with different scattering angles, energies, and momenta. Thousands of detectors are usually used to detect and measure the properties of the interaction. By merging the data provided by these detectors, it is possible to reconstruct the collision event and study their characteristics.

In these experiments, DAQ systems must handle extremely high data rates sourced from thousands of data channels [10, 11, 12]. These types of DAQ require cutting-edge technologies from a data processing perspective, as well as from data handling, networking, storage, and error recovery.

As an example of DAQ for a HEP experiment, in Figure 2.2 it is shown a simplified

1 400 Detector
Front−ends
(250 000 Channels)

150 Readout
Modules

S−Link
Multiplexer

64 optical
S−links

64 Readout buffer
16 PCs

4 Gigabit
Ethernet
switches

12 Event builder
& Recording

20 m

TCS CATCH

SMUX

TCS GeSiCA

10–100 m

40 MByte/s/link

150 m
100–160 MByte/s/link

4 x 512 MByte spill buffer per PC

10 m
100 MByte/s/link

Gigabit switch

10 m
100 MByte/s/link

5 km
70 MByte/s

event builder
and filter

Central Data
Recording
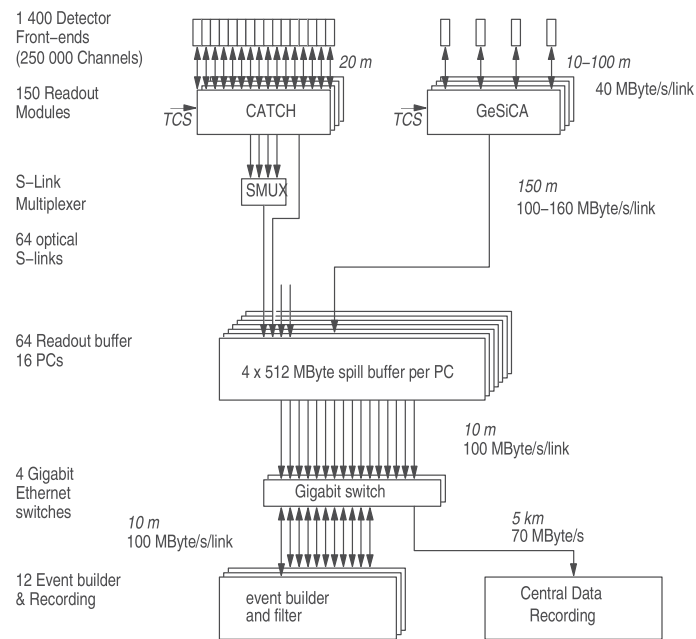
Figure 2.2: Simplified schema of the first implementation of the COMPASS DAQ, from [10].

architecture of the data acquisition system of COMPASS experiment at CERN. The characteristics and constituents of this system are explained in more detail in the following chapters.

The architecture implementations consist of several layers of hardware and firmware, organized in such a way that the data can then be used to completely reconstruct a specific event. However, it is not feasible to store the raw data from the detectors continuously, owing to the high amount of data generated per unit of time (TB/s) and the need of saving storage costs and offline data processing time. To avoid such a situation, the DAQ is equipped with a trigger control system (TCS), which decides both when storing and rejecting data. The decision about taking a "screenshot" of the detector state for storage can be generated at different stages of the system and with specific criteria depending on the stage. The trigger should be a very fast process, based on simple criteria, and should be used to provide a time reference for the readout of all detectors.

A common characteristic of all DAQ is that they can accept a limited number of events per unit of time. This number depends on the time response of the detector, sampling frequency, data size of the event, and event data transmission rate to the storage system.

The "dead time" (τ) [13] of the DAQ that is the time interval the system is busy after a trigger arrival, during which is insensitive to new events, gives an idea about the fraction of events that can be accepted by the system.

Without entering in details about the classification of the dead time model, considering that each time the trigger signal is generated the system will not be able to accept a new trigger (non-paralyzable) and calling m to the physics event rate:

- the time the DAQ is busy can be defined as: $(m \cdot \tau)$.

- the time when DAQ is free can be represented as: $(1 - \tau \cdot m)$.

Now, if we call n to the actual accepted event rate, it can be said that only a fraction
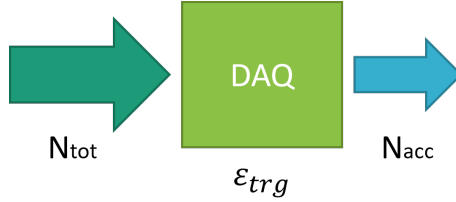
Figure 2.3: Actual physics events vs accepted events in the DAQ.

given by Equation 2.1 will be recorded by the DAQ.

$$n = \frac{m}{(1 + \tau \cdot m)} \tag{2.1}$$

This quantity will always be lower than the real event rate, owing to the nonzero processing or dead time.

Now, with these two factors $n = N_{acc}$ and $m = N_{tot}$, the efficiency (ϵ_{trg}) of the DAQ can be defined as the relationship between the accepted events and the total number of events, as shown in Figure 2.3. The efficiency is constrained by the dead time and event rate, as shown in Equation 2.2.

$$Efficiency = \frac{N_{acc}}{N_{tot}} = \frac{1}{(1 + \tau \cdot m)} \tag{2.2}$$

Then, depending mainly on the type and technology of the detector, each channel can have a different τ and raw data production rates. Therefore, it may be possible that the data transmission requirements for the detectors can be different as well. The bandwidth of the DAQ must be sufficient to handle all data, transmit the entire event, and store it for future offline processing.

The required bandwidth per channel is defined as:

$$BW_{CH} = N_{acc} \cdot S_e$$

where N_{acc} is the actual event rate running in the DAQ and S_e is the event size.

Similar to classic micro-architecture computing concepts [14], there are two approaches that allow minimizing the dead time or boost the time processing: one is

based on the parallelism and the other on a pipeline processing.

The strategy of the parallelism, shown in Figure 2.4a, is to have independent processing and trigger paths for each of the readout detector elements, as many as affordable. From the pipeline approach, shown in Figure 2.4b, the strategy is to segment the data path as much as possible to absorb the fluctuations. The processes can be reorganized in different steps. The use of FIFO buffers between steps allows steps with different latencies to be handled at different frequencies, processing those longer at higher clock frequencies, balancing the final total time in a more efficient way. As long as the FIFOs do not fill up, no data is lost, so the depth of the input FIFOs must be carefully chosen.

2.2 Triggered systems

In HEP experiments, it can be said that there is a trigger condition when the physics status meets the requirements for capturing the data of an interesting event. The trigger is a system built up by the electronics to indicate the occurrence of a desired temporal and spatial correlation in the detector signals [15]. This correlation is determined by an examination of some dedicated detectors that provide information about a characteristic signature that distinguishes the pursued event from others that occur at the same time. A time coincidence increases the probability that all particles may originate from the same event.

A trigger must pass the events under interest efficiently without permitting the DAQ to become swamped with data that are not relevant or with events that are similar to those of interest but not identical. The design of a trigger is strongly dependent on the scope of the experiments and must be built according to the physics environment (beam parameters, geometry, target, etc.).

In a trigger system, we can usually find several different types of detectors and techniques to use the information they provide. There are detectors with inclusive
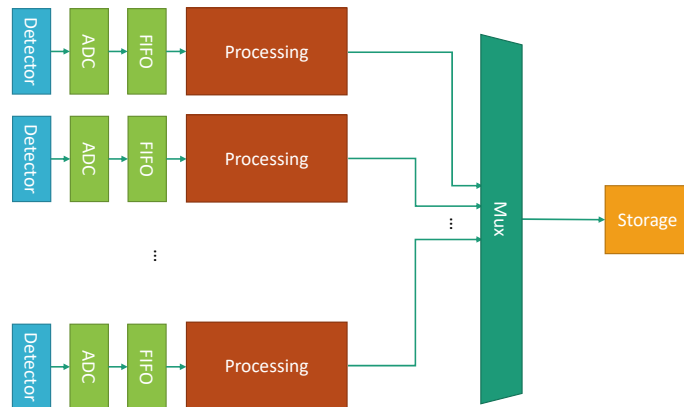
(a) Parallel processing approach.
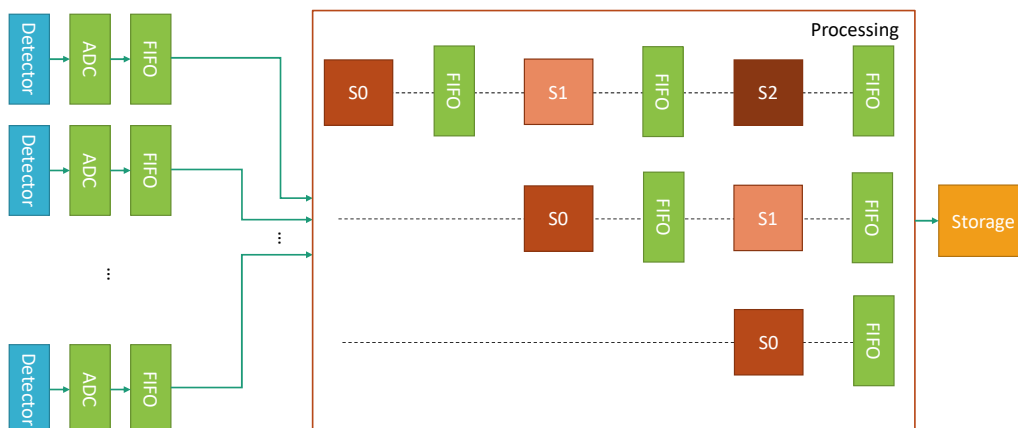
(b) Pipelined processing approach.

Figure 2.4: Different data processing approaches for minimizing the dead times.

information, meaning the trigger decision is affected in a positive result if the information of the detector is actively present. Opposite to the previous, we can find veto detectors, which means that detectors are used to explicitly exclude the event. This is very useful in those cases when different particles leave traces on the detectors. For example, muon triggers take advantage of the muon's ability to penetrate large amounts of matter before being absorbed, the muon trigger could consist of a massive absorber followed by a general purpose particle detector as a multiwire proportional chamber [16].

In big HEP experiments, the trigger system is typically constructed as a hierarchical structure. There is a first level, where the event sizes are small and the rates are likely high. Then, there is a second or High-Level Trigger (HLT) with lower rates but larger event sizes. Figure 2.5 shows a typical trigger system with a two level trigger scheme. First-level triggers or Level-1 (L1), in general, are tied directly to detectors and are typically very fast; however, the data are recorded once the trigger decision is made. It may happen that during data acquisition or during the dead time another event could be tagged as a valid trigger, but as it can not relaunch the signal, the event will be superimposed on the previous. During the L1 trigger decision time, all data from the detectors are buffered in the frontend electronics. This is the time needed for transmitting the data from the detector's subset to the place where the trigger decision is processed and constructed, and back to the frontends; and it is usually of the order of a few μs. Classically, an analog delay is used for compensating this trigger decision latency. Then, when the trigger decision arrives at the frontend, a buffered screenshot of the event is sent to the DAQ for storage or for being processed by the HLT stage.

An important parameter given by the experiment for the trigger is the expected event rate, which depends on the type of experiment (collider or fixed target) [17], luminosity of the beam, and energy. To process the events efficiently, the system must evaluate the trigger decision in a time shorter than the period between events. Nevertheless, because this period is an expected quantity shaped by a normal distribution, if the processing time
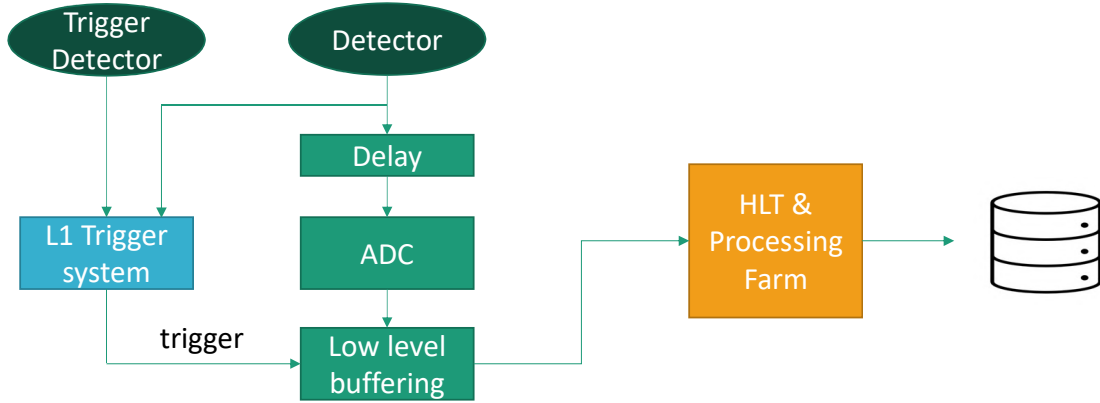
Figure 2.5: Typical trigger system with a two level triggering scheme.

is too marginal, there will be overlapping events, and an important part of the data will be lost. Then, the maximum trigger rate will be constrained by the limiting dead time of the system.

For large HEP experiments, there is another big challenge to deal with, which is the trigger synchronization between all the detectors. As these experiments can have times of flight lengths to traverse all the detectors longer than the expected time between events, so that there must be a very tight timing and synchronization of the trigger system to adjust the data acquisition of all detectors.

From Equation 2.2, the efficiency of the trigger DAQ is determined by dividing the number of events that pass the trigger by the number of actual events. The trigger efficiency must be sufficiently high at a low threshold to ensure a high number of events to provide sufficient statistics for physics. This efficiency is evaluated by considering the target that comes from the physical goals of the experiments by benchmarking the physical processes.

The measurement of trigger efficiency requires to have some overlapping triggers such that the efficiencies can be measured from the data. To understand the trigger efficiency, the data used as input to L1 should also be transmitted via the DAQ for storage, together with the event readout data. In addition, all trigger components, whether they were

responsible for the L1 trigger or not, should also be sent.

The L1 acceptance rate is then limited not only by the speed of the detectors' electronics but also by the rate at which the DAQ can harvest the data from them. The maximum L1 trigger acceptance rate is then given by the average time to read the data for processing by the higher-level triggers and the average steps in the HLT logic. In general, L1 triggers are simple arithmetic group operations of the detectors participating in decisions.

After L1 trigger decision based on the information of the detectors, the event rate can decrease to some order of magnitude lower than the physics rate. Then at the HLT stage an online reconstruction of the event is performed and a decision whether the event has passed the conditions based on full event information is done. In general, there are two philosophies for HLT. The first assemblies and combines a complete event into a HLT node for processing. The second, making profit of a high degree of data locality in the event processing, processes the data where it is available, so the penalty in time for moving the data from one node to another is minimized [18].

While lower-level triggers are typically built with analog modules or FPGAs for low-latency determination, high-level triggers are implemented in farms with several PCs, as this decision can be done with longer latencies than L1. HLT frameworks allow dynamic reconfiguration, thereby supporting the online remapping of any failing node or link. The HLT must unpack the information coming from the filtered data from L1, process it, construct the entire event, and decide whether the event has all the conditions for storing it and saving the data [19].

In target-based HEP experiments, a physics trigger consists of three subsystems: beam-defining elements to select beam particles crossing the target, veto detectors to reject events containing particles produced outside the target or outside the spectrometer acceptance, and specific detector systems that account for the particular physics case.

On the other side, for collider based HEP experiments, as the beam parameters

(expected event rate and particles bunch crossing time) are perfectly known, the physics trigger is mainly build directly with limited detectors data, without ad-hoc trigger detectors.

2.3 Trigger-less approach

Triggered DAQs are widely used and are mainstream in almost all HEP experiments. However, there are situations where, owing to the high event rates that occur in some specific scenarios or because of the high luminosity of the beams, event selection with conventional triggers becomes almost impossible. This difficulty lies in the fact that the times in which the data would have to be processed by conventional triggers would be prohibitive from the point of view of both the response of the detectors and the delays incurred in processing the signals from the frontend to the higher levels of the DAQ hierarchy. In these cases, a trigger-less DAQ with a different trigger paradigm is needed.

In a trigger-less DAQ, the main difference is that the detectors send a continuous data stream to the DAQ; the concept of L1 trigger is built in a different way, and it no longer exists as in the triggered approach. The detector frontends are responsible for making some type of pre-selection of events to avoid jamming the first levels of the DAQ transmission channels.

Advances in electronics and new technologies in programmable logic devices offer the possibility to make online analysis of signals obtained with high-speed ADCs, reshaping the pulses, and extracting features and signatures in real-time. Figure 2.6 shows a tipical block diagram of a trigger-less DAQ. Extracting the arrival time of the pulse and timestamping it, the systems can run without an event trigger. This free-running mode is based on the reconstruction of the events using the timestamp information. Event reconstruction is performed by selecting all the data corresponding to a certain event. The reconstruction can be performed online, reordering the data of each subdetector. This reordering is associated to the timing calibration of each of them according to the physical position
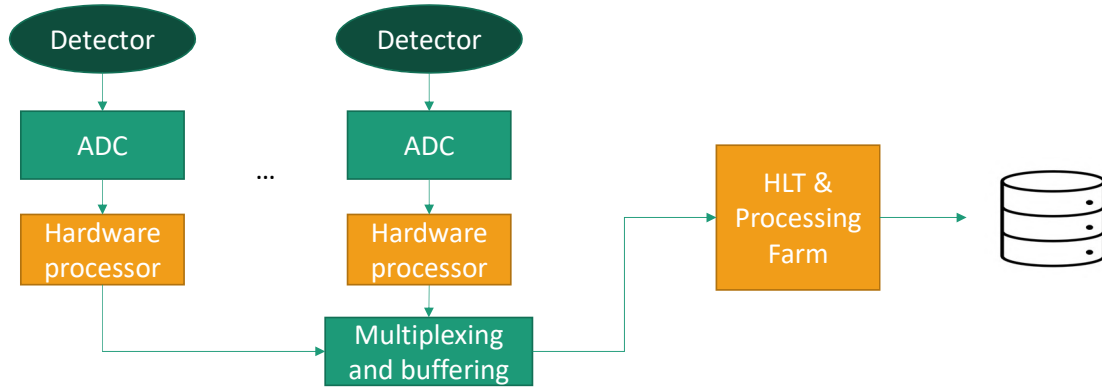
Figure 2.6: Typical trigger-less system, where the L1 level is supressed and replaced by a hardware processor.

where the particles collides or impinges.

2.4 DAQ & Trigger systems on HEP experiments

Data acquisition and trigger systems play important roles in particle physics experiments. They work as an interface between the detector and computer facilities for storing and processing data. The quality of the data for future studies depends on the design and maintenance of these systems and has a direct impact on the physics that can be extracted from this information [20].

To design, build, and operate DAQ and trigger facilities, a large number of challenges must be tackled. With the new generation of experiments, the energy and luminosity achieved in recent years have meant that the event rate has grown to increasingly demanding levels, making the data rates generated per unit of time very difficult to sustain, process, and consequently save for analysis [21]. In the largest experiments, the bandwidth of raw data can exceed 100 TB/s of data sourced from hundreds of thousands of detectors. In Figure 2.7 the event size and trigger rates of the main particle physics experiments for the Level 1 trigger or equivalent can be seen.

Figure 2.7: First level trigger rates and sizes for main particle physics experiments.

Although many efforts have been made since the first HEP experiments for having a common DAQ, each experiment has its own particularities and requirements, resulting in different needs for the event discoveries and processing of the data. Indeed, each experiment is constrained by the nature of the physics, the number of detectors, the number of channels, the structure of the data, the data integrity and robustness, and the bandwidth, among others.

As already mentioned in HEP, there are two common ways to study the interaction of accelerated particles: fixed-target and collider experiments. In the first, a particle beam is aimed at a proportionally large and stationary target. The target may be a chunk of metal, liquid, or gas contained in a flask. From the interaction of the particle beam and the target different subparticules and characteristics can be obtained. In the case of the collider setup, two particle bunch beams are aimed at each other and collide in a certain point at a well known moment. The reason for this last configuration setup is that significantly higher energy levels can be obtained. For each of them DAQ and trigger systems have different requirements, because of the different beam-on periods and frequencies at which the probability of occurrence of events exists.

In fixed-target experiments, random events are expected during a period of time t, during which the presence of a beam affects the target. DAQs must be prepared to buffer a large amount of data according to the average event rate, plus a guard for cases in which the average is exceeded. In collider experiments, there is a high probability of occurrence of events during a very short period of time (when the particle bunches are crossed), which is perfectly determined, so the DAQ can work in a pipelined approach but with a higher event acceptance operation.

In the following subsections, a brief review of the main current HEP experiments is presented to understand the main and particular challenges each experiment requires and the reason there is no industrial or universal approach for DAQ and trigger systems.

2.4.1 ATLAS

A Toroidal LHC ApparatuS (ATLAS) [22] is a CERN facility for studying the particles emerging from the proton-proton collisions at TeV levels. The high luminosity and increased cross-sections of the Large Hadron Collider (LHC) enabled high precision tests of Quantum Chromodynamics (QCD) [23], electroweak interactions, and flavour physics. Then, the search for the Standard Model Higgs boson was used as a benchmark to establish the performance of important subsystems of ATLAS. The detector also tracks and identifies particles to investigate a wide range of physics, from the study of the Higgs boson and top quark to searching for extra dimensions and particles that could make up dark matter.

ATLAS has the dimensions of a cylinder, 46 m long, 25 m in diameter, and sits in a cavern 100 m below ground. The ATLAS detector weighs 7,000 tonnes, similar to the weight of the Eiffel Tower.

In Figure 2.8 it is depicted the detector itself where it can be seen how it is composed as a many-layered instrument, designed to detect some of the tiniest and most energetic particles ever created on Earth. It consists of six different detection subsystems wrapped
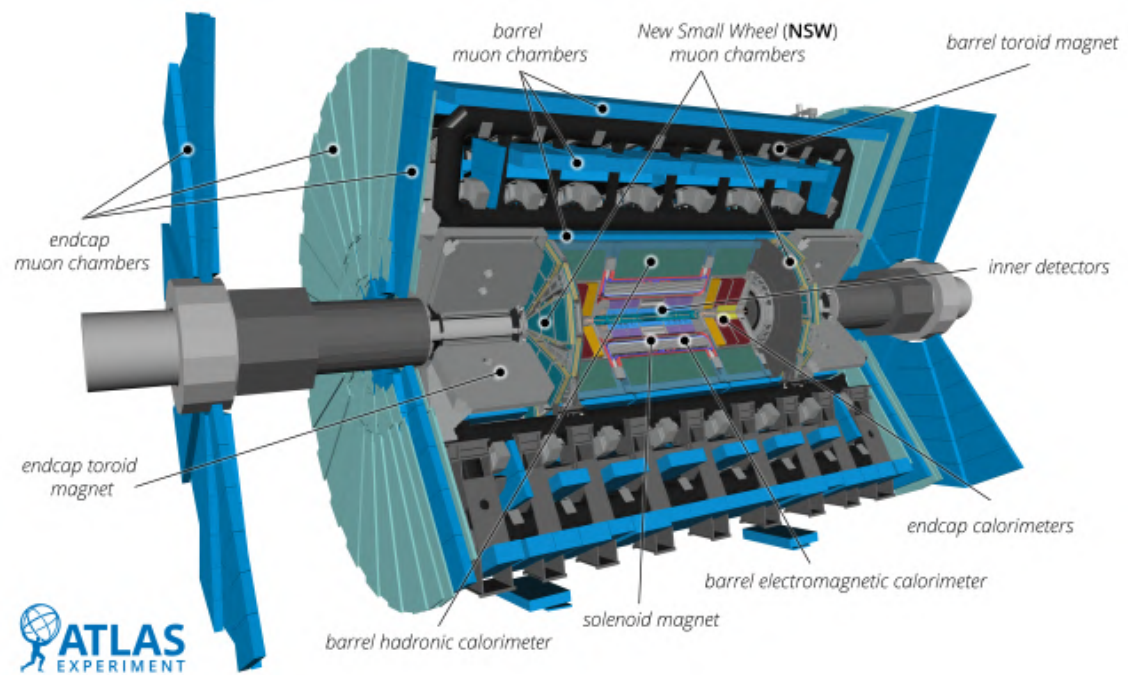
Figure 2.8: ATLAS detector schematic with the different subdetectors labeled, from [24].

concentrically in layers around the collision point to record the trajectory, momentum, and energy of the particles, allowing them to be individually identified and measured. A huge magnet system bends the paths of the charged particles so that their momenta can be measured as precisely as possible.

Beams of particles travelling at energies up to seven trillion electron-volts or speeds up to 99.999999% that of light coming from the LHC collide at the center of the ATLAS detector, producing collision and generating new particles that fly out in all directions. Over a billion particle interactions occur in the ATLAS detector every second, a data rate equivalent to 20 simultaneous telephone conversations held by every person on Earth. Only one in a million collisions are flagged as potentially interesting and recorded for further study.

For the LHC Run 3, the maximum rates at L1 and HLT are approximately 100 kHz and 1.5 kHz, respectively, [25]. The L1 trigger decision is built with data from the calorimeters and muon system, in a Central Trigger Processor (CTP). The resulting trigger signal is delivered to the tracking detectors, calorimeters, and muon system to indicate the datataking. The last two system detectors are read and assembly using the Front-End LInk eXchange (FELIX) and sent to the ReadOut System and through this to the Data Collection Network for HLT evaluation. In a last step, after the HLT the data is sent for permanent storage. Together, these constitute the DAQ system, as shown in Figure 2.9. They are based on commodity personal computer (PC) servers and a standard networking infrastructure. The readout system includes custom input/output cards for converting the detector frontend protocol data into standard network packets [26].

2.4.2 CMS

The Compact Muon Solenoid (CMS) detector [27], depicted in Figure 2.10, is a multi-purpose apparatus due to operate at the LHC at CERN.

The CMS Collaboration has a broad physics program, ranging from measurements
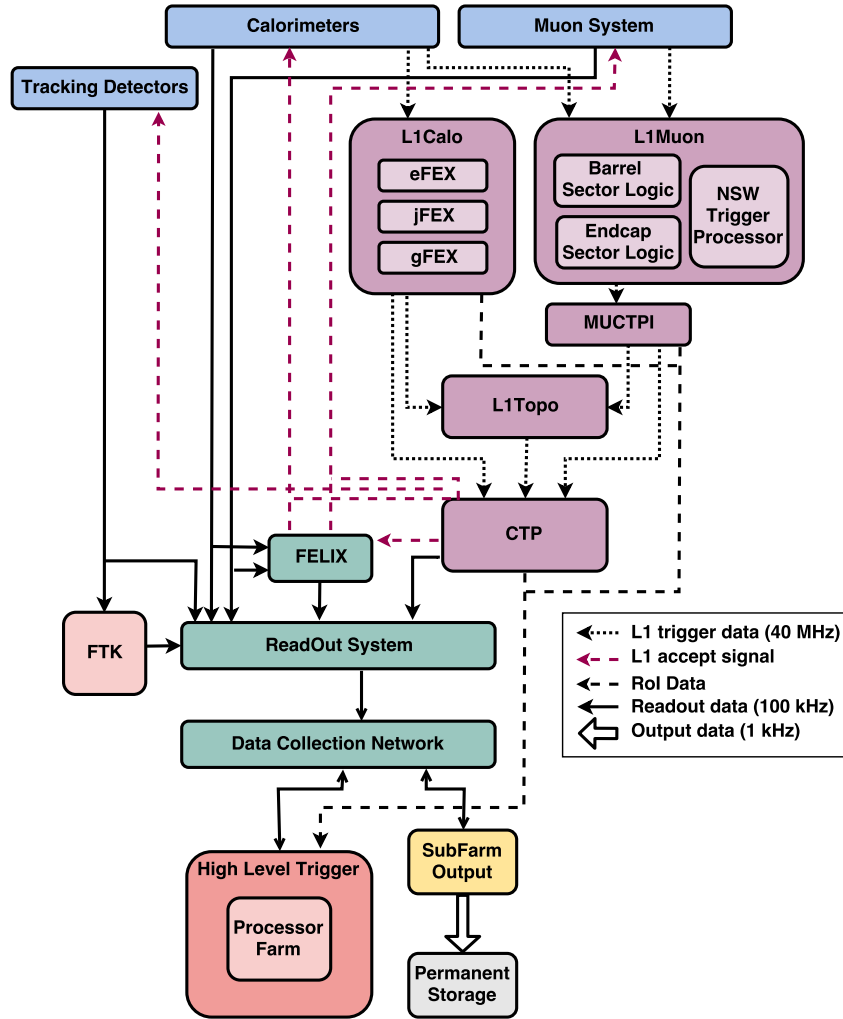
Figure 2.9: Block diagram of ATLAS TDAQ Phase-II, from [26].

CMS DETECTOR

Total weight   : 14,000 tonnes
Overall diameter : 15.0 m
Overall length   : 28.7 m
Magnetic field   : 3.8 T

STEEL RETURN YOKE
12,500 tonnes

SILICON TRACKERS
Pixel (100x150 μm) ~16m² ~66M channels
Microstrips (80x180 μm) ~200m² ~9.6M channels

SUPERCONDUCTING SOLENOID
Niobium titanium coil carrying ~18,000A

MUON CHAMBERS
Barrel: 250 Drift Tube, 480 Resistive Plate Chambers
Endcaps: 468 Cathode Strip, 432 Resistive Plate Chambers

PRESHOWER
Silicon strips ~16m² ~137,000 channels

FORWARD CALORIMETER
Steel + Quartz fibres ~2,000 Channels

CRYSTAL
ELECTROMAGNETIC
CALORIMETER (ECAL)
~76,000 scintillating PbWO₄ crystals

HADRON CALORIMETER (HCAL)
Brass + Plastic scintillator ~7,000 channels

Figure 2.10: CMS schematic with major features tagged, from [28].

of the standard model to the Higgs boson. The prime motivation of the LHC is to elucidate the nature of electroweak symmetry breaking for which the Higgs mechanism is presumed to be responsible and for studies of heavy-ion collisions. The program also includes searching for new particles, phenomena, and even extra dimensions in the universe.

The architecture of the CMS Data Acquisition system (DAQ) is shown schematically in Figure 2.11. As in the case of ATLAS, it is based on a two level trigger system where the L1 trigger is built with the detectors frontend information and then for a safe operation the data are sent through a high performance network to a computing service for event filter and storage.

The CMS trigger and DAQ systems [29] were designed to collect and analyze the

27
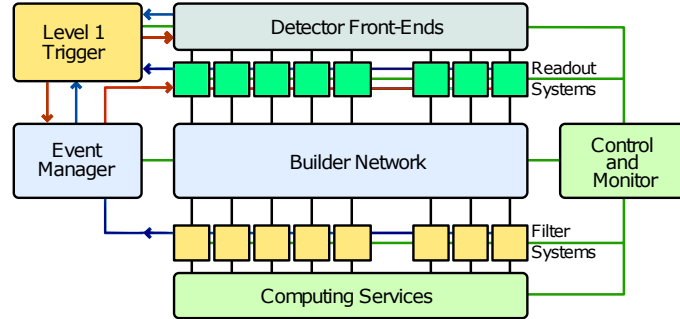
Figure 2.11: Simplified block diagram of the CMS DAQ, adapted from [29].

detector information at the LHC bunch crossing frequency of 40 MHz. The rate of events to be recorded for offline processing and analysis is on the order of a few 10^2Hz.

At the normal operative luminosity of Run 3, the LHC rate of proton collisions is approximately 20 per bunch crossing, producing approximately 1 MByte of zero-suppressed data in the CMS readout systems [30]. The first-level trigger is designed to reduce the incoming average data rate to a maximum of 100 kHz by processing fast trigger information coming from the calorimeters and muon chambers, and selecting events with interesting signatures. Therefore, the DAQ system must sustain a maximum input rate of 100 kHz for a data flow of approximately 100 GByte/s coming from approximately 650 data sources and must provide sufficient computing power for a software filter system, the HLT, to reduce the rate of stored events by a factor of 1000. In CMS, all events that pass the L1 trigger are sent to a computer farm (Event Filter) that performs physics selections to filter events and achieve an output rate in the order of 10^2 events per second [12].

The L1 trigger uses coarsely segmented data from the calorimeters and muon system while holding high-resolution data in pipelined memories in frontend electronics. The HLT has access to the complete readout data and can therefore perform complex calculations, similar to those made in the analysis off-line software if required for specially interesting events.

L1 trigger hardware is implemented in FPGA technology where possible, but ASICs

and programmable memory lookup tables (LUT) are also widely used where speed, density, and radiation resistance requirements are important. The allowed L1 trigger latency between a given bunch crossing and the distribution of the trigger decision to the detector frontend electronics, is 3.2 μs. The processing is then pipelined to enable a quasi-deadtime-free operation.

2.4.3 LHCb

Large Hadron Collider beauty (LHCb) [31] is an experiment located at the CERN's LHC with the special purpose of studying and precisely measuring the Charge Parity (CP) symmetry violation parameters in the B-B (Baryon-Baryon) quark system. The geometry of the detector is constructed as a forward single-dipole consisting of a vertex detector, tracking system, aerogel, and RICH detectors, electromagnetic and hadron calorimeters, and muon detector. A schematic of the detector is shown in Figure 2.12. The LHCb started taking data in 2010, and since then, it has had two main upgrades.

Originally, the DAQ architecture of the LHCb was organized as a multilevel triggered system. The main functional parts were a Fast Control and Timing for common clock distribution, synchronous to the accelerator for the frontend electronics and to the two levels of hardware triggers: Level-0 and Level-1. Other part was the frontend electronics for digitizing and buffering the data until the trigger latencies are processed and multiplexed before being passed to the DAQ system.

Then there were the Readout Units (RU), which worked as the multiplexer of the frontends and as a link to the Readout Network (RN). The RN provided the first services for the event-building as well. After this middle level there was the Sub-Farm Controller, which provided the interface between the RN and the higher level triggers running on a processor farm. Finally, the CPU farm where higher-level triggers are executed (Level-2 and Level-3). With this system, a data rate of 40 TB/s at the L0 trigger was carried to 20 MB/s at the last level for storage in data servers after the L3 trigger and event builder
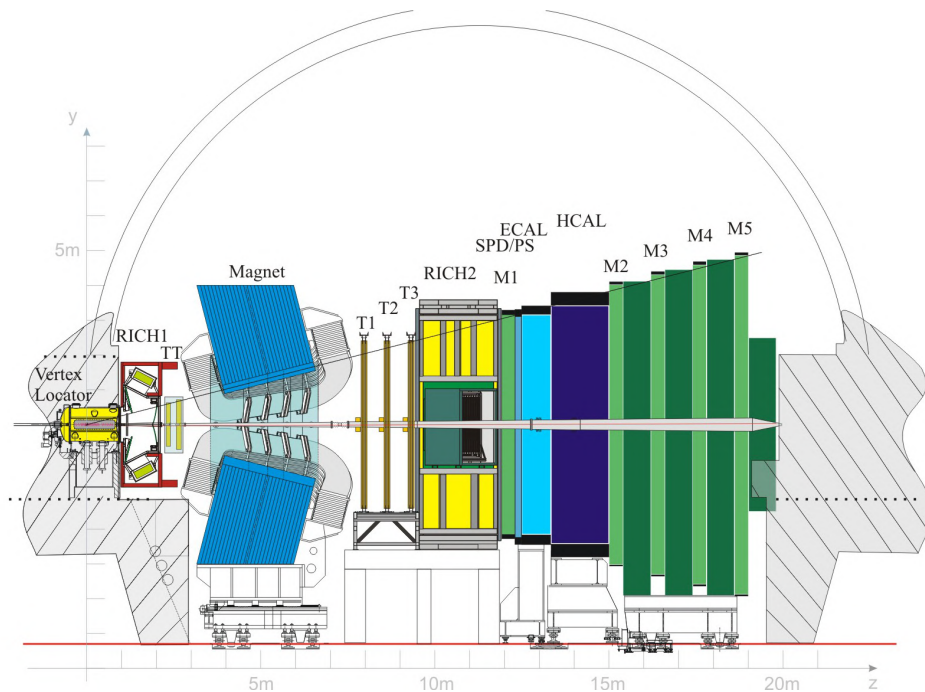
Figure 2.12: LHCb schematic, from [32].
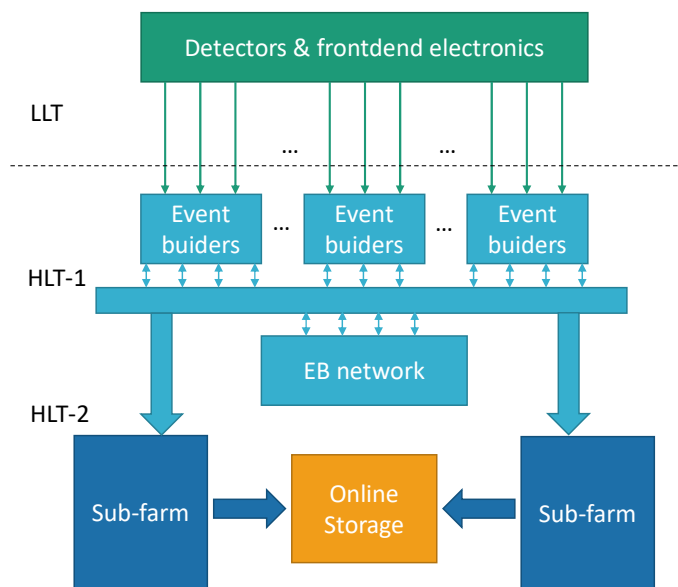
Figure 2.13: Simplified LHCb DAQ block diagram.

[33].

Owing to the increase in the energy and luminosity levels of the LHC Run 3, the LHCb trigger system had to be upgraded. The nominal output rate of the DAQ increased from 20 MB/s to 2 GB/s, scaling up every part of the trigger and DAQ system in the same levels. In Figure 2.13 there is a simplified block diagram of the DAQ with its main constituents grouped by colors. To process this amount of data per unit of time, two main parts of the architecture were upgraded: the low level trigger (LLT) and the HLT.

The LLT is essentially the L0 trigger of the original implementation, but runs in a free-running and trigger-less operation, selecting events according to the detector clustering information. This is possible because of a new readout system based on last generation FPGAs, capable of pre-processing and transmitting data with multi-gigabit transceivers. For the HLT there was a main change, the event building process is now made in a two-step approach, with a HLT-1 for a partial reconstruction and a HLT-2 for the final decision [34].

2.4.4 DUNE

The Deep Underground Neutrino Experiment (DUNE) is an experiment for neutrino science and proton decay studies. A beam of neutrinos will be fired from a detector placed 1300 km away at the Long-Baseline Neutrino Facility, Fermilab [35].

DUNE will pursue three major science goals: find out whether neutrinos could be the reason the universe is made of matter; look for subatomic phenomena that could help realize Einstein's dream of the unification of forces; and watch for neutrinos emerging from an exploding star, perhaps witnessing the birth of a neutron star or a black hole [36].

Before the final detectors commissioning, two prototype detectors were installed as an extension of CERN's North Area Beam Test Facility. Unlike DUNE, the prototypes are situated at the surface, and their dominant signal source is from cosmic rays rather than the beam delivered by the CERN Super Proton Synchrotron (SPS) [35]. The final
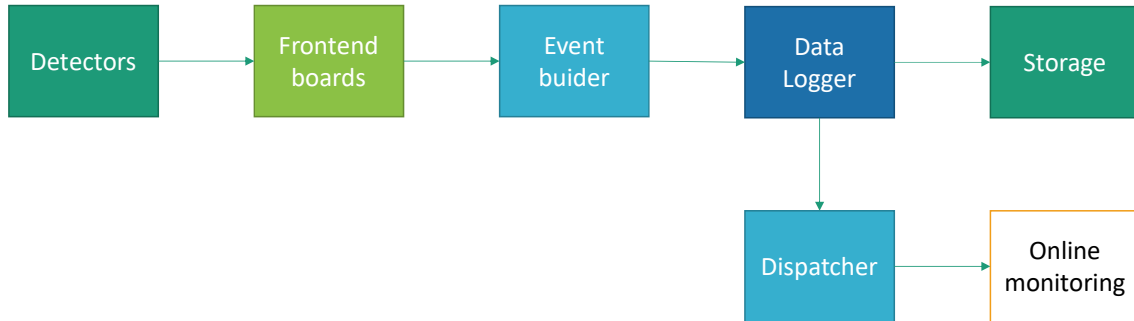
Figure 2.14: artDAQ data flow in DUNE hardware.

experiment is projected to be finished and ready to be fully operational in 2027 [37].

Because of the nature of the physics and detectors, the DAQ of the experiment will work in a trigger-less mode. Instead of maintaining a synchronous scheme for the data system for first-level triggering, it is proposed to tag the data at the very beginning of the data acquisition chain, in the detector's frontends with a timestamp, and process the data flow in the HLT stage. This scheme takes profits of the new technologies developed with the last generation of SoC-FPGAs composed of CPU + GPU + FPGA in a single chip, making it possible to implement complex algorithms embedded in CPU + GPU software for triggering while using the FPGA resources for L1 and data taking, in between other possibilities.

One of the DAQ proposals for the software for the data transfer, event building, run control, and event analysis is the artDAQ [38] developed by the Fermilab. The data flow of the artDAQ is shown in Figure 2.14.

2.4.5 COMPASS/AMBER

The Common Muon and Proton Apparatus for Structure and Spectroscopy (COMPASS) is located at the Super Proton Synchrotron (SPS) of CERN. The main objective of the experiment is the study of hadron structure and spectroscopy using high intensity muons and hadron beams [10].
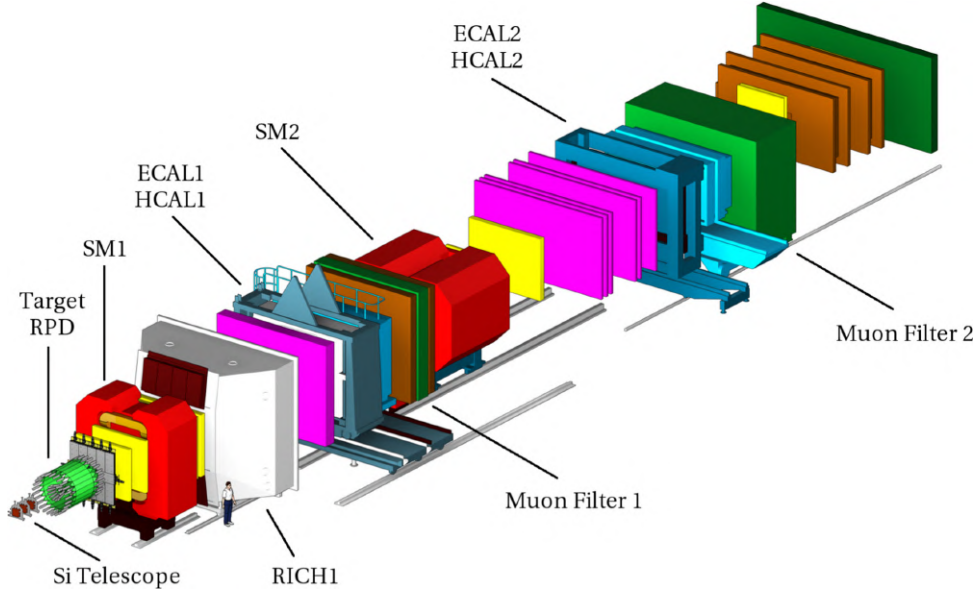
Figure 2.15: COMPASS spectrometer layout, from [39].

The layout of COMPASS is shown in Figure 2.15. It is organized as a 50 m longitudinal spectrometer, with a fixed target, and two other main parts, one for studying of the large angle scatterings and the other for the small angles scatterings.

Starting in 2023, COMPASS is formally called Apparatus for Meson and Baryon Experimental Research (AMBER) framed in the context of the Physics Beyond Colliders initiative. The first experiment projected for AMBER is the elastic muon-proton scattering process, using high-energy muons, to study the long-standing puzzle of the proton charge radius [40].

This thesis has been performed in the framework of the data acquisition system of COMPASS/AMBER. Therefore, this experiment and its DAQ are explained with more details in Chapters 3 and 4.

Chapter 3

The COMPASS Experiment

The **CO**mmon **M**uon and **P**roton **A**pparatus for **S**tructure and **S**pectroscopy (COMPASS) [10] is a fixed-target experiment situated at the M2 beamline of the Super Proton Synchrotron (SPS), at the North Area of the CERN. Figure 3.1a illustrates the CERN's accelerator complex and Figure 3.1b shows the SPS North Area beam facilities where COMPASS is located.

As shown in Figure 3.1b, the SPS protons beam coming from the left is forwarded to the T6 target from where the M2 beamline starts. The M2 was originally built as a high-energy, high-intensity muon beam with momenta of 100, 160, and 190 GeV/c produced from the decay of pions. The pions are originated from the hit of the SPS proton beam over a 500 mm Beryllium target (T6). For the COMPASS experiment, it was partially rebuilt to include a high-intensity hadron beam option, as well as the possibility of using low-intensity electron beams. When operating with a muon beam, the SPS works with a typical super-cycle composed of two 4.8 s bursts or spill structures as shown in Figure 3.2. Before being able to deliver spills, the SPS must be filled with proton bunches, accelerated to a certain energy, and then de-bounched after all particles circulate homogeneously in the accelerator [43]. The structure starts with the first proton injection, followed by a second injection in the SPS, after which it starts the extraction (beginning
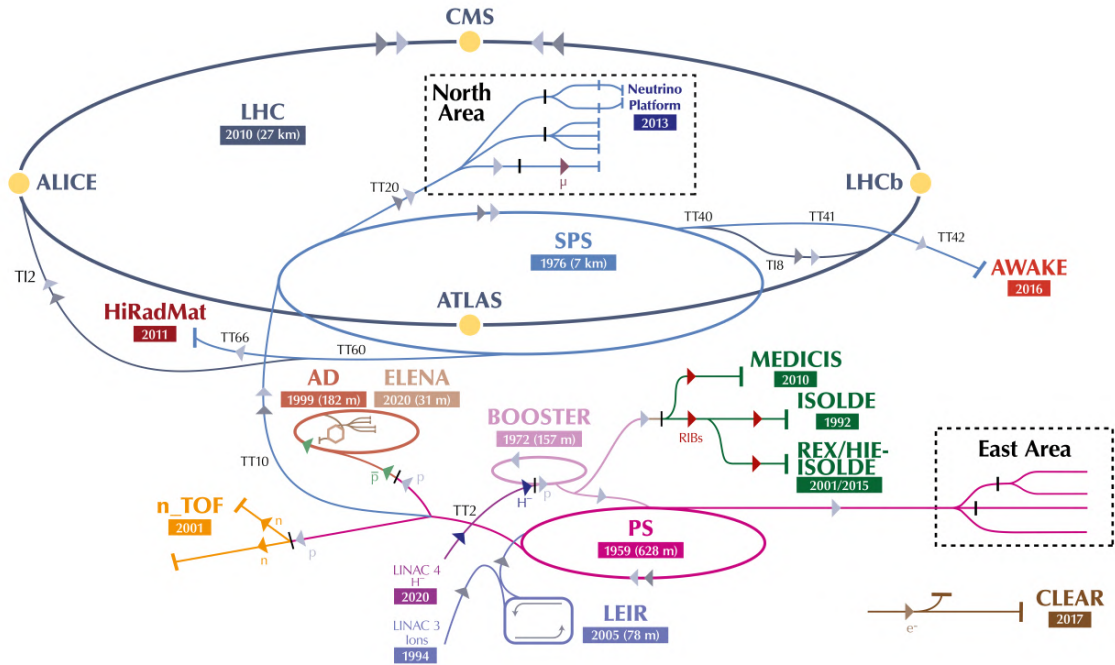
(a) 2022 CERN's accelerator complex, image from [41].
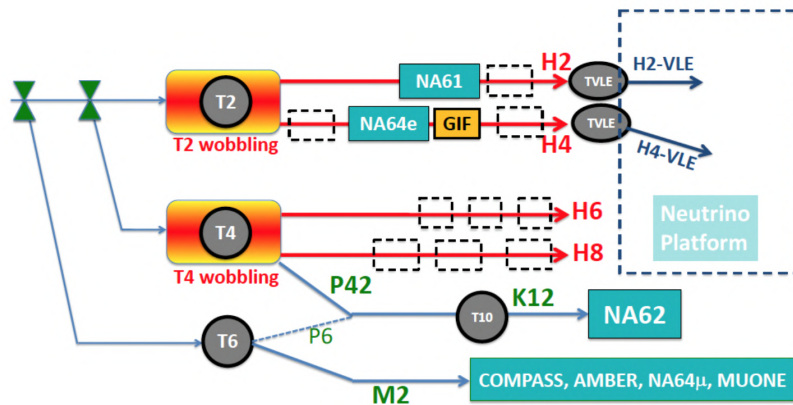
(b) SPS North Area beam facilities where COMPASS is located, image from [42].

Figure 3.1: CERN's accelerator complex and North Area beam facilities.
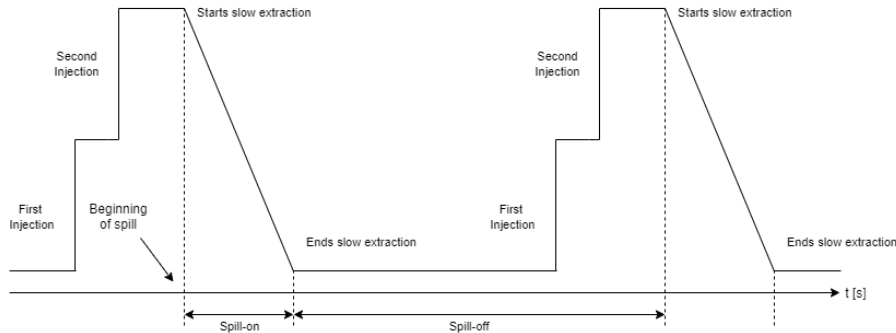
Figure 3.2: Timing of SPS supercycle beam intensity for COMPASS.

of the spill) once the particles are homogeneously distributed in the SPS ring.

The COMPASS scientific program was proposed in 1996 and approved by the CERN Research Committee in 1997. The main objective was to study gluon and quark structures and hadron spectroscopy using high-intensity muon and hadron beams [10]. The experiment was set up between 1999 and 2001, and finally, in 2001 the first commissioning run was performed.

The first phase of COMPASS (2002-2011 data taking period) aimed to study in detail how nucleons and other hadrons are made up of quarks and gluons. In this phase, by evaluating the spin structure with longitudinally polarized ^{6}LiD and NH$_3$ targets for semi-inclusive deep inelastic scattering (DIS), the nucleon spin structure, gluon polarization in nucleons, quark flavor decomposition of the nucleon spin, transverse spin, light meson spectroscopy, and baryon spectroscopy were studied.

A second phase started in 2012 to study the transverse and 3D structure of nucleons using Deeply Virtual Compton Scattering (DVCS), Hard Exclusive Meson Production (HEMP), Semi-inclusive DIS (SIDIS), and polarised Drell-Yan (DY) reactions [39].

Until the start of the LHC experiments, COMPASS was the experiment with the largest data-acquisition system at CERN. It is also a pioneer in adopting new detector technologies, such as MicroMegas, GEM detectors, and most recently THGEM photon detection, as well as in the development of readout systems based on FPGA.

36

The two-staged spectrometer of COMPASS is composed of numerous tracking detectors, particle identification, and calorimetry. To save the data produced by all detectors, a complex and sophisticated data acquisition system (DAQ) is required, which is an essential part of the experiment.

In the following sections, the spectrometer is presented, but the main focus is on an electromagnetic calorimeter called ECAL2, the detector on which the work of this thesis is based.

3.1 Spectrometer

The spectrometer is built as a two stages of similar sub-spectrometers, covering a wide range of kinematics while simultaneously exhibiting a large angular acceptance. The first stage is the Large Angle Spectrometer (LAS) and the second is the Short Angle Spectrometer (SAS).

The detectors are disposed over a 50 meters path in a longitudinal manner, and the exact disposition depends on the scientific program. The layout of the 2015 COMPASS spectrometer is shown in Figure 3.3.

Each of the stages is designed to perform particle identification, tracking, and energy measurement, and are equipped, individually, with a magnet and tracking detectors for charge and momentum measurements. In particular, the LAS is harnessed with a Ring Imaging Cherenkov Counter (RICH) and muon filters for particle identification. For photons energy measurement, undetectable by the tracking detectors, two electromagnetic calorimeters are used: ECAL1 for LAS and ECAL2 for SAS. The complete energy measurement is performed with two hadronic calorimeters, the HCAL1 and HCAL2, placed after each of the electromagnetic calorimeters, respectively. The so called muon filters are used for muon identification, and are filters for every particle except muons. Particle tracking includes gas electron multipliers (GEM) [44], scintillating fibers (SciFi) [45], and drift chambers (DC) [46].
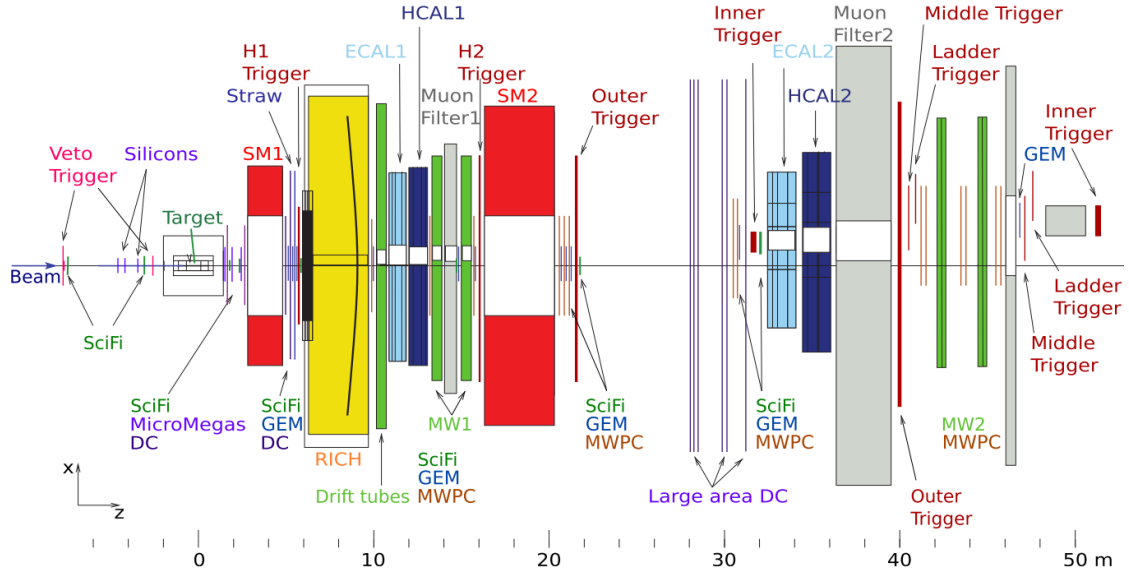
Figure 3.3: COMPASS spectrometer layout for hadron beam - 2015, from [47].

3.1.1 Tracking Detectors

COMPASS has more than 300 active planes for particle tracking purposes [48]. Diverse technologies of tracking detectors are used, covering specific requirements depending on the spatial coverage needs, particle flux, and region to cover. One of the most used criteria for classifying these detectors is according to their active area size. They are called very small, small, and large area trackers.

The class of very small-area trackers (VSAT) are detectors that can be placed innermost in the beam, covering a radial distance up to 2.5/3 cm. The detectors used are scintillating fibers (SciFi) and silicon micro strip detectors (SI) [49]. SciFi stations are used in the upstream of the target to track the incoming muon. The SI stations are also placed upstream of the target for precise measurement of the beam direction. Pixelized micro-mesh gaseous detectors (PMM) [50] and pixelized gaseous electron multipliers (PGEM) [51] are also used when high spatial resolution is required.

The small-area trackers (SAT) cover an area from 3 cm to 40 cm away from the beam.

38

They offer a good compromise between spatial and time resolutions against covered area. The detectors used for this are the micro-mesh gaseous detectors (Micromegas/MM) [52] and gaseous electron multipliers (GEM).

The large-area trackers (LAT) cover active areas of several m^2. For this purpose, the detectors belong to the gaseous-filled class, and are the drift chambers (DC) and multi-wire proportional chambers (MWPC) [53].

3.1.2 Magnets

The two dipole magnets define the stages of the spectrometer. The SM1 is the magnet used for measuring momentum of the tracks in the LAS region. Its average integrated magnetic field is 1 Tm and the bending is done in the horizontal plane. The second dipole magnet SM2 defines the SAS, it has a higher bending power than SM1 with a field integral of 4.4 Tm, hence allowing more accurate determination of high momentum tracks [54]. The main characteristics of the spectrometer magnets are shown in Table 3.1.

Table 3.1: Main characteristics of the spectrometer magnets.

	LAS (SM1)	SAS (SM2)
Aperture	2.0 x 1.6 m^2	2.0 x 1.0 m^2
Field integral	1 Tm	4.4 Tm
Angular acceptance	$\theta >$30 mrad	$\theta <$30 mrad
Momentum range	p <60 GeV/c	p >10 GeV/c

3.1.3 RICH

A RICH detector located at the LAS separates outgoing hadrons into pions, kaons, and protons. The detector is filled with C_4F_{10}. Particles passing through the detector's volume with a velocity greater than the speed of light in the gas produce Cherenkov radiation at an angle θ_C. The photons are collected through spherical mirrors, and when
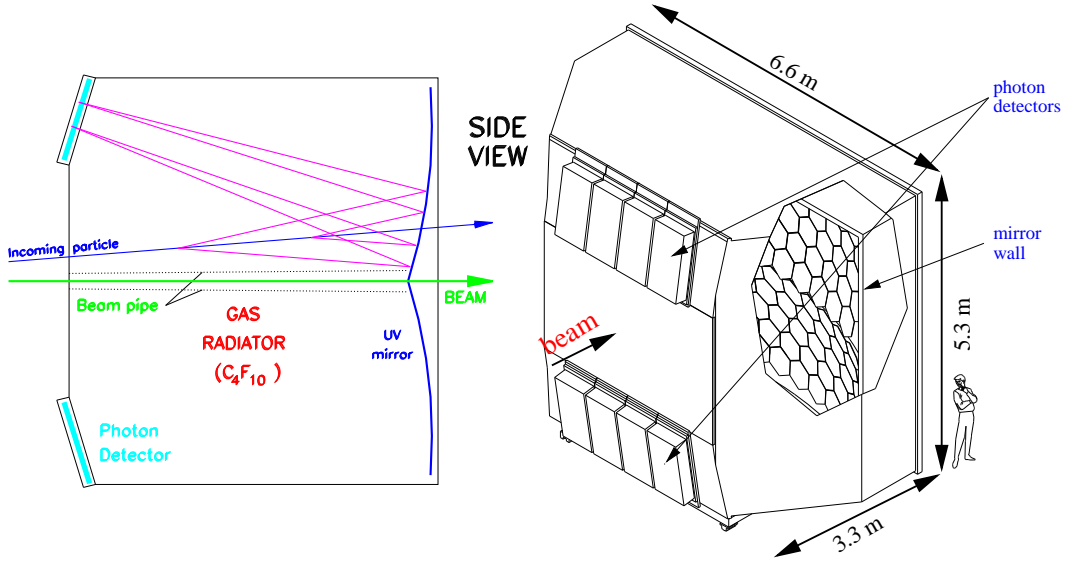
Figure 3.4: Artistic view of the RICH detector, based on [10].

they arrive at the same angle, a circle of the same radius is collected at the mirror's focus. The radius is correlated with the velocity, thus the momenta of the particle [10]. Figure 3.4 shows an artistic view of the COMPASS RICH detector.

3.1.4 Calorimetry

The COMPASS experiment is equipped with electromagnetic and hadronic calorimeters in both the stages of the spectrometer [39]. The calorimeters are used for measuring the energy of particles. The electromagnetic calorimeters (ECAL) measure the energy of electrons, positrons, and photons, and the hadronic calorimeters (HCAL) measure the energy of protons, kaons, and pions. In each stage the electromagnetic calorimeters are located upstream of the hadronic calorimeters. The calorimeters in the LAS are called ECAL1 and HCAL1, and in the SAS are called ECAL2 and HCAL2, respectively. Additional information regarding the electromagnetic calorimeters are presented in Chapter 4.

3.1.5 Trigger System

The spectrometer deals with a high flux of muons of approximately $2 \cdot 10^8$ per 4.8 s. It is not feasible to save information about all muon-proton interactions because the data rate produced in all detectors would be too high [55]. Instead, all frontend cards buffer the data until a decision on whether to save the data arrives. This decision signal is generated by the trigger system and relies on the fast detection of scattered muons to select interesting interactions.

3.2 COMPASS DAQ

The DAQ of COMPASS has been designed to withstand high trigger rates and large data flows and has been evolving since the start of the experiment [10].

It consists of several layers with a well-organized architecture using a pipelined approach. In the first layer, detectors and frontend electronics digitize and buffer the data continuously. The digitizers are mainly detector-specific and encapsulate the data in a predefined format according to the current data protocol. The overall architecture of COMPASS DAQ can be seen on Figure 3.5.

There are approximately 300000 channels in total. The data are assembled by concentrator modules called CATCH (COMPASS Accumulate, Transfer, and Control Hardware) [56], HGeSiCA (GEM and Silicon Control and Acquisition) [57], and GANDALF [58] based on Versa Module Europa IEEE 1014-1987 (VME) standard. These modules also receive the timing and trigger signals and perform the triggering to the frontends.

Data from 16 frontend cards are assembled on each of these boards and transmitted via S-LINK [59] to a readout buffer (Spill Buffer). This second layer was first performed on 30 readout buffer (ROB) computers to temporary buffer the data, and then migrated to custom FPGA solutions. This layer is responsible for the data buffering and subevent
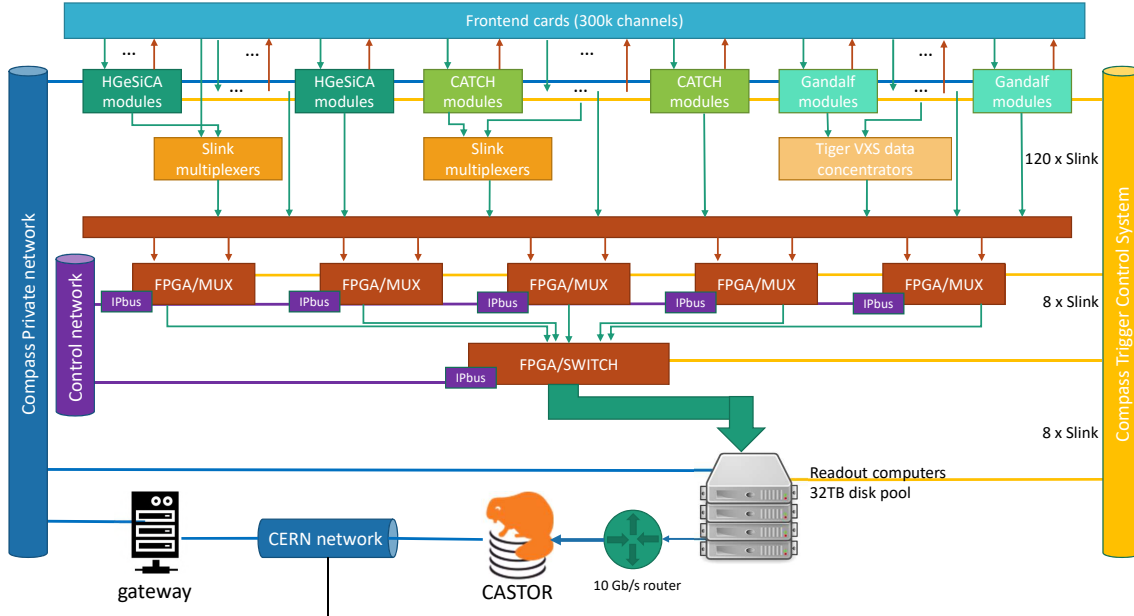
Figure 3.5: Schema of COMPASS DAQ and its main parts.

reconstruction.

Finally, in the third layer, subevents are sent to event builders computers that assembly the full events. The assembled events are stored temporarily on the event builder's local disks before being transferred to the CERN Advanced STORage manager (CASTOR) [60].

From the trigger prospective, the COMPASS DAQ was from the beginning organized as a two trigger levels hierarchy as shown in Figure 3.6a, where some specific detectors were used for building the L1 trigger in a separate system. The L1 trigger operates directly over the frontend electronics and data concentrators (GeSiCA and CATCH). From the concentrators the data are transferred to the event builder services, originally running in PCs but since the last implementation done entirely in FPGAs (FPGA/MUX-FPGA/SWITCH), and from there to the HLT processing farm, running in PCs. Finally, the data are stored in the CERN data storage services.

A new DAQ and Triggering scheme has been proposed for AMBER. Similar to the

(a) Triggered DAQ.

(b) Trigger-less DAQ.
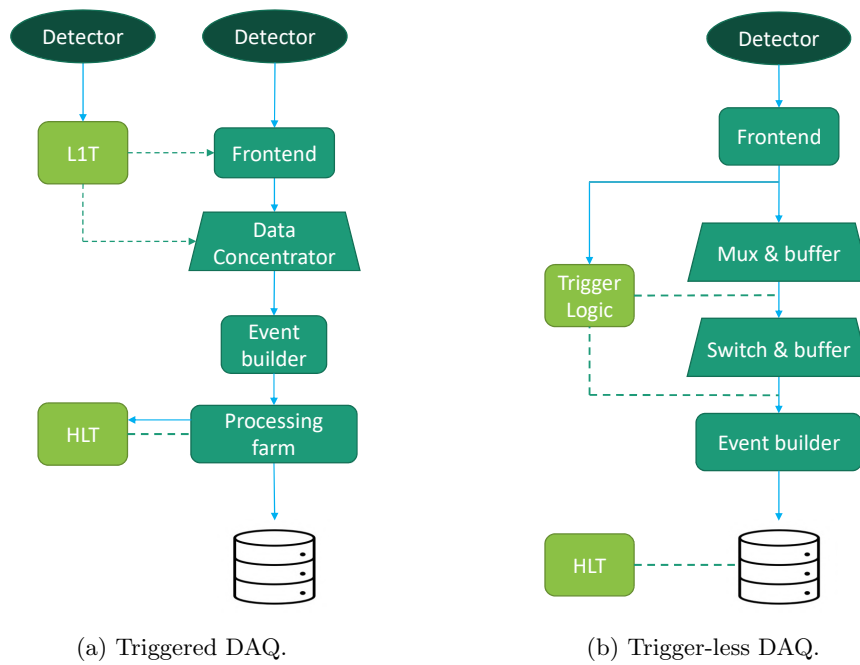
Figure 3.6: COMPASS/AMBER DAQ evolution.

DUNE proposal, the AMBER DAQ will work as a free-running system with two different operation modes. The first mode is a completely non-triggered data acquisition, which saves everything for further analysis. The second mode, shown in Figure 3.6b, includes a digital multilevel trigger, where L1 is built with the raw data from the detectors and several trigger levels implemented directly in FPGAs, and a last HLT operates directly over the stored data, before sending the data for permanent storage at CERN facilities.

3.3 AMBER – New Setup and Requirements

Originally, AMBER was proposed as a continuation of the COMPASS project [61], partially changing the scope of the original experiment, but using the same facilities and most parts of the instrumentation of COMPASS [10, 47]. Figure 3.7 shows a scheme of the last COMPASS setup used during the run 2022. Using this setup as a base, and depending on the physics program, some specific instrumentation needs to be upgraded or included. The upgrades in the instrumentation for Phase-1, to measure the electric mean square charge radius of the proton are (a) the DAQ/trigger system, (b) a high-pressure active time projection chamber (TPC) target, (c) the SciFi trigger system on the scattered muon, and (d) the silicon trackers on straight tracks.

A hydrogen-filled TPC acts as an active target [63] for muon-proton scattering and two silicon pixel detectors for the precise tracking of scattered muons. Inner tracking detectors (SciFi detectors and GEMs) will be used from the COMPASS spectrometer, and the bending magnet SM2 to measure the momentum of the scattered muons. The second muon filter will also be used to separate muons from secondary particles, and ECAL2 to detect photons created in radiated events.

For the proton radius measurement (PRM), the TPC was installed and filled with high-purity H_2 at a pressure level of some atm. The approach to be realized is the elastic scattering of muons at low momentum transfer, using a muon beam with an energy in the range of 100 GeV.

(a) COMPASS setup, upstream half.
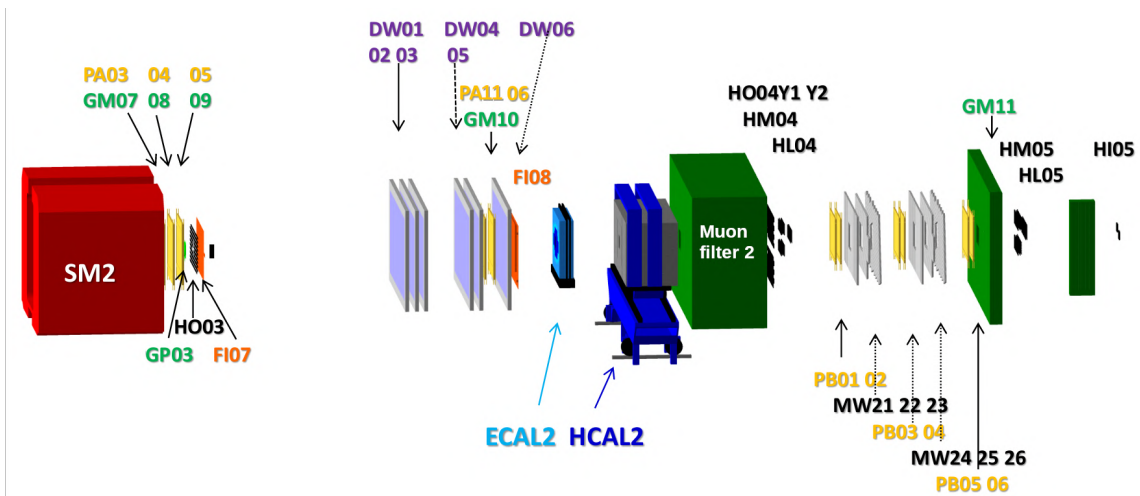
(b) COMPASS setup, downstream half.

Figure 3.7: COMPASS setup for year 2022, from [62].

(a) TPC location in beamline.
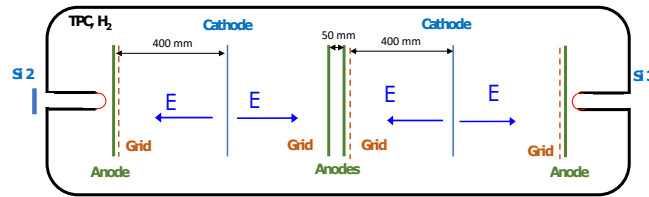
(b) Schema of the TPC.

Figure 3.8: TPC used as active target for the Proton Radius Measurement.

To measure the elastic muon-proton scattering with high resolution and efficiency at such a high beam energy, it is necessary to measure the muon scattering kinematics and recoiling proton. The technique employed is an active target, where hydrogen gas is at the same time the target and detector [64]. The ionization produced in the gas is recorded by operating the target as a TPC with appropriate high-voltage field and readout plates, as depicted in Figure 3.8.

The DAQ/trigger system operates in free-running mode. The rapid development of technology enables the possibility of migrating from classical triggered data acquisition to a scheme in which the detector subsystems provide continuous time-stamped data streams for real-time processing into the upper parts of the DAQ hierarchy. The evolution of COMPASS DAQ, starting from the very beginning, led to an architecture where most traditional computers were substituted with FPGA. This architecture, called intelligent FPGA-based Data Acquisition System (iFDAQ), was introduced for the first time in 2014. Since then, there have been successive improvements until having a very stable operation [65] and a fast recovery time after a crash. The next step in the iFDAQ is to adapt the

operative status for continuous data acquisition using a digital trigger system.

For this approach all the detectors must work in a free-running and trigger-less mode, in which the frontend must be able to digitize the analog signals and stream the data in a dead-time free mode to the DAQ [66]. Once in the DAQ domain, the data are split into two streams: one for buffering and the other for processing. Buffering is performed until a digital trigger condition is satisfied. The data are then multiplexed and reordered, depending on the channel latency, to reconstruct the entire event using the data of all detectors. The stream for processing continuously analyzes the data of the detectors involved in the trigger decision, when that condition is met, it indicates the buffering resources to release the data of the event and send them to storage in disks.

Besides the one mentioned in the previous paragraph, AMBER DAQ is prepared to work in two different regimes. One is pure untriggered data taking, where the data of all detectors are stored in disks for further analysis. The data can be pre-processed in the frontend electronics, hence raw data from the detectors or with some filtering in between can be transmitted to the DAQ. The other regime is enabled by a multi-level digital trigger. In this mode, the data and information from different detectors are sent to a trigger processor to analyze and determine the event occurring at multiple levels and with different granularity of information.

All frontends are synchronized with a global clock of 40 MHz, distributed by a timing control system (TCS). The TCS is also responsible for delivering framing information for data transfer inside the DAQ. The data are sent in frames or slices commanded by the TCS. The time slice is the coarse time at which the detectors should fit their data, as shown in Figure 3.9. Therefore, as each type of detector has its own characteristic time response, the imaging or sub-slicing depends on its technology and purpose. For the detectors in charge of tracking particles, the image time will be very fast, whereas for those in charge of measuring the momentum and energy, the time will be longer. The length of the image depends on the time resolution of each detector subsystem and hit rate. Thus, each event
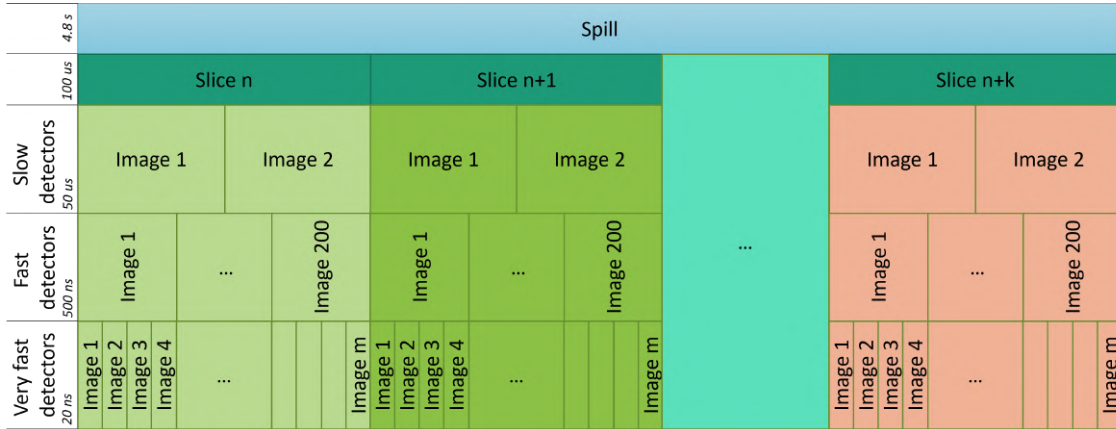
Figure 3.9: Time-slicing and imaging inside the on-spill time.

corresponds to the data from each and every detector combined in a single bunch, before being saved to the disk.

The new trigger system, as shown in Figure 3.10, can be represented as a two-stage system. The event generator, which receives a continuous data stream from all detectors engaging the trigger decision, and builds events of interest (EOI) conforming to the timing correlation of the hits. The other stage is the trigger processor, which can be a single or multi-level system, in charge of making the trigger decision in a hierarchical way and with the possibility of making fast decisions by groups of detectors. After a trigger decision is made, the information of the image in which it occurred is sent to the TCS. Through it, the images are distributed to the concentrator and buffering modules to assemble the entire event with data from all participant readout detectors. Both, the event generator and trigger processor are implemented in FPGA units, making the system the most versatile and adaptable to the different physics programs.
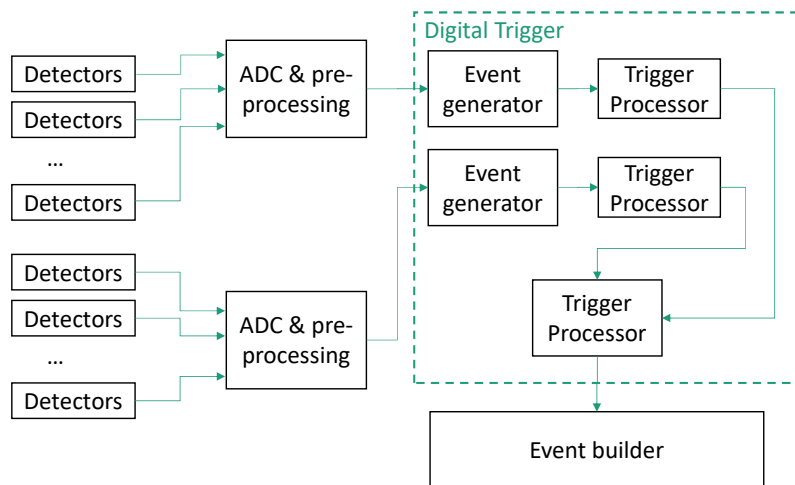
Figure 3.10: New digital trigger system.

Chapter 4

ECAL2 Readout

As mentioned previously, most part of the work of this thesis is focused on the hardware upgrade of the COMPASS ECAL2 electromagnetic calorimeter, in the framework of the development of a new trigger-less DAQ architecture for COMPASS/AMBER experiment.

The ECAL2 electromagnetic calorimeter is part of the small angle spectrometer (SAS) of COMPASS. It consists of 3068 calorimeter elements of three different types, depending on their radiation hardness, all of which have the same transverse dimensions $(3.82 \times 3.83 \text{ cm}^2)$. The whole ECAL2 wall has a dimension of $2.44 \times 1.83 \text{ m}^2$, covering angular ranges between 1.3 mrad and 39 mrad in the horizontal plane and between 1.3 mrad and 29 mrad in the vertical plane.

To clearly understand the complete chain in the readout system, first the basics of calorimetry and the details on the detector structure are introduced. After this, the current electronics for the frontend and digitizers is presented and lastly the free-running proposal is explained, including some of the most relevant details on the required upgrade.

4.1 Calorimetry in HEP

In the context of particle physics, calorimetry refers to the process in which particles to be measured are completely absorbed in a material and their energy transformed to a measurable quantity [16]. A calorimeter is built by a material where the interaction of an incident particle produces a shower of secondary particles with a progressive decrease on its energy until it is completely absorbed. The sum of all elementary losses builds up the calorimeter signal, which can be of ionization or scintillation nature. Due to the scope of this work, in this chapter is addressed only the high energy calorimetric processes, which produces showering; for low energy the calorimetry does not produce showers.

Although the calorimetry phenomena can be mainly generated both by electromagnetic and strong interactions, due to the nature of ECAL2 only the former type of interactions are presented here.

In addition to the energy measurement, other quantities and features can be extracted from a calorimeter, such as timing, impact position, particle direction, and identification.

4.2 Electromagnetic Calorimeter Showers

When a high-energy electron, positron, or photon hits an absorber material, it initiates an electromagnetic cascade producing independently secondary photons by bremsstrahlung, or secondary electrons and positrons by pair production. These particles give rise to others until they eventually fall below the critical energy and dissipate their energy by ionization and excitation.

Photons propagate deeper into the material and are finally absorbed by the photoelectric process.

Figure 4.1 shows a scheme of the electromagnetic shower after a high energy photon (γ) impinges an absorber material, producing first an electron/positron pair, and then
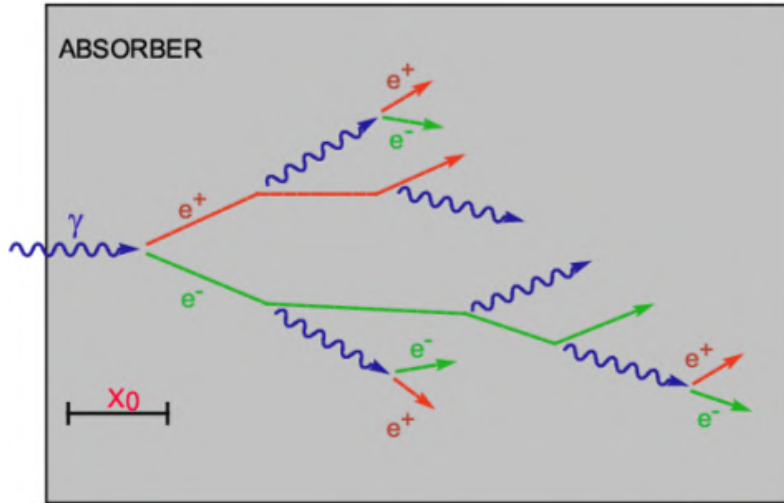
Figure 4.1: Electromagnetic shower schema, from [67].

starting with the cascade.

The energy lost by electrons and photon interaction with a material can be divided into two regimes depending on the energy carried by the particle and in the material.

In the case of lead and for energies lower than 10 MeV, electrons [68] lose their energy mainly due to collisions with the atoms and the molecular structure of the material, giving rise to ionization and thermal excitation. Photons lose their energy through Compton scattering [69] and photoelectric effect.

For the same material and energies larger than 10 MeV, the electron energy loss is mainly caused by bremsstrahlung and photon interactions will produce electron-positron pairs.

In Figure 4.2 there is the Feynman diagram of the interaction of an electron with a nucleus emitting one photon. When an electron or a positron passes near the electric field of a nucleus, it slows down and gets deflected, a fraction of its energy is directly converted into photons. The Feynman diagram shows the interaction of an electron with an initial energy and momentum (top left) interacting with a nucleus (A: weight of material,
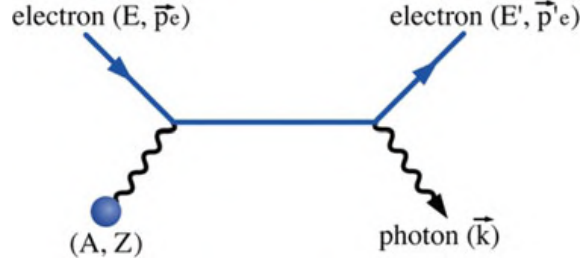
Figure 4.2: Radiation of a photon caused by bremsstrahlung effect, from [15].

Z: atomic number) producing as a result the electron with another energy (lower) and momentum state (lower speed) plus the radiation of a photon, thus satisfying the law of conservation of energy [15].

These secondary particles, in turn, produce other particles by the same mechanism, thus giving rise to a cascade (shower) of particles with progressively degraded energies. The number of particles in the shower increases until the energy of the electron component falls below the critical energy (ϵ).

An important parameter of electromagnetic showers is the radiation length (X_0), which represents the average distance (x) that an electron needs to travel into an absorber material to reduce its energy to $1/e$ of its initial state. It governs the rate at which electrons lose energy by bremsstrahlung, as follows:

$$\langle E(x) \rangle = E_0 e^{-x/X_0} \tag{4.1}$$

Similarly, a photon beam of initial intensity (I_0) traversing a block of material is absorbed primarily through pair production. After traveling a distance $x = 9/7X_0$, the intensity is reduced to $1/e$ of the original, as follows:

$$\langle I(x) \rangle = I_0 e^{-(7/9)x/X_0} \tag{4.2}$$

Figure 4.3 shows the Feynman diagram of the electron-positron creation when a photon with an energy greater than a few times $m_e c^2$ impinges to the scintillating material,

53
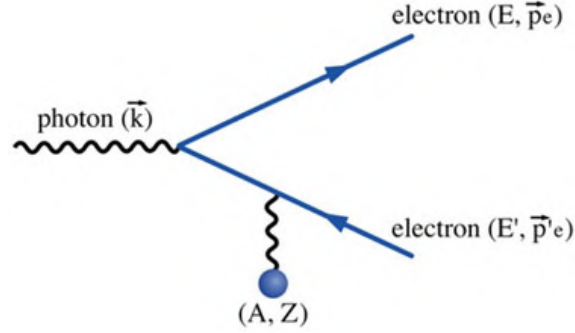
Figure 4.3: Electron-positron pair created by a photon under the electromagnetic field of a nucleus (A,Z), from [15].

where m_e is the mass of an electron and c the speed of light in vacuum. In the Figure 4.3, the positron is represented by the electron E' with opposite direction to the electron E.

4.2.1 Energy Measurement

The critical energy (ϵ) of a particle can be defined as the threshold energy at which the ionization loss per X_0 equals the electron energy E [70]. The definition is equivalent for bremsstrahlung.

From Equation 4.1:

$$\frac{dE}{dx} = \frac{E}{X_0} \qquad (4.3)$$

Then, the total track length (T_0) of the shower can be defined as the sum of all ionization tracks due to all charged particles in the cascade and is proportional to:

$$T_0 \propto X_0 \frac{X_0}{\epsilon} \qquad (4.4)$$

Thus, the term $\frac{X_0}{\epsilon}$ expresses proportionality to the total number of particles in the shower. Measuring the signal produced by the charged tracks of the cascade provides a measurement of the original particle energy (E_0).

The measurement in a scintillating material can be performed by detecting the light produced, or in a gas or liquid by collecting the produced charge. The calorimeters are then divided in two different types, according to with the energy they are designed to measure and the composition of the materials they are build up with.

4.2.2 Homogeneus Calorimeters

In these detectors, all the energy of a particle is deposited in the active material, which provides excellent energy-resolution properties. However, these calorimeters are less easily segmented laterally and longitudinally, which presents a drawback when position measurements and particle identification are required. Homogeneous calorimeters can be divided into four categories: Semiconductor, Cherenkov, Scintillator, and Noble-Liquid.

In detectors where the signal is collected in the form of light (Cherenkov and Scintillators), photons from the active volume are converted into electrons (usually called photoelectrons) by a photosensitive device such as a photomultiplier tube.

4.2.3 Sampling Calorimeters

Sampling calorimeters are usually built as a layered structure, in which one of the layers corresponds to the active material and the other to a layer to absorb part of the energy and to contain the shower in a limited space. Thus, they offer good spatial resolution and are suitable for particle identification. Nevertheless, owing to the sampling fluctuations produced by the absorber layers interleaved with the active layers, they often have a lower energy resolution than homogeneous calorimeters. In the case of scintillating sampling calorimeters, the photons are collected from the active material by mean of fibers that will carry the light to a device like in the case of homogeneous calorimeters.
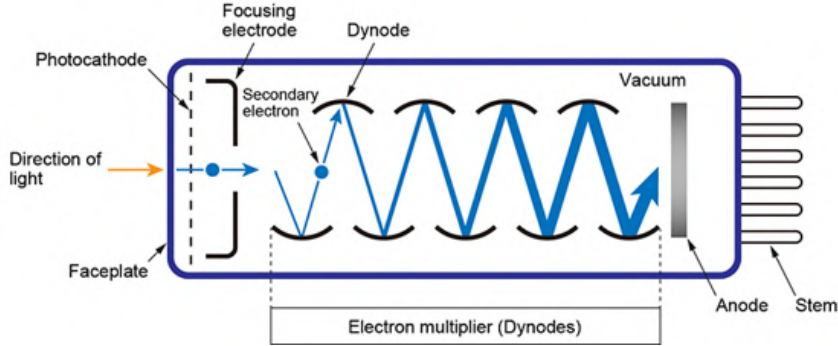
Figure 4.4: Schematic view of a photomultiplier tube, from [71].

4.2.4 Photomultiplier tube

A PMT is a light-sensitive detector wich multiplies the electric charges produced by incident light and converting it into a measurable current signal. The multiplication is done in multiple dynode stages. The sensitivity of these detectors is ideal for working with electromagnetic calorimeters.

In Figure 4.4 the schematic view of a PMT working principle (Hamamatsu 2021 [72]) is shown. The incident light ejects an electron in the photo-cathode as a consequence of the photoelectric effect. These electrons are directed to a dynode (electron multiplier), producing secondary electrons. A high voltage electrical field is applied in each dynode stage to guide and multiply the number of electrons in each stage. The electron cascade is collected by the anode producing a current that the electronics can read.

4.3 ECAL2 Electromagnetic Calorimeter

The ECAL2 modules are arranged in a 64×48 matrix, as shown in Figure 4.5. For data taking with hadron beams, the central hole with respect to the nominal beam directions was set to 2×2 modules. The outermost part of ECAL2 was equipped with 1332 TF1 lead glass [73] modules (GAMS). The intermediate part is filled with
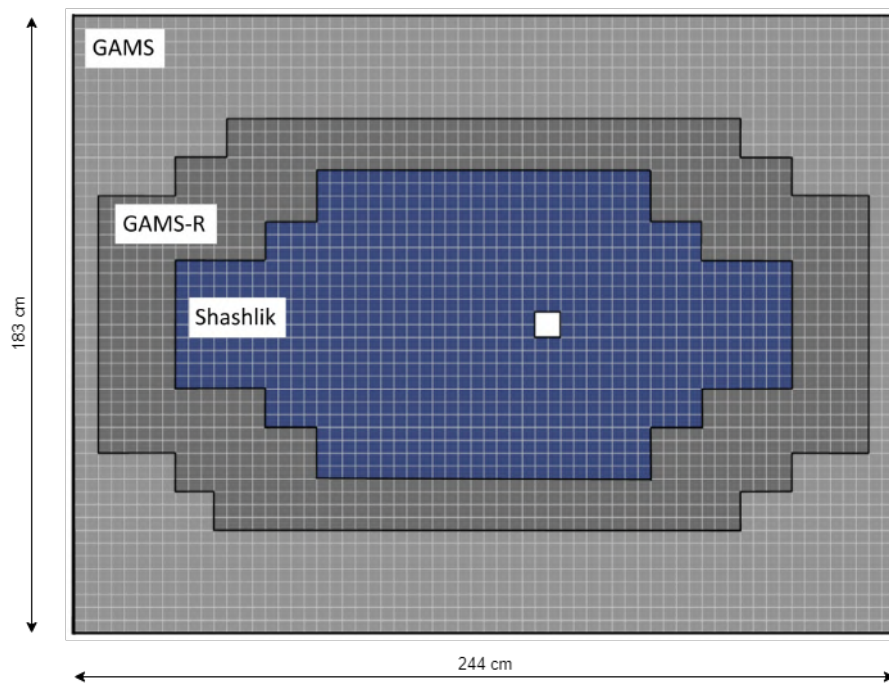
Figure 4.5: ECAL2 different scintillating technologies.
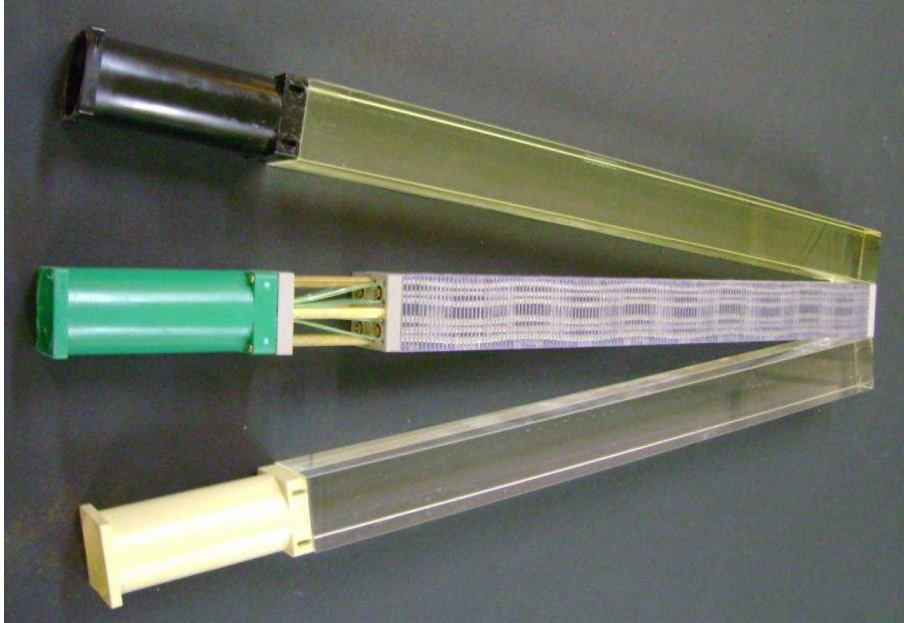
Figure 4.6: The 3 types of lead glass blocks: radiation hard GAMS-R (top), shashlik (middle) and GAMS (bottom), based on [78].

848 radiation-hardened modules (GAMS-R) made of TF101 material [74]. These two scintillators are of the type of the homogeneous calorimeters. The innermost part is equipped with 888 sampling calorimeters modules of Shashlik type [75]. The elements are composed of 154 double layers, each consisting of a 0.8 mm thick lead plate and a 1.55 mm thick scintillator plate. The photons from the Shashlik modules are collected by 16 wavelength-shifting light fibers and guided onto FEU84-3 photomultipliers [76, 77]. Figure 4.6 shows the three different types of scintillating elements used in the ECAL2.

In particular the Shashlik module is a 47 x 47 x 440 mm lead/scintillator sandwich, made of perforated lead and plastic scintillator plates. From the Figure 4.6, it can be seen the Shashlik composed of a layered structure with the fibres for collecting the light in the middle of the image.

Each scintillator block is viewed at one end by a photomultiplier tube (PMT),

converting light emitted during scintillation into an electrical signal and amplifying that signal to a useful level. Each PMT has an individual programmable high-voltage power source for polarization as in [79].

The output of each PMT is connected to an analog reshaping filter with approximately 64 ns of rise time (120 ns FWHM) that feeds the frontend readout electronics explained in Section 4.3.1.

A crucial parameter of ECAL2 is the gain of the photomultipliers, which depends on the polarization voltage. As the emitted Cherenkov light is proportional to the energy of the incoming particle, the gain should be constant for reliable measurements. To monitor and correct fluctuations of the gain, the scintillating elements are directly connected to external light sources. During the off-spill period when no beam occurs, light pulses are sent into all and each element, then the pulse signals are read out via calibration events. Light Emitting Diodes (LEDs) are used for generating the light pulses. The light is distributed by fibers directly to the scintillating elements. If the LED pulses carries a constant light intensity, the resulting amplitude signals must be constant, if not it means that the gain has changed. Figure 4.7 shows a typical calibration response from the COMPASS ECAL2 status interface. It can be seen that the region populated with shashlik elements has a lower response (dark blue) than the others This is because of their lead layered structure, being completely absorbent for LED light after the first scintillating layer. Rectangular clusters can be observed according to the mapping fibers of LEDs used for different regions.

4.3.1   ECAL2 Readout System

The data are digitized with a Mezzanine Sampling Card (MSADC) [80], composed of 32 analog channels with 12 bits resolution, and capable of working up to a sampling frequency of 40 MHz. The analog-to-digital conversion of the MSADC is based on four ADS5271 ADCs [81].
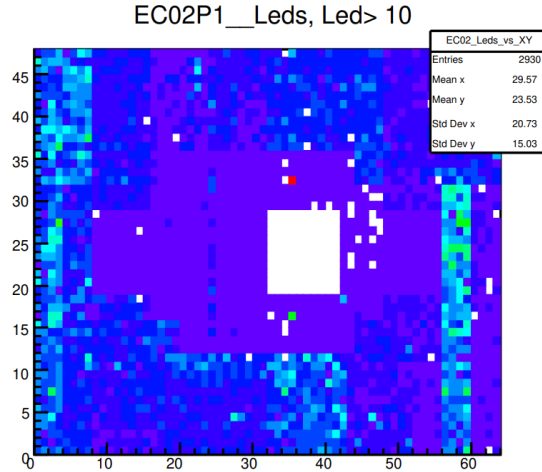
Figure 4.7: LED amplitudes per element for the ECAL2.

Four of these cards are mounted on a VME carrier card. This carrier card provides an USB interface for programming new firmware, a custom VME backplane, and also an interface for accessing to a HOTLink port [82].

The HOTLink interface is connected through an RJ-45 connector and provides the Timing Control System (TCS) information. It is used for slow-control and to interface with the GeSiCA module. Each GeSiCA module connects up to eight carrier cards.

The MSADC is operated as a 1:2 scheme, where two analog channels are connected to one input, sampling in an interleaved mode with a 180° phase shift to cover 16 channels, as if it were double frequency (80 MHz of effective sampling rate). The four MSADC cards are combined into a single carrier card to read 64 channels of the calorimeter. In Figure 4.8 a 9U VME carrier card with the MSADC boards can be seen. The carrier interfaces to 64 analog detector channels and provides power supply and data readout for the MSADC modules.

The carrier has a Virtex-4 FPGA for glue logic and to build the data packets for sending to the DAQ multiplexers and controlling the MSADCs. The same setup is used to read all 3068 ECAL2 channels, using 48 VME carrier cards and 192 MSADCs.
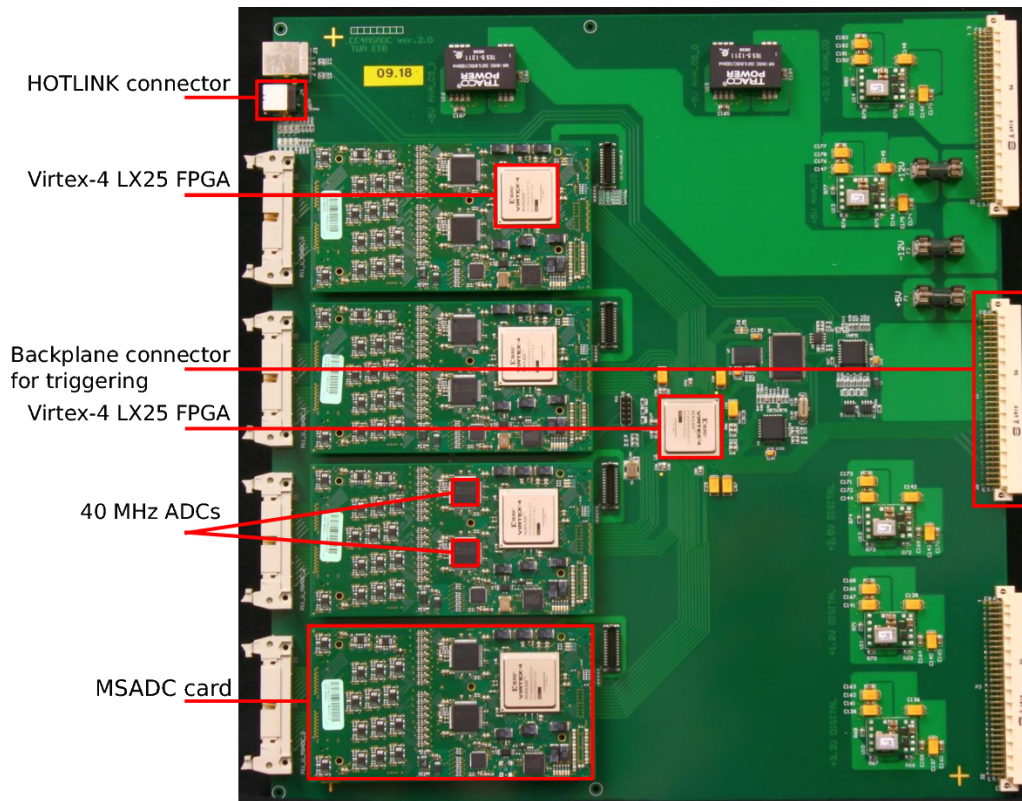
Figure 4.8: VME carrier card with four MSADCs mounted and main parts indicated, image taken from [47].

As the MSADC has four independent ADCs, it is highly probable to have the baselines of the ADCs working with different offset values. These values can be shifted by offsetting the voltage value with a special circuit using a digital-to-analog converter affecting individually each of the ADCs .

The detector connects to the analog input of the MSADC through a 34 position connector in a pseudo differential signalling scheme. Each channel connects with the MSADC to one of the lines with signal and with ground to the other. The inputs have an impedance of 50 Ω and feed a differential amplifier, which acts as a buffer before the ADCs.

Each MSADC board has also a Virtex-4 XC4VLX25 FPGA [83] to control, program, and read ADCs. The digitized data are transmitted to the FPGA over low voltage differential signaling (LVDS) lines with a bitrate of 480 Mbps working at Double Data Rate (DDR).

The FPGA firmware is in charge of making the interleaving, reading the channels at double frequency, and alternating the channels in each clock cycle. Special attention should be paid to the order of the channels after interleaving. When wrong, the shape of the assembled digitized signal will not be correct. Another task of the FPGA is to align the offset of the ADCs at those moments when there is no beam.

Once interleaved and aligned, the data are buffered into a FIFO memory to delay and store the data locally until a trigger condition is met. When a trigger arrives, a trace of 32 samples is copied into a second memory. This memory is read out sequentially by the carrier FPGA for each detector channel and sent to the DAQ.

From the perspective of data transmission to the carrier, each channel is formatted as a 32-bit data word and read out only if it passes a zero-suppression algorithm. This will fix a maximum trigger rate condition, owing to the maximum data transmission limits, as in Equation 4.5:

$$T_{MAXtrig} = N_{samples} \times N_{channels}/f_{sampling} = 32 \times 16 \ /80 \ \text{MHz} = 6.4\mu \ \text{s} \qquad (4.5)$$

For the standard configuration and for each event, a header, start-of-event, and end-of-event words are sent with the data. Each channel adds an additional header and sends half of the samples. As a maximum, there will be a number of bytes per event, as in Equation 4.6:

$$N_{bytes} = 4 \times (3 + (N_{samples}/2 + 1) \times N_{channels}) \tag{4.6}$$

In the standard configuration it is 1100 bytes. The MSADC transmits the data to the carrier at one byte per clock cycle at 40 MHz. This implies a maximum trigger rate of:

40 MHz/1100 = 36.36×10^3/s = 36.3 keps (kilo events per second).

4.4   Free-runnig Proposal

The COMPASS DAQ has been evolving since the beginning of its operation, being one of the first systems based on FPGAs and with outstanding features in terms of the data rate and stability with which the system could handle. The last of the upgrades on the DAQ is the free-running and trigger-less operation. Starting from 2023 in the same facilities there will have place the AMBER experiment. The AMBER proposal for the Proton Radius Measurement (PRM) [66] requires for its first run an operative free-running mode for the DAQ.

The PRM will be split into two data taking periods. The first with low intensity and the second with high intensity.

The rather low beam intensity of the first data taking period will allow the DAQ to work in a completely not-triggered mode and save all data to the disk for further offline analysis.

For the second data taking period, as the beam intensity will be an order of magnitude higher, a digital multilevel trigger is mandatory to reduce the amount of data rate to a level that is feasible to record on disks and with a reasonable offline computation

power. This will be the first time that the continuous iFDAQ and the new digital trigger system will be tested working in a free-running mode without losing any physics data.

For the first data taking period, where there will be no external trigger signalling and considering no data reduction algorithms, each time there is an event in the ECAL2 the DAQ should be able to handle:

$$3068 \text{ Ch } \times \text{ (12 bits } \times \text{ 32 samples)} = 1.178.112 \text{ bits} \tag{4.7}$$

And, considering a maximum event rate of up to 100 keps:

$$1.178.112 \times 100 \; 10^3 = 117.8 \text{ Gbps} \tag{4.8}$$

Thus, for a 4.8 s spill, the ECAL2 can generate up to $565, 5$ Gb of data that should be stored in disks.

Using the current digitizer board, but making an upgrade immediately after this board, we can drastically reduce the amount of data transmitted for storage. This upgrade is one of the main challenges addressed by this thesis work.

By adopting reconfigurable devices with high-processing capabilities, such as Ultrascale+ SoC-FPGAs, data can be reduced by real-time features extraction performed individually for all channels in parallel.

Extracting features, such as amplitude and time arrival, when the hit occurs in only one element allows reducing data transmission from 118 Gbps to 6.4 Mbps for the entire ECAL2 detector, as in Equation 4.9.:

$$(32 + 32) \times 100 \; 10^3 = 6.4 \text{ Mbps} \tag{4.9}$$

where 32 bits is the width of each feature.

To reach this goal, the ECAL2 electronics has been upgraded to work in a trigger-less mode with the development of new hardware needed for processing and evaluating all MSADC channels in real-time, detecting the pulses generated in the calorimeters,

extracting the features, and adding a global timestamp for recognizing the event number inside the DAQ dataflow.

In the following chapters, the requirements for this hardware, the proposal for giving solution to those requirements, and the implementations made at the COMPASS beam facilities are presented.

Chapter 5

The Mezzanine Sampling ADC Card for Trigger-less Operation

The first proposal made by the AMBER collaboration in [66] for the ECAL2, was to use the current frontend digitizer cards, but modifying the firmware in such a way that it processes the data to suppress the noise, detect the pulses, and measure the amplitude and time of arrival when the signal to noise ratio exceeds 12.04dB. For this purpose, a new carrier card for hosting the MSADC and processing the data in real-time was required.

5.1 MSADC Characteristics

The mezzanine sampling analog-to-digital converter board (MSADC) shown in Figure 5.1, is the core of the frontend for the data acquisition used in the ECAL2. The board was developed at the Technical University of Munich in 2007 for the first DAQ upgrade of COMPASS [84].

The printed circuit board (PCB) has a form factor of 70 mm × 130 mm and can be divided into an analog part, which holds the analog-to-digital conversion and signal conditioning, and a digital part with the FPGA, microcontroller, and digital circuitry for

control. The MSADC does not include power sources and must be powered from the outside using the correct power-on sequence. On the left side, up to 32 differential analog channels can be connected, independently buffered, and sampled using high-speed ADCs. The right side of the board has connectors for digital LVDS communication to higher levels in the readout chain, for the input of the clock and control signals, and for power.

Each ADC channel input is preceded by an AD8138 differential operational amplifier (OpAmp) to adapt the input impedance to the ADC and serves as a buffer stage. The OpAmps are powered with ± 5 V, and the input works in rail-to-rail mode, allowing both negative and positive signals to be handled. All 32 channels are independently biased to compensate for the possible offset of the baseline. The bias voltage is controlled by a digital-to-analog converter (DAC) that is accessible via a slow control interface.

Analog-to-digital conversion is performed using a Texas Instruments ADS527x ADC. This family provides eight differential input channels per integrated circuit to cover the 32 channels of the MSADC by using four of these devices. After being digitized with 12 bits resolution, the data of each channel are transmitted in serial LVDS at DDR in a single line. In particular in the MSADC the ADS5270 is used, which offers a maximum sampling frequency of 40 MHz. The channel configuration can be changed such that two ADCs can work in an interleaved mode, digitizing the data of a single analog channel. Thus, the sampling rate can be doubled to 80 MHz with an increase in effort from the firmware perspective. In addition to the four High-Speed ADCs, there is an ADS1258 single-ended ADC for slow control purposes. All analog inputs are mapped to a 120 pin Samtec Q-Strip connector. The analog input connector is on the bottom-left side of the board.

The LVDS interface of the ADCs maintains a low pin count for data transmission, while does not generate high power consumption. Nevertheless, because the frequency sampling can be as high as 40 MHz, the bit rate of the LVDS can reach 12×40 Msps = 480 Mbps. To handle the high bit rate, the digitizer has a Xilinx Virtex 4 FPGA that counts on ISERDES to receive data from all 32 channels at a frequency of 240 MHz.
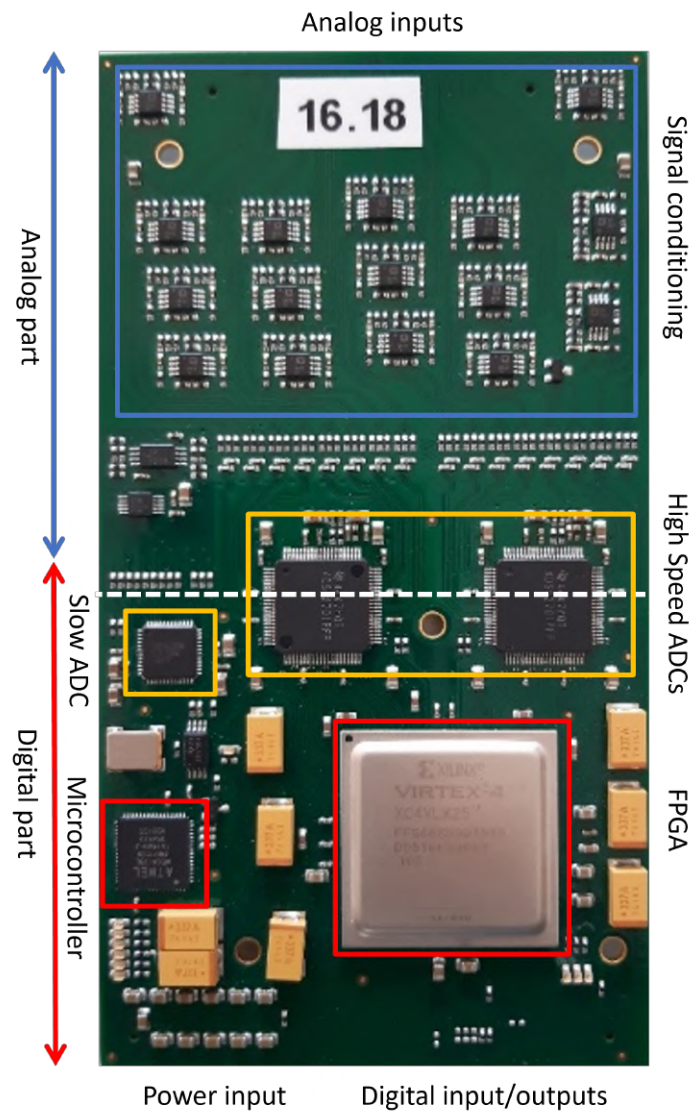
Figure 5.1: MSADC, different parts.

The device has sufficient logic resources for preprocessing, detecting word boundaries, and reconstructing and buffering data coming from the channels. For the interleaving process, there must be a fine adjustment with the clock distribution to all ADCs such that the FPGA has digital clock managers (DCM) capable of synchronizing and adjusting the phase and frequency of all clocks.

Communication with the next stages in the data flow is performed through digital input/outputs directly connected from the FPGA to a high-speed Samtec connector. High-speed data transmission is performed using 20 LVDS pairs. The Virtex-4 FPGA introduced IODELAY and ISERDES/OSERDES primitives for the cases in which a clock alignment must be performed owing to high-frequency data clocking, presenting a more stable and strong response. Using these cores, safe and high-speed data transmission to other FPGAs or programmable devices can be achieved. The LVDS lines can be reconfigured to work as single-ended signals if it is necessary to have more general-purpose input/outputs (GPIOs) or to connect to other serializers/deserializers. The FPGA banks for the LVDS are powered with 2.5 V, so in the case of using the GPIOs as single ended, from the other side of the communication, the device should be LVCMOS 2.5 V compatible. In addition to these 20 LVDS general-purpose channels, there are three more LVDS pairs dedicated for clocking and triggering. For communication with 3.3 V devices there are 25 single-ended IOs that can be used for general-purpose functions.

Because the MSADC is thought to operate in systems that include a large number of channels, the boards are equipped with a control and monitor subsystem for reliability purposes. The supplied voltages and temperatures are monitored in the critical parts. The six voltage domains are connected to the low-speed ADC. Subsequently, the PCB temperature is measured at three points to verify power dissipation and proper cooling of the parts.

There is an Atmel AVR microcontroller (uC) for all the slow control tasks, interfacing all devices via a serial peripheral interface (SPI). The uC is accessible from outside the

MSADC via either a universal asynchronous receiver/transmitter (UART) interface or 2-wire inter integrated circuit (I2C) interface.

The uC is also responsible for configuring the FPGA, loading the bitstream, and handling programming sequence signals. Both the uC and FPGA firmware can be stored in flash memory located in the PCB. The firmware is written in the FPGA after the power up. For programming, debugging, and control of the uC and FPGA there is also a Joint Test Action Group (JTAG) interface implemented.

5.2 Hardware for Testing

As FPGA devices are among the most complex integrated circuits, they use cutting-edge technology and architectural structures to achieve the flexibility and high performance that characterizes them. The design and implementation of systems including these devices have also become difficult tasks. In particular, designing power supplies for FPGAs is a nontrivial task [85]. It depends not only on the specific device but also on the final application in which it will be used. The power consumption in each of the voltage rails is tied to the amount of logic they feed, operating frequency, I/O characteristics, and high-speed serial transceivers used in the design, among others. In addition, FPGAs are typically powered by several different voltage and power domains.

Pure FPGA devices are, from the point of view of technology, a collection of configurable logic cells tightly tied together through a configurable mesh of connections. Logic cells are typically organized as a look-up table with a memory element as a flip-flop in its output. Any logic function can be realized by connecting several logic cells. Then, to make the designs synchronous, at least one clock tree must exist. This part of the FPGA electronics can also be very power demanding, depending in the quantity of clock domains, the frequency and the amount of logic each of them feeds. Accompanying the logic and clock tree, there are usually massive memory resources such as block random access memory (BRAM), FIFO, and digital signal processor (DSP) blocks. The core of

Table 5.1: Required voltages in the MSADC and for the Virtex-4 FPGA.

Voltage Domain	Voltage Level
V_{CCINT}	1.2 V
V_{CCOn}	3.3 V
V_{CCAUX}	2.5 V

the FPGA is powered by a low-voltage power domain called V_{CCINT} and its value and consumption are dependent on the electronic technology of silicon [86].

The input/output voltage rails are crucial to the operation of the FPGA, as they ensure that the device receives a stable and sufficient power supply for the interfaces. Furthermore, they play a key role in meeting the various industry standards and protocols, as the voltages must match the requirements of each of the implemented interfaces. The I/O interfaces are grouped into banks, each with its own characteristics, and require specific voltage domains to power the I/O logic. The voltage levels of these domains depend on the standards that the bank must comply with. The two commonly used voltage domains are V_{AUX} and V_{CCO}, which are responsible for providing power to the I/O logic and the output drivers, respectively. The power consumption of these domains depends on the number of I/Os used and the frequency at which they switch. In addition to the static DC power required for the banks, the dynamic power consumption is also an important consideration for the designer.

5.2.1 Power requirements for the MSADC

From the Virtex-4 data sheet [87], the voltages required for the FPGA to operate are those shown in Table 5.1.

Another very strict requirement for powering FPGAs is the order and monotonic ramp-up of the voltages of the different domains. A power-on matrix sequence must be followed to ensure a safe start; otherwise, the device may be damaged. The main reason

for the monotonic rise is the internal power of the reset circuit used to enable the most sensitive analog parts of the electronic, such as the phase-lock loop (PLL) controlled by a voltage-controlled oscillator (VCO) present in the digital clock managers of the FPGA in its logic. In addition, the power sources must be able to supply a large in-rush current owing to the large parasitic capacitance of the FPGA transistors and bulk capacitance required for the low-impedance power distribution network (PDN).

Although linear regulators seem to be the easiest solution for powering FPGAs, they are not a good choice because most of them turn on hard at startup and attempt to provide a regulated voltage within a short time (hundreds of microseconds) [88]. With high loads before voltage regulation, they quickly reach their current limit and begin to operate as a constant-current source. If, for some reason (very probable due to the power limitation of every power source), the input rail falls below the linear regulator under-voltage lockout circuit, it will turn off. This cycle will enter in a turn-on turn-off loop until the input power rail voltage increases and the linear regulator enter in the safe operation region [89].

The safest choice is to use a switching regulator, most of which have a built-in soft-start circuitry that minimizes the input and output surge current during startup while providing a monotonic voltage ramp-up. Subsequently, the FPGA in-rush current is reduced when the voltage increases slowly. In a capacitor $I = C \cdot dV/dt$, the first *sip* of the current is due to all capacitive behaviors of the circuitry, plus the bulk capacitors.

Therefore, a proper selection and sizing of the bulk capacitance is essential for the PDN design. The bulk capacitance must be able to supply the initial current demanded during startup and must maintain the voltage level in case of power transients. A common practice is to use ceramic capacitors because of their low equivalent series resistance (ESR) and equivalent series inductance (ESL), which leads to a better high-frequency performance. However, it is important to note that ceramic capacitors have a limited capacitance value and can be costly if high capacitance values are needed.

Table 5.2: Required voltages in the MSADC and for the analog circuitry.

Voltage Domain	Voltage Level
V_{Pbuf}	+5 V
V_{Nbuf}	-5 V
$V_{ADCanalog}$	+3.3 V

Therefore, electrolytic capacitors or tantalum capacitors can be used as an alternative for high-capacitance values, but their ESR and ESL should be taken into account during the design process. Moreover, the number of capacitors and their placement in the PCB are also important factors to consider. The capacitors should be placed as close as possible to the power pins of the FPGA and distributed evenly to minimize voltage drops and inductance effects. The selection and placement of decoupling capacitors are critical for high-frequency noise suppression, so proper care must be taken in their selection and placement to achieve the desired performance.

Then, the analog part of the MSADC will require the voltage domains needed for the analog input buffers and for the ADC reference to be according to Table 5.2.

The first power source design for the FPGA was made using a reference design from Texas Instruments, ADC08D1520RB [90], but by adapting the components and power rails to our needs. It is based on the LM26400 for the 2.5 V and 3.3 V and, the LM20242 for the 1.2 V. Both are integrated buck converters with soft-start and over-current protection, with a conversion efficiency of over 80% for the operation range.

For the analog part of the design, see Figure 5.1, due to the nature of the expected input signals, the most important topic is to have a stable and clean voltage level. This can be achieved using linear regulators to obtain an appropriate output noise level. The ADC has a 12 bit resolution over 2 V_{pp}, which means that the resolution in terms of the voltage levels is 488 μ V_{pp}. The +5 V is obtained with a TLV1117, which specifies an output noise of 150 μ V_{pp}. The -5 V is generated through a negative output power module LMZ34002R
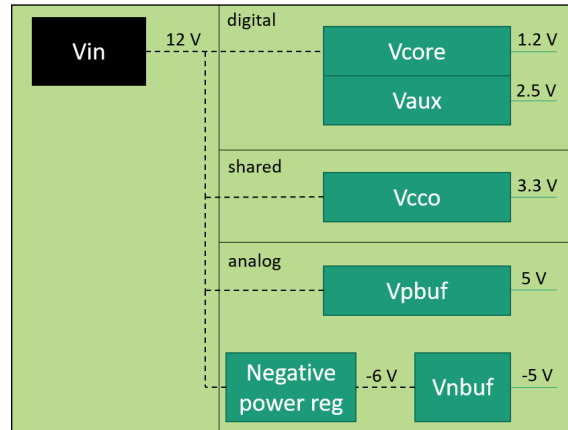
Figure 5.2: MSADC power schema.

and a negative linear regulator LT3015, and the noise output for this combination is on the order of 60 μ V_{pp}. In Figure 5.2 the different power domains required for the MSADC can be seen and in Figure 5.3 the schematics for the MSADC powering is shown.

5.2.2 MSADC Adapter Board

To interface the MSADC with an FPGA, we chose the FPGA mezzanine connector (FMC), commonly adopted on Xilinx, Altera, and Microsemi carriers, for their development boards. The FMC is under the ANSI/VITA 57.1 standard, is a small mezzanine form factor that enables easy customization of FPGA I/O (69 mm x 71-84 mm). It is prepared to operate with low power and can deliver up to 10 W at different voltages. It can map up to 80 differential pairs, 4 differential clocks plus two sourced by the carrier, 10 high-speed SEREDES plus two dedicated clocks, and I2C interface and JTAG signals between the most important.

The interface connector from the MSADC side is a samtec Q Strip High-Speed 180 position, and the 20 LVDS, 25 I/O, I2C signals, UART, clock, trigger, and JTAG signals are fully mapped to an FMC low-pin count (LPC) connector.

The MSADC connector has pins disposed in three different islands or subgroups:
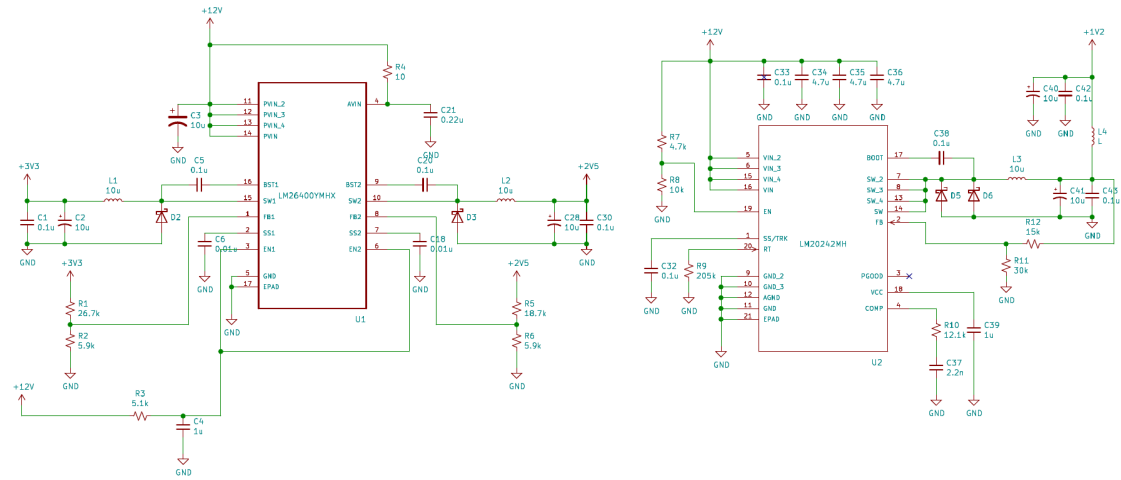
(a) Digital FPGA power domains.
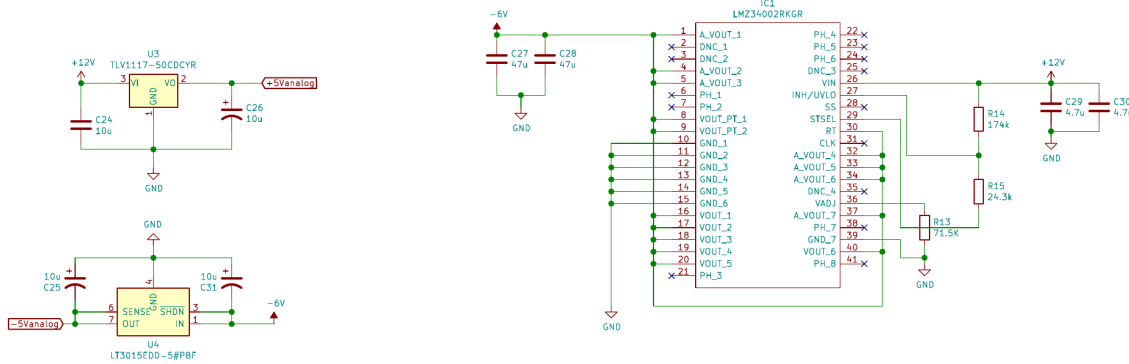
(b) Analog power domains.

Figure 5.3: MSADC power subsystem.

(a) Top side.

(b) Bottom side.
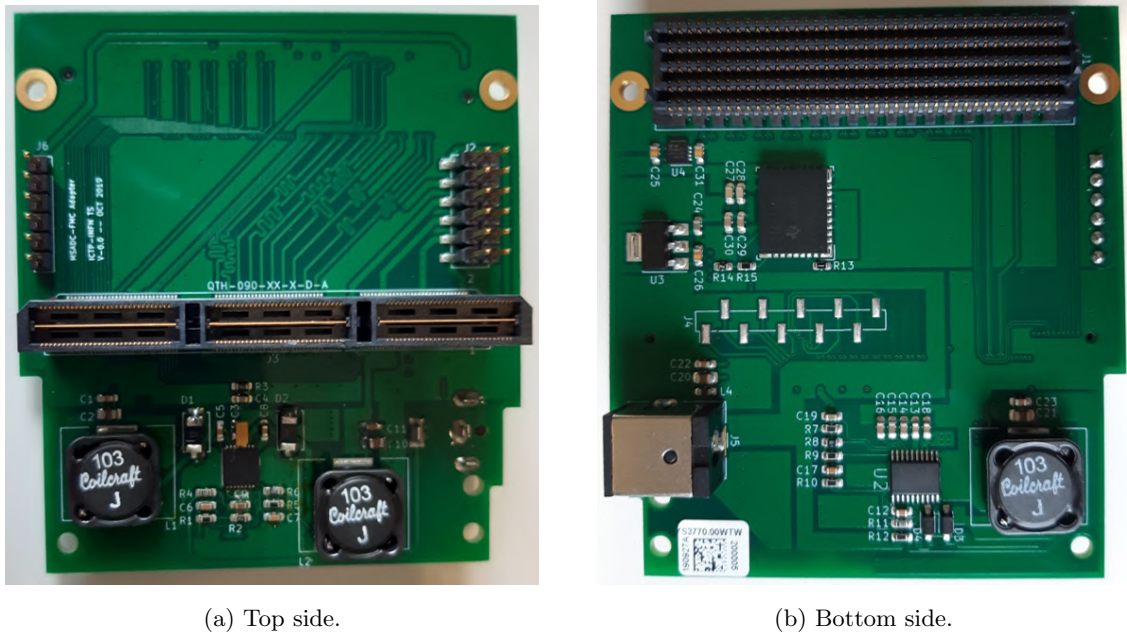
Figure 5.4: MSADC to FMC adapter board V0.0.

one is only for powering, the other for differential signals, and the last for I/O and specific-purpose signals.

LVDS lines are thought to carry high-speed signals, so the printed circuit board (PCB) layout was designed paying special attention to the signal integrity considerations. That is, the differential pairs should have 100 Ω of characteristic impedance, should be as short as possible, and have the same length. In Figure 5.4 the MSADC adapter board can be seen, with the QSH connector on top side, the FMC connector on bottom, and the power circuit disposed on both sides.

The first tests on the adapter board were performed without any board connected to verify the power source operation and stress only on the voltage rails.

After power testing, the MSADC was connected, and it was found that if the power supply had insufficient current for the initial load, the board did not start. When powered with 9 V the initial current has a peak of 3 A for a couple of milliseconds, after which

the current goes to the steady state down to 1.2 A (this last current value depends on the FPGA firmware)

For evaluating the MSADC using an FMC-based carrier card, we use the CIAA-ACC development board [91]. The board, shown in Figure 5.8, is based on Xilinx Zynq7030 SoC-FPGA and has populated a high pin-count (HPC) FMC connector, 1 Gb Ethernet, between others.

From the hardware point of view, to interface the MSADC with the CIAA using the Adapter Board, the first task is to configure the carrier to include the MSADC in the JTAG chain and then program the V_{aux} voltage to 2.5 V.

The power source and connections for the MSADC with the SoC-FPGA are different from those of the VME carrier card depicted in Figure 4.8. Therefore, the communication quality and data integrity were evaluated and characterized, taking into consideration the guidelines presented in [92].

The data from each channel of the ADCs are transmitted using a single LVDS line per channel, as illustrated in Figure 5.5. The FPGA integrated in the MSADC decodes each channel at a rate of 40 MHz × 12 bits = 480 Mbps, which is encoded in DDR and corresponds to a clock frequency of 240 MHz for ADCLK.

Once the data is interleaved, the data rate for each channel increases to 80 Mbps. To transmit this data outside the MSADC, an ODDR can be used to assign six LVDS per channel, which allows for sending the 12 bits per clock cycle, as depicted in Figure 5.6. This configuration enables the transmission of three of the 16 channels in parallel at a clock frequency of 40 MHz, utilizing the output DDR serializer. Nevertheless, for the initial implementation of the DAQ system, a continuous streaming mode was adopted, where one arbitrary and selectable channel was transmitted at a time.

A different approach was taken using the OSERDES, which is also integrated into the FPGA. The OSERDES is a parallel-to-serial converter that simplifies the implementation of high-speed synchronous interfaces, as shown in Figure 5.7. It includes all the necessary
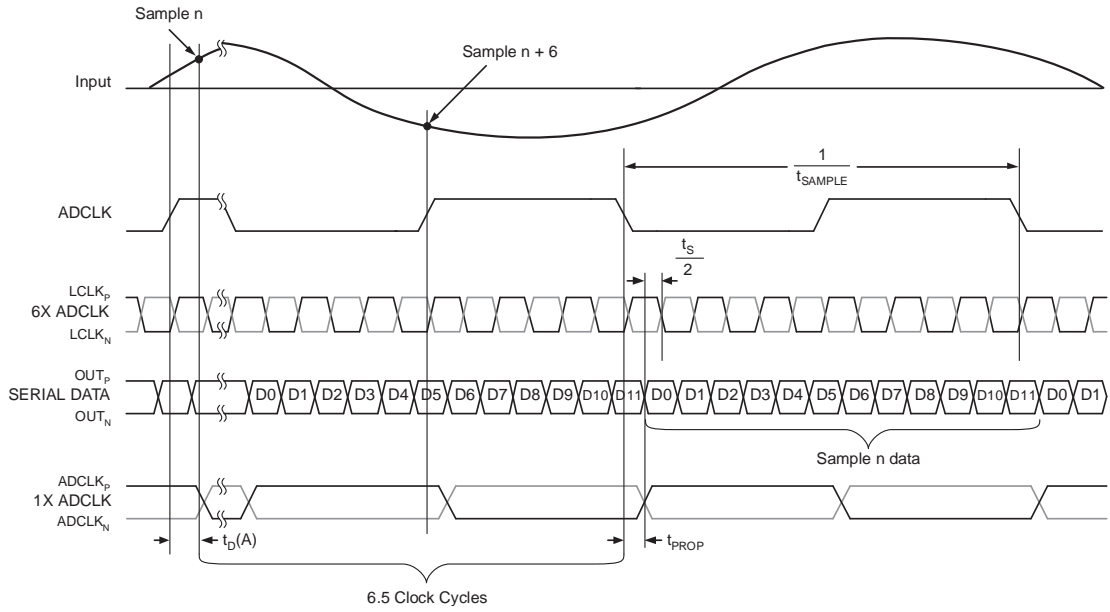
Figure 5.5: Serial data and clocking schema, less significant bit first.
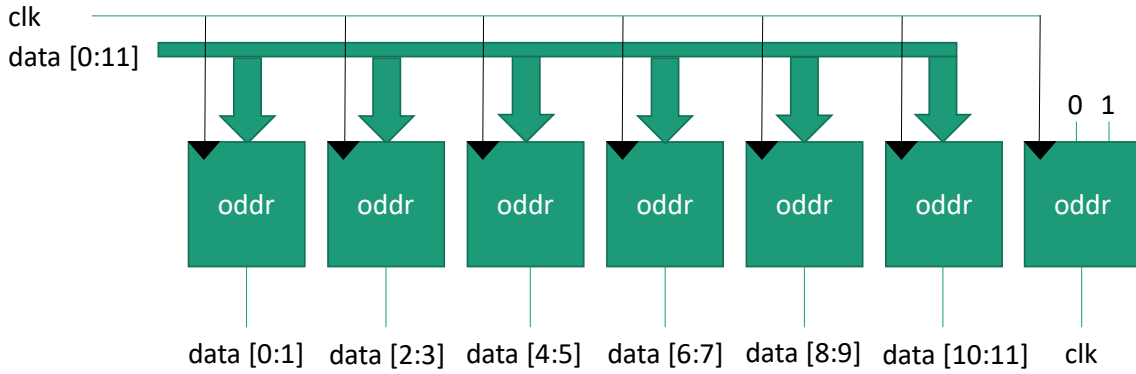
Figure 5.6: ODDR implementation for 12 bits data words.

78

(a) only master configuration.

(b) master slave configuration.
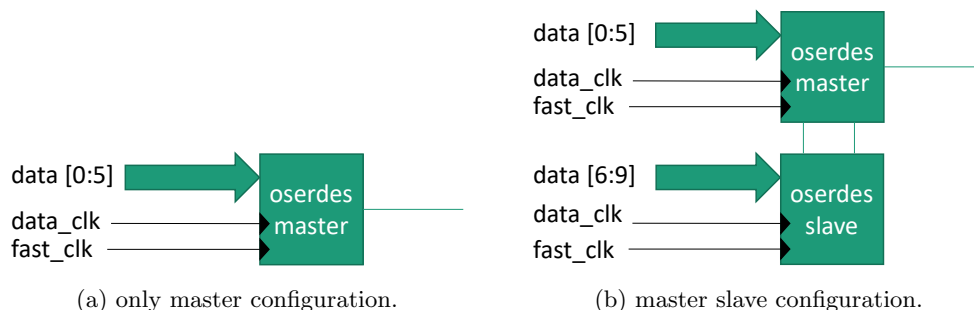
Figure 5.7: Virtex-4 OSERDES.

clocking and logic resources to function effectively and in a simple manner.

The OSERDES receives the data in parallel, from two to six bits in width, serializes the data, and dispatches the bits to the input/output buffer (IOB). The data are serialized least significant bits first and can work in a single data rate (SDR) or DDR. It has to be fed with two synchronous clocks, a fast clock for serializing the data and a low clock, which is at the speed of data updating. Both clocks must be phase aligned. OSERDES can be used to serialize up to 10 bits when used in the width expansion mode [83]. Thus, for transmitting 12 bits, it is required to use two OSERDES in master mode, with 18 LVDS, and it is possible to transmit nine channels in parallel with a fast clock of 120 MHz.

The OSERDES is designed to convert parallel data into serial data and transmit it at high speeds. It receives data in parallel, ranging from two to six bits in width, serializes it, and sends it to the IOB. The data are serialized starting with the least significant bit and can operate in single data rate (SDR) or double data rate (DDR) mode. The OSERDES requires two synchronous clocks: a fast clock for serializing the data and a slow clock, which operates at the speed of data updating. Both clocks must be phase aligned. The OSERDES can serialize up to 10 bits when used in width expansion mode [83]. To transmit 12 bits to the external world two OSERDES are used in the Virtex-4, one in master mode and one in slave.

To determine the maximum data rate that can be transmitted between the MSADC
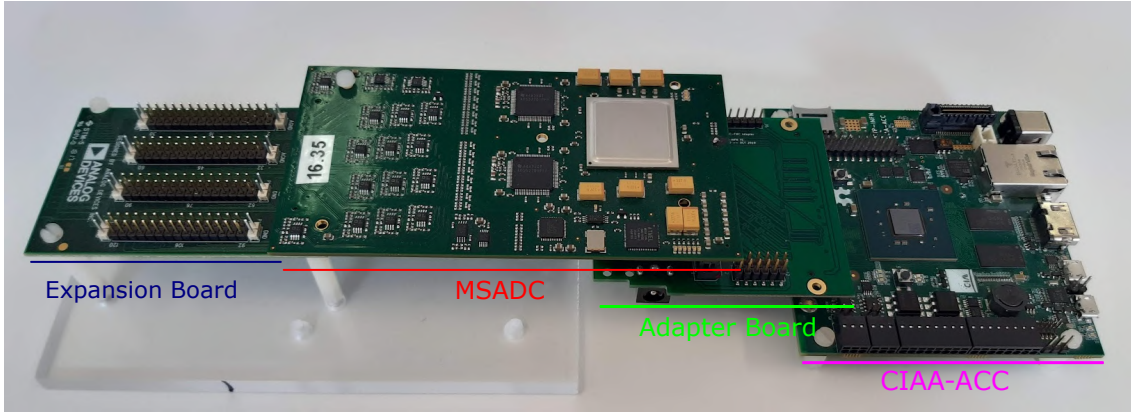
Figure 5.8: Data acquisition system using the adapter and input expansion boards.

and the SoC, two implementations were tested. The first implementation utilized the ODDRs in bus mode, which achieved a maximum data rate of 450 Mbps, corresponding to a clock frequency of 225 MHz. The second implementation used the OSERDES in an 8:1 configuration, which achieved a maximum data rate of 600 Mbps, corresponding to a clock frequency of 300 MHz on the serialized data. In both implementations, a known data sequence was transmitted and the clock frequency was increased until the data became corrupted, providing a measure of the maximum achievable data rate.

To interface with the analog channels, a separate board was necessary since the adapter board did not have such functionality. The Analog Devices board (ADZS-120ANA-SAM [93]) was chosen for this purpose due to its compatibility with the MSADCs QSH input connector, as depicted in Figure 5.8. However, the ADZS-120ANA-SAM has only inputs in the semi-differential mode, which means that one terminal of the MSADC input was mapped to ground. Despite these limitations, it was still suitable for testing and evaluating the analog channels of the MSADC.

5.2.3   MSADC Adapter Board Version 1.0

As a progression of testing, a new version of the prototype of the adapter board was designed and produced. The new board was intended to give a more robust mechanical support to the MSADC, providing a cooling fan in the ADCs region to try to reduce the temperature and test if the thermal noise can be reduced, and provide fully differential inputs to the analog connector.

Besides the mentioned characteristics, the most important part of the upgrades is the Power Distribution Network (PDN) redesign.

In general, a PDN has the following objectives:

- Feeds the system during periods of high current peaks transient, which are outside the capabilities of voltage regulators.

- Make a distribution of power (ground and power) with an acceptable level of noise.

These two characteristics can be simplified as follows: keep the impedance of the supply network below a value for the range of working frequencies of the system by means of a properly calculated network of capacitors. This task must be performed independently for each power domain voltage. These requirements may vary according to the specific current, spectral content, ripple tolerance, and working voltage.

When a current component passes through the impedance of the supply network, a noise voltage is generated, which can exceed the limiting specification of the ripple, causing the associated system to fail or malfunction. Noise can also be attributed to parasitic inductance present in components, which arises when there are sudden changes in current consumption, resulting in voltage surges given by the equation $V = L \cdot \frac{di}{dt}$.

At the time of design, to limit the noise voltage peaks, the largest transient current that can be established should be considered, because it is usually the one with the highest spectral content. To calculate the maximum impedance value of the PDN, the transient current and the ripple voltage must be taken into account. Typically, expert authors

Table 5.3: Capacitors (C) types.

Type	Value	Quant
Bulk	10 uF	1
Decoupling	0.1 uF	1

estimate the transient current to be 50% of the maximum current since this value is rarely specified in data sheets [94]. The formula used to calculate the PDN impedance is

$$Z_{PDN} = \frac{V_{dd} * \%ripple}{I_{tra} * 100}$$

.

The most critical power domains from a high-speed point of view are those related to V_{core} and V_{LVDS} because they are the ones whom must feed current at higher frequencies.

Analysis and redesign on the 2.5 V Power Domain

2.5 V - Original PDN analysis   The voltage domain that powers the Virtex-4 SERDES transceivers can support up to 800 Mbit/s when operating in DDR with LVDS, as per the datasheet. Therefore, it should be capable of providing a clean clock signal of up to 400 MHz in a worst-case scenario (Virtex-4 FPGA Data Sheet: DC and Switching Characteristics). To be conservative about the spectral content of the signal, we should consider the 3rd harmonic (1, 3, 5, 7, etc.), which means we need to account for a frequency of 5 x 400 MHz to ensure a clean digital signal at the highest frequency specifications.

From Figure 5.3a we can see that the MSADC Adapter Board has two capacitors on the 2.5 V, their values are showed in Table 5.3. The maximum current in this domain is $I_{max} = 3$ A. In Figure 5.9 the impedance profile vs. frequency is plotted. The capacitor models include the parasitic resistance and the parasitic inductance, the $Z_{PDN} = 0.05$ Ω.

2.5 V - PDN Redesign   By making as few changes as possible and using the same stack-up as the previous version, we can redesign the impedance profile to be closer to
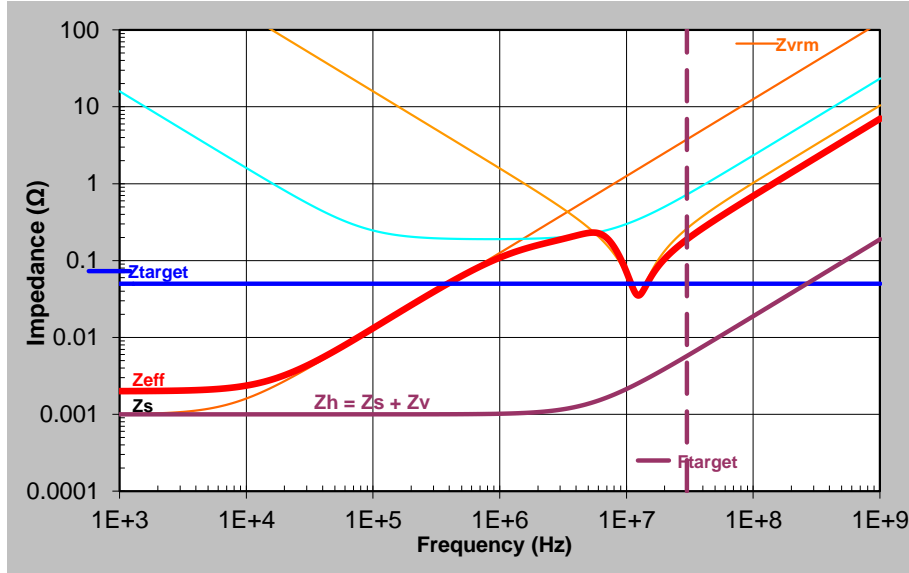
Figure 5.9: Analysis: Impedance vs. Frequency profile on the 2.5 V rail.

the ideal impedance by adding capacitors selected for their value, ESR, ESL, size, and technology (all decoupling capacitors are size 0603). Figure 5.10 shows the impedance calculated using the capacitors listed in Table 5.4a.

Analysis and redesign on the 1.2 V Power Domain

1.2 V - Original PDN analysis This voltage domain feeds the Virtex-4 V_{core} for the FPGA logic and some clocking resources, and it will demand spectral content up to the maximum frequency we operate inside the FPGA. As before, it must be considered up to the 3rd harmonic, $f = 2000$ MHz. The maximum current for this domain is $I_{max} = 2$ A and $Z_{PDN} = 0.036$ Ω. This power domain has also two capacitors of the same value to the 2.5 V domain, see Table 5.3 and the impedance profile as Figure 5.9 but, with a lower target impedance value.

1.2 V - PDN Redesign Similarly to the 2.5V domain, in order to maintain the PDN impedance below the target value throughout the entire operating range, the PDN for the

83

Table 5.4: Modified capacitor values on 1.2 V and 2.5 V rails.

(a) Modified C net on 2.5 V.

Type	Value	Quantity
Bulk	10 uF	1
	100 uF	1
Decoupling	4700 pF	1
	0.01 uF	3
	0.022 uF	1
	0.047 uF	1
	0.1 uF	1

(b) Modified C net on 1.2 V.

Type	Value	Quantity
Bulk	10 uF	2
	47 uF	1
	100 uF	1
Decoupling	4700 pF	3
	0.01 uF	3
	0.022 uF	2
	0.047 uF	2
	0.1 uF	1
	0.47 uF	1
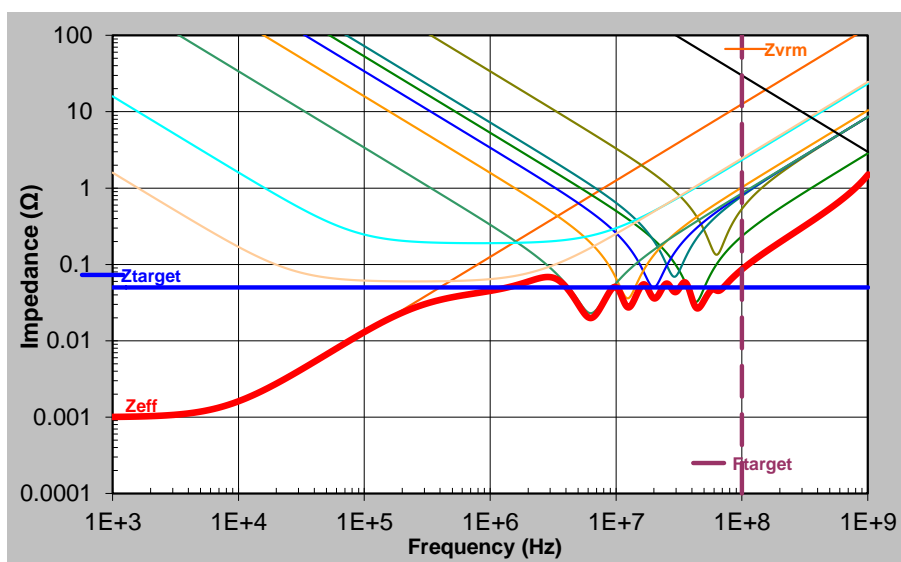
Figure 5.10: Redesigned 2.5 V: Impedance vs. Frequency profile.

1.2V domain was redesigned using 0603 capacitors of different values and technologies, listed in Table 5.4b.

Together with the power sources redesign, it was also modified the profile of the QSH connectors to a lower size to increase the robustness and, in between others the Adapter board can be fed by the 12 V of the FMC connector. The 5 V generation was modified from the previous design, and the new version was based on a buck converter to reduce the voltage from 12 V to 6 V in a more efficient manner, and then with an LDO from 6 V to 5 V, as shown in Figure 5.11.

5.3   Summary

The Mezzanine Sampling Analog-to-Digital Converter (MSADC) is a versatile and reliable digitizer board used in the ECAL2 data acquisition frontend. Developed for the COMPASS DAQ, the MSADC offers various features and characteristics for efficient analog-to-digital conversion.

With a compact form factor of 70 mm × 130 mm, the MSADC supports up to 32 differential analog channels, each buffered and sampled independently using high-speed ADCs. The analog channels are equipped with AD8138 OpAmps for impedance adaptation and buffering, and each channel has individual biasing controlled by a digital-to-analog converter (DAC). The MSADC utilizes Texas Instruments ADS527x ADCs, providing eight differential input channels per IC. The digitized data, with 12-bit resolution, is transmitted via serial LVDS at double data rate (DDR). The board incorporates a Xilinx Virtex-4 FPGA for high-speed data preprocessing, synchronization, and communication through LVDS connections.

The FPGA integrated with the MSADC decodes each channel at a rate of 40 MHz × 12 bits, resulting in a data rate of 480 Mbps. After a two ADC channels interleaving process, the data rate for each composed channel increases to 80 Mbps. To transmit this
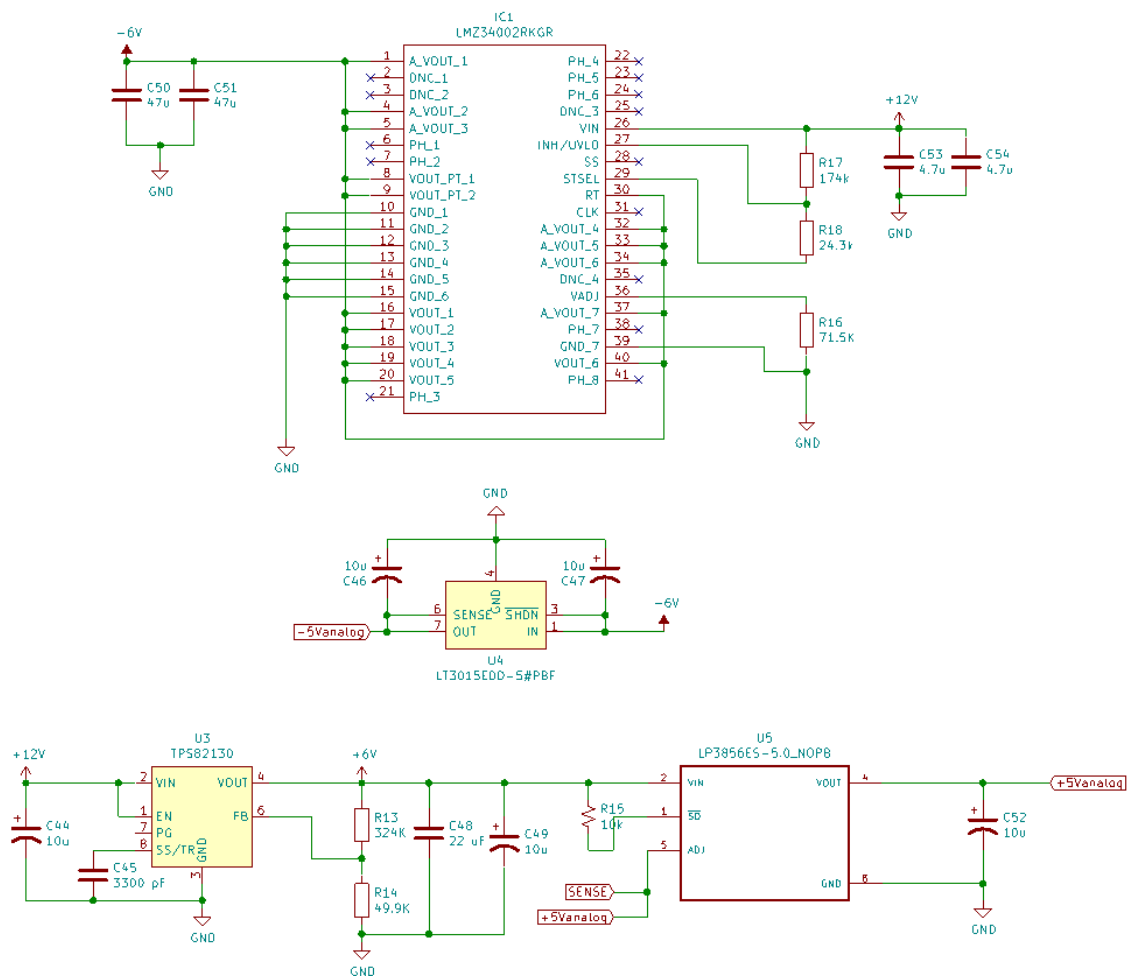
Figure 5.11: Schematic of new 5 V generation.
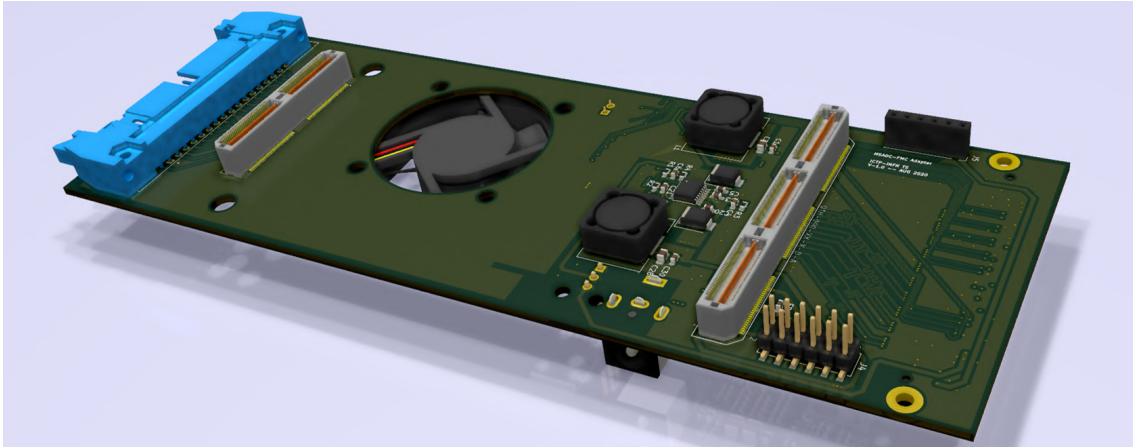
(a) Rendering of the Adapter board V1.0.
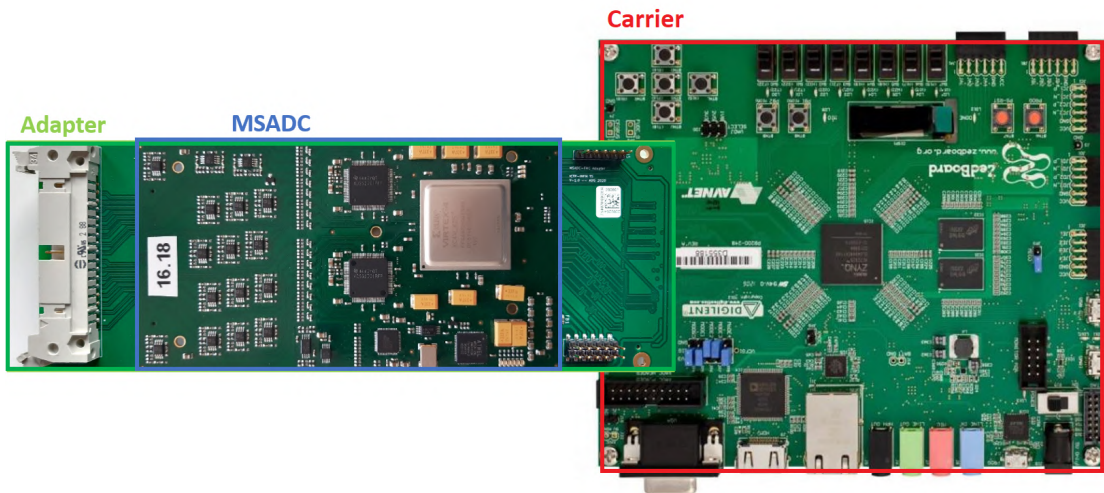
(b) Adapter board V1.0 using the Zedboard carrier.

Figure 5.12: Adapter Board V1.0.

87

data outside the MSADC, either ODDRs or OSERDES can be used.

For evaluating the MSADC an adapter board was developed. This board populates all the interfaces from the QSH connector to a standard FMC connector and the powering for all power domains required by the digital and analog parts of the digitizer. With all the MSADC input/outputs populated in the FMC connector, it is possible to use the MSADC with any commercial carrier board including this connector.

Two versions of the adapter board were produced. The first version was designed for programming the MSADC and testing the power source system. The second version was developed, redesigning the power sources taking into consideration the signal integrity. This new adapter board enables the measurement of signals in fully differential mode. Figure 5.12a shows a rendering of the second version of the board, in Figure 5.12b a Zedboard with the MSADC mounted on it can be seen.

Using the adapter boards, two implementations were tested to determine the maximum data rate achievable between the MSADC and the SoC. The first implementation using ODDRs achieved a maximum data rate of 450 Mbps, whereas the second implementation using OSERDES in an 8:1 configuration achieved a maximum data rate of 600 Mbps.

The development of the MSADC adapter board provided a robust platform for testing and characterizing the performance of the MSADC, using any SoC-FPGA commercial carrier board including an FMC connector [95, 96, 97]. In addition, the adapter enables the possibility to reuse the MSADCs for digitizing data from other detectors besides the ECAL2.

Chapter 6

SoC-FPGA Frontend Carrier Card

Based on the requirements of the AMBER proposal, the frontend for the new data acquisition system should read the 16 channels of the MSADC, process all of them in real-time to extract the most important features, and send them using high-speed serial protocol channels to the DAQ.

As mentioned in Section 4.4, according to the Equations 4.7 and 4.8, without considering any suppression algorithm, for each spill the DAQ must handle up to 565,5 Gb of data coming from the 3068 channels of the ECAL2.

In contrast, the free-running approach allows us to operate in at least two distinct modes, each with its own independent channel event detection.

In the first mode, whenever a pulse is detected, a raw data packet consisting of 32 samples is transmitted for subsequent offline analysis. This approach enables a more detailed examination of the pulse during post-processing stages.

In the second mode, the data is analyzed in real-time, and instead of transmitting the entire pulse, only the pertinent characteristics of the detector signals are extracted. This approach serves as a data compression technique, reducing the amount of data that needs to be transmitted. By extracting relevant features in real-time, the overall data transfer requirements are minimized while still preserving the essential information needed for

further analysis.

The compressed data are then sent to a digital trigger processor, also working in free-running mode, and the trigger decision is made at several levels depending on the granularity of the information [66]. The trigger processor is part of a digital trigger system (Figure 3.10), where the data from the detectors are analyzed online to decide whether to store the data in long-term memory. Detectors and conditions that participate in the decision can be externally selected and programmed.

The overall data rate involved for the ECAL2 is as follows:

$$3068 \text{ Ch} \times (12 \text{ bits} \times 80 \text{ Msps}) = 29457.8 \text{ Gb/s}$$

This rate cannot be handled by COMPASS DAQ [65]. Therefore, data reduction is performed at the beginning of the frontend. In this manner, the readout system detects events, extracts only the relevant features of the signal of the detector, and sends them to a higher hierarchy level of the data flow path, saving storage and network resources.

6.1 New SoC-FPGA Frontend Carrier Card

As part of the readout hardware update for the free-running and trigger-less DAQ of AMBER [98], we designed and developed a new SoC-FPGA based Front-end Carrier Card (FFeCCa) for reading and processing the raw data from the MSADC. Thus, the existing ECAL2 digitizer board can be reutilized by changing the firmware to allow free-running mode.

The most demanding power-processing-capable and reconfigurable devices for data processing are FPGAs. Nevertheless, at the moment of the design, the question arises of whether it is appropriate to consider using a plain FPGA or an SoC-FPGA. When designing new custom hardware with applications framed in specific hard processing and real-time tasks, it may be a good idea to use a plain FPGA, giving the most reconfigurable possibility at the best costs. FPGA design and programming are often

difficult tasks that require long development, testing, and validation times. For this reason, many system designs using FPGAs include digital signal processors (DSP) or microcontrollers for tasks that are not time critical or not require intensive parallel processing. In this case, it is very likely that both devices access external memory and interchange information between them. SoC-FPGA devices combine FPGA resources, multicore processors, on-chip memory, and dedicated processors within a single chip. They are interconnected with high-performance data buses, resulting in reduced circuit board size and power consumption. The latest SoC-FPGA generation integrates SDRAM access from both the FPGA and processors through high-performance snooped buses, enabling efficient cache coherency control and utilizing a snoop control unit (SCU). This integration improves system compactness, performance, and power efficiency while allowing to use only one main DDR system memory without major drawbacks.

Using an SoC-FPGA facilitates hardware implementations and designs because the layout routing tends to be simpler owing to the smaller quantity of traces to equalize and match between integrated circuits.

The addition of application processors in the frontend electronics, together with the FPGA, also allows to include Edge Computing concepts [99] at the very beginning of the data flow path, in a simpler way, allowing the possibility of embedding complex algorithms in the processors for post-processing the data after intensive and fast FPGA pre-processing.

Moreover, the design of a carrier card for reading and powering the MSADC, from the perspective of modularity, can be considered as a purely embedded and fixed design with all its components soldered on the same board or as a System on Module (SoM)-based design. There are no specific rules or guides to determine whether one is better than the other. The main topics for making adoption decisions are budget, expandability and versatility, design complexity and support, and the need for mechanical robustness in terms of vibrations and thermal stress.

Among all the motivations, one of the most important arguments resides in the fact that designing PCBs for FPGAs or SoC-FPGAs with embedded memories is not an easy task. The implementation of PCBs with SoCs, DDR4, and high-speed data transceivers usually not only requires a high level of expertise and knowledge in signal and power integrity but also the need for field solver software to verify the correct design of all transmission lines and power delivery systems. The number of layers on a high-performance PCB is proportional to the miniaturization and pin density of the integrated circuits, and to the number of voltage domains required by the devices [100]. Another crucial factor influencing the decision to design a carrier using an SoM or a soldered device is the aspect of obsolescence and the lifetime of its main components, specifically the SoC-FPGA and digitizer board in this case. It is essential to take into account the available budget and the demanding processing requirements when choosing compatible devices for the same PCB hardware. Additionally, selecting components with a relatively long lifetime, established presence in the market, and supporting documentation and reference designs is crucial for long-term viability and ease of development.

In addition, as the testing of the hardware in general is an important part of the validation process, the design must carefully address this topic. To make the process easier and to have an interface for connecting the devices to a monitor or external programmer, the debugging interfaces should be considered and properly designed. Serial high-speed interfaces are preferable to parallel interfaces. Serial interfaces consume less power than the second, and at high frequencies, the synchronization and clocking of parallel interfaces is a very complicated task. Serial interfaces and Multi-Gigabit Transceivers (MGT) are much safer for transmitting information because the clock can be directly recovered from the data itself.

A desirable aspect in hardware design is the ability to easily redesign it when if there is a need for similar applications. From a licensing perspective, the best option is to go through an Open Hardware License.

Table 6.1: Resources comparison between used FPGA devices, MSADC Virtex-4, CIAA-ACC Zynq-7030 and FFeCCa ZU9EG.

Device	Logic Cells	Max. Distributed RAM (kb)	BRAM (kb)	DSP	Processors	Max User I/O
XC4VLX25	24192	168	1296	48	N/A	448
Z7030	125000	157	9300	400	2xCortex A9	382
ZU9EG	600000	8800	32100	2520	4xCortex A53	570

6.1.1 SoC-FPGA Selection

Considering all the above-mentioned topics, it was decided to go for a modular design based on a carrier to host the MSADC and a System-on-Module (SoM) based on SoC-FPGA. Besides the embedded SoC-FPGA, the SoM includes the power sources and DDR4 memories. The SoC-FPGA family with the best cost-benefit ratio for the application addressed in this work with all the necessary requirements is the Xilinx Zynq Ultrascale+. In Table 6.1 there is a comparison between the resources of XC4VLX25 (MSADC), Z7030 (CIAA-ACC), and the ZU9EG device selected for the new carrier card.

Among the Zynq Ultrascale+ devices, we can find three different subfamilies, as shown in Table 6.2. These devices are catalogued based on their resources and target applications. The main differences reside in the inclusion of a video codec on the EV devices, mainly for real-time video processing and fast encoding, and in the amount of logic and FPGA resources available for EG devices. The CG devices are the "low end" of the Zynq family, as well as the cheapest.

The Zynq Ultrascale+ family has a projected end-of-life later than 2035, which is the projected lifetime of the previous Zynq-7000 SoC family. Zynq Ultrascale+ EG devices are part of the high-end portfolio products of Xilinx, thus opening a new spectrum of on-site hard-processing possibilities, an almost mandatory requirement for the real-time features

Table 6.2: Xilinx Zynq Ultrascale+ Devices.

	CG Devices	EG Devices	EV Devices
Application Processor	2-Core ARM Cortex A53 MPCore up to 1.3GHz	4-Core ARM Cortex A53 MPCore up to 1.5GHz	4-Core ARM Cortex A53 MPCore up to 1.5GHz
Real-time Processor	2-Core Arm Cortex-R5F MPCore up to 533MHz	2-Core Arm Cortex-R5 MPCore up to 533MHz	2-Core Arm Cortex-R5 MPCore up to 533MHz
Graphics Processor		Mali-400 MP2	Mali-400 MP2
Video Codec			H.264/H.265
Programmable Logic	81K–600K Logic Cells	81K–1143K Logic Cells	192K–504K Logic Cells
DSP Slices	Up to 2520	Up to 3528	Up to 1728
Block RAM	Up to 32.1	Up to 34.6	Up to 11.0
UltraRAM	Up to 18.0	Up to 36.0	Up to 27.0
GTH 16.3Gb/s	Up to 24	Up to 44	Up to 24

extraction on multichannel frontends.

The real-time processing of parallel data streams require a high quantity of logic resources, DSP, Block RAM (BRAM) and UltraRAM on the FPGAs for parallel signal processing in the detection and amplitude extraction of the pulses generated in the ECAL2.

BRAM and UltraRAM are large blocks of fast memory embedded in FPGA logic. These two hard blocks can be used for short-term data buffering in the FPGA when performing real-time processing or as long FIFOs while waiting to transmit the collected data.

In Ultrascale+ devices, the BRAM in Figure 6.1 can be used in different operation modes, as follows:

- **Synchronous Operation**: Each memory access, read, and write is controlled by a clock. All inputs, data, address, and write enable are registered. The data output is always latched and retained until the next operation. An optional output data pipeline register allows for higher clock rates at the cost of an extra latency cycle. During a write operation, the data output can be made to reflect the previously stored data, newly written data, or remain unchanged. There is independent reset control of the output latches and registers.

- **Asynchronous Operation**: Data outputs can also be set or reset asynchronously.

- **True Dual-port Operation**: The block RAM has two completely independent ports, A and B, which share only the stored data.

- **Simple Dual-port Operation**: One port is dedicated to a write port, and the other to a read port. Consequently, the data width can be extended to 72 bits for the 36 Kb full block RAM or 36 bits for the "split" 18 Kb Block RAM [101].

On the other hand, UltraRAM is a dual-port synchronous 288 Kb RAM with a fixed configuration of 4,096 deep and 72 bits wide. Ports A and B share the same clock signals.

Port A   **Port B**
clk   clk
wr_ena   wr_ena
addr   BRAM   addr
rd_data   rd_data
wr_data   wr_data
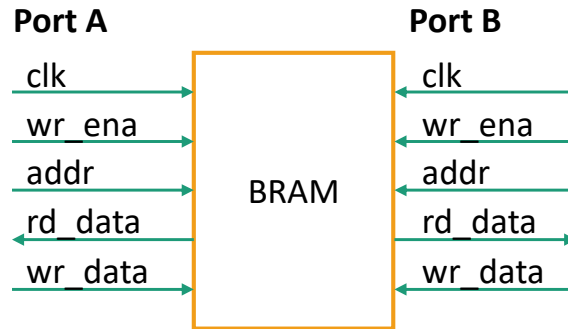
Figure 6.1: Block RAM Ports.

Within a single cycle of the external clock, Port A operation is always completed before Port B operation. All columns of UltraRAM can be connected using fabric routing to create memory arrays of up to 360 Mb in the largest devices. When there is a need for large buffering, this memory is more efficient than BRAM, both for the largest capacities and the possibility of dynamic low-power modes when they are not used [102].

Regarding the cost for the available SoMs in the market with a Zynq Ultrascale+ EG, it was found a company offering a complete SoM solution at a lower cost than the single SoC. This, together with the aforementioned reasons, is the best option for an SoM-based solution. As an example, when writing this thesis the Zynq Ultrascale+ ZU4EG was found on the electronic market with a reference price of 1060.80 €, VAT included [103], while a SoM with same SoC with 2 GB of DDR4 from the company Trenz had a price of 516.46 €, VAT included [104].

The Trenz TE080X SoM family comes in a 5.2 x 7.6 cm form factor, with the possibility of selecting the SoCs with different FPGA characteristics and processing resources.

For interfacing with a carrier board, the SoM has four board-to-board razor beams, a low-profile terminal/socket strip, 0.5 mm pitch, and a 5 mm stack height [105]. The connector offers a bandwidth of 20 GHz for an insertion loss of 3 dB, as shown in Table 6.3. This bandwidth is sufficient to carry the signals of the Ultrascale+ MGT, which has

Table 6.3: Connector bandwidth at a 3dB of insertion loss.

Signaling	Speed rating
Single-ended	13.5 GHz / 27 Gbps
Differential	20.0 GHz / 40 Gbps

a specified maximum operative bit rate of 16.375 Gb/s when working at 8.15 GHz in DDR mode over an LVDS line.

Figure 6.2 shows the main components and connections of the SoM as a block diagram. The SoC has 48 high-density (HD) and 156 high-performance (HP) I/Os, and 65 multi-use I/Os (MIO) populated in board-to-board (B2B) connectors. The SoM is then equipped with four SDRAM DDR4-2400 chips and can populate up to 8 GBytes of memory. The memories are connected to the Zynq Processing System (PS) DDR controller via a 64-bit data bus. The maximum data transmission rate from the PS to the memory is 2400 MB/s, according to the specifications [106].

For booting, configuration, and operation, the SoM can be configured with 512 Mbyte SPI flash memories. These memories are connected to the PS. It also contains a 2 Kbit serial EEPROM for general purposes and a MAC address, which can be accessed by the I2C interface.

For clocking purposes, the SoM relies on a Si5338 high-performance low-jitter clock generator. The IC can synthesize any frequency for each of the four output drivers of the device. The frequency range goes from 0.15 to 710 MHz when the outputs are configured to work in LVDS. This low-jitter programmable oscillator is fundamental for GTR and GTH transceiver operation.

For high-speed data transmission, the SoM has eight serial transceivers populated to the B2B connector. Four of them are PS-GTR transceivers, used for DisplayPort, Ethernet SGMII, PCIe, SATA, or USB3.0. The remaining are connected to the GTH, with the possibility of working with several protocols with a line rate support of up to
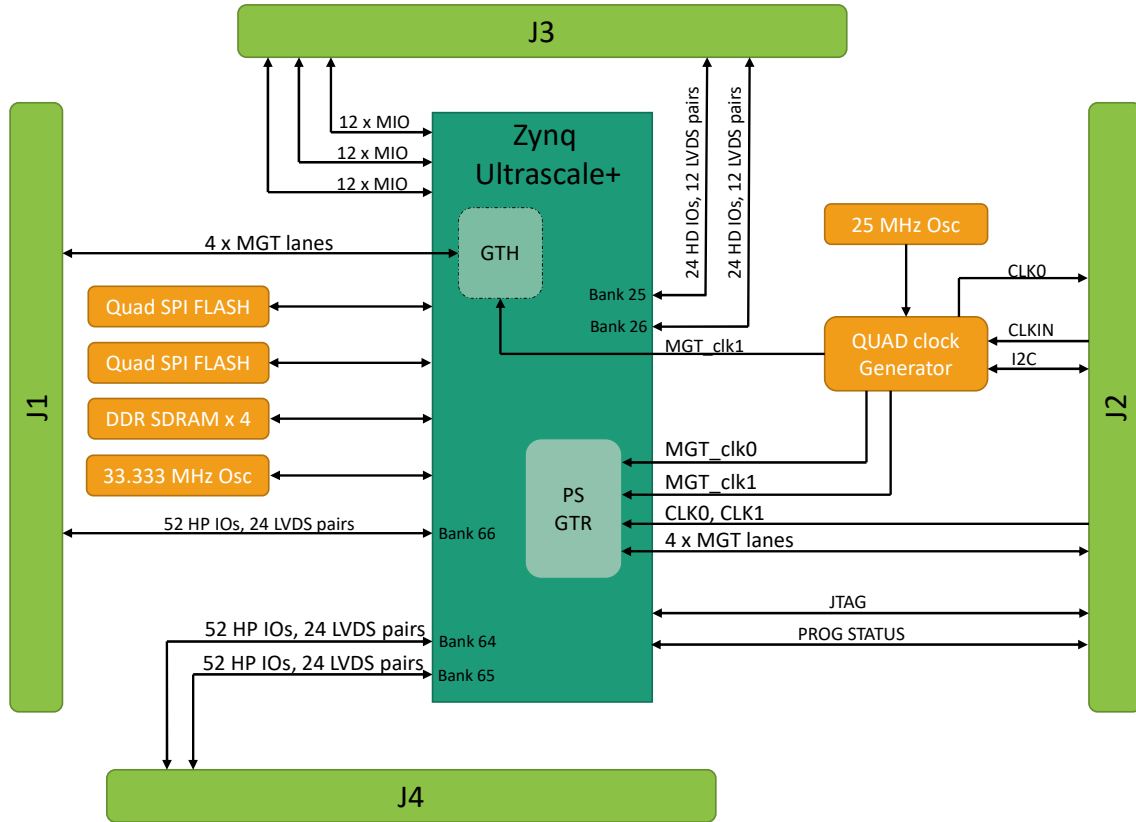
Figure 6.2: TE0803 SoM Block Diagram.

16.375 Gb/s.

It is recommended a 3.3 V Power supply with a minimum current capability of 3 A for system startup. Then, all the voltage rails and power domains are generated into the SoM in the right sequence and with the correct load capacity.

The positive aspects of adopting an SoM solution can be summarized as follows:

- Encapsulate the complexity of the SoC-FPGA along with essential common components to facilitate the design of large hardware systems, avoiding the risk of errors related to critical interconnections, such as those between the SoC-FPGA and DDR memories.

- Standardize a board to board connection to facilitate hardware upgradeability, partial replacement, and interchangeability.

- Allow the HW designer to concentrate the efforts on the application specific aspects, relying on an error-free partial system module.

- Availability of several affordable commercial solutions and design support.

6.1.2 Carrier Card Design

The FFeCCa board developed for hosting both the SoM and the MSADC is shown in Figure 6.3. The interfaces are indicated in yellow, the SoM in red, and the MSADC in green.

As introduced in Section 3.3, the MSADC digital interface is operated through 20 LVDS pairs for high-speed data transmission, plus two LVDS for triggering and clocking, adding to 25 single-ended signal input/outputs for general purposes. Special attention must be paid when connecting and routing high-speed signals between the MSADC and SoM. The integrity of the data depends not only on the PCB layout but also on the connections, voltage standards, and banks of the FPGAs. The best option is to connect all 22 LVDS, including the clock, to the same SoC-FPGA bank. Thus, less effort might have to be made when synchronizing data inside the FPGA; when routed in the same bank, all flip-flops have the same clock delay, and the probability of error due to a different clock path delay decreases. For the implementation, all 22 LVDS pairs were routed to the HP bank 64 and mapped to the J4 B2B connector. The 25 GPIOs were mapped to Bank 25 of the SoC, which corresponds to J3 in the SoM.

The COMPASS DAQ and, in general, the DAQ for HEP experiments require high-data-rate transmission channels, which are often accomplished by optical interfaces. For high-speed communication, the board counts on 4 small form-factor pluggable (SFP+) interfaces that were connected to the GTHs transceivers mapped to an MGT on J1. Although SFP+ foresees the use of copper cables, the best performance when dealing
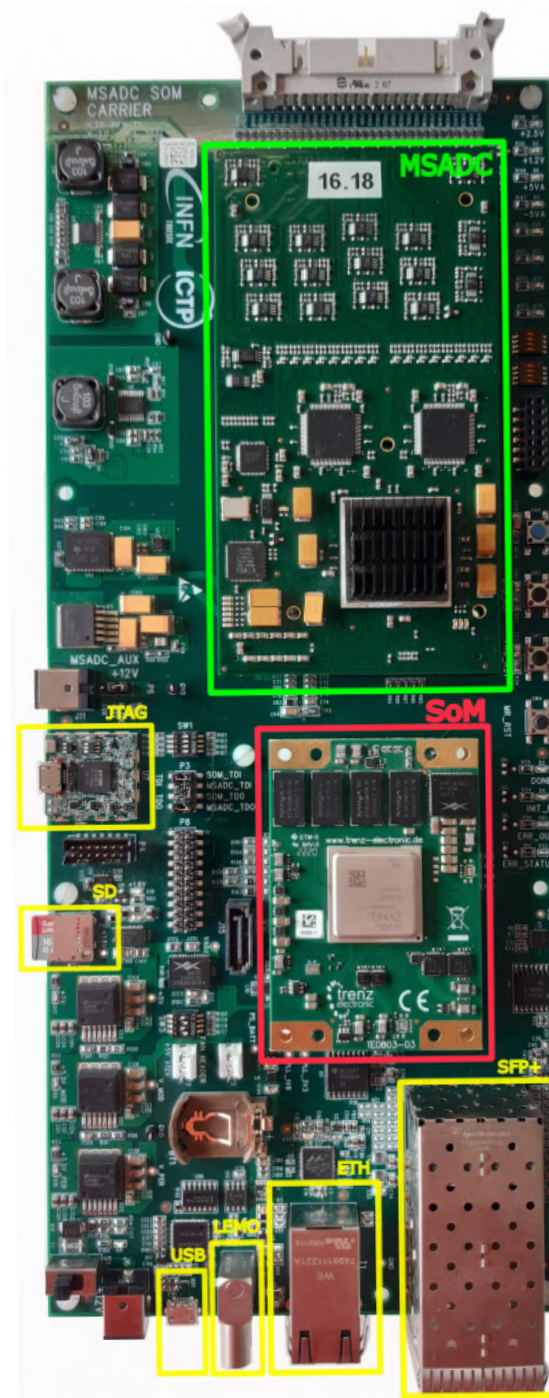
Figure 6.3: First prototype of the FFeCCa carrier card.

with long distances is achieved when using fibers [107]. The SFP+ interface can operate at a data rate of up to 10 GB/s, which is sufficient for transmitting and receiving data inside the DAQ network [98].

To connect the carrier to a Local Area Network (LAN) or the Internet, the board has a Gigabit Ethernet interface. The interface is connected to the PS system by means of a reduced gigabit media-independent interface (RGMII) to a Marvell 88E1512 interface IC [108]. The transceiver implements the Ethernet physical layer portion of the 1000BASE-T, 100BASE-TX, and 10BASE-T standards. The connection to the SoM was performed through the MIOs on J3.

For debugging purposes the carrier counts on a USB to UART bridge. The bridge is based on a Silicon Labs CP2108 IC. The IC provides a solution to a USB2.0 full-speed function controller, USB transceiver, and four UARTs, although only two of these are connected, Processing System (PS) UART_0 and UART_1. The interfaces work as a virtual solution with a royalty-free implementation with all drivers provided by Silicon Labs. The connection to the SoM was also performed through the MIOs on J3.

Many interfaces are done with LEMO connectors, there is a LEMO input and output. The input is implemented as a 3.3 V low-voltage transistor-transistor-level (LVTTL), as shown in Figure 6.4, and the output according to the nuclear instrumentation module (NIM) standard [109]. Both were mapped to Bank 25 on J3.

To configure and debug the FPGA and microcontroller of the MSADC and the SoC of the SoM, the carrier board counts on a joint test action group (JTAG) [110] programmer that is compatible with Xilinx electronic design automation (EDA) tools. To ensure all possible connections inside the JTAG chain, an interconnection header was added, as shown in Figure 6.5. The JTAG daisy chaining allows to have one device to be connected or both, but not both individually. With the interconnection header, it is possible to choose which device will be inside the JTAG chain or connect both devices simultaneously. In JTAG, Virtex-4 is identified with a unique code. With JTAG, it can be programmed, and
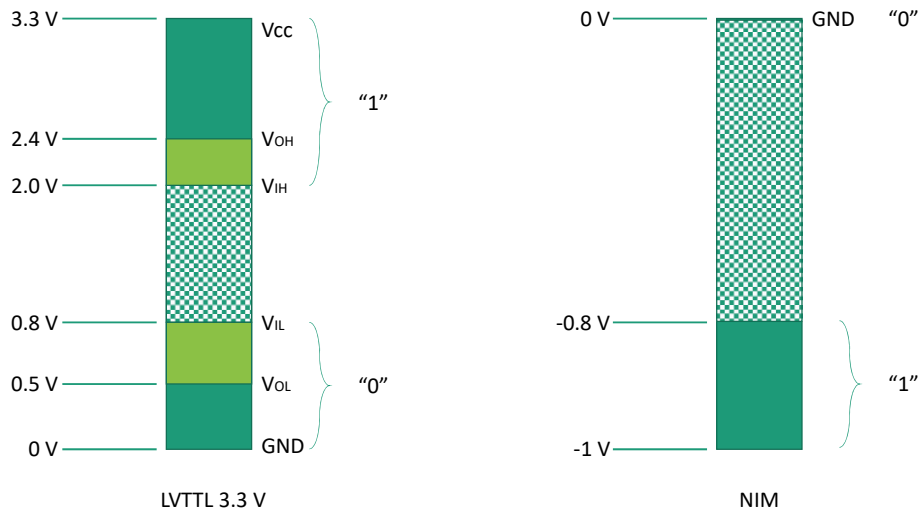
Figure 6.4: LVTTL 3.3V and NIM logic states.

if there is a need for debugging, Xilinx offers the ChipScope core to be embedded in the design, allowing access to specific registers selected by the user. On the other hand, as the SoC is built by two different logic devices, a PS and a programmable logic (PL), the solution of JTAG is by means of a debug access port (DAP), presenting both the PS and PL as different devices inside the chain.

When using devices from the Zynq family, the programming and configuration of both the processors and FPGA are performed through the PS. For programming and nonvolatile data storage, the design has an SD card holder connected to the PS of the SoC. From this storage, the entire operative system and FPGA firmware can be booted with the possibility of adding a logic partition for data storage. If the requirements for data storage are insufficient, a serial advanced technology attachment (SATA) interface can be used. The board has a standard SATA connector mapped to the SoC PS-GTR and all power rails for any SATA device.

The carrier board has a form factor of 312×131 mm, which is sufficient for hosting all connectors, power sources, MSADC, and SoM. It was designed on a 6 layers stack-up process with controlled impedance for all critical high-speed lanes.
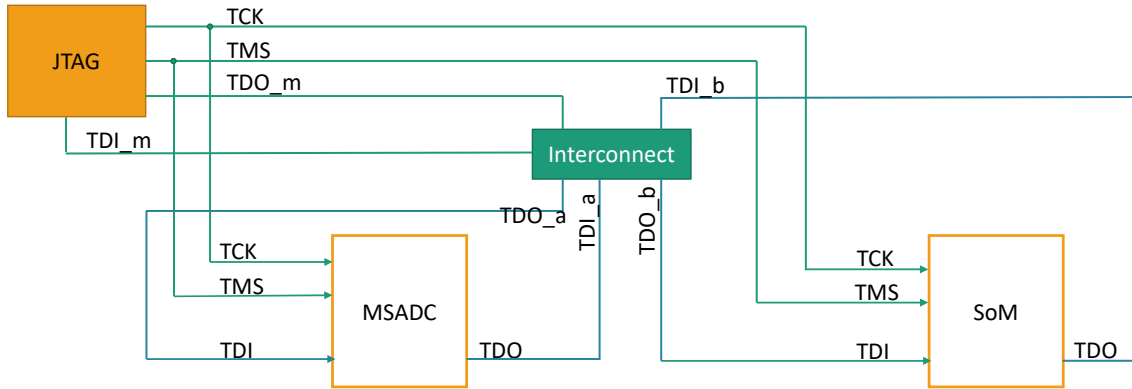
Figure 6.5: JTAG daisy chaining solution for MSADC and SoM.

In high-speed PCB design, it is crucial to understand that the traditional concept of ground, as commonly applied to DC and slow-speed signals, is no longer valid. The complexities of high-speed signals require a different approach to ground design.

There is often confusion between the circuit ground, earth ground, and chassis ground, which can lead to misinterpretations. In the context of high-speed design, the focus is on ensuring a reliable return current path for fast transients. This return current path can be assigned to any power plane, depending on the path with the lowest impedance.

It is important to note that if there are discontinuities in the return planes, such as splits or gaps, the impedance of the transmission line will also experience a discontinuity. This can result in signal integrity issues and signal reflections, leading to degraded performance.

When designing a PCB for high-speed applications, adherence to certain rules becomes crucial. One of the most important rules is to establish a robust return current path for each signal. This involves careful consideration of the target impedance and ensuring length equalization of the lines within a specific domain.

By following these principles, designers can mitigate signal integrity problems and maintain the integrity of high-speed signals on the PCB. Proper ground design, along with impedance control and signal path equalization, plays a vital role in achieving reliable and

```
Number of layers: 13
Total thickness = 1637 um
```

NN	Layer Name	Type	Usage	Thickness um	Er
1	SOLDERMASK_TOP	Dielectric	Solder Mask	20	3.3
2	TOP	Metal	Signal	35	<Auto>
3	PP_2116	Dielectric	Substrate	130	4.4
4	INNER_LAYER_2_GND	Metal	Plane	18	<Auto>
5	CORE	Dielectric	Substrate	500	4.7
6	INNER_LAYER_3	Metal	Signal	18	<Auto>
7	PP_7628	Dielectric	Substrate	195	4.7
8	INNER_LAYER_4	Metal	Signal	18	<Auto>
9	CORE	Dielectric	Substrate	500	4.7
10	INNER_LAYER_5_PWR	Metal	Plane	18	<Auto>
11	PP_2116	Dielectric	Substrate	130	4.4
12	BOTTOM	Metal	Signal	35	<Auto>
13	SOLDERMASK_BOT	Dielectric	Solder Mask	20	3.3

1637 um

Figure 6.6: Stackup used for the Carrier Board.

Table 6.4: Design parameters for controlled impedance transmission lines domains.

Z0 Domain	Type	Outer - Pwr		Pwr - Inner	
		Width	Distance	Width	AirGap
100Ω	Diff	150 μm	200 μm	150 μm	150 μm
50 Ω	Single	200 μm		450 μm	
85 Ω	Diff	200 μm	150 μm	200 μm	120 μm

optimal performance in high-speed PCB designs.

Specifically, the lines with controlled impedance design were (i) with 100Ω differential: analog inputs, LVDS Data, LEMO inputs, and SFP+; (ii) with 50Ω single-ended: LEMO output, RGMII, SDIO; and (iii) with 85Ω differential: USB and SATA. Table 6.4, shows the layout track dimensions according to the designed stackup for the PCB, shown in Figure 6.6. As the stackup is symmetrical, Table 6.4 shows the transmission line parameters for the TOP or BOTTOM layers vs. the GND or PWR layers, and the GND or PWR vs. Layer 3 or Layer 4.

For verifying the parameters of the microstrip lanes the next formulas based on

Wheeler's equation [111] were used:

$$Z_0 = \frac{\eta_0}{2\pi\sqrt{2(E_r + 1)}} \cdot ln\left(1 + 4 \cdot \frac{h}{w_{eff}}(X_1 + X_2)\right) \tag{6.1}$$

where,

$$W_{eff} = W + \frac{t}{\pi} \cdot \left(\frac{4e}{\sqrt{\left(\frac{t}{h}\right)^2 + \left(\frac{t}{w\pi + 11 \cdot t\pi}\right)^2}}\right) \cdot \frac{E_r + 1}{2E_r} \tag{6.2}$$

$$X_1 = \frac{4h}{W_{eff}} \cdot \left(\frac{14E_r + 8}{11E_r}\right) \tag{6.3}$$

$$X_2 = \sqrt{\frac{4h}{W_{eff}}^2 \cdot \left(\frac{14E_r + 8}{11E_r}\right)^2 + \frac{E_r + 1}{2E_r} \cdot \pi^2} \tag{6.4}$$

In Figure 6.7 there are depicted the main parameters for the microstrip and stripline transmission lines.

$\eta_0 = 120\pi$

$h = substrate\ height$

$w = trace\ width$

$E_r = substrate\ dielectric\ constant$

$t = trace\ thickness$

$s = spacing\ between\ traces$

According to the IPC-2141A [112] for the asymmetric striplines ($h_1 \neq h_2$) the impedance value can be estimated as,

$$Z_0 = \frac{80}{\sqrt{E_r}} \cdot ln\left(\frac{1.9(2h_a + t)}{0.8w + t}\right) \cdot \left(1 - \frac{h_2}{4h_1}\right) \tag{6.5}$$

The differential version for the microstrip and stripline can be verified with equations 6.6 and 6.7 respectively:

(a) Microstrip.

(b) Differential Microstrip.

(c) Stripline.

(d) Differential stripline.
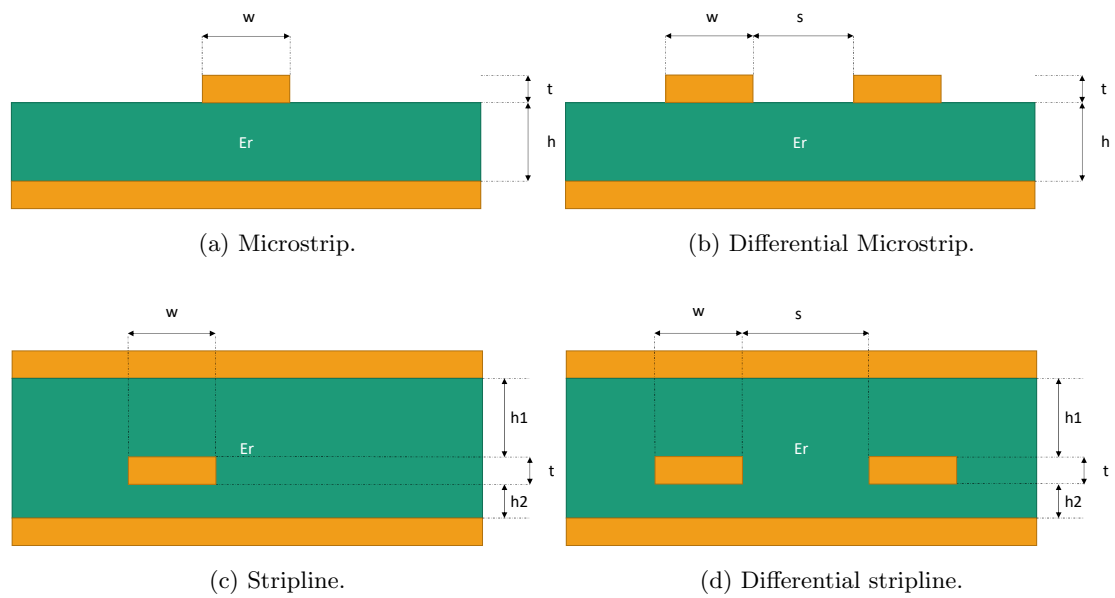
Figure 6.7: Parameters for designing characteristic impedance Z_0 according to the geometry and materials.

Table 6.5: Measured voltages from the MSADC power subsystem.

Voltage Domain	Nominal	Measured
+5Va_MSADC	+5 V	+5.03 ± 0.02 V
-5Va_MSADC	-5 V	-4.99 ± 0.02 V
+2.5Vd_MSADC	+2.5 V	+2.51 ± 0.03 V
+1.2Vd_MSADC	+1.2 V	+1.18 ± 0.03 V
+3.3Vd_MSADC	+3.3 V	+3.31 ± 0.02 V

$$Z_{diff} = 2Z_0 \cdot \left(1 - 0.48e^{-0.96\frac{s}{h}}\right) \tag{6.6}$$

$$Z_{diff} = 2Z_0 \cdot \left(1 - 0.37e^{-2.9\frac{s}{h_1+h_2}}\right) \tag{6.7}$$

6.1.3 FFeCCa Testing

After the production and reception of the first set of FFeCCa boards, the first task was to test the power system and remaining functionalities. The power system is divided into two subdomains: MSADC and SoM power. The MSADC is powered using the same design as in the Adapter Board V1_0. All the voltage domains operate in the specified ranges, as shown in Table 6.5.

The SoM requires, in principle, a single 3.3 V power source for startup and PS operation capable of delivering a transient current of at least 3 A. Then, for the peripherals and communication between the SoM and the MSADC, there are three more power sources, 5 V, 3.3 V, and 1.8 V. All the power sources were tested, met the requirements and operated as expected.
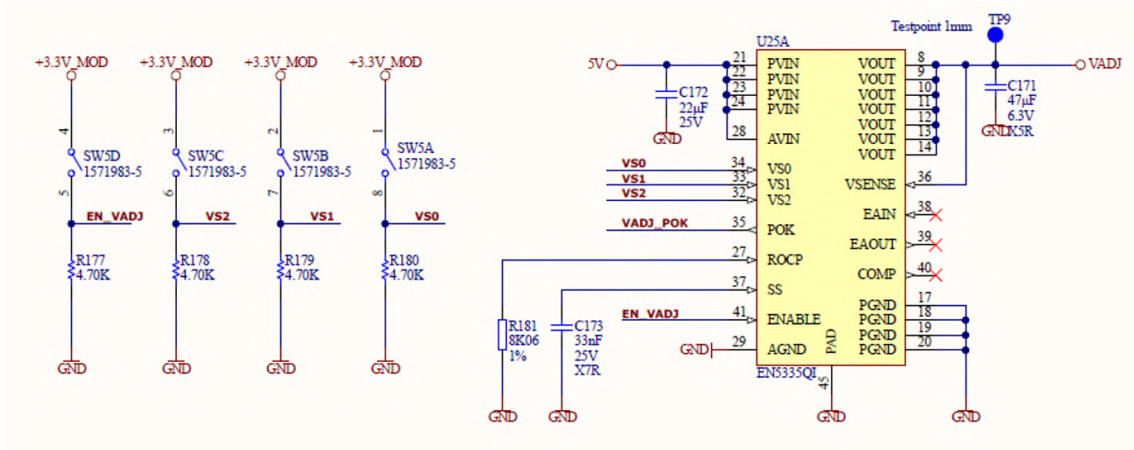
Figure 6.8: VCCO programmable power domain for the FPGA banks.

In the initial version of the carrier design, a fixed voltage of 1.8 V was set for $VCCO_n$ or V_{Adj}, which supplied power to the FPGA banks connected to the MSADC for the LVDS lanes and GPIOs. Although this configuration was not problematic, it was decided to incorporate a programmable power source for that voltage rail to allow for voltage level adjustments in the interface FPGA banks if necessary. For this purpose, the EN5335QI IC from Enpirion [113] was selected as part of the design. The decision was based on reference designs provided by Trenz, specifically from their carrier design TEBF0808 [114].

The chip integrates the power MOSFET and inductors; therefore, minimum components are required for the design, which can deliver up to 10 W of continuous power and achieve seven pre-programmed voltage levels. Figure 6.8 shows a schematic of the programmable voltage regulator and switches for controlling the output voltage levels and enabling operation, as shown in Table 6.6.

The LEMO interfaces were tested according to the standards to which they belong. LEMO input is prepared for working with a 100 Ω load and input voltages over 2.8 V for a logic "1" or high state. For the implementation, an LVDS buffer is used, so that it can be directly connected to a differential FPGA input. The LEMO output operates according

Table 6.6: EN5335QI switches positions for voltage programming.

Output Voltage	VS0	VS1	VS2
0.8 V	0	1	1
1.2 V	1	0	1
1.25 V	0	0	1
1.5 V	1	1	0
1.8 V	0	1	0
2.5 V	1	0	0
3.3 V	0	0	0
reserved	1	1	1

to the NIM standard when loaded with 50 Ω.

To test the USB and Ethernet functionalities, customized designs were developed by utilizing Xilinx examples for the UART and embedding a TCP server on the SoM. These designs enabled the testing of the USB-UART and Ethernet interfaces.

The testing process involved connecting the carrier board to a PC. Through these interfaces, the operation was evaluated at maximum speeds. The USB-UART interface successfully operated at a baud rate of 115200 bps, while the Ethernet interface operated at 1 Gbps. Both interfaces performed as expected, demonstrating their proper functionality during the testing phase.

Two different approaches were employed to conduct the testing of the SFP+ channels. The first approach involved using an SFP+ loopback connector, which enabled the evaluation of each channel individually. For the second approach, an SFP+ cable was used to perform inter-channel testing.

The SFP+ module is directly connected to the GTH transceivers, offering the possibility of using a specialized IP Core tester provided by Xilinx. This IP Core enables the testing, evaluation, and monitoring of the SFP+ operation. It encompasses various

essential resources, such as pattern generators and checkers, which are implemented within the FPGA. Additionally, it grants access to the configuration and status registers of the GTH transceivers through a dynamic reconfiguration port (DRP) [115]. The DRP can be accessed at runtime using the JTAG interface, enabling dynamic reconfiguration of the transceiver settings and parameters.

The core is called Integrated Bit Error Ratio Tester (IBERT) [116], and is a proprietary and licensed Xilinx IP core. The TE0803 SoM, used for the tests is based on the XCZU4EG, a device included in the free Vivado Webpack edition. Thus, it is possible to instantiate IBERT and test the interfaces without paying any fees for implementation. In addition, Trenz offers a prebuilt IBERT project for TEBF0808, which was adapted for this carrier. The only modification made for the FFeCCa was an I2C remapping, required to program the external low-jitter oscillator included in the SoM. This remapping was necessary due to a different position in the B2B connectors.

The testing procedure for evaluating the bit error ratio (BER) involves stressing the lanes with a known sequence of codes at a speed of 12.5 Gbps. The evaluation is performed by assessing the error at all possible combinations of voltage and phase steps in the GTH configuration. This means that the test systematically varies the voltage and phase settings to analyze the system's performance across different configurations and determine the bit error ratio in each of the combinations. The result of that test is plotted as an eye diagram.

The eye diagram scan provides important metrics such as the open Unit Interval (UI) percentage and open area, which are indicative of signal quality. The UI represents the time interval within a period that is used to evaluate errors in the signal. From Figure 6.9 it can be seen that all four channels are suitable for working at a maximum speed of 10 Gbps with approximately 55% of the UI in the safest operative area (blue) for 12.5 Gbps. Considering the use of a 3 dB attenuator for rates exceeding 10 Gbps during the tests and the planned operation of the SFP+ at 6.25 Gbps, the margin for reliable signal
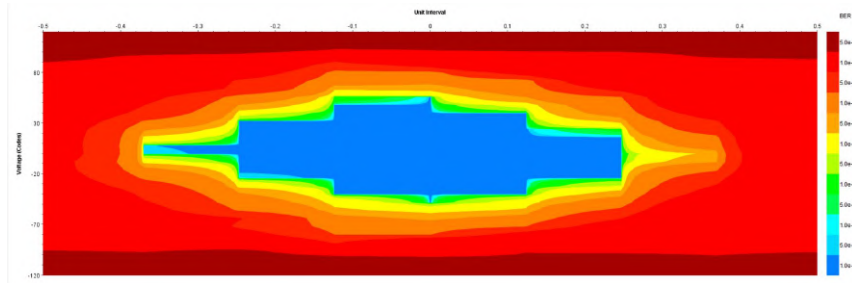
(a) CH0 loopback.
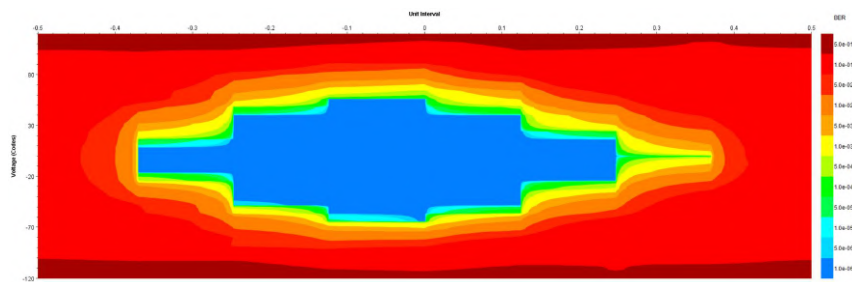
(b) CH2 loopback.
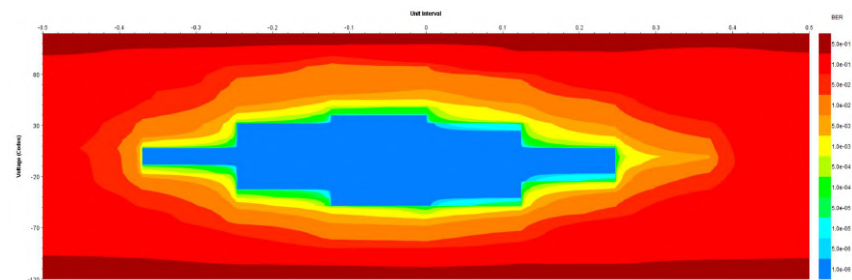
(c) CH2 loopback.
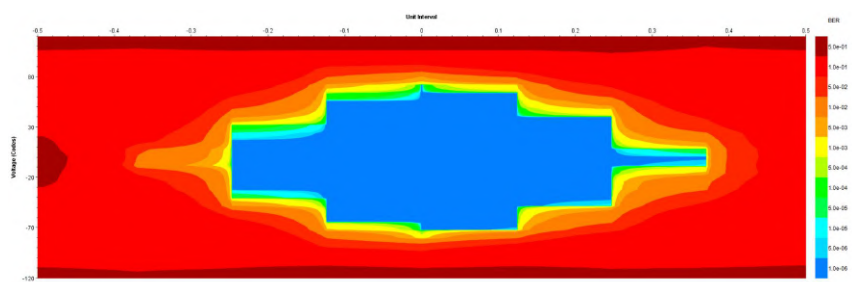
(d) CH3 loopback.

Figure 6.9: IBERT loopback tests on each of the SFP+ channels.

111

transmission becomes even more significant. This means that the results obtained during the tests provide a greater level of safety and assurance for operating at 6.25 Gbps.

A second version of the carrier board was designed with the inclusion of some new characteristics. In between the inclusions it can be mentioned a connector for a cooling fan (+5 V and +12 V), two PMODS, 8 dip switches for identification, 4 user LEDs and a micro tact-switch.

6.2 Summary

The FFeCCa board underwent comprehensive validation by rigorous testing of all the power sources, interfaces, and characteristics. This thorough validation process confirms that the board effectively and efficiently fulfills the requirements set forth by the AMBER proposal. Thus, providing a reliable solution for the project's upgrade on the new free-running data acquisition system.

The FFeCCa board integrates the existing ECAL2 digitizer board (MSADC) with a high-performance SoM based on the Xilinx Zynq Ultrascale+ MPSoC family. The SoM can handle real-time signal processing, perform data feature extraction, and implement protocols for consistent integration with the DAQ of the experiment.

The FFeCCa design was driven by considerations such as cost-effectiveness, ease of redesign, and adaptability to future projects. By adopting a SoM-based approach with a standardized interface, the system allows for flexibility in selecting the SoC device and the desired memory capacity while optimizing production costs. Additionally, selecting a SoM simplifies complex tasks such as SDRAM routing, power source design, power-on-matrix sequence, and precise clock generation.

To accommodate the MSADC, SoM, high-speed serial interfaces, power generation, and cooling the FFeCCa PCB was implemented in a large form factor. This approach ensures that production costs remain low.

The project can be found at `https://gitlab.com/brunovali/ffecca_MSADC`.

Chapter 7

Firmware and Software, Development and Implementations

The FFeCCa board, schematically depicted in Figure 7.1, although it possesses the necessary hardware components for operate safely and efficiently, to function as a free-running frontend for the ECAL2 is dependent on the logic design and its programming.

The firmware encompasses the implementation of necessary algorithms, control logic, and communication protocols to enable data acquisition, signal processing, and transmission.

The SoM works as the central processing unit that orchestrates the operations of the FFeCCa board and facilitates interactions with external devices.

The MSADC, which is also mounted on the FFeCCa board, requires as well to be programmed before it becomes operational. The programming process configures the MSADC to perform the desired analog to digital conversion according to the specific requirements of the ECAL2 system.

Once the firmware is configured and programmed into the FPGAs and processors, the DAQ can begin operating. Some tasks and functionalities are common to most DAQ systems based on SoC-FPGAs, regardless of the type of processing or ADC being read.
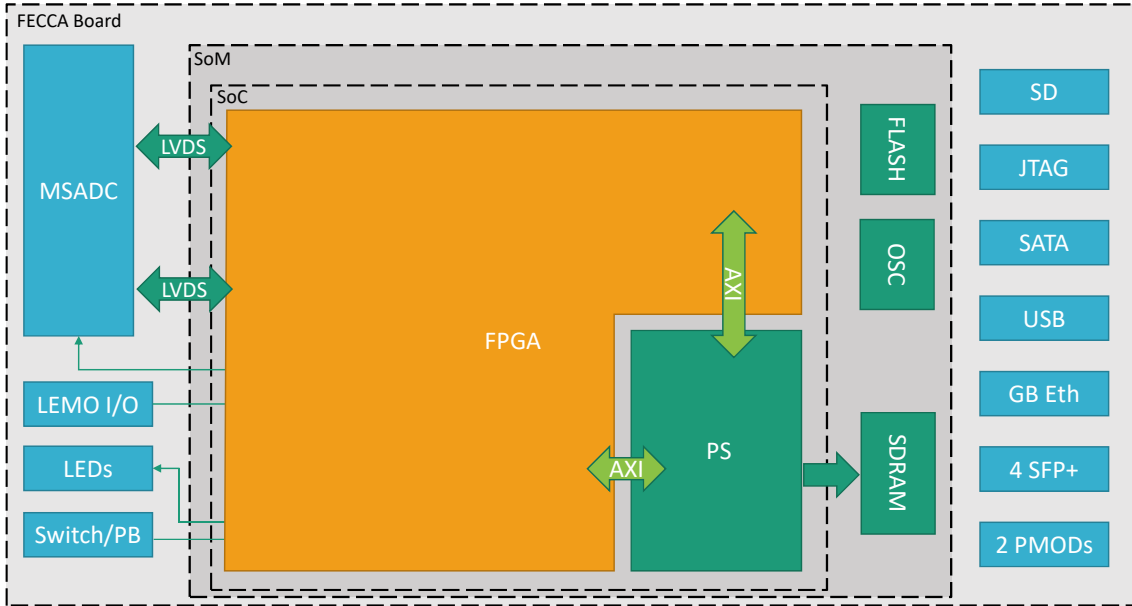
Figure 7.1: Hardware resources of the FFeCCa with the MSADC and SoM mounted.

From the firmware and software perspective, to speed up the developments it is useful to have a base starting point for covering those common tasks and functionalities. By this way saving time when starting a firmware project and avoiding rethinking the architecture each time there is a need for implementation. Thus, a complete framework for working with SoC-FPGA-based DAQs was developed. The framework is composed of software, firmware, and all the IPs required for interfacing with multichannel ADCs, packetizing the data, receiving it from a PC, and plotting it in a straightforward manner.

The value and effectiveness of the FFeCCa board reside not only in its hardware design but also in the logic design and programming, plus the software that enables a fully operational DAQ system. Collaboration between hardware and software components is essential for harnessing the potential of the FFeCCa board and executing its intended application in the ECAL2 system. In addition, an Open Software Framework for DAQ working together with FFeCCa offers the possibility of fast and easy adoption.

This chapter presents the framework and different implementations of the firmware

for the DAQ, together with the base software. Each implementation is tuned according to the scope functionality (testing or ADC data reading), number of channels, and type of approach (triggered or trigger-less).

7.1   Open Framework for SoC-FPGA DAQ based platforms

A complete framework for SoC-FPGA based DAQ systems is presented in this subsection, the Open Framework for SoC-FPGA DAQ (OFSODA). The solution is based on a modular design, allowing the operation of multichannel ADC systems and presenting minimum effort when migrating to different ADC systems and between diverse FPGA families. It is thought to efficiently use the resources within the FPGA, maximizing performance and efficiency. By providing optimized and validated IPs, it ensures that FPGA resources are effectively utilized, enabling the system to handle high-speed data processing and transmission.

The framework allows for easy customization and expansion to accommodate the requirements. New features and functionalities can be easily integrated into the framework without disrupting the overall system architecture.

Modularity allows for reuse in different designs, which saves time and effort during development. This also makes maintenance and updating easier. Thus, it is also easier to scale the design up or down and simplify the testing of the developments by identifying and isolating the errors more quickly.

Figure 7.2 shows a simplified block diagram of the framework. As a template, it comprises all the necessary sub-systems for reading ADC channels in a free-running mode from a PC connected to the same local area network. All the Very High Speed Integrated Circuits Hardware Description Language (VHDL) cores, the firmware for the micro-processor (μP) and the Python scripts for interfacing with the system from an external PC were developed and tested.

Initially, the framework was deployed and tested using a CIAA-ACC board.
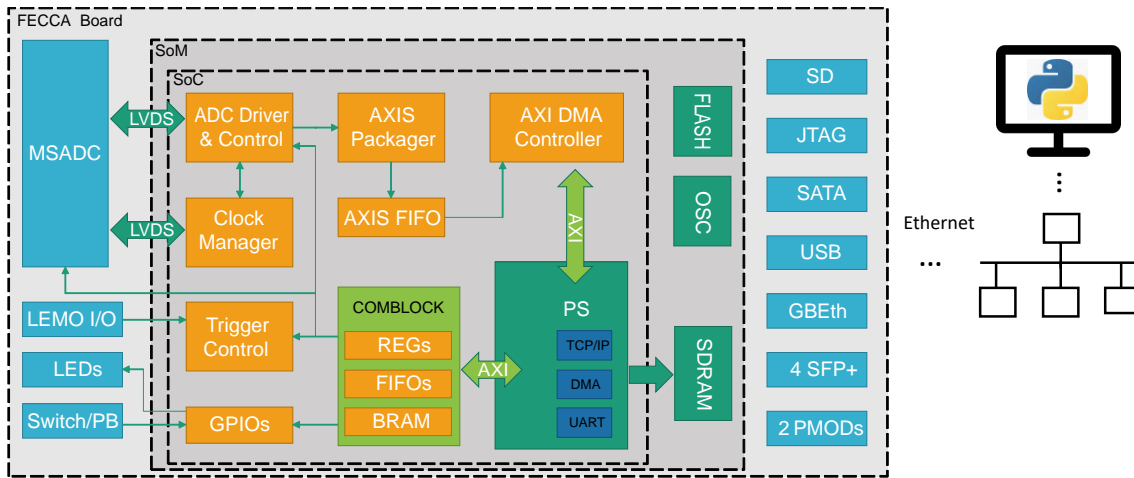
Figure 7.2: Block diagram of SoC-FPGA Framework.

Nevertheless, the Framework was also ported to a ZedBoard [117], based on the Zynq-7000 series and to the FFeCCa, which is based on Zynq Ultrascale+ architecture.

7.1.1 FPGA Firmware

The FPGA VHDL descriptions of the Framework are organized as shown in Figure 7.2. When possible, the cores were deployed such that all hardware generated by the design tools is inferred from the VHDL descriptions, thus ensuring the highest possible portability.

To control, configure, and read the status of the FPGA design, is used a custom IP block developed jointly by the MLab-ICTP group and INTI, called the COMmunication BLOCK (ComBlock) [118]. The ComBlock offers well-defined interfaces (registers, RAM, and FIFOs) to users of the Programmable Logic (PL), thereby avoiding the complexity of the bus provided by the Processor System (PS), which is AXI [119] in the case of the Ultrascale+ SoC-FPGA. It provides five interfaces for the user in the FPGA side, which can be individually enabled and customized, as depicted in Figure 7.3. From the μP's perspective, the ComBlock can be used to provide an AXI interface to FPGA IP cores,
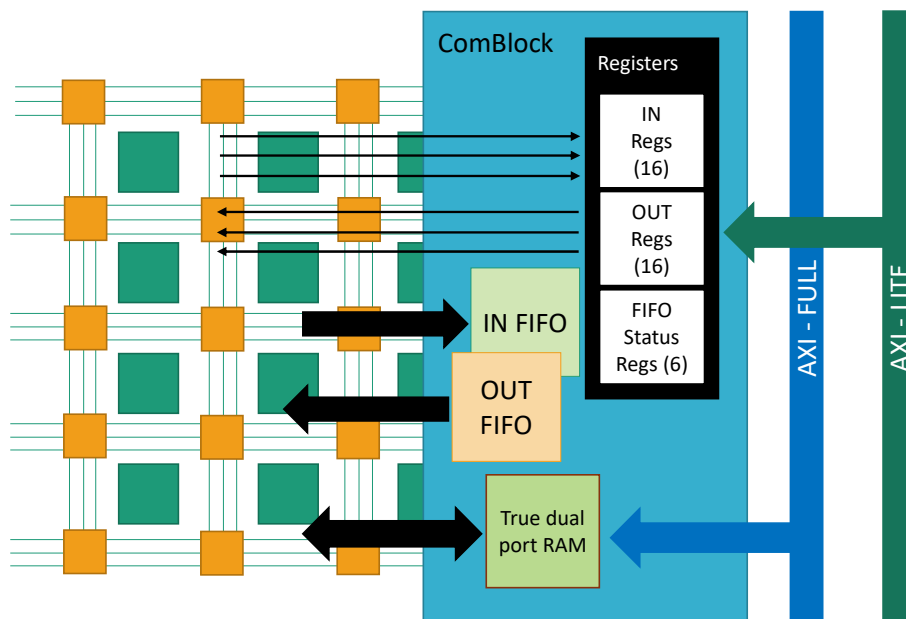
Figure 7.3: Communication block (ComBlock).

similar to if there were peripherals in a microcontroller. Along with the ComBlock IP core, the firmware for interfacing with the μP it is also provided.

In the implementation of the base framework, only the registers are instantiated, as they are used for interfacing at a slow speed.

For control three registers are used, associated with the following parameters: number of samples to read, active channels, and global reset. The ADC system reset and trigger source are mapped together in one register address.

The status registers are related only to the ADC driver and are: error code and valid channels.

The configuration registers are related to the data sourcing: ADC data or synthesized data.

In Figure 7.4b it is shown the control, status and configuration implementations for the system.

Figure 7.4c shows the clocking management, where a Mixed-Mode Clock
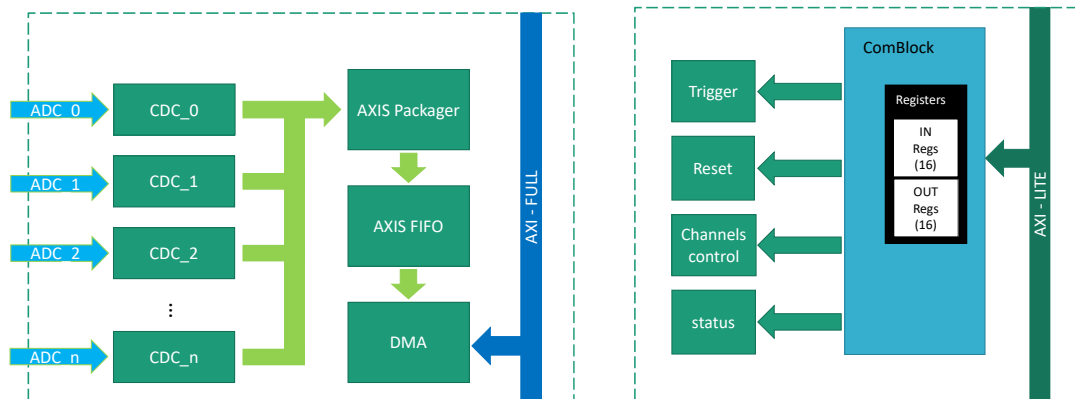
(a) Data handling.

(b) Control and status.
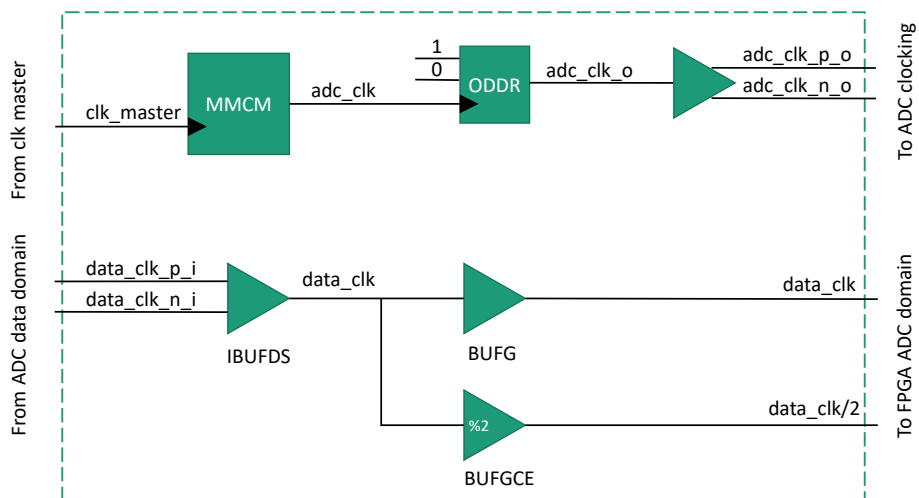
(c) Clock managing.

Figure 7.4: Detailed descriptions of the Framework IPs.

Management [120] (MMCM) is sourced with a clock from outside the core. This MMCM re-targets the clock frequency according to the requirements of the ADC (sent to the top-right side) using an ODDR through a differential output buffer. At the bottom, there is a clock coming from the ADC system, which will be synchronic to the serialized data. Thus, the entire design is synchronic with a main clock, with a frequency value that will be adjusted according to the ADC and other system requirements.

After being decoded with the ADC driver IP, see Figure 7.2, the data coming from the ADCs are sent to a small FIFO, used as Clock Domain Crossing, to capture the data in case there is some deviation in the phase of the clock coming from the ADCs. The clock period for reading the data from these FIFOs must be always faster or equal to the data period they are stored. As all clocks are derived from the main clock, there should not be frequency divergences between the clock domains. From Figure 7.4a it can be seen that the reading of CDC FIFOs is performed at the AXI Stream (AXIS) packager clock. The packager reads all FIFOs and packages the data in the AXIS protocol and saves them in a larger AXIS FIFO. The decision to use the last FIFO is to profit from the DMA controller IP offered by Xilinx. In this way, it can be launched the data reading and move it from the FPGA to memory without intervention and release the processor for other tasks, while moving the data to PS memory for transmission through Ethernet. This last task is done in the μP firmware implementation.

7.1.2 Processor Firmware

The SoC-FPGA μP firmware is based on the FreeRTOS Xilinx [121] implementation and lwIP [122] library for the TCP/IP when using Ethernet. The design is prepared to receive the configuration and control parameters from a PC using Ethernet. The firmware of the processor is in charge of handling the initial configuration of the FPGA part, and performing a reset to the entire system. Then implementing a TCP server, it receives the packages from port number 1000 and interprets them to configure the system and launch

the data acquisition.

The limit on the data packet length is 4.194.304 data points, because of the Xilinx AXI DMA implementation.

For data acquisition, as the ADC handling is made directly into the FPGA domain, the μP will be in charge of indicating the reset to all FPGA FIFOs, invalidating all its cache lines with a *flush* [123], launching the data take indicating the AXIS packager to start sending data to the AXIS FIFO, and launch the DMA operation. Once DMA is completed, the data packet is sent to the PC, and the TCP server is ready to attend another call.

The framework can also be used with the Universal DMA (UDMA), a remote control suite for interfacing a PC with custom logic in a SoC-FPGA, as explained in [124]. From the perspective of a PC the UDMA can be seen as a direct interface with the ComBlock resources. Thus, when using the system with the UDMA compared to the AXI DMA, the only constraint in the implementation is the number of samples to transmit. This is because of the limitation on the BRAM instantiated with the ComBlock for buffering, which depends on the SoC-FPGA resources.

7.1.3 Interface and Control Software

To interface with a PC and facilitate the data acquisition Python scripts were developed and implemented. These scripts allow for selecting the channels to retrieve and choosing the data source, either from internally synthesized signals or from the ADCs. In addition, for testing the communication the scripts check the integrity of the data when the FPGA synthesizer is selected.

The script for configuring the system and launching the data taking is called **adc_read.py** and allows the user to customize the board IP address and port, select the channels to take data, amount of data and source in a simple way.

The script, for a two ADC channels streamer implementation, accepts the following

options:

python adc_read.py −a [IP] −p [PORT] −f [FILE_NAME] −c1 [CHANNEL1]
                    −c2 [CHANNEL2] −s [SAMPLE_QUANTITY] −t −ti −r

where:

- **-a**: IP address of the board, default is 192.168.10.10.

- **-p**: Port of the service for data retrieving, default is 1000.

- **-f**: Output file name for the data read, default is *samples* and generates two different files, one in binary and the other in ASCII.

- **-c1**: Channel 1 select, is hexadecimal and must be in the range 0:F, default is 0.

- **-c1**: Channel 2 select, is hexadecimal and must be in the range 0:F, default is 0.

- **-s**: Quantity of samples to retrieve, it can be in numbers or with the specifiers k or M for kilo or Mega, default is 1024. The maximum value is 4M data points.

- **-t**: For test mode enable, in this mode a +1 step increment for each sample is generated inside the ADC board.

- **-ti**: For test mode enable, in this mode a +1 step increment for each sample is generated inside the SoC.

- **-r**: Applies an entire system reset.

The other script is a simple plotter, for fast verification of the data in a visual mode. For plotting the data is only necessary to call the script as:

python plot_simple.py file_name.bin

If no file name is provided then it will search for a file called *samples.bin*, according to the default file name of the script for launching the data taking.

7.2 LVDS Testing and Charaterization

The maximum data rate of the communication between the MSADC and the Soc-FPGA is a critical topic in the design since they must be able to sustain the raw data rate from the ADCs for processing at the SoM level. The LVDS lines (channels) are driven according to the output characteristics of the Virtex-4 (XC4VLX25) when the data is sourced from the MSADC. According to the DC and switching characteristics specified in the XC4VLX25 datasheet [87], LVDS is capable of working up to 800 Mb/s if configured in DDR mode for the lowest speed grade device. However, the actual speed may be lower due to poor signal integrity in the transmission lines or a suboptimal PDN design. As discussed in Section 5.2, there are 20 LVDS connected to the FPGA for the interface, and at least two of them must be used for clocking and synchronizing the data. Therefore, the maximum theoretical data rate can be calculated as follows:

$$18 \ \times 800 \text{ Mbps} = 14400 \text{ Mbps} \tag{7.1}$$

When working at 80 Mhz, each MSADC channel stands a raw data rate of:

$$12 \text{ bits} \times 80 \text{ MHz} = 960 \text{ Mbps} \tag{7.2}$$

For the 16 channels the overall data rate is:

$$960 \text{ Mbps/ch} \times 16 \text{ ch} = 15360 \text{ Mbps} \tag{7.3}$$

The raw data rate from the ADCs surpasses the transmission capacity of LVDS outputs to the external SoC-FPGA, as shown by Equation 7.3 and Equation 7.1. Therefore, it is crucial to determine the actual maximum data rate of the MSADC. Without this information, it is not possible to determine the number of channels that can be transmitted to the SoM or assess the need for data compression and the required compression rate.

To determine the maximum data rate that the MSADC can handle when transmitting data to the external SoC-FPGA through LVDS outputs, a specialized test was conducted. The purpose of this test was to assess data integrity and stress the outputs.

The test procedure involved gradually increasing the transmission frequency of a known number sequence until errors started to occur. The number sequence was generated using a free-running counter, which would reset after reaching its maximum value. The counter was driven by a programmable clock frequency, and both the data and clock signals were transmitted together.

This testing was performed within the MSADC itself, as depicted in Figure 7.5a. By incrementing the transmission frequency until data corruption was observed, the maximum achievable data rate of the MSADC's LVDS outputs could be determined.

During the data reading process, the OFSODA was utilized. Specifically, the block diagram depicted in Figure 7.5b was employed for the ADC Driver and clock generation.

The clock frequency responsible for data generation was controlled from the SoM side using a clock generator referred to as *data_clk*. Within the MSADC, this frequency was multiplied by a factor of 4 to achieve serialization of the data, resulting in a faster clock referred to as *fast_clk*.

Figure 7.5c illustrates the timing diagram that showcases the interaction between the clock and data within the MSADC implementation. The generated data was directly fed to the OSERDES alongside the *data_clk* and *fast_clk* signals, required for the serialization process.

The maximum data rate without corruption was characterized for both the Adapter Board and the FFeCCa. The Adapter Board was able to work up to a frequency of 300 MHz while the FFeCCa was able to work up to 340 MHz. A ramp signal generated at 85 MHz and serialized at 340 MHz (680 Mbps) on the MSADC was successfully received in the FFeCCa and sent to the PC without any data corruption. However, when the same ramp signal was generated at 87.5 MHz and serialized at 350 MHz (700 Mbps), data
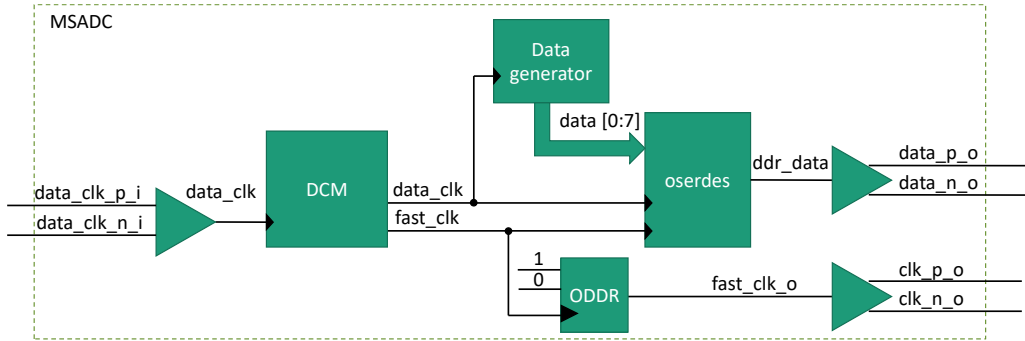
(a) MSADC VHDL block diagram for data generation.
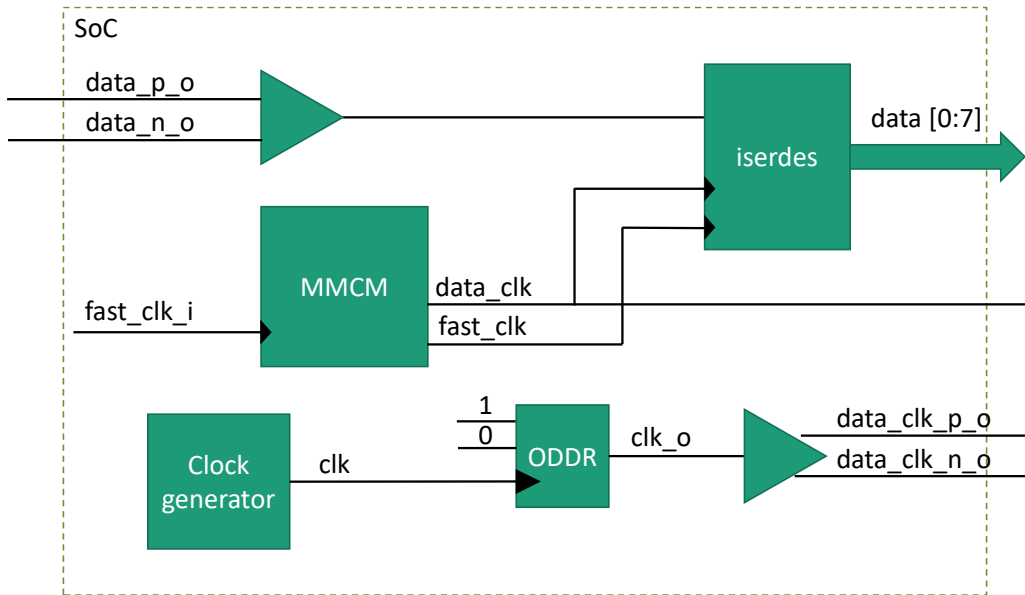
(b) ADC driver block diagram for the data decoding.
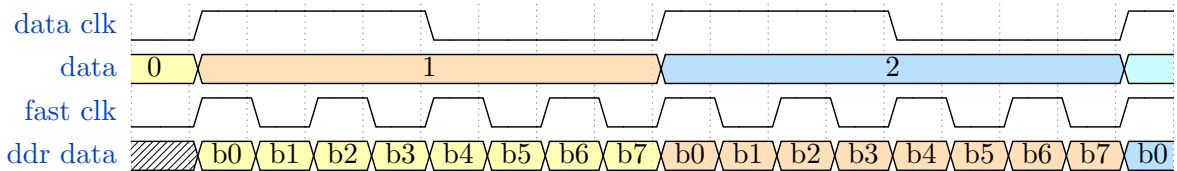
(c) Data and clocks in the serdes scheme.

Figure 7.5: Block diagram and signals in the serdes scheme.

124
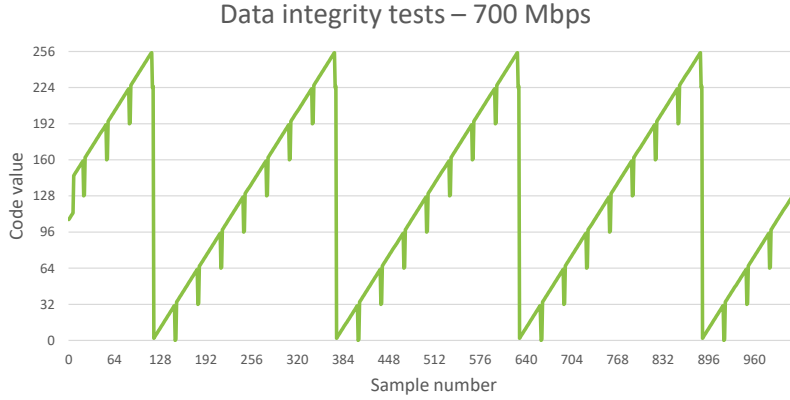
Figure 7.6: Frequency limit when data corruption starts to be evident.

corruption was observed, as shown in Figure 7.6. Additional information will be provided in Chapter 9. It should be noted that the data are asserted in the data taking script once received and the plots are used for visual representation purposes.

When trying to implement the LVDS lines in the MSADC FPGA design, it was found that there are 4 of 20 lines that can't be routed as differential pairs, so there are actually not 18 LVDS free for routing, but 16.

The overall data transmission capabilities for the MSADC-SoM (FFeCCA) is:

$$DR_{max} = 16 \times 680 \text{ Mbps} = 12.2 \text{ Gbps} \tag{7.4}$$

From Equation 7.3 the raw data from the ADCs is 15.36 Gbps.

From the tests is concluded that for sending the 16 channels from the MSADC to the SoM, it has to be done data compression in a factor of at least:

$$C = \frac{DR_{raw}}{DR_{max}} = \frac{15.36 \text{ Gbps}}{12.20 \text{ Gbps}} = 1.26 \tag{7.5}$$

7.3 DAQ Implementations

From the firmware point of view the DAQ implementations can be grouped as:

a) MSADC firmware

b) SoC firmware (FPGA descriptions + **C** code)

To ensure a consistent description of the implementations, we will first provide an overview of the base MSADC project. The implementations for the SoC were done using the OFSODA as base design. Subsequently, it will be discussed the specific variations based on the number of channels and transmission schemes. This approach intends for a systematic and organized presentation of the different implementations.

7.3.1 MSADC Firmware

The MSADC firmware is an adapted version of the current implementation of the ECAL2 readout [125]. As a brief description, the MSADC original firmware was in charge of: configuring the ADCs, reading the 16 channels in parallel, making a small pipeline buffer and sending the last 32 samples each time there was a trigger.

In Figure 7.7 the new firmware implementation is presented as a block diagram. The main clock is provided by the SoM system and is fed directly to the clock and reset generator block. From this, all subsystem's clocks are generated using Digital Clock Managers (DCMs) to have fine control of the frequencies, phases and starting sequences. There will be at least three different frequencies, 40 MHz for feeding the ADCs, 80 MHz for transmitting synchronous with the effective data rate, and 5 MHz for the configuration of the ADCs. As all the clocks are derived from the same source, there will be no frequency drift between them. This last characteristic is important, due to the full synchronous approach is minimized the need for data buffering for cross-domain clock purposes, saving lot of resources.
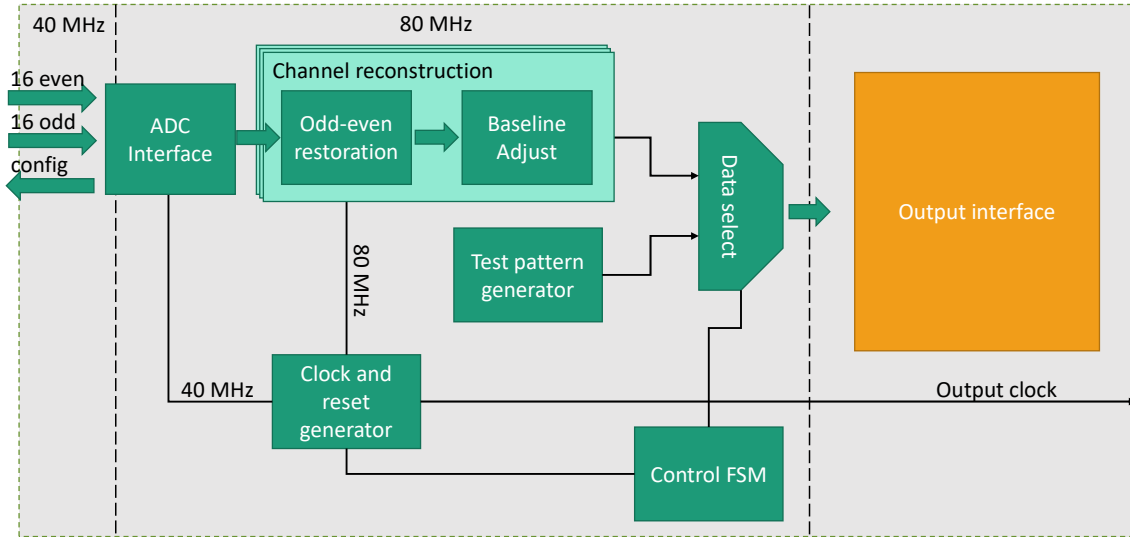
Figure 7.7: Block diagram of new MSADC FPGA firmware for free-running mode.

The *ADC Interface* core is responsible for configuring all 32 channels of the 4 ADCs, ensuring that the received data is correctly paralyzed and framed according to the methodology described in [126]. Each of the 16 analog channels is read by two logic ADC channels, which are referred to as Even and Odd channels. The Even channels are fed with a 0° clock phase, while the Odd channels are fed with a 180° clock phase, both from the same 40 MHz clock. The *ADC Interface* core interleaves the logic data channels, combining them as if they were sampled at double frequency by a single logic channel.

However, since the logic channels can have different offset levels due to belonging to different physical ICs, there is an *odd-even restoration* level performed after the interface IP to balance any discrepancies in the baseline levels. Following this, there is a *Baseline adjust* to bring all channels to a common base level of 50 counts.

Then, the *data select* block is responsible for multiplexing and selecting the source of the data. By default, the data is streamed from the ADC chain. However, it is also possible to connect the input of the multiplexer to a test pattern generator, which enables checking the data integrity in the communication between the MSADC and the SoM.

Table 7.1: MSADC base design resources utilization.

Resource	Utilization (%)	Available
REGISTERS	1030 (9)	10275
LUT	1247 (5)	21504
BRAM	3 (4)	72
BUFG	11 (34)	32
DCM	4 (50)	8
IO	108 (24)	448
SERDES	16 (4)	448

Following the interleaving process and up to this point, the logic works with an 80 MHz clock frequency.

The base MSADC firmware utilizes various hardware resources, as shown in Table 7.1. This table provides an overview of the resource utilization in the base MSADC firmware, indicating the utilization of different resources such as LUTs, register (FFs), BRAM, DCM, BUFGs, SERDES and IOs.

The base design is then customized to meet the specific requirements of each implementation by incorporating the output streamer scheme. This enables the transmission of multiple channels based on the transmission capabilities of the physical constraints, such as available bandwidth or interface protocols.

7.3.2 Two Channels Streamer

The first implementation of the DAQ working in free-running mode was deployed for two channels. As before, the firmware is presented into two domains, one related to the MSADC and the other to the SoC, as shown in Figure 7.8.

In the MSADC implementation (Figure 7.8a), the process of selecting specific channels follows the *Channel Reconstruction* stage. A two-channel selector is employed to choose any two independent channels from the 16 available ADC streams. Additionally, a
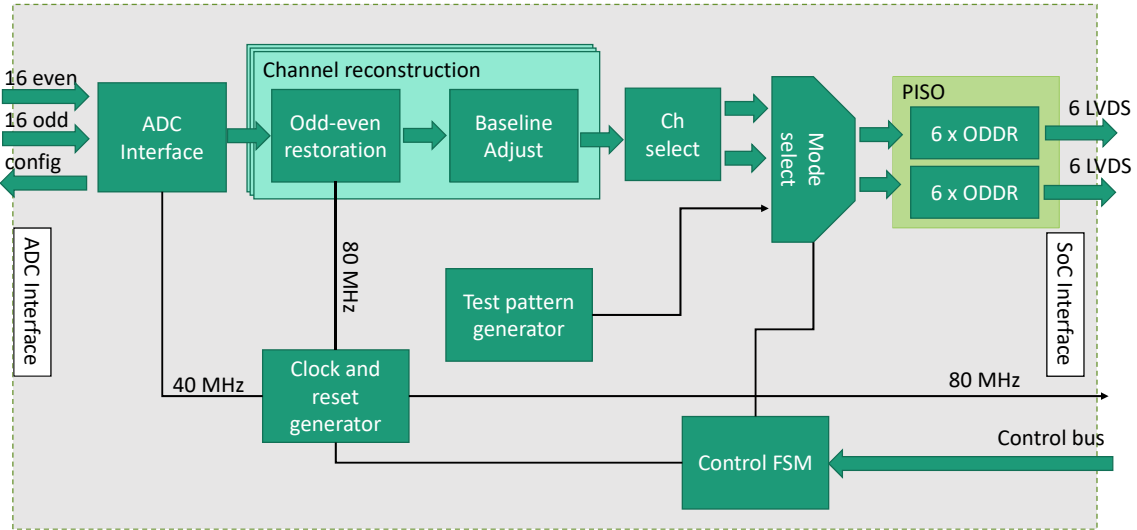
(a) MSADC VHDL block diagram for 2 channels streamer.
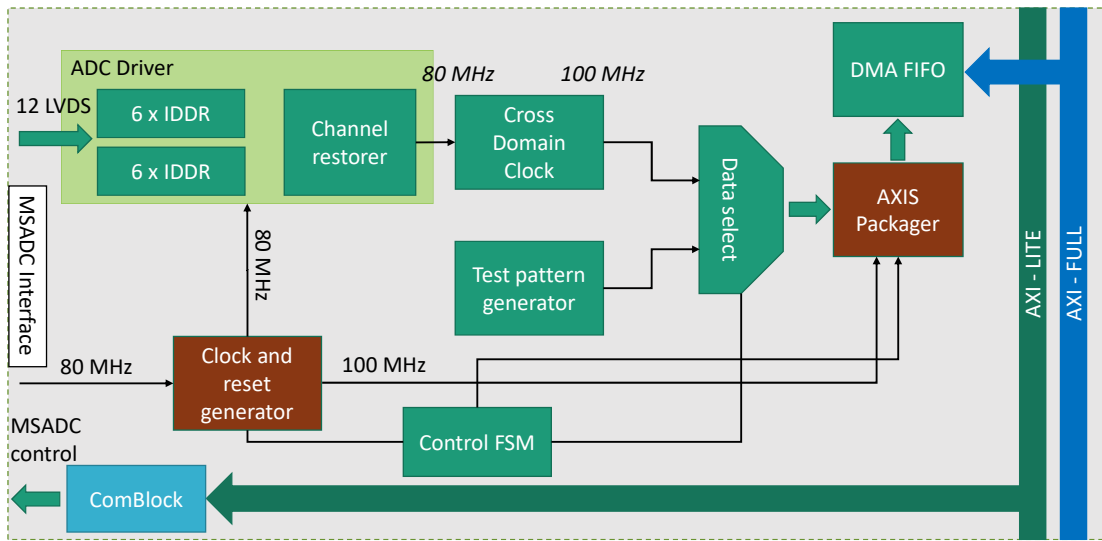
(b) SoC FPGA block diagram for 2 channels free-running mode.

Figure 7.8: Two channels streamer firmware implementation.

mode-select multiplexer is included in the signal path to determine the data source. The resulting output is then transmitted using six ODDRs per channel, utilizing LVDS lines for transmission to the SoM.

The configuration parameters for data sourcing, reset, channel 1, and channel 2 are received from the SoM via single-ended GPIOs of the MSADC joined in a Control Bus.

The SoM implementation can be summarized into three main parts: the ADC Driver plus the MSADC control interface, the clock and reset generator, and the bulk data transfer section.

The *ADC Driver* is in charge of receiving the data streams from the MSADC, aligning the bits from the IDDRs and assembling the 12 bits words at 80 MHz. There is a cross-domain clock FIFO for synchronizing the data from the ADC clock domain (80 MHz) to the system clock (100 MHz). To control the MSADC the ComBlock is used, connected to the AXI-LITE interface, providing an interface with the μP and acting as a slow control interface.

The *Clock and Reset Generator* is responsible for generating the 40 MHz main MSADC clock, receiving the 80 MHz data clock, and resynchronizing that last clock with the received data.

The bulk data transfer section is done using the OFSODA facilities, using the DMA resources.

7.3.3   Eight Channels Streamer

For the eight channels implementation as in Figure 7.9, it was changed the channel transmission scheme. Now, instead of using the ODDR-IDDR scheme, the SERDES approach was used.

To overcome the limitation of the Ultrascale+ SERDES architecture, which only supports serializing words of 8 bits in DDR mode, a gearbox circuit was implemented in the MSADC. This circuit allows for the transmission of 8 channels of 12 bits each into 12 SERDES channels of 8 bits. The *gearbox* is part of a Parallel Input Serial Output (PISO) core, as shown in Figure 7.9a. The PISO core also handles the transmission of an initialization data sequence, which is necessary for adjusting the bit framing boundaries.

This sequence is always sent after a reset to resynchronize the channels or, anytime under request. The synchronization process is done at OSERDES-ISERDES pair connection and is required from the paralyzer (SoM side) to adjust the received 8 bits boundaries. The OSERDES are fed with the data clock and a fast clock of four times the data frequency for serializing the data in DDR. As the data frequency is 80 MHz, the fast clock frequency for serialization is 320 MHz and is forwarded to the SoM to recover the data.

In the SoM design, each of the ISERDES is accompanied by a *bitslip* core for readjusting the bit-0 word frame alignment. The *bitslip* core is enabled during the synchronization stage and aligns automatically the 8-bits word in two clock cycles. Once the data is aligned the bitslip remains active, readjusting all the data flow that the ISERDES paralyze. Then to reconstruct the original 12 bits words, there is the *inverse gearbox* for going from 12 channels of 8 bits to the 8 channels 12 bits width.

This implementation allows for transmitting half of the channels of the MSADC to the SoM in a continuous mode and without making any compression. The design is prepared for free-running implementations.

7.3.4 Sixteen Channels, buffered and triggered

The 16-channel buffered and triggered approach expands the base MSADC design by incorporating 16 FIFOs. These FIFOs operate in parallel to buffer the data from all 16 channels, enabling continuous and synchronized data reading. The system is triggered externally, ensuring synchronized data acquisition across all 16 channels. As shown in Figure 7.10, the IP Core is in charge of buffering the 16 channels, processing the trigger signal, and transmitting the data to the SoM through a single data interface channel. The system continuously buffers data from all 16 channels in parallel and stores it in a 4096-sample FIFO. The number of samples to buffer before triggering can be programmed at any time to meet the specific requirements of the experiment. This approach facilitates efficient data collection and ensures accurate synchronization across all channels.
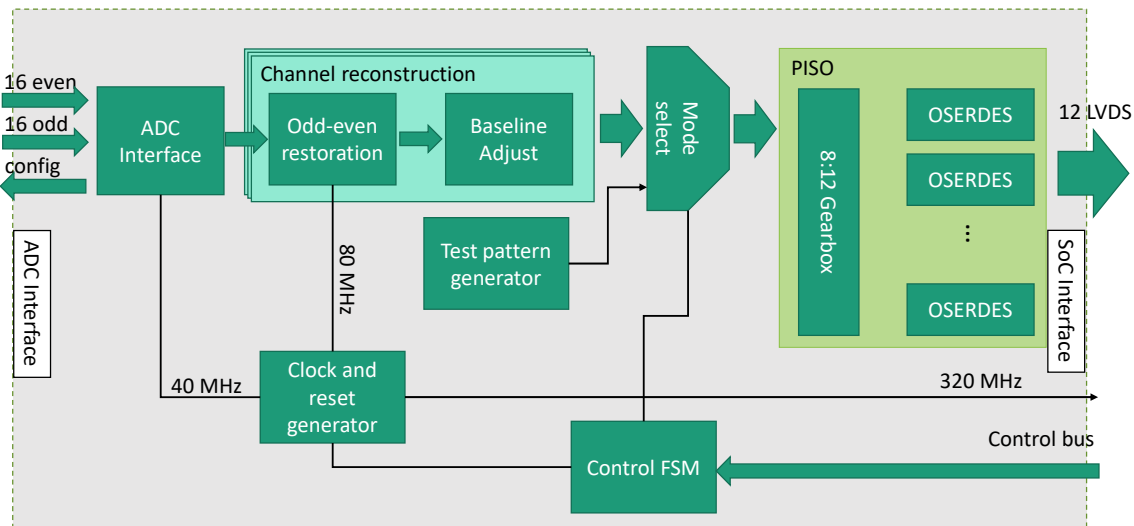
(a) MSADC VHDL block diagram for 8 channels streamer.
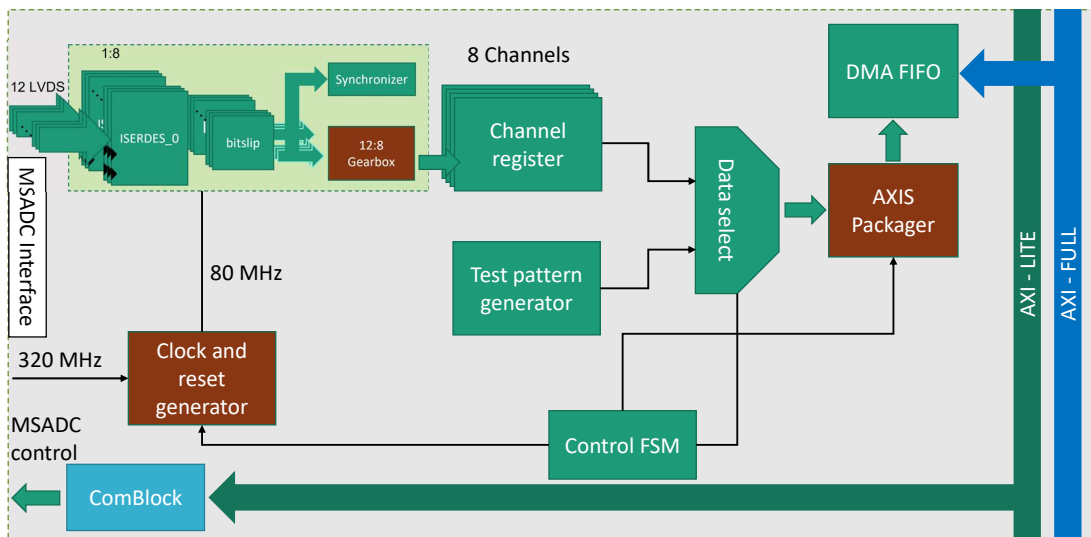
(b) SoC FPGA block diagram for 8 channels free-running mode.

Figure 7.9: Eight channels streamer firmware implementation.
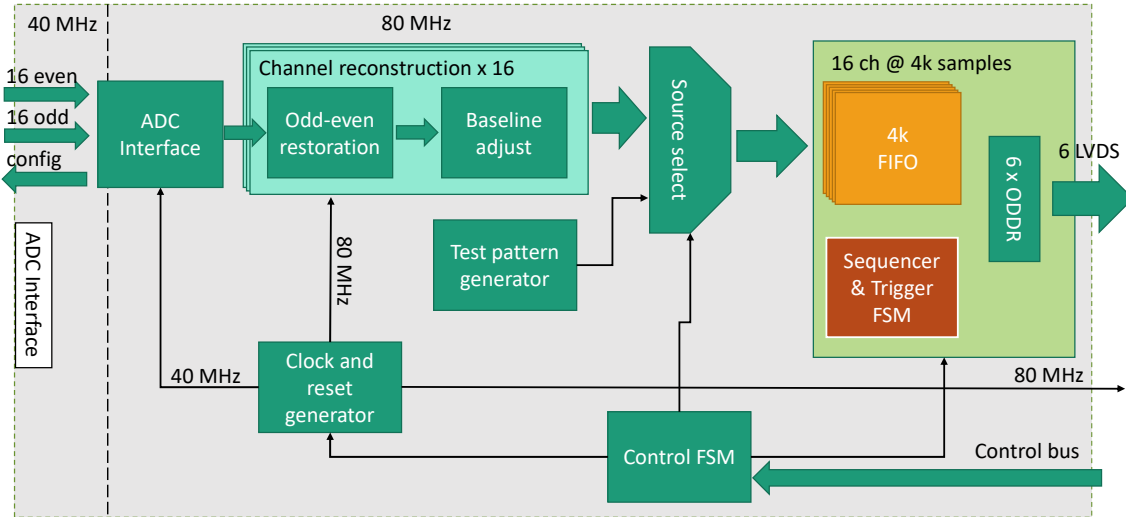
Figure 7.10: MSADC sixteen channels 4096 samples implementation.

Internally, when a trigger signal arrives, after a programmable number of samples the IP Core stops the buffering process and sends the buffered data to the SoM. The number of samples to transmit is always constant and equal to 4096, the number of samples after triggers allows for the selection of the region of interest (4096 samples window) in the acquired data.

After completing the buffering, all the data are sent sequentially from FIFO 0 to FIFO 15 through a single MSADC-SoM interface channel, similar to the 2-channel implementation.

From the prospective of the SoM, it is used the same design of the 2-channel streamer, with a fixed packet length of $4096 \times 16 = 65536$ samples and changing only the interpretation of the decoded data once in the PC domain.

Since the data now comes from 16 channels, the interpretation of the decoded data is different from the 2-channel streamer. The data needs to be demultiplexed to separate the samples from each channel and then combined into a single data stream for each channel.

This process is done in software once the data is received in the PC domain. The

software reads the decoded data packets, demultiplexes the data, and combines it into a single data stream for each channel, which can then be analyzed and processed as required.

7.3.5   Sixteen Channels Streamer

As introduced in subsection 7.2, it is not possible to transmit the 16 channels from the MSADC to the outside in raw mode as a continuous data stream. To transmit all the data from the MSADC to the SoM, there is the need for some type of compression. From Equation 7.5, this compression must have a rate at least 1.26 times in the case data are transmitted at the highest frequency stand by the LVDS lines (360 MHz).

The use of a lossless compression technique to preserve data characteristics is indeed a necessary requirement for real-time analysis of signals shape, pulse detection, and amplitude measurement. This is because any loss of data during compression could result in inaccurate or incomplete analysis results.

Huffman coding, a well-known technique for lossless compression [127], employs a variable-length prefix coding algorithm. It assigns shorter codes to frequently encountered symbols and longer codes to less frequently encountered symbols. This approach optimizes the representation of data, resulting in efficient compression while preserving the original information.

The development of the CODEC scheme was a collaborative effort with the University of Warsaw, and was optimized for the statistics of the detector. The compression is carried out using a two-stage approach. In the first stage, a derivative is applied to the signal, reducing the symbol length to the differences between consecutive samples. While the second is performed by applying a Huffman codification to the 64 most probable values, as in Table 7.2. The statistics of the ECAL2 response showed that the best input for the Huffman codification is the first derivative.

For example, for a typical trace with pulses and noise as in Figure 7.11a, the histogram of values for the first and second derivatives are shown in Figures 7.11b and
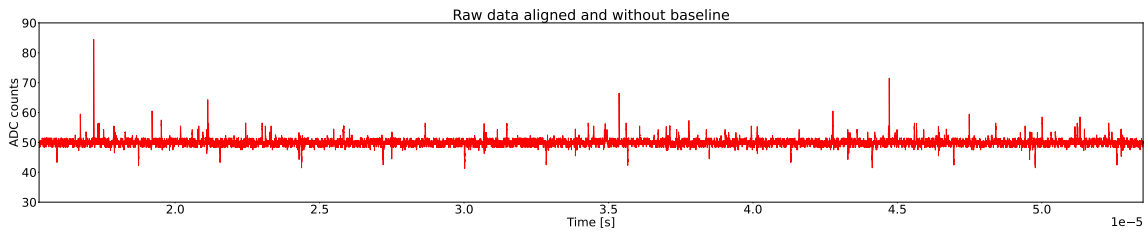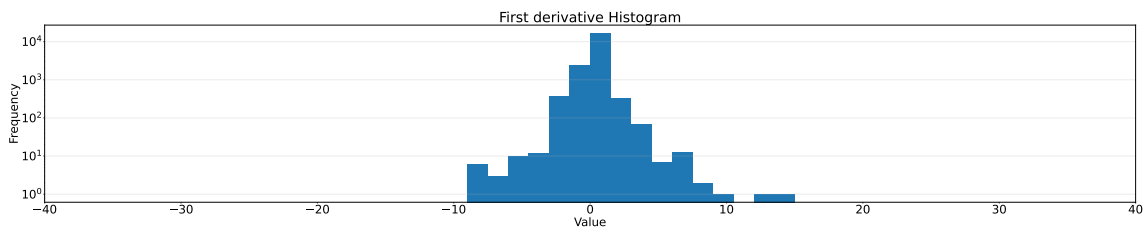
(a) Raw data.

(b) First derivative histogram.
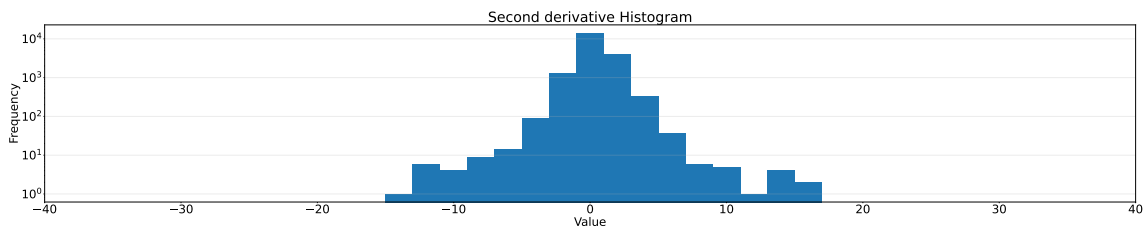
(c) Second derivative histogram.

Figure 7.11: Raw data trace and histograms of the first and second derivative.

7.11c. From the figures can be concluded that the first derivative performs a better first-stage code compression, as all histogram bins are distributed closer to zero than the second derivative. For the values out of the 64 codes, the raw data plus an identifier is transmitted, incurring a penalty but these values are expected with a low probability of occurrence.

In the MSADC base design, the multiplexer feeds the encoder and the output is serialized using ODDR's. The encoder's outputs work at a 200 MHz clock frequency for a safer margin operation. Each channel is serialized in a single ODDR and is connected through this to a differential buffer configured to work in LVDS.

135

Table 7.2: Huffman codes assignation.

Value	Codeword	Value	Codeword	Value	Codeword	Value	Codeword
-32	1111101010	-16	111101110	0	00	16	111110011
-31	1111101011	-15	11101010	1	100	17	111110100
-30	1111101100	-14	11101011	2	11001	18	1111110100
-29	1111101101	-13	11101100	3	110110	19	1111110101
-28	1111101110	-12	11101101	4	1110001	20	1111110110
-27	1111101111	-11	11101110	5	1110010	21	1111110111
-26	1111110000	-10	11101111	6	1110011	22	1111111000
-25	1111110001	-9	1101110	7	1110100	23	1111111001
-24	1111110010	-8	1101111	8	11110000	24	1111111010
-23	1111110011	-7	1110000	9	11110001	25	1111111011
-22	111101000	-6	110100	10	11110010	26	1111111100
-21	111101001	-5	110101	11	11110011	27	1111111101
-20	111101010	-4	11000	12	111101111	28	11111111100
-19	111101011	-3	1011	13	111110000	29	11111111101
-18	111101100	-2	010	14	111110001	30	11111111110
-17	111101101	-1	011	15	111110010	31	111111111110
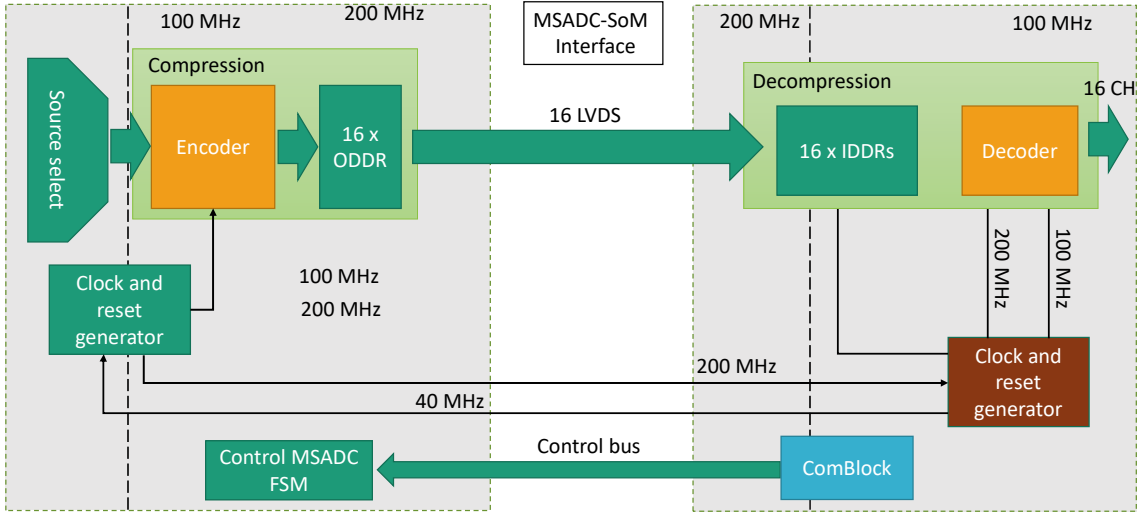
Figure 7.12: Interface between MSADC and SoM, and CoDec implementation.

During the codification process, each value is assigned a unique code consisting of a sequence of 0's and 1's. These codes are carefully designed to ensure that no code is a prefix of another code, allowing for unambiguous decoding. When transmitting the compressed data sequentially, the receiver can recover the original values in the correct sequence. To prevent error accumulation and facilitate data processing, the compressed data stream is organized into packets of a predefined length.

On the side of the SoM, the decoder IP core for 16 channels was added to the design of the base implementation. The ADC driver in this case is composed by: an IDDR per channel, the decoder and the clocking scheme. At the output of the decoder there is a continuous data stream that is caught by a CDC FIFO, used to read the data at the AXI Stream clock. From the decoder is expected a mean clock data rate of 80 MHz per channel, but from a 100 MHz clock domain, while the AXI Stream clock is 100 MHz.

The codec scheme is channel-independent and the compression rate depends only on the code assignment based on the signal statistics of each channel. To synchronize all channels within a 100 MHz (period) window, a very simple but effective mechanism is implemented. Taking into consideration that all channels are launched at the same time

and based on the fact that every 12.5 ns there is new data in all channels, the logical *and* of all FIFO *not empty* flag is used to signal a new synchronized data from the ADCs. This is crucial for ensuring synchronicity, as the time of arrival is one of the features that must be provided by the feature extraction process.

7.4 ECAL2 Prototype Setup in AMBER Pilot Run

During 2021-2022 there were pilot runs for the Apparatus for the Meson and Baryon Experimental Research (AMBER) experiment in the M2 beamline, the same as COMPASS.

The main objective of these pilot runs was to test the new instrumentation, as well to detect, measure, and identify the different particles that can be generated after muon-proton collisions using this instrumentation, and to obtain the first data for analysis. For this purpose, the spectrometer was slightly modified, as well as some parts of the DAQ and trigger systems.

As the beam time was assigned to AMBER, there was complete availability for deciding the moments when there was a beam by AMBER users. If there was a need for access to the beam area for hardware installation or intervention, it was completely destined for the AMBER team.

For testing the FFeCCa carrier and its hardware/firmware implementations, the AMBER collaboration assigned a space at the end of the beamline for installing an ECAL2 prototype.

7.4.1 ECAL2 Prototype

The prototype of the ECAL2 used for testing was built as a closed and complete module of 5×5 elements. The prototype is assembled as a miniature ECAL2, where each element consists of: a shashlik-type scintillator, an FEU 84-3 PMT, and a programmable Cockcroft–Walton high-voltage power supply to polarize the PMT.

(a) Photo-multipliers side.

(b) Fibers side.
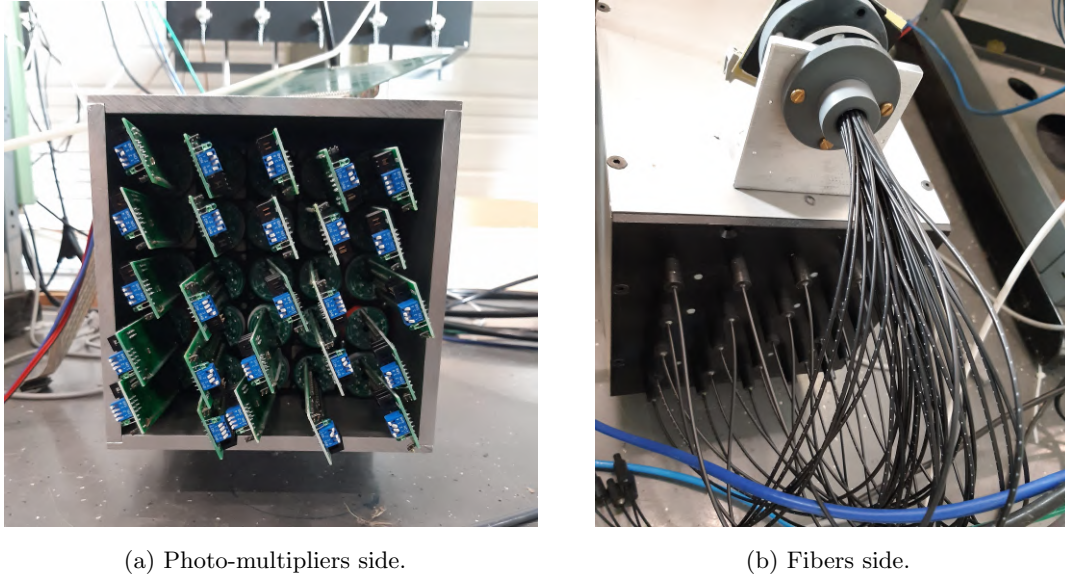
Figure 7.13: ECAL2 prototype pictures from both sides, on the left it can be seen the photo-multipliers connected to the scintillators, and on the right the fibers for LED pulsing, coming from the led pulser on top and being inserted on each of the elements.

For controlling and adjusting the slope of the high-voltage polarization value, the Cockcroft-Walton power supply incorporates a Serial Peripheral Interface (SPI) connected to a digital-to-analog converter (DAC), allowing for the soft start required by PMTs. Figure 7.13a shows the twenty-five PCBs of the Cockcroft-Walton power supplies, where it can be observed a blue switch for assigning each of them the address on the SPI bus. Similar to the ECAL2 detector, all elements have a fiber connected to a common LED pulser module to evaluate the operating status of the calorimeter, as shown in Figure 7.13b.

7.4.2 DAQ Readout in Beam Area

The new frontend for the ECAL2 was tested using a prototype placed at the end part of the beam downstream, in parasitic mode and without intervention of the ECAL2

wall and without disturbing any other measurement, as shown in Figure 7.14a.

As depicted in Figure 7.14b the setup is composed by: (1) FFeCCa Board, (2) Shaper, (3) High Voltage Power Supply Control Board, (4) ECAL2 prototype and (5) 100 V Power Supply. Additionally, there is a Personal Computer, a $\pm 5V$ Power Supply and a Gbit Ethernet Switch.

In Figure 7.14c, a picture taken from the downstream end looking upstream is shown, the detector was aligned with the beam to have higher pulse rates.

Figure 7.15 shows a block diagram of the DAQ prepared for this setup. For the ECAL2 prototype, there are two different interfaces: one connected to the Shaper block, connecting all the channel outputs, and the other to the high-voltage power supply control system (HVPSCS) for PMTs polarization. The Shaper output was connected directly to the FFeCCa board, and the main purpose of the shaping stage was to readapt the PMT pulses to a more situable signal. The filter has a rise time of 64 ns and a Full Width at Half Maximum (FWHM) of 120 ns, whereas the time to peak of the FEU84-3 PMT is on the order of 18 ns, with a FWHM of 20 ns [128]. In our system, this means that there are between four and five samples of rise time after the shaping filter (the sampling period is 12.5 ns). The COMPASS TCS provides information to the DAQ about the moments at which the beam is On through the Begin of Spill (BOS) block, connected directly to the FFeCCa by a LEMO input. Polarization of the PMT is achieved through the HVPSCS, as well as LED driving for the operative test of the system when there is no beam.

For safe data acquisition, readout, and control, we prepared a dedicated DAQ PC with Linux OS, all development tools, as in [129] and some specific debugging and programming tools. The entire system is connected to CERN's general-purpose Internet network using a GB Ethernet switch. Therefore, we can access the DAQ PC, HVPSCS, and FFeCCa board directly through a Secure Shell (SSH) via the Linux Public Login User Service (LXPLUS) [130]. The DAQ PC was used as a local readout service with a 2 TB disk space. It is used to launch the data-taking while uploading the data to EOS Open
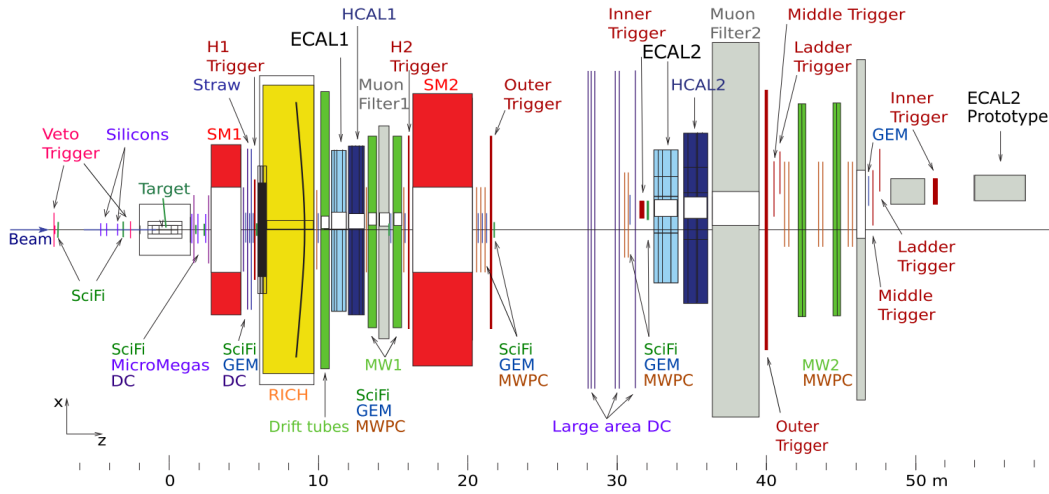
(a) Location of the ECAL2 prototype into the beamline.

(b) Front view of the setup.

(c) Back view of the setup.

Figure 7.14: Detectors setup for the 2021-2022 AMBER Pilot Run and pictures of the ECAL2 prototype setup at the end of the beamline.

141
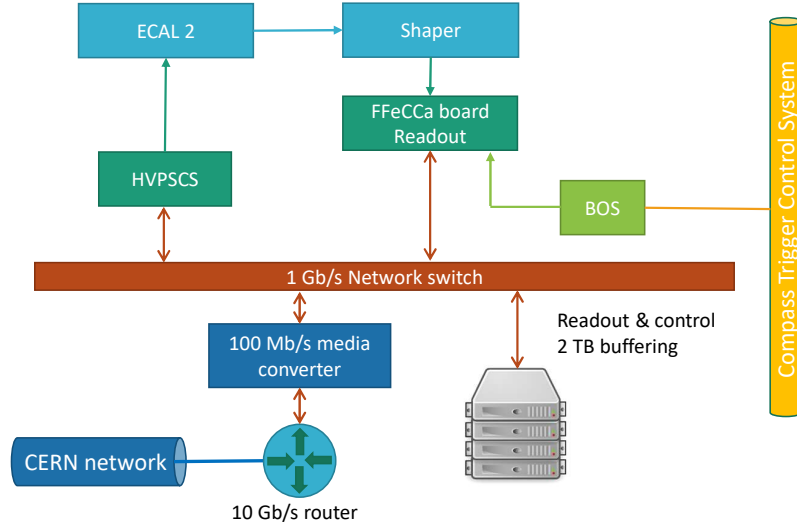
Figure 7.15: ECAL2 prototype DAQ installed for 2021 AMBER pilot run.

Storage [131]. Owing to the limited bandwidth of the upload link to the CERN network, the only limitation in data acquisition is the disk space. When 90% of the disk space was full, data acquisition was stopped automatically until the data were uploaded to CERN storage services.

With this setup, more than 2 TB of data was collected from the ECAL2 prototype. The results are presented in Chapter 9. Both, the hardware and firmware implementations were successfully tested and validated.

Based on these tests the FFeCCa-based DAQ system was confirmed for the ECAL2, and the production of additional 20 FFeCCa boards were ordered to cover 100 Channels of the detector, required for the Proton Radius Measurement on 2023/2024. In Chapter 10 more detailes are provided

7.5 Summary

In this chapter, an open framework for SoC-FPGA-based data acquisition (DAQ) systems is presented. The framework addresses the complexity of hardware development

and provides a modular design that allows for easy integration of new data acquisition frontend electronics. It includes FPGA firmware, processor firmware, and interface/control software components.

The FPGA firmware is designed using VHDL and is organized into several modules that handle data processing, control, and clock management. A custom IP block called ComBlock is used for communication between the programmable logic (PL) and the processor system (PS). It offers well-defined interfaces for configuring and reading the FPGA design, simplifying the interaction between the two components.

The processor firmware is based on FreeRTOS Xilinx and lwIP library for TCP/IP communication. It handles the initial configuration of the FPGA, system reset, and data acquisition control. It also establishes a TCP server for receiving commands from a PC and launching data acquisition.

Python scripts are developed for interfacing with the PC and facilitating data acquisition. These scripts provide a user-friendly interface for selecting channels, data sources, and other parameters. Additionally, a simple plotter script is provided for visualizing the acquired data.

Different implementations are discussed, including two-channel, eight-channel, and sixteen-channel streamers. The sixteen-channel buffered implementation involves FIFOs for buffering and an external trigger for synchronization. Lossless data compression techniques, specifically Huffman coding, are utilized to compress and transmit data effectively when working in a free-running mode. In Figure 7.16, a comparison of all the implementations is shown. It can be observed that the two-channel streamer utilizes the least amount of resources. However, when considering the number of channels implemented, it is also the least efficient in terms of resource utilization, while the most resources consuming implementation is the sixteen channels without the codec. The most efficient from the MSADC side per channel is the eight channels streamer, and from the SoM side the sixteen channels without the compression. Nevertheless, this last

143

(a) MSADC: FPGA total usage comparison.

(b) SoC: FPGA total usage comparison.

(c) MSADC: FPGA per channel usage comparison.

(d) SoC: FPGA per channel usage comparison.
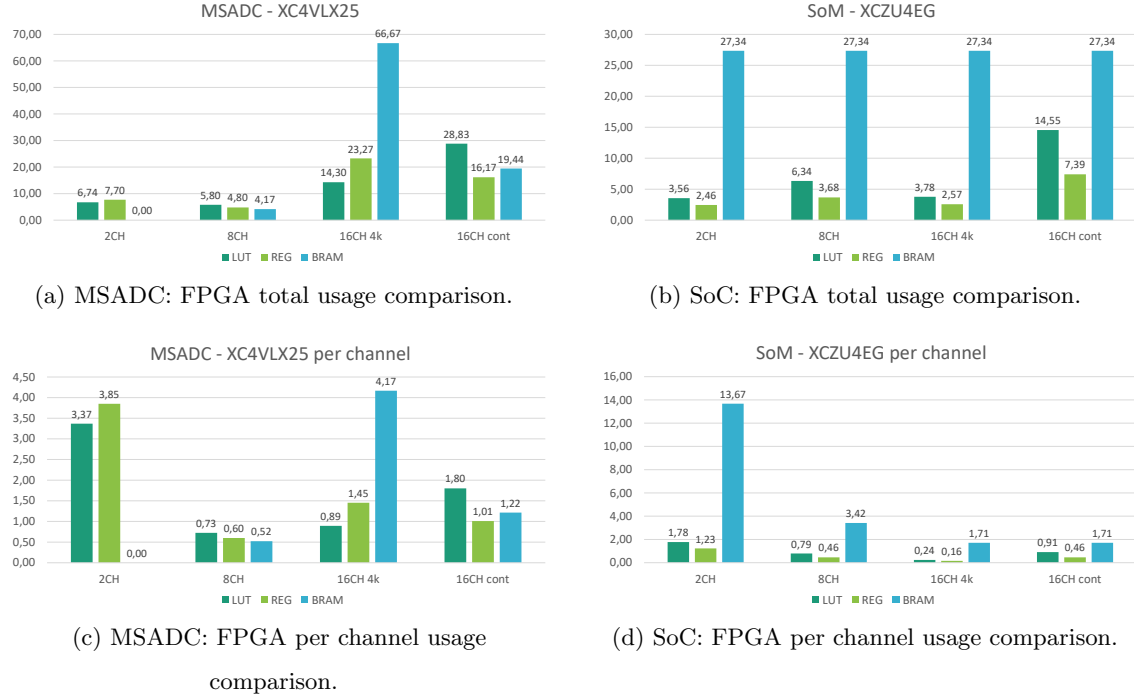
Figure 7.16: Comparison of resources usages between the different implementations in percentage.

implementation is not feasible to be used in the free-running mode, so it is excluded from the final evaluation.

In conclusion, the eighth channel streamer is the most efficient implementation. However, because of the loss of half of the channels, the preferred option is the sixteen-channel streamer with the encoder, despite requiring more than double the resources in both FPGAs. When evaluating the SoM implementation per channel, the resource utilization is nearly the same for both options, which encourages the implementation of sixteen channels.

During October and November 2021, the AMBER experiment conducted a pilot run at the M2 beamline to test new instrumentation and collect initial data. The FFeCCa DAQ with an ECAL2 prototype was successfully tested in parasitic mode, and data acquisition

144

was performed without errors. Based on this validation, twenty FFeCCa boards are under production to be integrated into the COMPASS/AMBER DAQ system.

The OFSODA project can be found at:`https://gitlab.com/brunovali/ofsoda`.

Chapter 8

Digital Pulse Processing

The data features extraction in real-time, specifically the amplitude and time of arrival extraction, is done by online digitally processing the signals coming from the detectors. A digital pulse processor based on Finite Impulse Response (FIR) filters for detecting and measuring the amplitude of the pulses is presented in this chapter. The time of arrival will be extracted in a coarse grain way, given by a free-running counter at the frequency of the ADC.

The method is introduced, explained and developed for an X-ray spectroscopy data set, for the decay of Fe_{55} into Mn using a custom Silicon Drift Detector, as in [132]. As for generating the filter coefficients the method relies on the mathematical model of the pulses and the intrinsic noise of the channels, for generating filters for the ECAL2 pulses the only change we have to do is to replace the pulse model and the noise parameters.

The main reason for presenting the digital pulse processor using an X-ray dataset instead of the ECAL2 pulses is due to the reference provided by the spectroscopy lines of K_α and K_β of the decay, allowing the possibility for calibration and comparison to other methods.

8.1 Related Works and State of the Art

Several digital pulse processing techniques have been introduced for spectral measurements over the years.

In 1993, Jordanov and Knoll [133] presented a real-time DPP using a moving average technique built on high-speed programmable logic devices (PLD) and fast-TTL integrated circuits. This implementation included a conventional quasi-Gaussian analog shaper after the CSA. In 1994, they extended their work by introducing fast recursive digital algorithms implemented on a personal computer (PC) for the synthesis of symmetric triangular and trapezoidal pulse shapes, thereby replacing the traditional analog pulse shapers [134]. Later that year, Jordanov V. et al. [135] implemented digital shaper algorithms on dedicated hardware by avoiding the use of a PC for offline pulse processing.

Guzik Z. and Krakowski T. [136] presented a full set of recursive algorithms based on the Z-transform for trapezoidal pulse shaping with pole-zero cancellation for exponentially decaying input pulses. The complete system was implemented on an FPGA, and included energy reconstruction, baseline restoration, trigger generation, and event acceptance. The use of this approach is limited because of the complexity of deriving recursive formulas for different input pulse shapes generated by various pulse detection systems.

Sajedi S. et al. [137] proposed an FPGA-based non-linear recursive filter design for high-rate pulse feature extraction in nuclear medicine imaging and spectroscopy. Real data were obtained directly from the pre-amplifier of the detection system. It was then fitted offline using the least-squares curve fitting method in PC to obtain the deterministic pulse model. The pulse shape model was then used to generate look-up tables (LUT) and implement non-linear recursive filters. The main disadvantage of this system is the large usage of memory elements. The aforementioned methods disregard the noise present in the system.

In contrast to the progression of recursive IIR methods, non-recursive FIR-based digital signal processing methods have been independently developed for pulse-height

analysis. In 1996, Gatti E. et al. [138] introduced a method for calculating the FIR filter coefficients for nuclear spectroscopy with time-domain constraints and the uncorrelated noise present in the signal. The filter was obtained by solving a set of linear equations derived by expressing the filter shape and equivalent noise charge as modified Fourier sine series. Later, Gatti E. et al. [139] modified the method and incorporated experimental noise by estimating the noise power spectral density of data obtained from an analog shaper in the absence of pulses.

In 2002, Riboldi et al. proposed a numerical approach based on the least mean squares method [140] to calculate the optimum FIR filter coefficients. In 2004, Gatti E. et al. presented the fully formalized method [141] named DPLMS. A drawback of this method is that it directly estimates noise using the sampled data stream by assuming that no correlated noise is present. In 2007, Riboldi S. et al. [142] extended the DPLMS method by addressing the correlated noise. Additionally, there are several publications that detailed FIR-based digital pulse shaping systems by utilizing the DPLMS method implemented on FPGAs [143, 144, 145] and SoC-FPGAs [146].

Considering the mentioned contributions, it can be noted that the DPLMS optimization method improves the SNR of the output pulse. However, the application of this method requires the knowledge of the real characteristics of the noise and an accurate mathematical model for the noiseless pulse. This chapter presents an effective procedure to evaluate the model, characterize the noise present in the system, and include this information along with other constraints in the DPLMS method.

8.2 X-ray Spectroscopy Detection System

Particle detectors are central devices in X-ray spectroscopy. They are available in different technologies, such as gaseous ionization detectors, silicon drift detectors (SDD), photodiode detectors, and photomultipliers. Among these, interest in SDD for single-photon detection has been constantly growing since its introduction by Gatti and

Rehak in 1983 [147, 148]. Owing to their intrinsic low noise and ability to operate with high photon rates, they are widely used in X-ray spectroscopy.

X-ray photon detectors generate a small amount of electric charge for each absorbed photon. This charge is proportional to the energy of the photon and produces a very small and short current pulse, which typically requires amplification and filtering before the analysis. The first amplification stage is commonly performed using a charge sensitive amplifier (CSA) that integrates a small charge, producing a relatively large voltage step [149, 150]. This voltage step is further amplified and filtered using a pulse-shaping amplifier (PSA), which produces a semi-Gaussian pulse ready for digitization.

A typical single-photon detection system in X-ray spectroscopy consists of a detector, CSA, PSA, ADC, and DPP for pulse amplitude measurement, as shown in Figure 8.1.
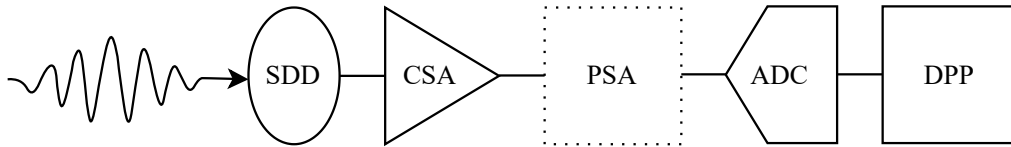
Figure 8.1: Block diagram of a typical single-photon detection system showing the incident photon on the silicon drift detector (SDD), CSA, optional pulse shaping amplifier (PSA), analog-to-digital converter (ADC), and digital pulse processor (DPP).

In these photon detection systems, the major electronic noise contributors are the CSA and the leakage current of the detector. The first CSA was proposed in 1956 by Gatti [151]. Subsequently, continuous modifications have been made to improve the SNR [152, 153].

An idealized noiseless CSA output pulse can be described by an exponential upward step-like pulse, expressed as follows:

$$V(t) = \begin{cases} 0, & t \leq t_0; \\ A(1 - e^{\frac{-(t-t_0)}{\tau}}), & t > t_0; \end{cases} \tag{8.1}$$

where t_0 is the pulse arrival time and τ is the exponential rise time of the CSA, which is limited and determined by the non-zero charge integration time.

The CSA integrates not only the charge produced by the absorbed photons but also the leakage current of the detector. Owing to this small constant leakage current, the CSA produces a ramp with a constant slope. The complete output signal of the CSA can be modeled as a superposition of the ramp, ideal pulse, and noise, as follows:

$$V(t) = \begin{cases} B_0 + B_1 t + n(t), & t \leq t_0; \\ A(1 - e^{\frac{-(t-t_0)}{\tau}}) + B_0 + B_1 t + n(t), & t > t_0; \end{cases} \quad (8.2)$$

where B_0 denotes an arbitrary offset, B_1 denotes the angular coefficient corresponding to the constant slope of the baseline ramp, and $n(t)$ is the noise component. For digitized signals, the CSA output can be rewritten by replacing the continuous time variable t with the discrete index i, which expresses time in units of sampling periods, as follows:

$$x_i = \begin{cases} B_0 + B_1 i + n_i, & i \leq t_0; \\ A(1 - e^{\frac{-(i-t_0)}{\tau}}) + B_0 + B_1 i + n_i, & i > t_0; \end{cases} \quad (8.3)$$

The parameters A and B_0 are then expressed in ADC value, t_0 and τ are expressed in units of sampling periods, and B_1 in ADC value per sampling period.

8.2.1 Experimental data

For this part of the work, an experimental dataset from a typical X-ray fluorescence experiment is considered. The data were obtained by digitizing the signal from a low-noise CSA coupled with a SDD-based single-photon detection system [152, 154, 155, 156] without a PSA. The dataset contains 2929 segments sampled at 40 Mhz with a 12-bit ADC. Each segment is 512 samples long and contains a single-photon pulse, as shown in Figure 8.2. This dataset was taken under normal operating conditions; thus, it includes real noise. A trigger system and a circular buffer allowed capturing the traces with the

pulses starting around the 200th sample.
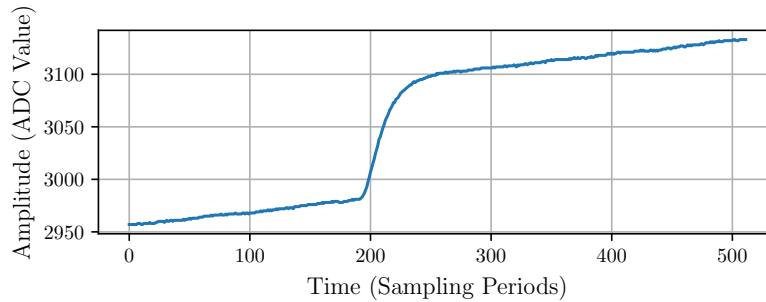
Figure 8.2: Typical experimental single-photon pulse at the output of the charge sensitive amplifier (CSA).

One characteristic of this dataset is that all photons present a different offset value. This was caused by the background slope and the random arrival times of the photons [157, 158].

8.3 Digital Pulse Shaping

In traditional pulse processing systems, the output of the CSA goes through a shaping stage, which improves SNR and converts step-like pulses to pulses suitable for subsequent digital acquisition and signal processing.

A typical $CR - (RC)^n$ analog pulse shaper amplifier consists of one differentiator followed by n integrators to produce a semi-Gaussian output pulse [159]. This type of analog shapers can be replaced by modern digital shaping systems, which offer several advantages [160, 161]. In these systems, the CSA output is directly digitized using a fast ADC and is immediately processed by a customized DPP. The pulse shaping can be digitally implemented in a more controlled way than with an analog circuit. An ideal DPP produces the most accurate and precise amplitude measurement of the CSA output.

A simplified block diagram of a DPP [162] for high-resolution X-ray spectroscopy
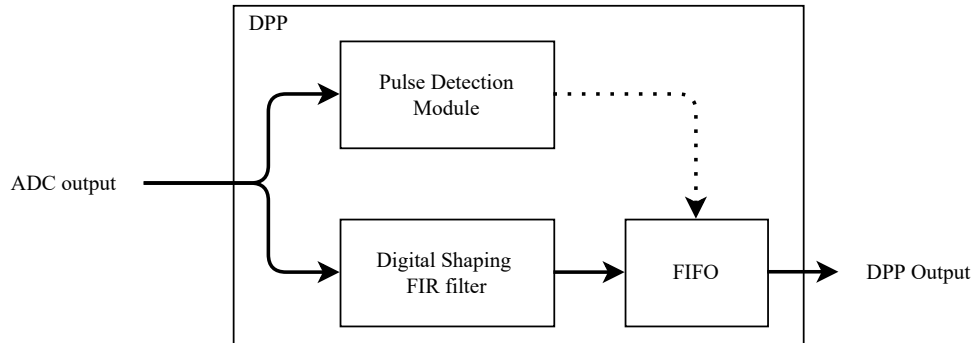
Figure 8.3: Simplified block diagram of the digital pulse processing unit.

is shown in Figure 8.3. It consists of two separate data paths from the same channel: one for the precise detection of photon arrival and the other one for shaping the input pulse based on an FIR filter. The pulse detection module detects the arrival time and decides when to retrieve the pulse amplitude. The module also controls a FIFO to store the amplitude from the digital shaping filter output at the correct sampling time. The FIFO allows asynchronous storage of the amplitudes of pulses, that typically occurs at random times, and a synchronous regular reading by the system hosting the DPP.

The digital shaping filter should fulfill the following requisites:

1. Be independent of any offset

2. Be independent of any constant background slope

3. Optimize the SNR, according to real noise characteristics

4. Generate a flat-top to mitigate the uncertainty of the pulse arrival time detection.

An additional requirement regards the time resolution of the filter, which strongly depends on its length. If two photons are separated by less than the filter integration time, the filter may not be able to properly process each one. For high photon-rate regimes, it is essential to make the shortest possible filter without significantly sacrificing the filtering

capabilities. Taking this aspect into account, the length of the filters considered in this part of the study has been fixed to 80 taps at 40 Msps.

8.3.1 Trapezoidal FIR Filter

As a starting point, a trapezoidal output filter approximation is implemented to measure the pulse amplitude [163]. Figure 8.4 shows the filter coefficients and the output pulse corresponding to an experimental input pulse, such as that in Figure 8.2.
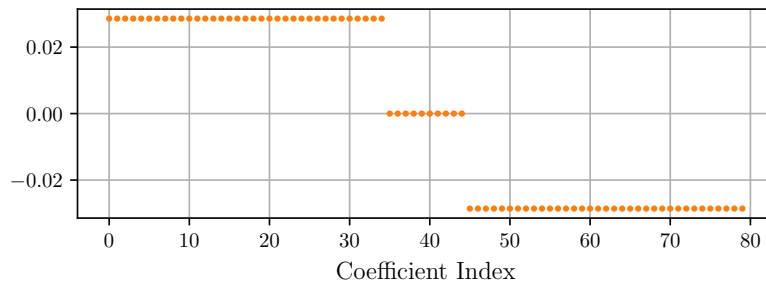
(a) Trapezoidal FIR coefficients.
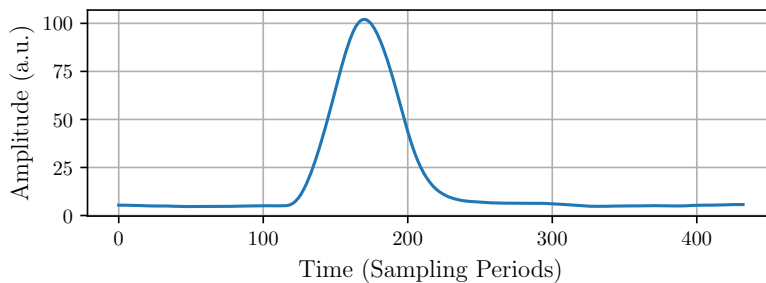
(b) Output.

Figure 8.4: Trapezoidal FIR coefficients (a) and the output pulse (b) corresponding to an experimental input pulse like that of Figure 8.2.

The underlying idea of this filter is that the amplitude of the pulse can be calculated by waiting for the CSA output to settle within an acceptable error and then subtracting from it the baseline before the arrival of the pulse. To attenuate the white noise, a simple

average before and after the pulse allows more precise estimation of the pulse amplitude. The number of positive, null, and negative filter coefficients respectively corresponds to the parameters t_R, t_{FT} and t_F. The t_R positive coefficients of the filter compute a moving average and determine the rise time of the output pulse. Their value is constant and equal to $1/t_R$. The t_{FT} central null coefficients define the time waited for the output pulse to settle within an acceptable error, and the duration of a nearly flat-top of the output pulse. Finally, the t_F negative coefficients compute another moving average and determine the fall time of the output pulse. Their value is constant and equal to $1/t_F$. These three parameters are bounded by the condition $t_R + t_{FT} + t_F = 80$, since the considered length of the filter has been fixed at 80. Regarding the FIR filter output, it can be seen that there is some top flatness that could reduce the error in the amplitude measurement. But it can also be observed that the output of the filter presents an offset that introduces an error in the amplitude measurement. This error increases with the background slope, but it can be corrected by modifying the FIR filter, as explained in Subsection 8.3.2.

8.3.2 Geometrically Derived FIR Filter

To correct the above mentioned error due to the background slope, we perform a study on the geometry of the pulse. A typical photon pulse with its geometrical features is shown in Figure 8.5. The height of the two points in the middle of the segments indicated with t_R and t_F correspond to the average height computed over those segments. The trapezoidal filter computes the difference between these average values as an estimation of the pulse amplitude A, but we can see that this difference is $A + D$ instead of the expected true value A.

The amplitude error D due to the background slope is related to $\tan \alpha$, as follows:

$$\tan \alpha = \frac{D}{\frac{1}{2}t_R + t_{FT} + \frac{1}{2}t_F} \tag{8.4}$$

The value of $\tan \alpha$ can be estimated using the least-squares method considering the t_R
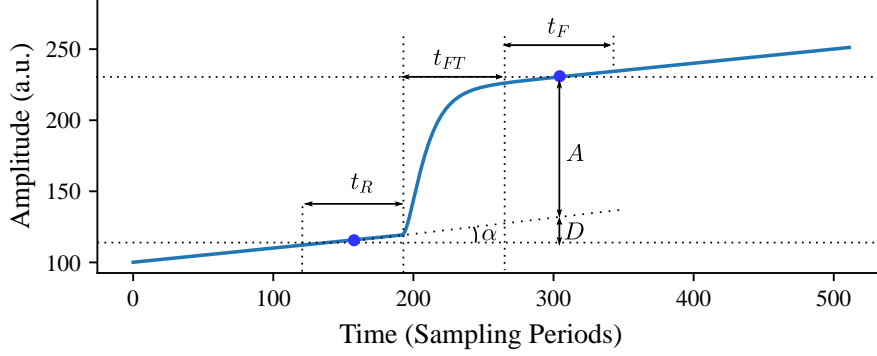
Figure 8.5: Typical photon pulse with its geometrical features highlighted. The two points in the middle of t_R and t_F segments corresponds to their average values.

samples before the arrival of the pulse. A closed-form expression for estimating $\tan \alpha$ can be written as follows (see Appendix 11.2 for details):

$$\tan \alpha \approx \sum_{i=0}^{t_R-1} -6 \left(\frac{1 + t_R - 2i}{t_R^3 - t_R} \right) x_i \tag{8.5}$$

From equation (8.4) and (8.5), the error D is estimated as

$$D = \sum_{i=0}^{t_R-1} -6 \left(\frac{1 + t_R - 2i}{t_R^3 - t_R} \right) \left(\tfrac{1}{2} t_R + t_{FT} + \tfrac{1}{2} t_F \right) x_i \tag{8.6}$$

and the correct amplitude A is then calculated as follows:

$$A = \frac{1}{t_F} \sum_{i=t_R+t_{FT}}^{t_R+t_{FT}+t_F-1} x_i - \frac{1}{t_R} \sum_{i=0}^{t_R-1} x_i - \sum_{i=0}^{t_R-1} -6 \left(\frac{1 + t_R - 2i}{t_R^3 - t_R} \right) \left(\tfrac{1}{2} t_R + t_{FT} + \tfrac{1}{2} t_F \right) x_i \tag{8.7}$$

By simplifying and rearranging the previous expression, we can show that the amplitude can be computed as a linear combination of the sequential data x_i with constant coefficients:

$$A = \sum_{i=t_R+t_{FT}}^{t_R+t_{FT}+t_F-1} \frac{1}{t_F} x_i + \sum_{i=0}^{t_R-1} \left[-\frac{1}{t_R} + 6 \left(\frac{1 + t_R - 2i}{t_R^3 - t_R} \right) \left(\tfrac{1}{2} t_R + t_{FT} + \tfrac{1}{2} t_F \right) \right] x_i \tag{8.8}$$

155

It can be seen that the pulse amplitude A can be continuously evaluated by an FIR filter whose coefficients are described as follows:

$$c_i = \begin{cases} \frac{1}{t_F}, & \$0 \leq i < t_F; \\ 0, & t_F \leq i < t_F + t_{FT}; \\ -\frac{1}{t_R} + 6\left(\frac{1+t_R-2i}{t_R^3-t_R}\right)\left(\frac{1}{2}t_R + t_{FT} + \frac{1}{2}t_F\right), & t_F + t_{FT} \leq i < t_F + t_{FT} + t_R; \end{cases} \qquad (8.9)$$

The central null coefficients determine the nearly flat-top region of the output pulse. Figure 8.6 shows these *geometrically derived* (GD) FIR coefficients with the parameters $t_R = 35$, $t_{FT} = 10$, and $t_F = 35$, and the output pulse obtained with this filter is applied to an experimental pulse. As expected, this GD FIR filter suppresses the offset and background slope of the input pulse.

8.4 Data Analysis and FIR Filter Optimization

As described in Section 8.3, it is desirable that the shaping filter output pulse has the highest possible SNR and a flat-top to mitigate the uncertainty of the photon arrival time. These two main conditions directly contribute to achieving an optimal energy resolution [164]. To satisfy these conditions, based on the analysis of the experimental data, an accurate mathematical model for the input pulse was defined (Subsection 8.4.1) and the noise characterized (Subsection 8.4.2).

The input pulse model is essential because (i) it allows the noise characterization by correctly separating the stochastic component (noise) from the deterministic signal and (ii) it contributes to a correct calculation of the FIR filter coefficients to determine a flat-top at the output.

The adapted DPLMS method, based on the pulse model and the characterized noise, is presented in subsection 8.4.3.
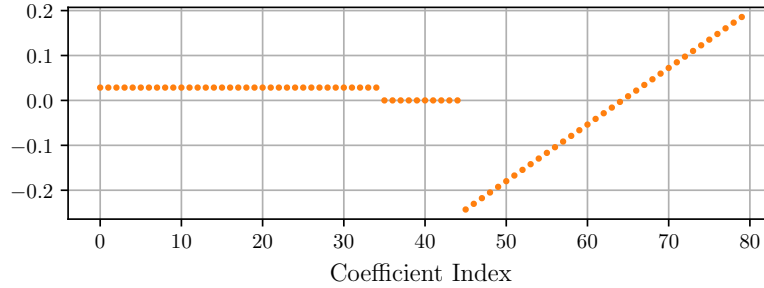
(a) GD FIR coefficients.
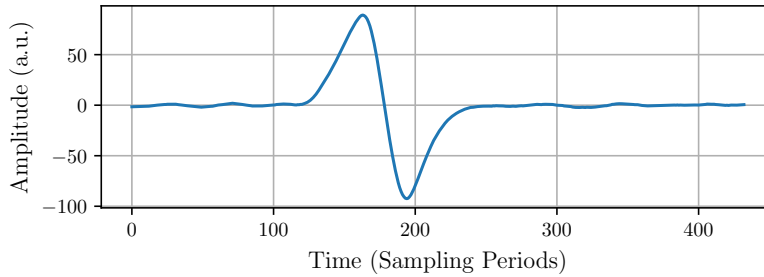
(b) Output.

Figure 8.6: GD FIR coefficients (a) and the output pulse corresponding to an experimental input pulse (b).

8.4.1   Pulse modeling

The model parameters A, B_0, B_1, t_0, and τ of the deterministic noiseless input pulse described in Equation (8.3) are estimated numerically by fitting the model to the experimental data. A typical experimental pulse with a fitted model and corresponding residuals are shown in Figure 8.7. The residuals would correspond to the stochastic component and should be considered as the noise $\{n_i\}$.

The residuals plot in Figure 8.7 shows a relatively large spike around the starting point of the pulse, which indicates inaccurate modeling. Therefore, we propose a bi-exponential pulse model with the same number of parameters, described by equation (8.10).
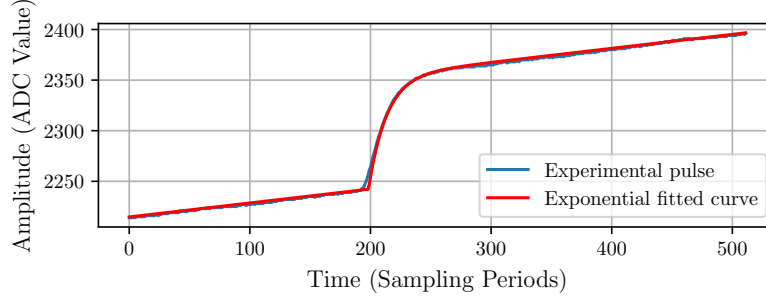
(a) Exponential model fitting.
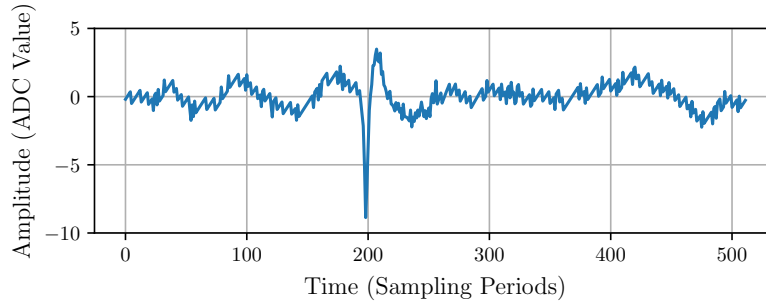
(b) Residuals.

Figure 8.7: Exponential model fitting (a) with its corresponding residuals (b).

$$x_i = \begin{cases} B_0 + iB_1 + n_i, & \text{i} \leq t_0; \\ A\left(1 - 2e^{\frac{-(i-t_0)}{\tau}} + e^{\frac{-2(i-t_0)}{\tau}}\right) + B_0 + iB_1 + n_i, & \text{i} > t_0; \end{cases} \quad (8.10)$$

This model is a heuristic model that can be analytically derived from some assumptions about the transfer function of the CSA (see Appendix 11.1 for details). The result of the fitting with the bi-exponential model is shown in Figure 8.8. The improvement can be observed in the residuals, which do not present evident artifacts.

Table 8.1 shows a comparison of both models using the mean quadratic residuals, peak-to-peak residuals, and Akaike information criterion [165] evaluated over all fitted pulses. The calculated values of these indicators confirm that the bi-exponential model is
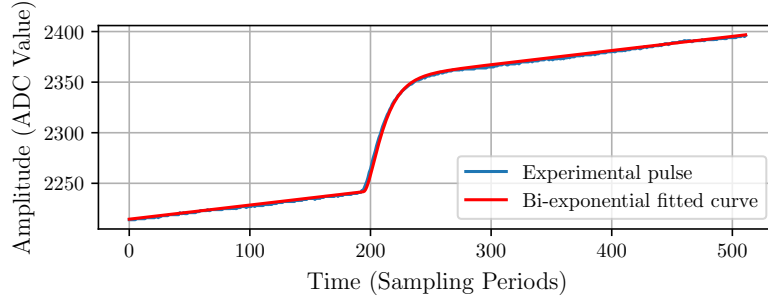
(a) Model fitting.
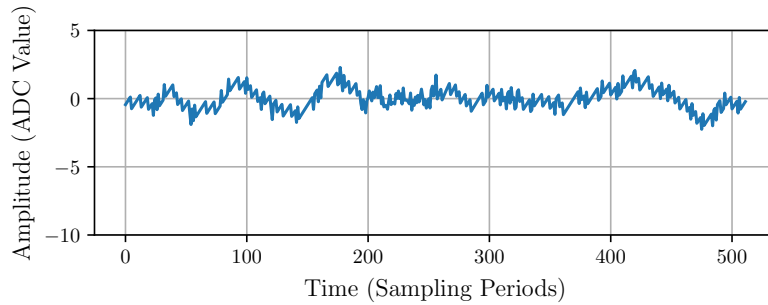
(b) Residuals.

Figure 8.8: Bi-exponential model fitting (a) and corresponding residuals (b).

significantly more accurate than the exponential one.

Table 8.1: Pulse models comparison.

	Exponential Model	Bi-exponential Model
Mean quadratic residuals	6201	5914
Mean peak-to-peak residuals	13.7	6.6
Mean Akaike information criterion	1720	1397

Figure 8.9 shows the distributions of all fitted model parameters when using the bi-exponential model. The average values of the fitting parameters along with their standard deviations are presented in Table 8.2. The parameter B_0 is a vertical offset

(a) Amplitude in ADC steps (A).

(b) Arrival time (t_0).

(c) Exponential time (τ).

(d) Slope (B_1).

(e) Offset (B_0).
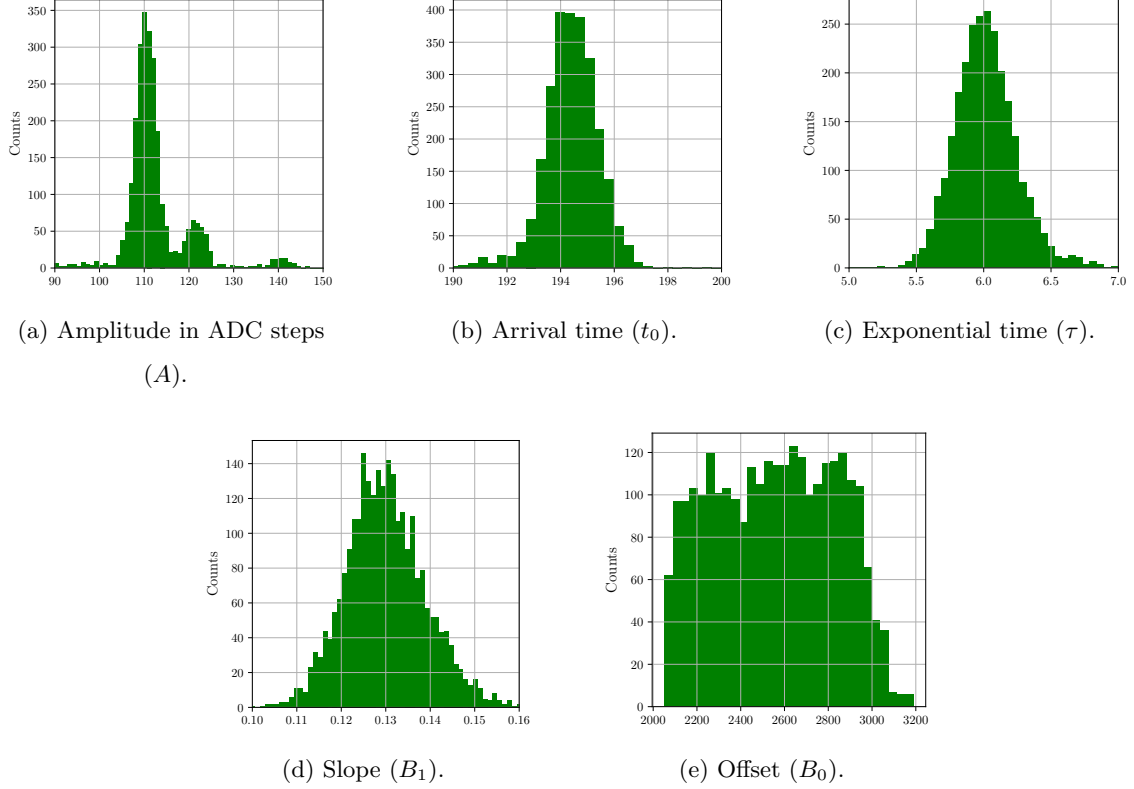
Figure 8.9: Histograms of the fitted parameters corresponding to the biexponential model.

that randomly changes from pulse to pulse, and is distributed rather uniformly. The arrival time t_0 and the slope coefficient B_1 are also stochastic variables that change from pulse to pulse but closely follow Gaussian distributions. In contrast, the mean value of τ is the estimate of the only parameter that characterizes the ideal pulse shape, and is assumed to be equal for all photons.

Since the amplitude of a photon pulse is proportional to the photon energy, the histogram of the fitted amplitudes in Figure 8.9a represents the energy spectrum of the detected photons that, in this study, corresponds to transition lines of Manganese (Mn). The two main peaks correspond to the lines K_α and K_β respectively at 5890 eV and 6490 eV [166], and the third small peak (around 140) corresponds to 90-degree Compton

scattered photons.

Table 8.2: Mean values and standard deviations of the fitted model parameters.

Parameter	Mean Value	Standard Deviation
Slope (B_1)	0.13	0.01
Arrival time (t_0)	194.3	1.4
Exponential time (τ)	6.02	0.24
Offset (B_0)	2548	272

8.4.2 FIR Input Noise Characterization and Output Noise Estimation

Based on the fitting of the pulses and the residuals calculation, we can proceed to define the noise at the output of the filter. Let y be the convolution of an input signal x with a k-tap FIR filter,

$$y_j = \sum_{i=0}^{k-1} c_i \, x_{j-i} \tag{8.11}$$

Assuming that x is the noise at the input, a statistical description of the noise at the output y is needed. Hence, the variance of y, denoted by σ_y^2, can be written as

$$\sigma_y^2 = \left\langle (y - \langle y \rangle)^2 \right\rangle = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} c_i \, c_j \underbrace{\langle x_i - \langle x_i \rangle \rangle \, \langle x_j - \langle x_j \rangle \rangle}_{\text{Covariance Matrix } V_{i,j}} \tag{8.12}$$

In this case, the noise in the experimental data is considered stationary. Moreover, the autocovariance matrix becomes the normalized autocorrelation function (ACF) when the data $\{x_i\}$ is standardized such that the mean $\langle x_i \rangle$ is 0 and the standard deviation σ_x is 1; in this case, equation (8.12) can be rewritten in terms of the ACF as

$$\sigma_y^2 = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} c_i \, c_j \, ACF(|i-j|) \tag{8.13}$$

where the normalized ACF is estimated from the experiemntal data $\{x_i\}$ as follows:

$$ACF(j) = \frac{\sum_{i=1}^{N-j} x_i \, x_{i+j}}{\sum_{i=1}^{N-j} x_i^2} \tag{8.14}$$

Here N is the maximum number of consecutive samples available in the residuals. From the experimental dataset, the normalized ACF of each segment was calculated using equation (8.14) where $N = 512$ and $\{x_i\}$ are the model fitting residuals. Then, all ACFs estimated on each segment were averaged to be later used in equation (8.13). Figure 8.10 shows the first 80 values of the normalized average ACF.
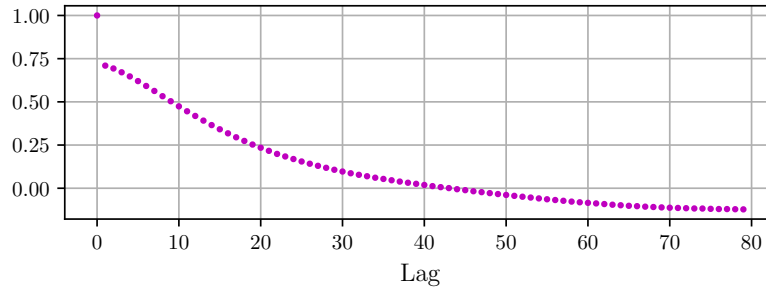
Figure 8.10: Normalized average autocorrelation function estimated from the residuals of the fitted photon segments.

The normalized average autocorrelation function in Figure 8.10 shows an abrupt change, from 1 with $lag = 0$, to about 0.7 with $lag = 1$, and from there it follows a smooth decay to slightly negative values from $lag = 40$ onward. The first abrupt change would correspond to a white noise component, whereas the smooth decay would correspond to relatively low frequency components of the noise spectrum.

8.4.3 Adapted DPLMS Filter Optimization

The original DPLMS method considers a number of constraints and a set of corresponding weights. These constraints define the objectives of the optimization, and their corresponding weights determine their relative relevance. In this way it is possible to reach a trade-off among goals that cannot be all fully satisfied. These constraints were adapted taking into account the characteristics of the experimental system. Based on the optimum digital shaping filter requisites, described in Section 8.3, four constraints are defined. One constraint removes the constant offset in the signal. Another one removes the background ramp due to the leakage current. The SNR is maximized using a constraint that minimizes the variance at the output of the filter in the presence of noise, as described in equation (8.13). Finally, the flat-top is determined by a constraint that minimizes the error between the amplitude of an ideal input pulse and the convolution of the filter with that input pulse model.

The weights of each constraint can be arbitrarily set between zero and infinite. By adjusting the relative values of the different weights, it is possible to obtain diverse trade-offs among competing requirements.

To be immune to the offset introduced by the term B_0 of the pulse model in equation (8.10), it is enough that the k-tap FIR filter coefficients $\{c_i\}$ comply with the following constraint:

$$\sum_{i=0}^{k-1} c_i = 0 \tag{8.15}$$

The ramp slope given by the angular coefficient B_1 will also introduce a bias in the pulse amplitude measurement. To cancel this effect we impose the following constraint:

$$\sum_{i=0}^{k-1} c_i \, i = 0 \tag{8.16}$$

Another source of error when measuring the amplitude of the pulse is biven by the pulse arrival time detection method. Since the value of the pulse amplitude is captured

at the output of the FIR filter after a fixed time from the pulse detection, any error in the determination of the pulse arrival time will be automatically transferred to the sampling time of the filter output. By holding the amplitude value for a determined time, a flat-top is created in the output pulse effectively compensating for the error in the estimation of the photon arrival time. This condition is expressed by the following constraints on every output y_j of the flat-top region,

$$y_j = \sum_{i=0}^{k-1} c_i \, x_{j-i} = A, \qquad j \in [t_R, t_R + t_{FT} - 1] \tag{8.17}$$

where x is a pulse modeled with equation (8.10) without noise, and A is its amplitude.

To improve the amplitude measurement, the noise in the filter output needs to be minimized. This can be achieved by decreasing the filter output variance described by equation (8.13). By considering this requirement, and the constraints expressed in equations (8.15), (8.16), and (8.17), we define the following quadratic cost function:

$$\Psi(c_0, c_1, \ldots, c_{k-1}) = \alpha_1 \left(\sum_{i=0}^{k-1} c_i \right)^2 + \alpha_2 \left(\sum_{i=0}^{k-1} c_i \, i \right)^2 + \alpha_3 \sum_{j=t_R}^{t_R + t_{FT} - 1} \left(\sum_{i=0}^{k-1} c_i \, x_{k+j-i} - A \right)^2 +$$

$$\alpha_4 \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} c_i \, c_j \, ACF_{|i-j|}$$

$$\tag{8.18}$$

The significance of each constraint is determined by the relative values of the weights $\{\alpha_j\}$ associated with each corresponding quadratic term. By minimizing the function Ψ for a given set of weights $\{\alpha_j\}$ it is possible to obtain an optimal set of coefficients $\{c_i\}_{opt}$, that is

$$\{c_0, c_1, \ldots, c_{k-1}\}_{opt} = \underset{\{c_0, c_1, \ldots, c_{k-1}\}}{argmin} \ \Psi(c_0, c_1, \ldots, c_{k-1}) \tag{8.19}$$

If a weight α_i is set to zero, then the associated constraint is completely ignored, whereas in the limit where the weight approaches infinite, the constraint tends to be fully satisfied

by the optimization, as in the case of the Lagrange's multipliers method. There are no formulated rules to obtain the best weights $\{\alpha_j\}$, and these are manually adjusted after multiple trials.

An optimized set of coefficients was generated by minimizing the quadratic cost function Ψ (equation (8.18)). The minimization was performed using numerical software optimization routines, where the function Ψ was set by carefully selecting the values of the weights $\{\alpha_i\}$. Figure 8.11 shows the generated 80-tap FIR filter and the filtered output pulse corresponding to an experimental input pulse.
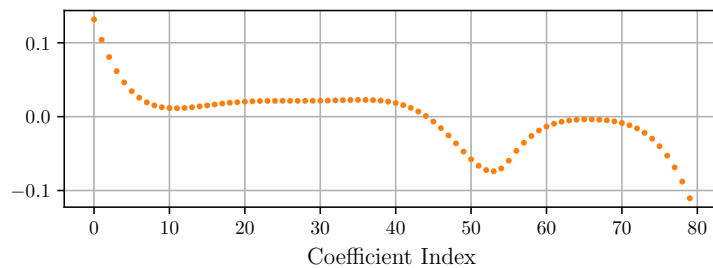
(a) DPLMS coefficients.
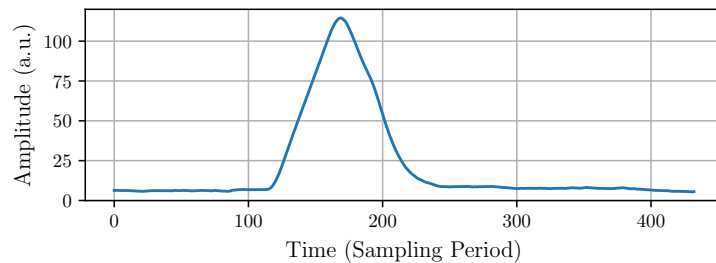
(b) Pulse output.

Figure 8.11: DPLMS FIR coefficients (top) and the corresponding output after being applied to an experimental pulse (bottom).

8.5   Comparison of the Described Methods

Four methods were described and applied to obtain the energy spectrum from the same experimental dataset of single-photon pulses.

Three methods are based on FIR filtering for pulse amplitude measurement, Subsections 8.3.1, 8.3.2 and 8.4.3. The other method consists in fitting each single pulse trace with a model where the amplitude is one of the fitting parameters.

In all cases, the histogram of the amplitudes estimates the energy spectrum of the detected photons. Each histogram has been approximated using a weighted sum of three Gaussian distributions. The two largest peaks correspond to X-Ray fluorescent photons, and the smallest one to Compton scattered photons (see histogram of amplitudes in Figure 8.9). Given that the two main peaks correspond to the K_α and K_β transition lines of Mn, whose energies are respectively 5890 eV and 6490 eV, it is possible to calibrate the system [166] establishing a linear correspondence between amplitude expressed in ADC channels and energy expressed in eV.

Table 8.3 shows the full width at half maximum (FWHM) obtained with each method with its corresponding uncertainty. The FWHM for the 90-degree Compton scattered photons is not considered due to insufficient statistical representation in the dataset. The background slope introduces an offset error in the measured amplitude. This error is corrected in the GD FIR and fitting methods, allowing a simple one-point calibration and making the filter immune to possible slope variations after calibration. On the other hand, the trapezoidal FIR method requires a two-point calibration process to correct the background slope error. The DPLMS FIR results have been achieved by emphasising energy resolution, placing the slope-error correction in a lower priority.

The best results in terms of energy resolution have been obtained with an FIR filter optimized with the adapted DPLMS method. This method is the only one that considers the specific noise in the dataset and simultaneously allows control of the top flatness of the output pulse. In order to obtain the best results, we have relaxed the weight of the

Table 8.3: Comparison of energy resolutions with different methods to estimate the energy spectrum.

Method	FWHM K_α [eV]	FWHM K_β [eV]	Slope-error correction
Trapezoidal FIR	207 ± 3	247 ± 17	no
GD FIR	286 ± 4	316 ± 16	yes
Fitting[†]	267 ± 4	288 ± 17	yes
DPLMS FIR	202 ± 2	233 ± 12	no

[†]These results correspond to the histogram of the amplitudes obtained by fitting all available photon traces.

slope error correction.

The fitting of individual photon pulses is a numerically heavy procedure to obtain the pulse amplitude. Although this method is not suitable for online data processing, it is expected to provide the most precise spectrum. The best results however were obtained with the DPLMS FIR method which outperformed the fitting procedure by about 20% in terms of energy resolution.

8.6   ECAL2 Pulse Model and Noise Characterization

As introduced in Chapter 4 the ECAL2 converts the energy of electrons, positrons, and photons into an electric pulse. These pulses are then reshaped by a first-order low-pass filter to readapt the signal properties for being read with the MSADC. Figure 8.12 shows a typical pulse from the ECAL2 captured by the COMPASS DAQ after the arrival of a trigger decision (blue dots).

To study the ECAL2 pulses a dataset from a COMPASS run of 2012 was selected. The dataset corresponds to events derived from the interaction between a fixed target and a 160 GeV muon beam. In this study and for a general approach, the pulses are not

Figure 8.12: Typical sampled pulse of the ECAL2 and fitted curve.

tagged by the (x,y) position or type of scintillation material. In this way, the analysis is performed according only to the shape of the pulse.

The pulse information is obtained by analyzing a data set containing around 1.2 million segments, corresponding to triggered events. As these pulses are 32 samples long, the frequency information is very limited and may be not enough for applying the the DPLMS method. 32 samples means a frequency resolution of:

$$\Delta f = \frac{f_s}{N} = \frac{80MHz}{32} = 2,5 \text{ MHz} \tag{8.20}$$

This length does not allow the study of low-frequency spectral content. As a big part of the noise will be located in that region the dataset is not appropriate for including the noise information in the DPLMS method. For this purpose, the database generated during the data taking of the AMBER Pilot run with our setup, as in Section 7.4 will be used. The dataset contains data traces of up 4M data points without interruption per channel (52 ms length).

$$\Delta f = \frac{f_s}{N_{MAX}} = \frac{80MHz}{4.194.303} = 19.07 \text{ Hz} \tag{8.21}$$

As in principle there is no clue about the low-frequency noise of the detector chain, the best option is to have the best resolution. The low-frequency noise will be related to

Table 8.4: Noise statistics on different channels of ECAL2 prototype.

Stat	ch0	ch1	ch2	ch3	ch4	ch5	ch6	ch7
μ	49.43	49.40	49.73	49.45	49.80	49.49	49.41	49.63
σ	1.37	1.19	1.21	2.83	0.74	0.75	0.74	0.74
Stat	ch8	ch9	ch10	ch11	ch12	ch13	ch14	ch15
μ	49.41	49.45	49.75	49.52	49.44	49.58	49.30	49.64
σ	0.72	0.73	0.71	1.52	0.81	0.91	0.89	1.69

the base level of the signals and can influence the pulse detection method and amplitude calculation.

For each of the channels the mean value and the standard deviation of the noise was estimated. Table 8.4 shows both values, for all channels and for data taken with the prototype setup installed at the end of the COMPASS beamline, when there was no beam and with the power supply on. The dataset is the same used to adjust the compression algorithm of Section 7.3.5 and, to build the statistics for Figures 7.11b and 7.11c.

8.6.1 Pulse Modeling

Once the typical pulse is identified, a suitable model and the correct value of the parameters need to be determined.

Without making any consideration from the shaper filter order, the first approximation for the mathematical model was obtained by proposing it heuristically. From the bell-shaped distribution of the signal, its smoothness and slow tail decay a model corresponding to a second order semi-Gaussian analog filter is proposed.

Considering that the experimental signals are the superposition of an ideal signal with a secondary smaller pulse plus noise, we can assume that the residuals of individual fittings represent the actual noise in the trace. Taking all this into account, the model is

defined as follows

$$F(t, t_0, \tau, \beta, a, k, t_1) = \begin{cases} \beta, & t < t_0 \\ \beta + a \left(\frac{e(t-t_0)}{2\tau} \right)^2 e^{-\frac{(t-t_0)}{\tau}}, & t_0 \leq t < t_0 + t_1 \\ \beta + a \left(\frac{e(t-t_0)}{2\tau} \right)^2 e^{-\frac{(t-t_0)}{\tau}} + \\ ka \left(\frac{e(t-t_0-t_1)}{2\tau} \right)^2 e^{-\frac{(t-t_0-t_1)}{\tau}}, & t_0 + t_1 \leq t \end{cases} \qquad (8.22)$$

where β represents the constant offset due to the dark current of the PMT [167] plus the amplifier DC offset value and other factors. The amplitude is denoted by the parameter a, the arrival time by t_0, and the exponential time by τ. The parameter k defines the relative amplitude of the second pulse with respect to the main pulse. t_1 defines the secondary pulse arrival time as a delay relative to t_0. From a graphical representation, all parameters are marked in Figure 8.12.

Fitting of the pulse model was performed using a subgroup of pulses, the most representative of the system. The pulses were filtered according to their shapes and minimum amplitudes: all pulses with their barycenter in the range between 12 and 15 and an amplitude value greater than 100 ADC counts above the baseline will be in the group. After the pulse selection, from the 1.2 million traces of the data base, only about a 10 % of the pulses remained. Figure 8.13 shows the superposition of all pulses filtered with according to the previous constraints.

Then, the selected pulses were used to fit the model parameters, the result of the values being those of Table 8.5.

The results of the fitting are also depicted in Figure 8.14. The procedure for the fitting was done based on the Differential Evolution algorithm [168], which for our scope gives the best results with the lowest execution times. To fit the pulse, we enter the exploration ranges for each of the parameters so the algorithm can converge in an efficient way. The only parameter that was fixed as a constraint was "N = 2".

The figures show the distribution of all the parameter values as histograms and their mean values at the top. The mean amplitude distribution starts at 100, and the mean
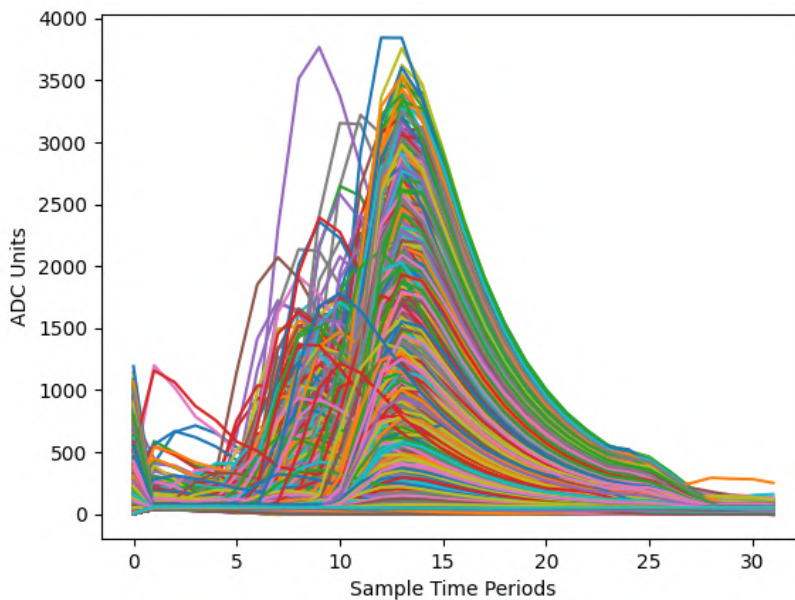
Figure 8.13: Filtered pulses overlay.

Table 8.5: Best parameter values of bi-exponential model fitting.

Parameter	Value
N	2
t_0	8.43
α	671.50
τ	1.77
β	48.23
k	0.11
t_1	7.91

171

(a) Amplitude (α).

(b) Arrival time (t_0).

(c) Exponential time (τ).

(d) Offset (β).

(e) Reflected pulse (k).

(f) Reflected pulse time (t_1).
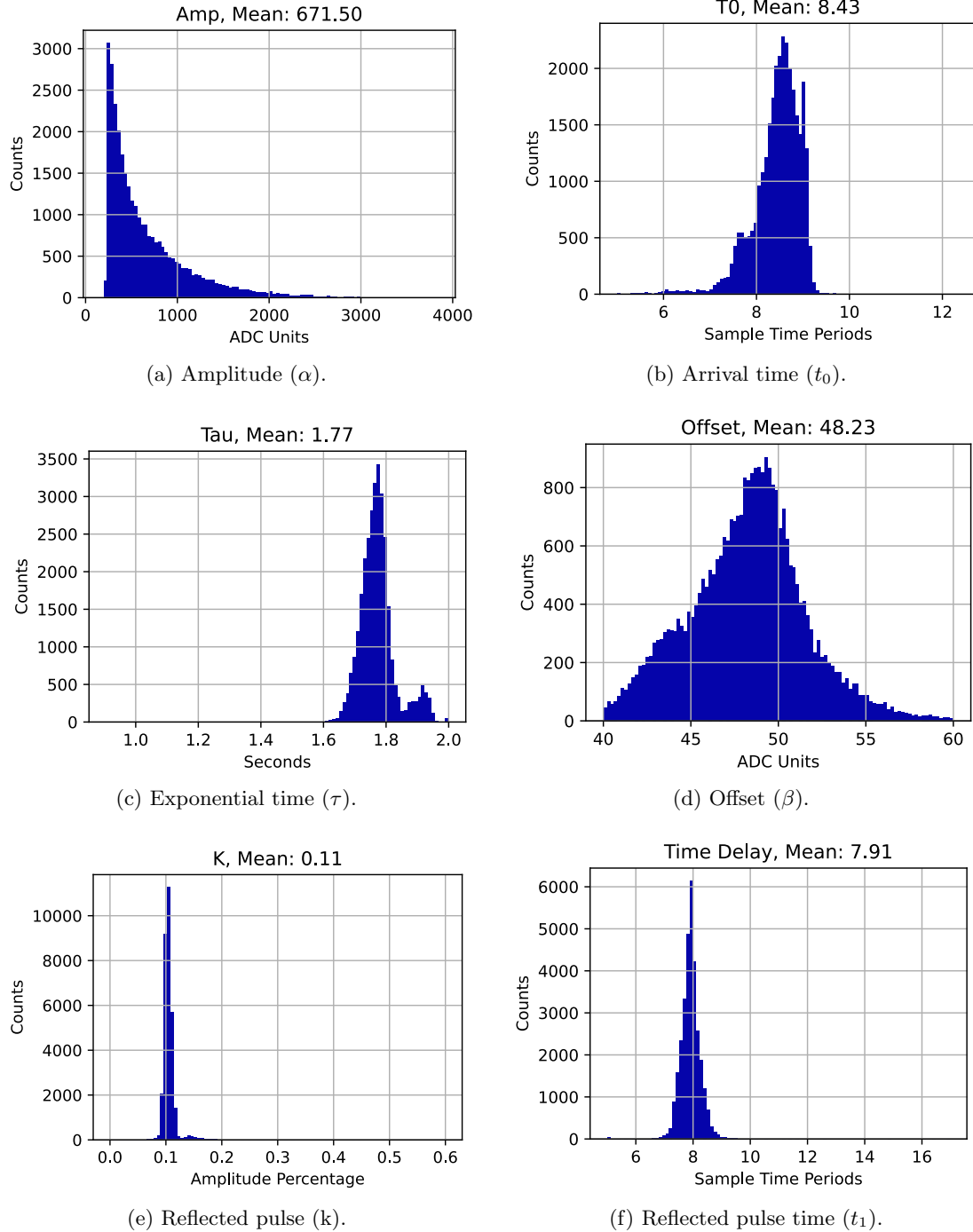
Figure 8.14: Histograms of the fitted parameters corresponding to the biexponential ECAL2 pulse model.

172

(a) Residuals.

(b) Autocorrelation values.
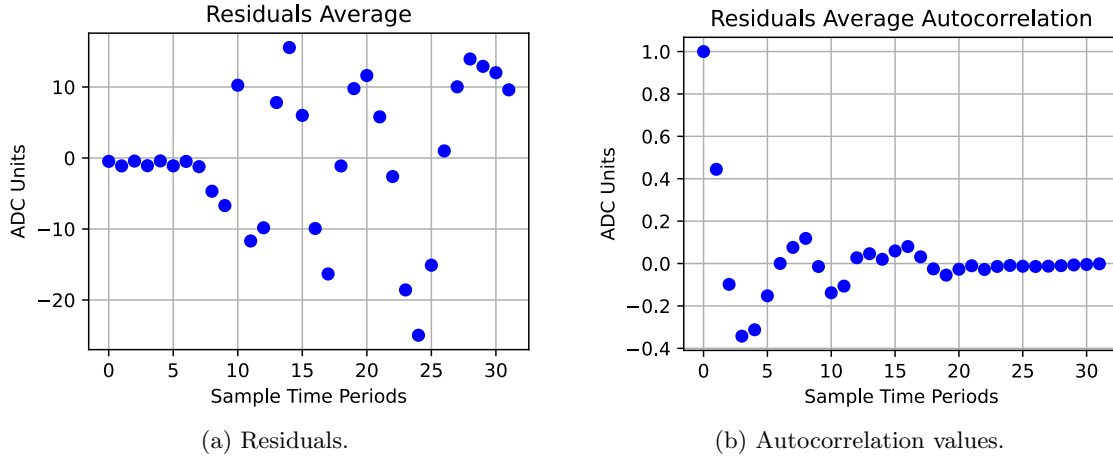
Figure 8.15: Residuals of the fitted parameters corresponding to the biexponential ECAL2 pulse model.

is calculated as the average of all pulse values. The t_0 is the arrival time of the pulses, calculated as the sample number in which the pulse starts inside the 32 samples trace. τ represents the exponential time of the pulse, and the two peak distributions can be based on two different types of pulses owing to the difference in the particles that started the scintillation process or to some discrepancies in the shaping filter response. For the offset, as each channel is digitally corrected to have a base line of 50 counts, it can happen that they are not properly aligned owing to a voltage drift on the analog channels itself. As the parameters related to the reflected pulse (k and t_1) are relative to the main pulse, they have a much more concentrated distribution than the others, as shown in Figures 8.14e and 8.14f.

The values that represent the fitting bounty are related to the residuals (Figure 8.15a), obtained as the mean difference between the fitted model and the real pulses.

8.7 Digital Pulse Processor Implementation

The DPP operates continuously, processing digitized signals from the channels and extracting pulse features in real-time as they arrive.

As discussed in Section 8.3, the amplitude of the pulses is measured using an FIR operator. The FIR coefficients are generated using the DPLMS method, and have a length of 20 taps (sampling periods), which is sufficient since more than 99% of the pulse energy is contained within this length, according to the 5τ rule [169].

Alternatively, and similar to the case of Section 8.3.1, symmetrical trapezoidal coefficients with the same length and a zero zone of 4 taps can be used to achieve noise reduction.

Now with the model of the pulse already fitted to the data base, for obtaining a set of coefficients and similarly to the Equation 8.18, the next equation is used for the minimization:

$$\Psi_{ECAL2}(c_0, c_1, \ldots, c_{k-1}) = \alpha_1 \cdot \sum_{j=1}^{t_R+t_{FT}+t_F-(k-1)} \left(\sum_{i=0}^{k-1} c_i \, x_{k+j-i} - A \right)^2 +$$

$$\alpha_{1a} \cdot \sum_{j=t_R+t_F}^{t_R+t_F+t_{FT}-1} \left(\sum_{i=0}^{k-1} c_i \, x_{k+j-i} - A \right)^2 + \alpha_{1b} \cdot \left(\sum_{i=0}^{k-1} c_i \, x_{k+j_{peak}-i} - A \right)^2 + \qquad (8.23)$$

$$\alpha_2 \cdot \left(\sum_{i=0}^{k-1} c_i \right)^2 + \alpha_3 \cdot \sum_{j=0}^{k-1} \sum_{i=0}^{k-1} c_i \, c_j \, ACF_{|i-j|} + \alpha_4 \cdot \left(\sum_{i=0}^{k-1} c_i - T_i \right)^2$$

Where, $j_{peak} = t_R + t_F + \lfloor \frac{t_{FT}}{2} \rfloor$, with $\lfloor . \rfloor$ is the floor function. As before A stands for the condition of Equation 8.7, ACF for the 8.14 and $\{T_i\}$ the trapezoidal template of length $t_R+t_{FT}+t_F$.

The parameters are slightly different from those of Subsection 8.4.3 because of the different shapes of the pulses and because of the addition of a new constraint, but the methodology is the same.

After the coefficient minimization procedure, the FIR is applied to the best-fit pulse model (Figure 8.16a), and the output of the filter is like the one depicted in Figure 8.16b.

From Figure 8.16 it can be seen that the filter output remains flat during 5 samples, time enough for absorbing the pulse detection uncertainties (constrained to ±1 sample period).

Once the coefficients are obtained what remains is to constrain them to operate as integers of 16 bits width, by this way enlarging the 12 bits resolution to 18 bits at the FIR's output.

8.7.1   Finite Impulse Response Design

The FIR implementation was done using the FIR Compiler IP of Xilinx [170]. The configuration of the Compiler IP depends on the specific requirements of the application and has diverse options. It offers several different architectures, each of them has its own advantages and disadvantages in terms of latency, area, and power consumption. The quantization process of the coefficients introduces quantization errors, which can affect the filter's performance [171]. Higher resolution coefficients can help reduce quantization errors and improve the accuracy of the filter's response. If the input data has sufficient resolution, the higher-resolution coefficients can help shape the filter response more accurately, reducing artifacts or distortion in the output signal. The overall resolution of the output will still primarily depend on the resolution of the input data.

For the implementation of our FIR neither the area nor the power consumption is an issue, the preference is for the lowest latency, as it will be operating for real-time tasks and must provide the output the fastest and accurately as possible. The filter was generated using a datapath of 12 bits with a multiply and accumulate (MAC) with single rate frequency architecture, and with the parameters and characteristics as in Table 8.6a, using Vivado 2019.1.3 and for the part XZCU4EG.

The results of Table 8.6b are only for the FIR filter; for complete DPP results, we need to add the resources for the detection and the logic for the synchronization and buffering of the detected data.

The implementation of a single-channel Digital Pulse Processor (DPP) can
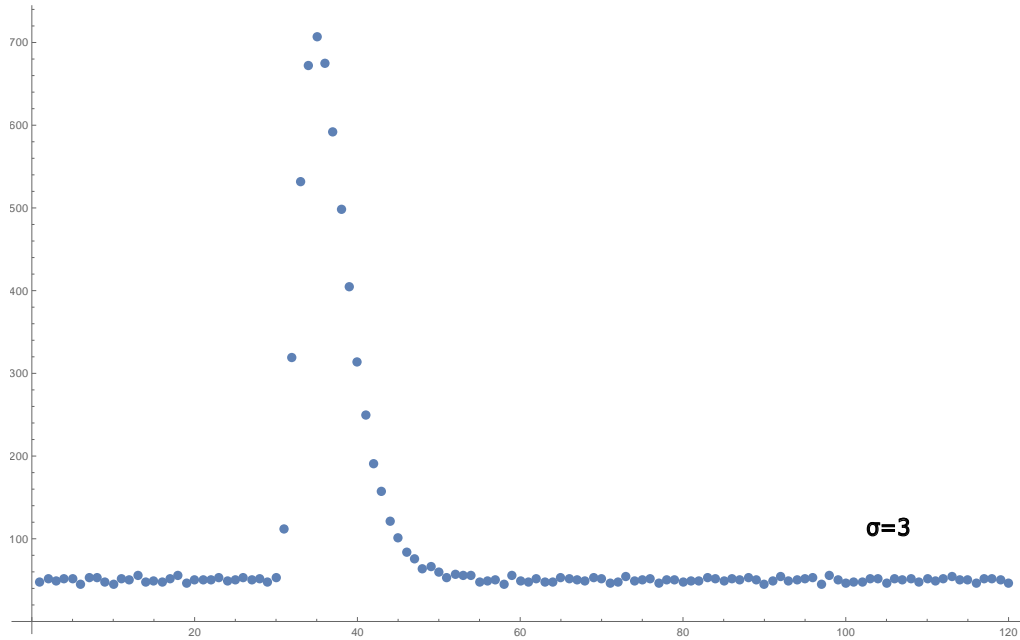
(a) Pulse model plus noise.
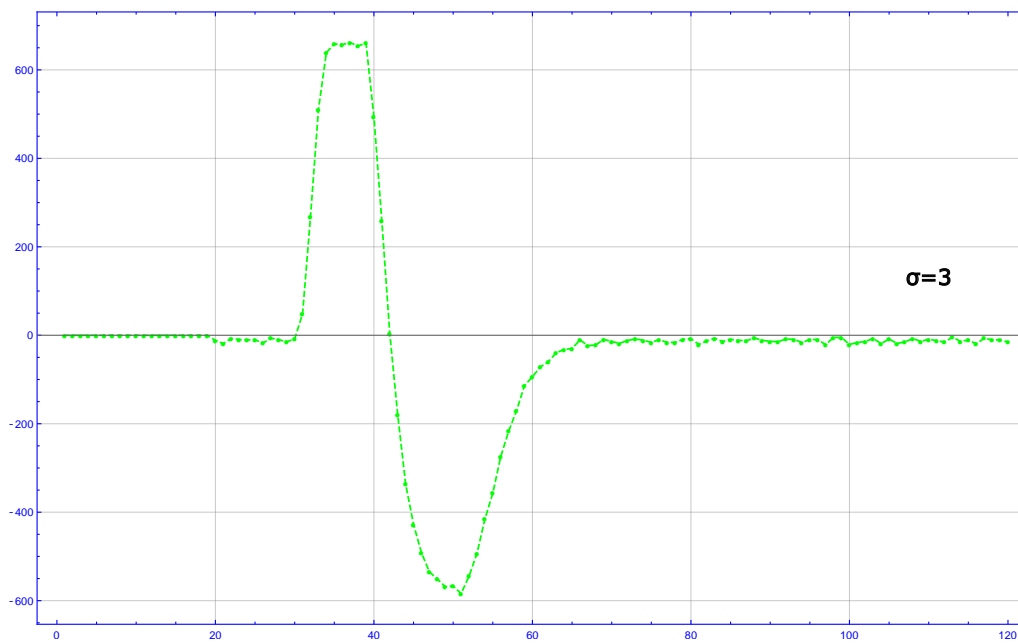
(b) FIR output.

Figure 8.16: Pulse model for best-fit parameters and FIR output.

Table 8.6: FIR implementation characteristics and results for XCZU4EG.

(a) Implementation characteristics.

Parameter	Value
Filter type	Single rate
Clock frequency	80 MHz
Input data width	12 bits
Number of coefficients	20
Coefficient width	16 bits
Rounding mode	Full precision
Output width	18 bits
Cycle latency	17
Filter architecture	Systolic MAC

(b) Implementation results.

Resource	Utilization (%)
LUT	232 (0.42%)
REGISTER	409 (0.21%)
DSP	10 (1.37%)
BRAM	0 (0.00%)

be depicted in Figure 8.17, which showcases the different components and their interconnections.

The input of the DPP receives the raw ADC data stream, which is then divided into two parallel branches. The lower branch is dedicated to pulse detection, while the upper branch is responsible for amplitude measurement.

The pulse detection is achieved by applying a weighted differentiation technique, which is empirically adjusted based on the signal statistics. The differentiator IP allows for programmable coefficient values to adapt the output to the channel's dynamic range. It is pre-loaded with standard and pre-calculated values by default.

The amplitude measurement is performed with the generated FIR filter. The IP delivers a continuous data stream, where the output can be interpreted as an uninterrupted filtered data flow.

A finite state machine within the DPP handles various tasks such as configuring and
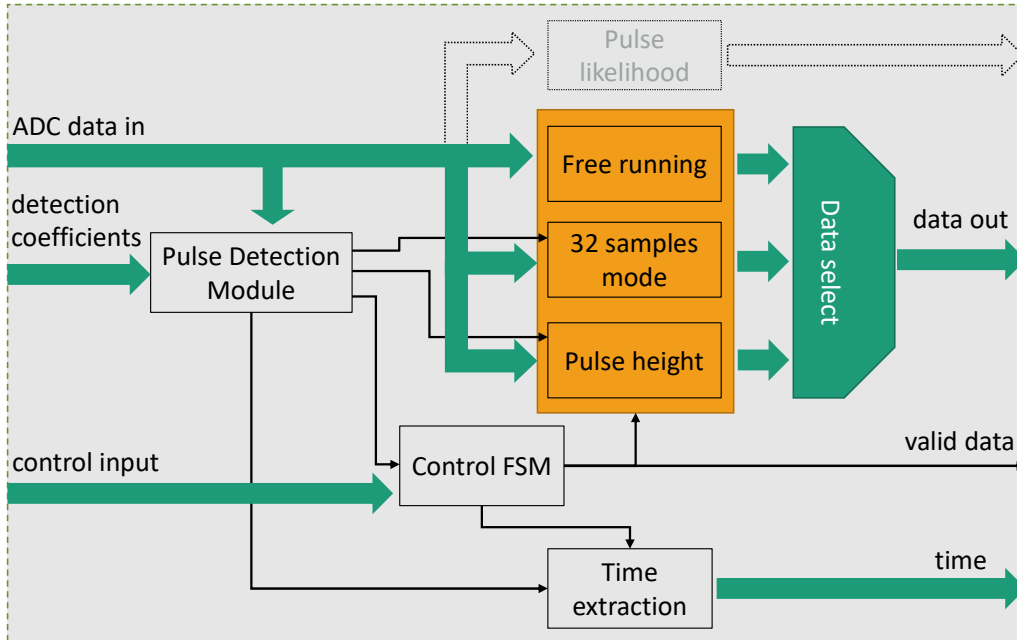
Figure 8.17: DPP IP implementation with its main features.

selecting the operation mode, synchronizing the data to the buffers, managing the data valid signals, and selecting the data output source. Additionally, the DPP can provide both raw data and the derivative at the output, accommodating different operating modes to meet AMBER requirements and very useful when debugging the systems.

The IP block in charge of detecting the arrival of the pulse uses 6 DSPs in addition. To extract the timing information of the pulse, a 16-bit counter has been implemented. This counter operates as a free-running counter at a frequency of 80 MHz. It is reset at the beginning of each slice, which has a duration of 100 μs as in Figure 3.9. The synchronization and reset of the counter are controlled by the Trigger Control System (TCS), which provides the necessary framing information. The specific handling of pulses that start or end at the boundary of an old/new slice is still under consideration by the AMBER collaboration. However, for the initial implementation, the pulse will be sent within the slice in which it was detected.
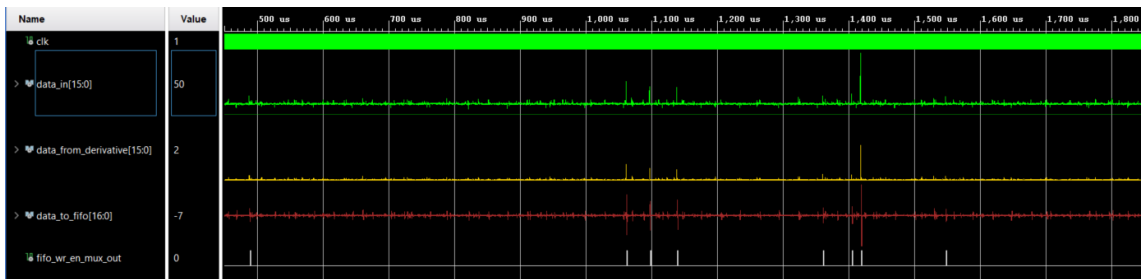
(a) Long trace of DPP real signals from ECAL2.
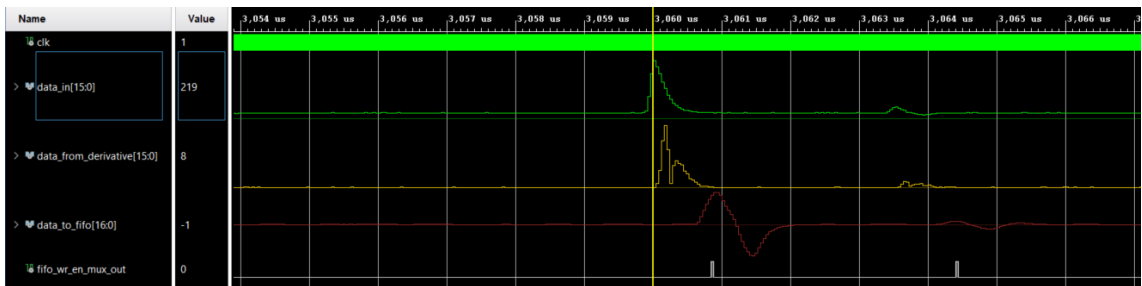
(b) Detected and measured pulses.

Figure 8.18: ECAL2 prototype signals processed with the DPP for amplitude extraction.

Table 8.7: DPP FPGA implementation results for XCZU4EG.

Resource	Utilization (%)
LUT	348 (0.50%)
REGISTER	513 (0.28%)
DSP	16 (2.20%)
BRAM	0 (0.00%)

In Figure 8.18 there is depicted a raw data stream from the ECAL2 prototype with a DPP working and some of it main signals plotted.

For a single channel, the whole DPP with its main building block uses the resources as in Table 8.7.

The DPP block design shown in Figure 8.3 can be enhanced by incorporating an IP block to provide a pulse likelihood index. This index is based on the research described in [172], which allows for the estimation of whether a pulse exhibits the desired shape characteristics relevant to the physics being studied. The study was conducted for simulated pulses under different signal to noise ratio conditions, and it is based on the Pearson correlation of the incoming signal with the pulse model template. By analyzing the pulse shape, it becomes possible to differentiate between genuine pulses of interest and spurious signals that may lead to false detections. This pulse shape information can be utilized by the trigger processor to make informed decisions about the reliability of the data provided by the DPP. Incorporating the pulse likelihood index helps enhance the digital trigger system's overall accuracy and efficiency.

The incorporation of the pulse likelihood IP into the DPP can be achieved by computing a simplified Pearson correlation index, which requires significantly fewer logic resources compared to the original method while maintaining excellent performance; the implementation results for the XCZU4EG are shown in Table 8.8. However, it should be noted that the simplified correlation index introduces a latency of 1095 clock cycles, in

Table 8.8: Single channel resources utilization for Pearson Coefficient Index (PCI) and the Simplified Pearson Coefficient Index (SPCI) for XCZU4EG.

Resource	PCI Utilization (%)	SPCI Utilization (%)
LUT	21,349 (11.12%)	22,718 (11.83%)
REGISTER	21,524 (12.22%)	23,468 (13.33%)
DSP	120 (16.48%)	53 (7.28%)
BRAM	0 (0.00%)	0 (0.00%)

contrast to the 17 clock cycles required for amplitude and time extraction.

Consequently, when using the correlation index (likelihood) in the system, it becomes necessary to implement a mechanism for tagging past events and associating the corresponding data with the event under study. This mechanism ensures that the data required for providing the correlation index is properly synchronized and available for processing, despite the latency introduced by the simplified correlation index.

8.8 Summary

In the first part of this chapter, a procedure is presented for optimizing a set of finite impulse response filter (FIR) coefficients for digital pulse amplitude measurement. An optimized filter was designed using an adapted digital penalized least mean square (DPLMS) method. The effectiveness of the procedure was demonstrated using a dataset from a case study of high-resolution X-ray spectroscopy based on single-photon detection and energy measurements.

By applying the optimized filter, significant improvements were achieved in the energy resolutions of the Kα and Kβ lines of the Mn energy spectrum. Specifically, the energy resolutions were improved by approximately 20% compared to the reference values obtained by fitting individual photon pulses with the corresponding mathematical model.

These results highlight the potential of this approach to enhance the accuracy and precision of pulse-amplitude measurements in nonuniform and stochastic pulsatile systems, such as high-resolution X-ray spectroscopy.

The second part discusses a study of ECAL2 pulses using a dataset from a COMPASS run in 2012. The dataset consists of pulses originated from subparticles generated by the interaction between a fixed target and a 160 GeV muon beam and does not include information about the position or type of scintillation material. The analysis focuses solely on the shape of the pulses. The pulse information was obtained by analyzing a dataset containing approximately 1.2 million segments from COMPASS DAQ. However, owing to the limited length of the pulses (32 samples), the frequency resolution is insufficient for studying low-frequency spectral content. To address this limitation, a different dataset was obtained from COMPASS data-taking with a setup using the ECAL2 prototype and an ad hoc DAQ based on the FFeCCa.

The noise statistics for each channel of the ECAL2 prototype, including the mean value and standard deviation of the noise, were estimated. The results are presented in Table 8.4. The noise was also used to optimize the FIR coefficients with its autocorrelation function.

A bi-exponential model was proposed to model the pulses. The model consists of a second-order semi-Gaussian pulse and considers the superposition of an ideal signal with a secondary small pulse and noise. Model parameters were determined by fitting a subgroup of representative pulses using a Differential Evolution algorithm. The fitted parameter values are listed in Table 8.5 and their distributions are shown in Figure 8.14.

Overall, this chapter demonstrates the practicality and benefits of the proposed procedure for optimizing FIR filter coefficients in digital pulse-amplitude measurements. The positive outcomes obtained in the case study indicate the potential for improving the performance of energy measurement to any type of pulse if the model is known and noise is characterized.

Chapter 9

Experimental Results

The FFeCCa as frontend electronic, including the codec scheme and the digital pulse processor, is a powerful and versatile solution for capturing and processing data from multiple channels in parallel. A thorough testing and evaluation process was conducted to ensure its reliability and assess its performance, thereby providing valuable insights into the system's capabilities.

The first part of this chapter presents part of the testing performed to characterize the interface between the MSADC and the SoM. The testing process was designed to verify the data integrity, speed, and resource utilization of the system. The MSADC interface with the SoC-FPGA hardware was tested by incorporating a pattern generator IP in the firmware that can be multiplexed with the MSADC driver output, allowing the generation of a known data sequence, which serves as a reference for evaluating the integrity of data transmission. This step is crucial for validating the ability of the system to reliably capture and transmit data and evaluate the need for a compression scheme.

Then, the system integration and details of the implementations for the final version of the DAQ system are presented. Resource utilization analysis revealed that the system efficiently utilized the available hardware resources and optimized the performance. This efficiency is crucial for maximizing the system capabilities and ensuring scalability for

future applications.

Finally, the metrics of the different operating modes are presented, revealing the advantages of the features extraction method compared with all previous approaches, including the current method used in the ECAL2 system.

9.1 FFeCCa: Maximum data rates between MSADC and SoC-FPGA

The FFeCCa DAQ system was tested using a two-step scheme to ensure data integrity and to evaluate its performance when using the MSADC. The SoC-FPGA firmware incorporates an IP to generate a known number sequence that can be multiplexed with the MSADC driver's output. This was selected to validate the data integrity from the CDC stage up to the receiver on the PC side, as shown in Figure 7.8. On the MSADC side, a similar data generator IP is used to transmit a known data sequence to the SoM and validate it on the PC side.

In addition to validating the ADC reading scheme and configuration, the system was tested using various signal shapes generated by an arbitrary bench waveform generator. These tests were successful, indicating that the system can handle different input signals without encountering any issues.

A specific test was conducted to determine the maximum data rate that the MSADC sustains without causing data corruption. A free-running counter was fed to a serdes configured to serialize the data in a DDR 8:1 scheme. The frequency of the counter+serdes was gradually increased, and the number of incorrect codes received on the PC side was evaluated. The clock frequency control was maintained from the SoM side, whereas the data were generated within the MSADC at the same frequency. The clock frequency was multiplied by four and forwarded to the SoM using an ODDR to mitigate clock-loading problems and to readjust the phase modifications made in the fabric. This fast clock also

Table 9.1: Error rate vs MSADC-SoM transmission frequency in the FFeCCa.

Frequency (MHz)	Wrong Codes	Errors (%)
330	0	0
340	0	0
350	31	6
360	39	8
370	50	10
380	243	47
390	512	100

fed the OSERDES for bit serialization, as shown in Figure 5.7.

These comprehensive tests ensured the validation of the FFeCCa + MSADC DAQ system, covering data integrity, ADC reading capabilities, and the maximum sustainable data rate without data corruption. The results obtained from these tests provided valuable insights into the performance of the system and helped to optimize its operation.

The results of the tests conducted on the FFeCCa + MSADC DAQ system are summarized in Table 9.1. The table illustrates the performance of the system at different clock frequencies, indicating the point at which data reliability starts to degrade. From the presented data, it can be observed that the system maintains reliable data transmission up to a clock frequency of 340 MHz. Beyond this threshold, the integrity of the transmitted data begins to diminish.

With these tests, we were able to start working in collaboration with a group from the University of Warsaw to develop a compression technique for transmitting the 16 channels in a lossless manner. The maximum operating frequency is crucial because of the compression rate dependence between the channel capacity and raw data from the ADCs.

9.2 System integration

The ECAL2 elements, as described in Section 4.3, require polarization with a High Voltage (HV) power source that incorporates an initial soft start mechanism to ensure safe operation. Owing to variations in the scintillating material properties, aging, and other factors, each channel exhibits unique characteristics. Therefore, it is essential to have an independent channel voltage control system that can be programmed accordingly.

An SPI interface was utilized to enable the configuration of the PMT high-voltage power supply. The PMTs cannot be powered without proper programming. For this purpose, part of the hardware subsystem from another project [173] based on the CIAA-ACC board was reused. To facilitate interfacing with a PC, the UDMA [174] was implemented with slight modifications to automate the ramp-up and ramp-down voltage levels of the PMTs.

The UDMA Command-Line Interface (CLI) depicted in Figure 9.1 provides a user-friendly interface for programming and controlling the high-voltage power supply of the PMTs. In addition, a specific command was introduced into the UDMA to control the driver of the LED pulser. This driver activates an LED array connected to 25 optical fibers, which are in turn linked to the scintillating material cages. It is worth mentioning that the driver for the LED pulser requires a voltage level translator to enable its activation and subsequent powering of the LED array. The activation of the LED pulser allows for periodic testing of the ECAL2 elements response.

To enable online data acquisition at the COMPASS beam area, a dedicated PC was prepared as the readout control and buffering system. Equipped with the required hardware and software, the PC serves as a central hub for data acquisition, buffering, and processing, ensuring efficient handling of high data rates. The PC also provides the resources for giving access to debugging and programming all the boards (MSADC, FFeCCa and CIAA for the high voltage power supply control). This setup enhances data collection while prioritizing personnel safety in the beam area.

186
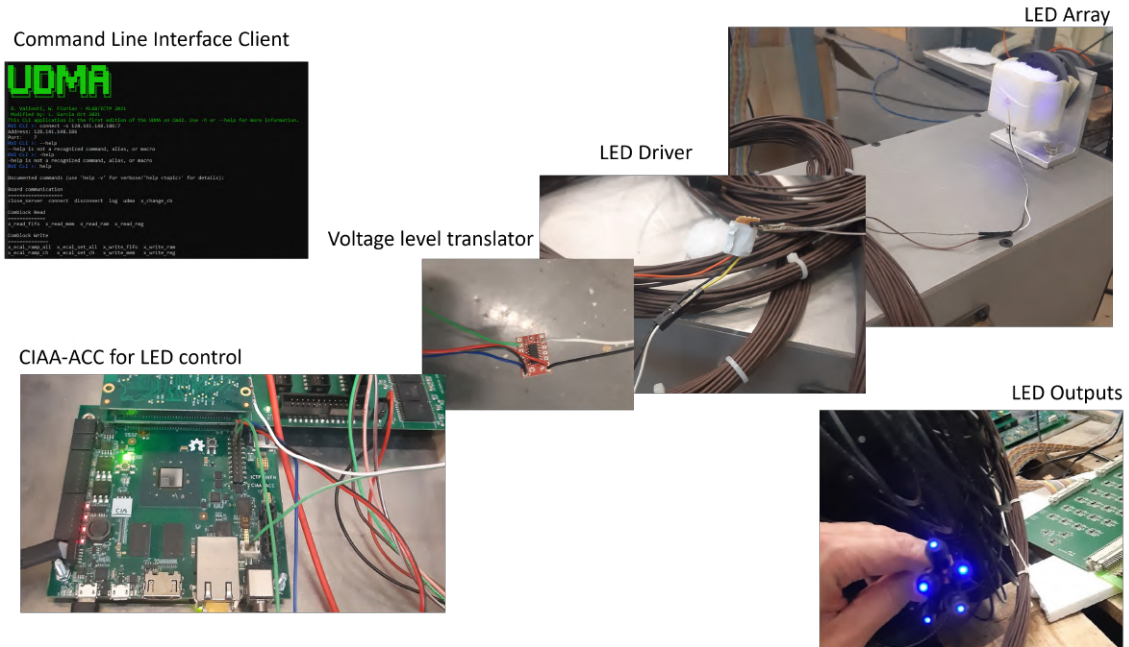
Figure 9.1: UDMA CLI client, LED pulser, and not connected fibers.

In that setup, as introduced in Section 7.4 the two channels streamers of Section 7.3.2 were implemented. In the context of the AMBER pilot run, the intensive testing of the system with different PMT regimes (polarization settings) and in various operational modes (with beam and without beam) aimed to study and characterize the noise behaviour of the system under different conditions.

The noise in each mode can have distinct characteristics and sources. For instance, in the absence of a beam, the noise can arise from various sources such as electronic components, thermal noise, and environmental interference. By analyzing the noise patterns and levels in this mode, it is possible to assess the baseline noise performance of the system and identify any potential issues or sources of interference.

During beam operations, additional noise sources may come into play. High-energy particle beams can induce ionizing radiation, which may generate additional noise and interfere with the signal readings [175]. This radiation-induced noise can vary depending

on factors such as beam intensity, beam energy, and beam profile. By studying the noise characteristics during beam operations, it is possible to evaluate the system's robustness and its ability to handle and distinguish signal events from noise in the presence of radiation.

Testing the system with different PMT polarization regimes is crucial because it allows for the evaluation of the system's response and performance under varying voltage settings. By systematically adjusting the PMT voltages and studying the resulting noise patterns, it becomes possible to optimize the voltage settings to achieve the best signal-to-noise ratio and overall performance for each channel. From previous studies, the calibration factor in % must be in the range between 0-40 %, if greater then the gain is too high and the PMT output saturates the MSADC inputs.

Figure 9.2 shows the mean value and standard deviation of different channels of the ECAL2 prototype when operating in beam area with a beam-on condition, during particle spill-on time. It can be observed that even if the noise values are almost constant in between the channels, the noise, which is proportional to the standard deviation of the signals, increases together with the polarization values of the PMTs. It can also be observed that some of the channels are noisier than others, in our case channel 15 and channel 3 for certain polarization values.

No significant differences were found in the error distribution during the beam-on/beam-off conditions.

From Figure 9.3, the raw data of channel 4 operated during different beam conditions can be observed. When there is beam, on Figures 9.3a and 9.3c, random pulses coming from the beam can be seen. These pulses indicate the presence of signal events caused by the interaction of the beam with the ECAL2 elements. On the other hand, when there is no beam, as shown in Figures 9.3b and 9.3d, only noise is present without signal events.

The histograms provide a visual representation of the distribution of the recorded data. In both cases, the histograms are plotted with a tap size of two, which represents
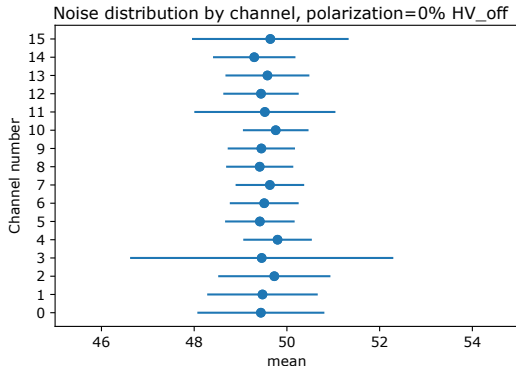
(a) HV power off, PMT polarizing V 0%.
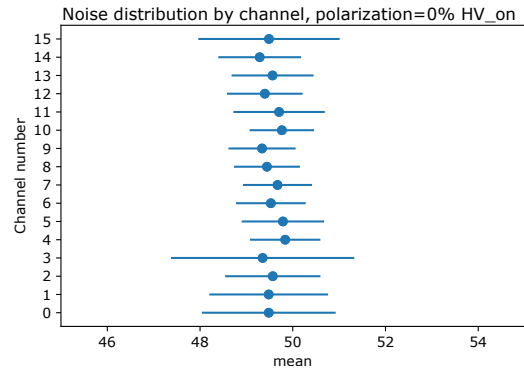
(b) HV power on, PMT polarizing V 0%.
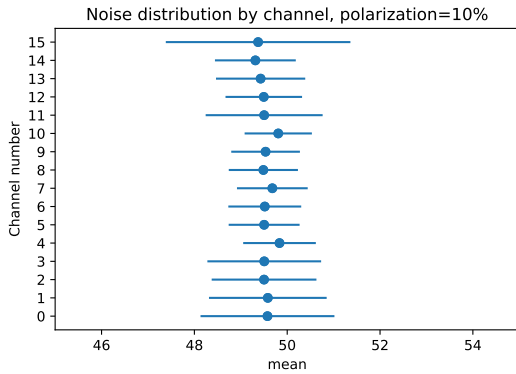
(c) PMT polarizing V 10%.
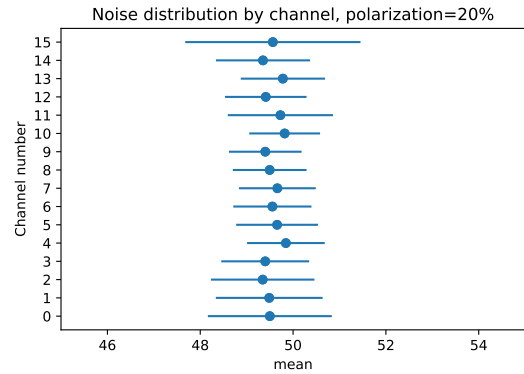
(d) PMT polarizing V 20%.
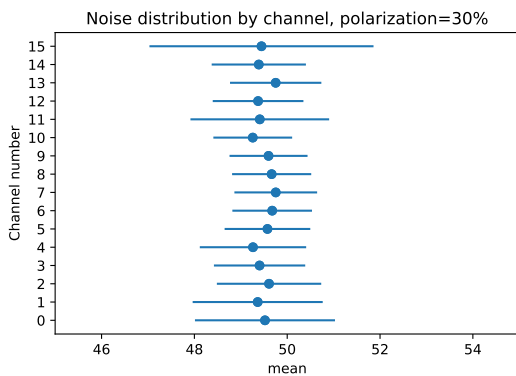
(e) PMT polarizing V 30%.
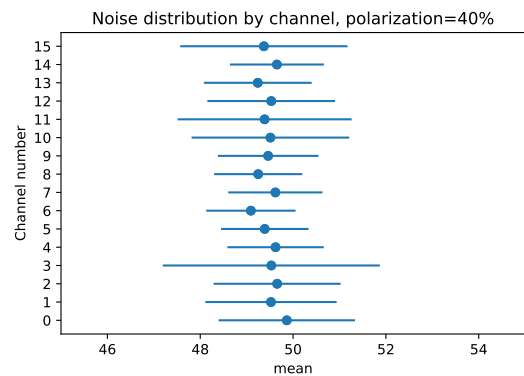
(f) PMT polarizing V 40%.

Figure 9.2: Mean baseline values and standard deviation of all channels for different PMT polarization regimes.

189
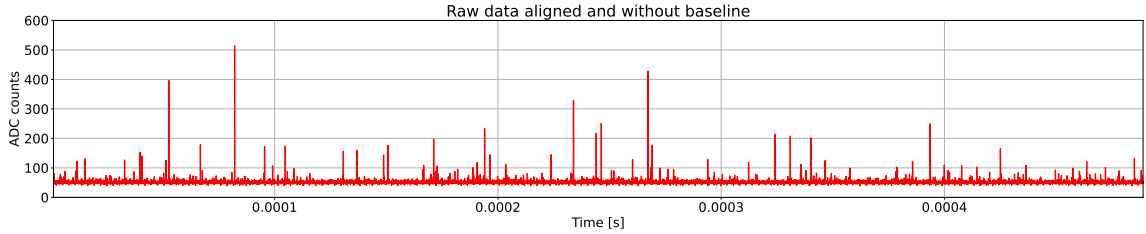
(a) Channel with 40% polarization index during beam on.
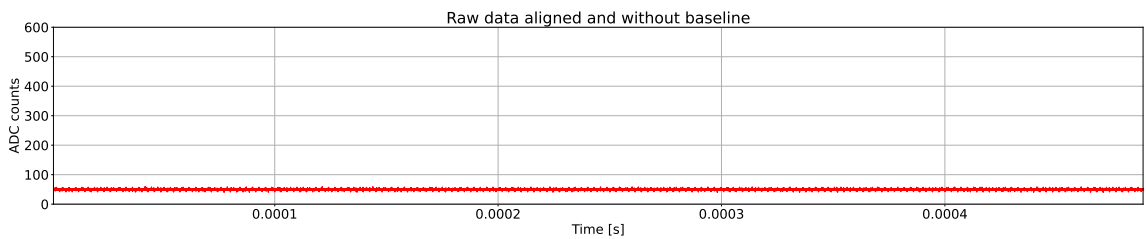
(b) Channel with 40% polarization index during beam off.
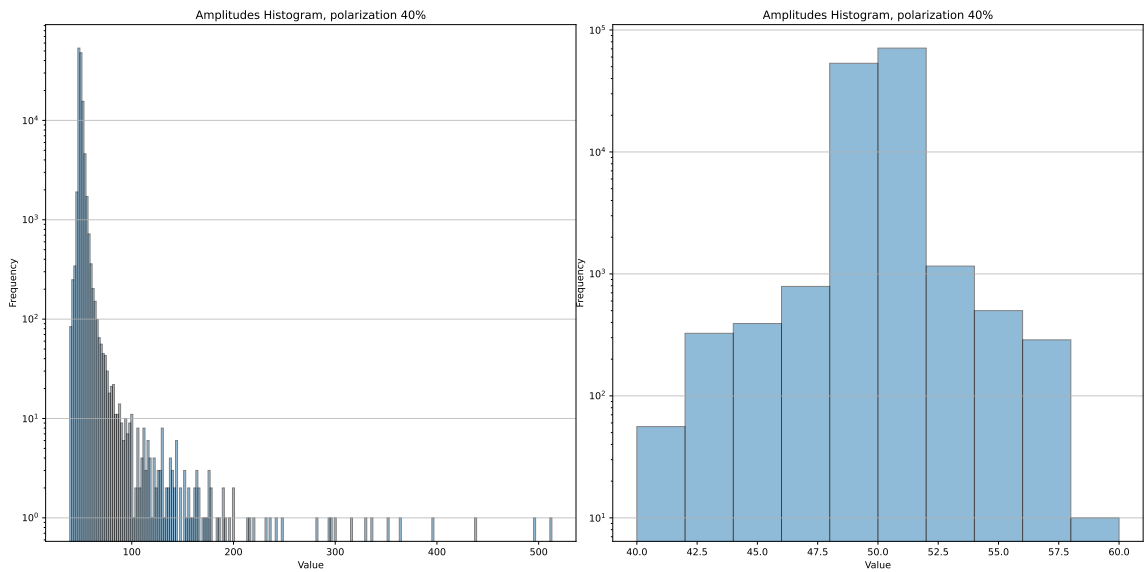
(c) Channel with 40% polarization index during beam on, histogram.

(d) Channel with 40% polarization index during beam off, histogram.

Figure 9.3: Raw data signals from a single channel during different beam conditions.

190

the grouping of data values into bins. The y-axis of the histograms is displayed in a logarithmic scale to emphasize the distribution of low-occurrence events, such as baseline plus noise and the event pulses.

9.2.1 Integration

With all the main developed IP cores the system integration was implemented as in Figure 9.4, understanding the IP Cores as:

1. MSADC

   - ADC Drivers

   - Channels restorer

   - 16 channels encoder

2. SoC-FPGA

   - clock managers

   - 16 channels decoder + CDC FIFOs

   - 16 Digital Pulse Processor

   - AXI Stream Packager

From the MSADC perspective, the design implementation uses the resources as in Table 9.2, while for SoC-FPGA the implementation of the integration involves the resources as in Table 9.3.

In the implementation of the system, it is assumed that the statistics of the ECAL2 signals remain constant. This assumption is necessary because the assignment of Huffman codes and the coefficients of the DPP filters is based on the statistical characteristics of the signals.

Table 9.2: Resources utilization for XC4VLX25: Sixteen channels encoder implementation for free-running mode.

Resource	Utilization (%)	Available
LUT	6205 (28)	21504
REGISTER	3478 (16)	21504
BRAM	14 (19)	72
DSP	0 (0)	48
IO	168 (37)	448
BUFG	11 (34)	32
DCM	4 (50)	8

Table 9.3: Resources utilization for XCZU4EG: Sixteen channels decoder implementation with DPP incorporated for free-running mode.

Resource	Utilization (%)	Available
LUT	12778 (14.55)	87840
REGISTER	12990 (7.39)	175680
BRAM	32.50 (25.39)	128
URAM	16 (33.33)	48
DSP	416 (57.14)	728
IO	50 (19.84)	252
BUFG	5 (1.42)	352
MMCM	2 (50.00)	4
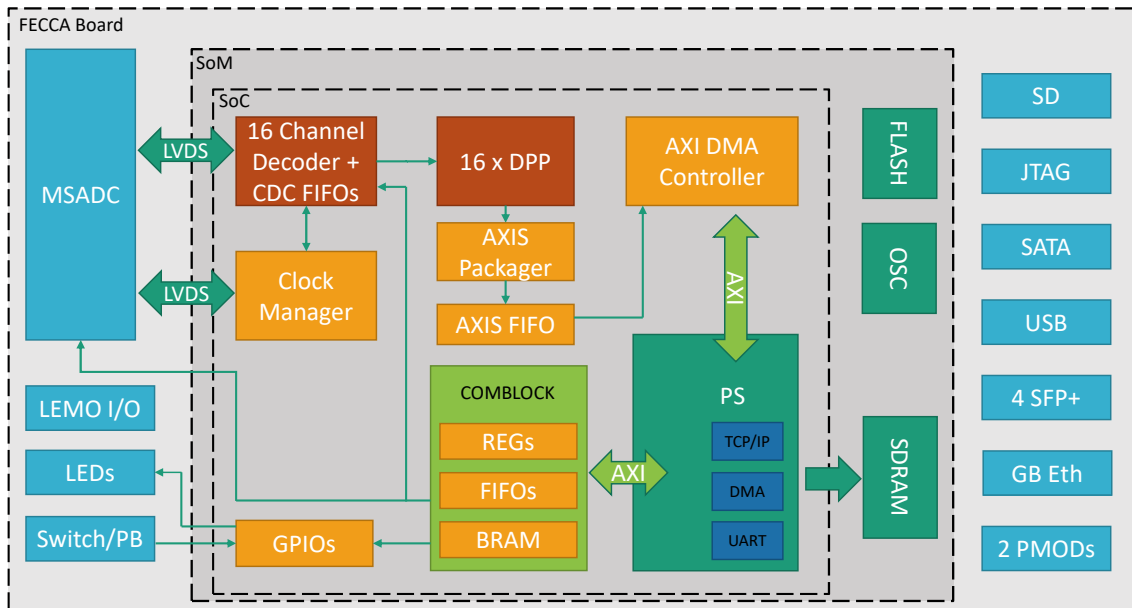
Figure 9.4: System integration.

However, if the statistics of the signals change the assigned Huffman codes are no longer optimal for compression, and the coefficients of the DPP filters may not provide optimal signal processing performance.

To adapt the system to the changing statistics of the signals, readjustments of the Huffman codes and the coefficients of the DPP filters are required. This readjustment process involves analyzing the new statistical characteristics of the signals and recalculating the optimal coding and filtering parameters accordingly. By performing these readjustments, the system can work in a free-running mode for other pulse-like detectors.

However, it is important to note that the readjustment process requires thorough analysis and validation to ensure that the updated Huffman codes and filter coefficients effectively capture the new signal characteristics.

9.2.2 Performance

The system performance of the features extraction implementation demonstrated significant improvements with respect to the current triggered approach in terms of the amount of transmitted data, leading to savings in the processing time and storage resources. Table 9.4 presents an overview of the performance in various operating modes, highlighting the advantages of the features extraction method. The data presented for both the zero suppression and features extraction techniques are based on a maximum event rate of 100 kHz and represent the lower limit scenario where the event occurs in only one channel.

The operation modes and their corresponding functionalities can be summarized as follows:

- **Raw Data**: In this mode, the system operates in free-running mode, continuously streaming raw data from all detectors at the ADC sampling frequency. No data processing or selection is performed, and the entire data stream is transmitted.

- **32 Samples**: When a trigger signal is received, the system captures and transmits 32 samples from each channel. This mode provides a fixed number of samples for further analysis and reduces the amount of transmitted data compared to the raw data mode.

- **32 Samples with Zero Suppression**: Similar to the previous mode, this mode captures and transmits 32 samples from each channel upon receiving a trigger signal. However, an additional threshold value is applied to each channel as a zero suppression technique. Only channels that exceed the threshold value are transmitted, effectively suppressing channels with low amplitudes or noise.

- **Features Extraction**: In this mode, the system performs data processing on the captured samples. The Digital Pulse Processing algorithm analyzes the channels

194

Table 9.4: System performance of different implementation approaches for the free-running mode.

Approach	Data per event	Mean data rate per channel	Total data rate (100 ch)
Raw Data	-	960 Mbit/s	96 Tbit/s
32 Samples	384 bits	38.4 Mbit/s	3.84 Tbit/s
32 Samples with Zero Suppression	384 bits	38.4 Mbit/s	38.4 Mbit/s
Features Extraction	34 bits	3.4 Mbit/s	3.4 Mbit/s

and extracts features such as amplitude and time information. The amplitude is represented with 18 bits, and the time of arrival with 16 bits.

The last two methods mentioned in the items indicate specific scenarios in which only one element in the ECAL2 wall is hit.

The system has been designed to support free-running and continuous streaming, as requested by the AMBER collaboration. However, in order to transmit the 16 channels at a rate of 960 Mbit/s (15360 Mbit/s), a dedicated implementation for the connection with the DAQ is required. The FFeCCa board's SFP+ transceivers are capable of providing rates of up to 10 Gbps each. This opens up the possibility of implementing a 40 Gbit interface, which would allow for the transmission of raw data from all 16 channels.

The extraction of the time of arrival and the amplitude of the pulses means a saving of up to three magnitude orders of data when compared to the triggered to the 32 Samples non-zero-suppressed approach. And, about one order of magnitude when compared to the 32 samples with zero-suppression.

By adopting a finer granularity and using the concept of an image to represent the timing, the time of arrival can be effectively represented with 5 bits, resulting in significant resource savings. Combining the amplitude and timing information in this way allows for

the usage of 21 bits against 384 bits compared to the zero suppression technique. This translates into a data compression factor of 18.28 times. However, further discussion with the AMBER collaboration is necessary to determine if the image framing information can be obtained from the TCS (Trigger and Control System) for the proposed approach.

Chapter 10

Conclusions and Future Work

This thesis focuses on the development of a trigger-less readout system for the ECAL2 detector of the COMPASS/AMBER experiment. With this system, the detector can operate in a free-running mode, significantly reducing the amount of transmitted data and improving offline processing time.

To provide a solution to the system, I developed a hardware platform (FFeCCa) that reuses the current digitizer board (MSADC) of the detector and includes a SoM based on Ultrascale+ Zynq MPSoC for the online data processing. The trigger-less and free-running operation is based on real-time data features extraction. For this purpose, a novel digital pulse processor (DPP) was developed and implemented by considering the pulse model and channel noise characteristics of the ECAL2 detector. Extensive data analysis was performed, and various methods were evaluated for pulse amplitude extraction.

The integration of the FFeCCa board with the DPP involved meticulous firmware implementations, hardware developments, and a pilot run to collect data for developing a lossless compression, study the noise, and validate the system operation.

Since one of the core devices of the ECAL2 DAQ is the MSADC, the first task in the workflow of this research was to characterize the maximum data rates of the interfaces between the MSADC and the SoM. This is because one of the constraints for the AMBER

upgrades is that the digitizing stage for the frontend of ECAL2 should remain the same, as the resolution and sampling frequency of the ADCs are sufficient for the characteristics of the signal shapes coming from the detectors. Based on the original firmware of the MSADC-FPGA, additional IP cores were implemented for interfacing with the SoM. To enable the MSADC board to interface with any commercial FPGA development board featuring an FMC connector, custom hardware has been developed and produced (MSADC Adapter Board).

After multiple tests and iterative cycles of firmware development, it was determined to consolidate the main FPGA IP Cores, μP **C** code and Python interface Script into an unified framework. The framework (OFSODA) is specially designed for working with multichannel ADC DAQ systems based on SoC-FPGA devices. This significantly expedites the development process and aims at reducing development times and enhance overall productivity. The framework is open and released under a BSD-3 license [BSD-3].

To facilitate hardware interfacing and establish direct communication between the FPGA and a PC, the Universal DMA (UDMA) was used. The UDMA was developed as a simple Command Line Interface (CLI) client, running on the PC side and interacting with a TCP server operating on the SoC-FPGA device. UDMA consists of a set of Python packages and **C** libraries for μP, and is licensed under a GPL-3 license [GPL].

Using OFSODA as the base design for the SoC-FPGA, several multichannel acquisition systems were implemented, and both triggered and free-running approaches were successfully conducted. For the free-running mode, three different approaches were designed and validated: a) two-channel arbitrary and independently selectable, b) eight-channel fixed, and c) sixteen-channels with a lossless encoder. For the triggered mode: a) two-channels up to 4 million data points or b) sixteen-channels with a buffering length of 4096 data points per channel.

Using both UDMA and OFSODA, it is possible to configure the systems and manage slow-control tasks from a remote-access PC.

During a pilot run, the FFeCCa was subjected to various measurements to assess its functionality and reliability. The platform exhibited excellent performance, meeting the required specifications for ADC conversion, signal processing, and data transmission. The measurements validated the board's ability to handle a large number of channels.

The final implementation of the FFeCCa also involved the integration of a codec with the DPP. The codec allows for efficient compression of the acquired data, allowing the transmission of all ADC channels between the MSADC and the SoM. The DPP amplitude extraction is based on a custom Digital Penalized Least Mean Squared method adapted to the ECAL2 pulse model and intrinsic noise characteristics, offering the possibility of adjusting each channel of the system according to its statistics. The DPP also offers the possibility of working in the legacy COMPASS operating mode (triggered) or in a free-running mode, sending the raw data as received from the ADCs. The time extraction is provided by a free-running counter, offered as a coarse grain inside the microseconds time slice projected by the AMBER proposal. In addition, there is the possibility of including a pulse-likelihood index for providing informed decisions about the nature of the pulses to the trigger processor.

The FFeCCa platform has proven to be an essential component for the free-running DAQ of AMBER experiment. The multichannel high-performance signal processing capabilities makes FFeCCa an efficient and reliable platform for precise and fast measurements in HEP experiments. The inclusion of a DPP in free-running mode allows for saving processing time and storage resources, contributing to accurate event reconstruction and data analysis. Additionally, the modular hardware design of FFeCCa allows for flexibility in hosting different front-end readouts and SoM families. Furthermore, the developed hardware is released under an open hardware license, fostering collaboration and innovation in the field.

10.1   Final Remarks and Future Work

An alternative implementation for the 16-channel streamer system is an 8-channel streamer without the codec. Although this approach lacks half of the channels compared to the 16-channel implementation, it offers a unique advantage over the encoder scheme.

In the encoder scheme, there is a potential risk of data loss if the pulse rate exceeds the expected 100 kHz events over a certain period or if the statistics of the channels change drastically. Although the probability of data loss is low, there is still a possibility, and once the data are lost, they cannot be recovered. This limitation poses a challenge in scenarios in which the event rate is higher than anticipated.

By contrast, the 8-channel approach, which transmits the raw ADC data, avoids the risk of data loss. By bypassing the encoding process, it ensures that all data are transmitted without any loss, even in high-event-rate situations. This makes it a viable solution for scenarios involving unexpectedly high pulse rates.

However, it is important to note that neither the 16-channel encoder system nor the 8-channel streamer system can coexist in a single implementation. This is because of the sharing of hard block resources between these two approaches. Therefore, the choice between the 16-channel encoder system and the 8-channel streamer system depends on the specific requirements and constraints of the application.

While the DPP has been initially presented for X-Ray applications, the final deployment will be for the ECAL2. In order to optimize the performance of the DPP and adapt it to the ECAL2 requirements, calibration data is required.

This involves collecting data from known sources or reference signals and comparing them with the output of the DPP. Through this comparison, necessary adjustments can be made to align the DPP's response with the desired performance.

The calibration process typically involves determining the gain, energy resolution, and timing response of the DPP. Gain calibration involves establishing a conversion

factor between the input signal amplitude and corresponding digital output value. Energy-resolution calibration quantifies the ability of the DPP to accurately measure the energy of the detected signals. Timing calibration focuses on determining the time delay and resolution of the output of the DPP.

To perform the calibration, dedicated signals, such as pulses with precise timing characteristics or special datasets referenced with calibrated data, must be used. These calibration datasets will be processed by the DPP, and the resulting measurements compared to the expected values. Based on this comparison, the calibration coefficients or correction factors can be determined and applied to optimize the performance of the DPP.

As part of future work, a mandatory task is the commissioning of the frontends into the AMBER DAQ system. This involves the installation and connection of all boards to the 100 central channels of the calorimeter, as shown in Figure 10.1. The green channels indicate the ones to be replaced in the ECAL2 wall.

To accomplish commissioning, the frontends need to be connected to the COMPASS private Local Area Network (LAN) for slow control using Ethernet communication. Additionally, they must be connected to the DAQ system through optical SFP+ interfaces for data transmission and to the Trigger Control System (TCS).

Although the high-speed interface is still under development, the FFeCCa board SFP+ transceivers were tested and characterized, showing a safe behavior for the connection with the DAQ (3.25 Gbps). To bridge the interface between the frontends and the AMBER multiplexers, a media converter will be used.

During the commissioning process, special attention needs to be given to the implementation of the AMBER communication protocol. Both the FFeCCa and the FPGA/MUX (Multiplexer) subsystems must be compatible with this protocol. Figure 10.2 shows a scheme of the proposed FFeCCa firmware and hardware. After DPPs, the inclusion of the AMBER packager is followed by the GTH Transceiver IP connected to
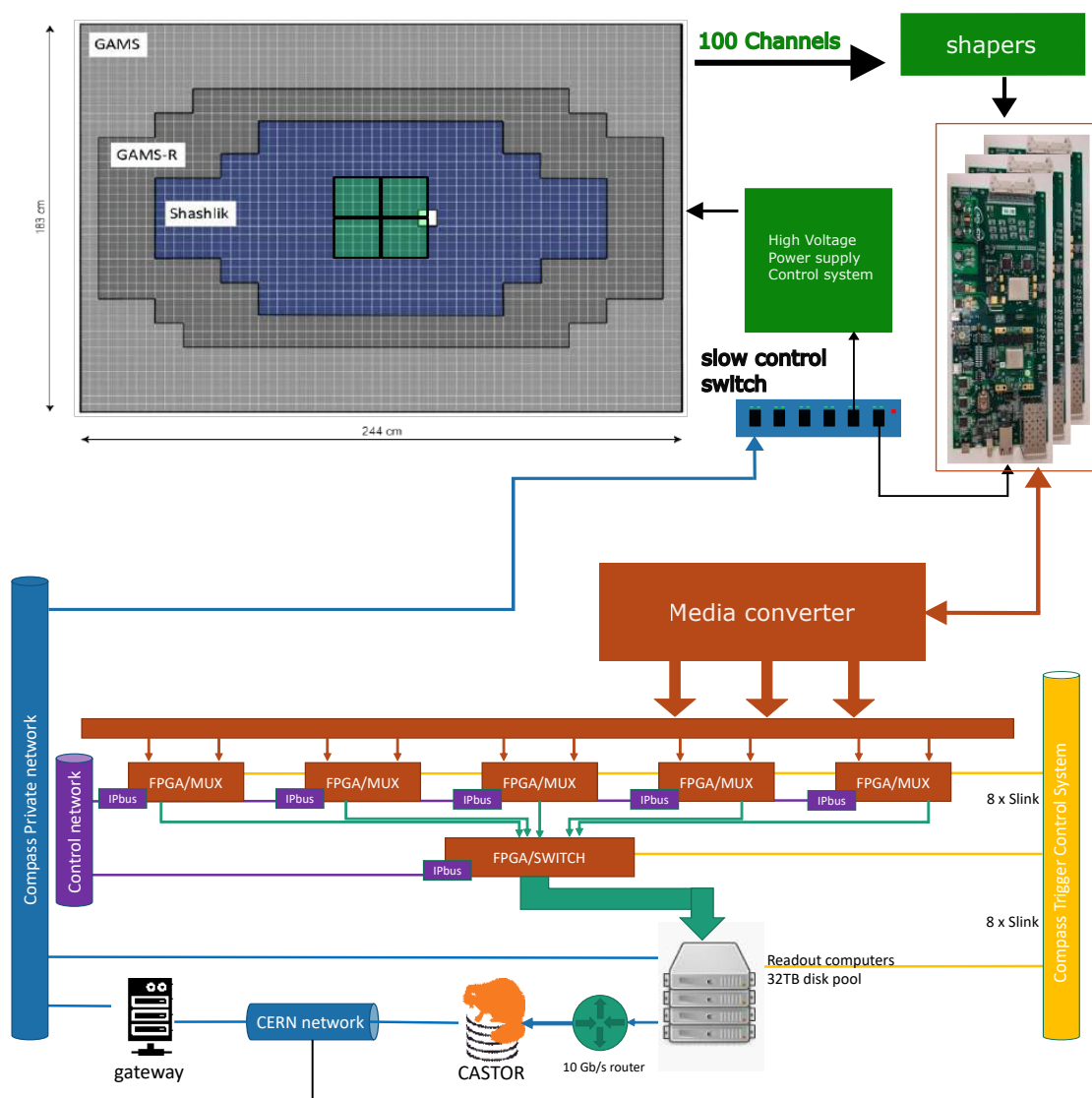
Figure 10.1: Commissioning proposal for the frontend electronics into the AMBER DAQ.
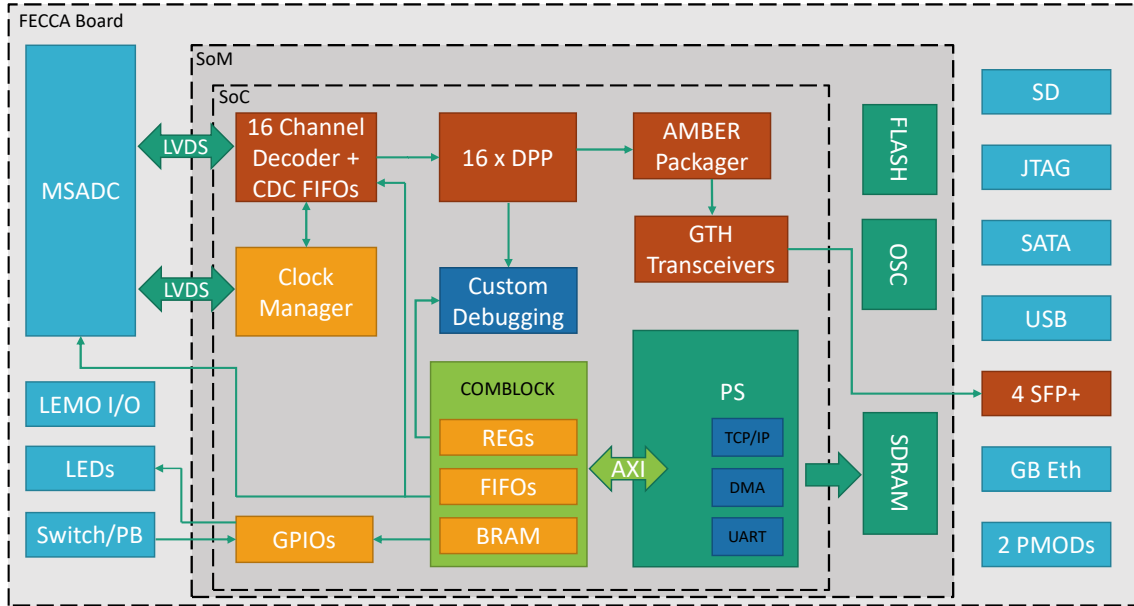
Figure 10.2: System integration for the commissioning.

the SFP+ interface.

Furthermore, the High Voltage Power Supply Control system, responsible for controlling and monitoring PMTs state, must be connected to the COMPASS private network. This connection will also utilize the COMPASS private LAN and Quasar [176], an OPC server implementation that integrates the high-voltage Power Supply Control system with the Detectors Control System of the experiment.

The commissioning process involves meticulous setup, configuration, and testing of the frontends as well as establishing proper communication and data transfer with the AMBER DAQ system. This is a critical step towards the successful integration and operation of the FFeCCa frontends within the larger experimental setup, ensuring reliable data acquisition and control of the calorimeter system.

By completing the commissioning process and establishing the necessary connections and protocols, the FFeCCa frontends can be fully integrated into the AMBER DAQ system, enabling efficient data acquisition and control of the ECAL2 calorimeter.

Chapter 11

Appendix

11.1 Analytical derivation of the proposed bi-exponential ideal pulse model

The proposed heuristic bi-exponential pulse model,

$$S(t) = A \left(1 - 2e^{\frac{-(t-t_0)}{\tau}} + e^{\frac{-2(t-t_0)}{\tau}} \right), \qquad t > t_0 \tag{11.1}$$

is essentially a one-parameter model depending on the exponential time τ which completely determines the pulse shape. The amplitude A is a scale factor and t_0 is a time offset corresponding to the starting point of the pulse. This model can be analytically derived by considering the following transfer function:

$$H(s) = A \left(\frac{1}{s \left(1 + s\tau_1\right) \left(1 + s\tau_2\right) \left(1 + s\tau_3\right)} \right) \tag{11.2}$$

The inverse Laplace transformation of $H(s)$ is,

$$\widetilde{H}(t) = \mathcal{L}^{-1}\{H(s)\} = A \left(1 - \frac{\tau_1^2\, e^{\frac{-(t-t_0)}{\tau_1}}}{(\tau_1 - \tau_2)(\tau_1 - \tau_3)} \frac{\tau_2^2\, e^{\frac{-(t-t_0)}{\tau_2}}}{(\tau_1 - \tau_2)(\tau_2 - \tau_3)} \frac{\tau_3^2\, e^{\frac{-(t-t_0)}{\tau_3}}}{(\tau_1 - \tau_3)(\tau_3 - \tau_2)} \right) \tag{11.3}$$

The heuristic model can be obtained by taking the limits when $\tau_1 \to 0^+$ and $\tau_3 \to 2\tau_2$, and assuming $t > t_0$ and $\tau_2 = \tau$, as follows:

$$S(t) = \lim_{\substack{\tau_1 \to 0^+ \\ \tau_3 \to 2\tau_2}} \widetilde{H}(t) = A \left(1 - 2e^{\frac{-(t-t_0)}{\tau}} + e^{\frac{-2(t-t_0)}{\tau}} \right), \qquad t > t_0 \tag{11.4}$$

11.2 Continuous estimation of the angular coefficient of a background ramp

Let us consider the error between a point y_i of a set of experimental data $\{(x_i, y_i)\}$ and the corresponding value of its linear approximation with coefficients a and b,

$$\epsilon_i = y_i - (ax_i + b) \tag{B.1}$$

Assuming a regularly spaced discrete abscissa and $x_i = i$ we can write

$$\epsilon_i = y_i - (ai + b), \qquad i = 1, 2, 3, \dots \tag{B.2}$$

Now, the total error E can be defined as the sum of all squared errors,

$$E = E(a, b) = \sum_i (y_i - (ai + b))^2 \tag{B.3}$$

where the sum is extended to n consecutive samples. E is a function of the two variables a and b. Having a positive quadratic form, in its minimum the partial derivatives of E with respect to a and b must both be zero, then

$$\begin{cases} \frac{\partial E}{\partial a} = 0 \\ \frac{\partial E}{\partial b} = 0 \end{cases} \Rightarrow \begin{cases} -2\sum_i (y_i - ai - b)i = 0 \\ -\sum_i (y_i - ai - b) = 0 \end{cases} \tag{B.4}$$

After expanding all terms in previous expressions we obtain

$$\begin{cases} \sum_i i y_i - a \sum_i i^2 - b \sum_i i = 0 \\ \sum_i y_i - a \sum_i i - nb = 0 \end{cases} \tag{B.5}$$

By solving the above system of equations and considering that

$$\sum_{i=1}^{n} i = (1+n)\frac{n}{2} \tag{B.6}$$

and

$$\sum_{i=1}^{n} i^2 = \frac{1}{6} n (1+n) (1+2n) \tag{B.7}$$

the angular coefficient a can be defined as a linear combination of the n data values y_i as follows:

$$a = \sum_{i=1}^{n} -6 \frac{(1+n-2i)}{(n^3-n)} y_i \tag{B.8}$$

The above expression shows that the estimation of the angular coefficient a is a linear combination of the last n data samples and a set of constant coefficients. The angular coefficient can then be continuously evaluated by means of an FIR filter.

References

[1] Xilinx. *Zynq UltraScale+ MPSoC Data Sheet: Overview (DS891)*. English. AMD-Xilinx. Nov. 2022. 42 pp.

[2] V. González et al. "Data Acquisition in Particle Physics Experiments". In: *Data Acquisition Applications*. Ed. by Zdravko Karakehayov. Rijeka: IntechOpen, 2012. Chap. 11. DOI: 10.5772/48463. URL: https://doi.org/10.5772/48463.

[3] "Pixel Detectors: From Fundamentals to Applications". In: Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. Chap. 2, pp. 25–128. ISBN: 978-3-540-28333-1. DOI: 10.1007/3-540-28333-1_2. URL: https://doi.org/10.1007/3-540-28333-1_2.

[4] J.L. Schmalzel and D.A. Rauth. "Sensors and signal conditioning". In: *IEEE Instrumentation & Measurement Magazine* 8.2 (2005), pp. 48–53. DOI: 10.1109/MIM.2005.1438844.

[5] G. Kasprowicz W. M. Zabołotny et al. "FPGA and Embedded Systems Based Fast Data Acquisition and Processing for GEM Detectors". In: *Journal of Fusion Energy* 38.3 (Aug. 2019), pp. 480–489. DOI: 10.1007/s10894-018-0181-2. URL: https://doi.org/10.1007/s10894-018-0181-2.

[6] B T Hjertaker et al. "A data acquisition and control system for high-speed gamma-ray tomography". In: *Measurement Science and Technology* 19.9 (July

2008), p. 094012. DOI: 10.1088/0957-0233/19/9/094012. URL: `https://dx.doi.org/10.1088/0957-0233/19/9/094012`.

[7] Xinrui Zhang et al. "BRAM-based asynchronous FIFO in FPGA with optimized cycle latency". In: *2012 IEEE 11th International Conference on Solid-State and Integrated Circuit Technology*. 2012, pp. 1–3. DOI: `10.1109/ICSICT.2012.6467891`.

[8] In: *Pixel Detectors: From Fundamentals to Applications*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. Chap. 3, pp. 129–199. ISBN: 978-3-540-28333-1. DOI: `10.1007/3-540-28333-1_3`. URL: `https://doi.org/10.1007/3-540-28333-1_3`.

[9] Herwig Schopper Christian W. Fabjan. In: *Particle Physics Reference Library - Volume 2: Detectors for Particles and Radiation*. Springer Cham, 2020. Chap. 1, pp. 1–4. ISBN: 978-3-030-35318-6.

[10] P. Abbon et al. "The COMPASS experiment at CERN". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 577.3 (2007), pp. 455–518. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/j.nima.2007.03.026`.

[11] G. Aad et al. In: 15.04 (Apr. 2020), P04003. DOI: `10.1088/1748-0221/15/04/P04003`. URL: `https://dx.doi.org/10.1088/1748-0221/15/04/P04003`.

[12] The CMS Collaboration et al. "The CMS experiment at the CERN LHC". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08004–S08004. DOI: `10.1088/1748-0221/3/08/s08004`. URL: `https://doi.org/10.1088/1748-0221/3/08/s08004`.

[13] Glenn F Knoll. *Radiation detection and measurement*. John Wiley & Sons, 2010.

[14] David A. Hennessy John L. Patterson. *Computer Architecture: A Quantitative Approach*. 3rd ed. Morgan Kaufmann Publishers, 2003. ISBN: 1558607242.

[15]   Richard Clinton Fernow. *Introduction to experimental particle physics*. CUP. Cambridge University Press, 1986. ISBN: 052130170X.

[16]   Herwig Schopper Christian W. Fabjan. *Particle Physics Reference Library - Volume 2: Detectors for Particles and Radiation*. Springer Cham, 2020. ISBN: 978-3-030-35318-6.

[17]   Helmut Burkhardt. "Accelerators for Particle Physics". In: *Handbook of Particle Detection and Imaging*. Ed. by Ivor Fleck et al. Cham: Springer International Publishing, 2021, pp. 161–183. ISBN: 978-3-319-93785-4. DOI: `10.1007/978-3-319-93785-4_7`. URL: `https://doi.org/10.1007/978-3-319-93785-4_7`.

[18]   V V Gligorov and M Williams. "Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree". In: *Journal of Instrumentation* 8.02 (Feb. 2013), P02013–P02013. DOI: `10.1088/1748-0221/8/02/p02013`. URL: `https://doi.org/10.1088/1748-0221/8/02/p02013`.

[19]   Volker Lindenstruth and Ivan Kisel. "Overview of trigger systems". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 535.1 (2004). Proceedings of the 10th International Vienna Conference on Instrumentation, pp. 48–56. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/j.nima.2004.07.267`. URL: `https://www.sciencedirect.com/science/article/pii/S0168900204015748`.

[20]   Rainer Bartoldus, Catrin Bernius, and David W. Miller. *Innovations in trigger and data acquisition systems for next-generation physics facilities*. 2022. DOI: `10.48550/ARXIV.2203.07620`. URL: `https://arxiv.org/abs/2203.07620`.

[21]   Andreas Hoecker. "Physics at the LHC Run-2 and Beyond". In: (2017), pp. 153–212. DOI: `10.23730/CYRSP-2017-005.153`. arXiv: `1611.07864`. URL: `https://cds.cern.ch/record/2236645`.

[22] *ATLAS: technical proposal for a general-purpose pp experiment at the Large Hadron Collider at CERN*. LHC technical proposal. Geneva: CERN, 1994. DOI: `10.17181/CERN.NR4P.BG9K`. URL: `https://cds.cern.ch/record/290968`.

[23] Christine Sutton. *Quantum chromodynamics*. 2016. URL: `https://www.britannica.com/science/quantum-chromodynamics`. accessed: 18.10.2022.

[24] Riccardo Maria Bianchi and ATLAS Collaboration. "ATLAS experiment schematic illustration". 2022. URL: `http://cds.cern.ch/record/2837191`.

[25] Mizukami, Atsushi. "ATLAS Level-1 Endcap Muon Trigger for Run 3". In: *EPJ Web Conf.* 245 (2020), p. 01002. DOI: `10.1051/epjconf/202024501002`. URL: `https://doi.org/10.1051/epjconf/202024501002`.

[26] *Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System*. Tech. rep. Geneva: CERN, 2017. DOI: `10.17181/CERN.2LBB.4IAL`. URL: `https://cds.cern.ch/record/2285584`.

[27] P. Aarnio et al. *CMS, The Compact Muon Solenoid, Technical Proposal*. English. Tech. rep. CERN/LHCC 94-38, LHCC/P1. European Laboratory for Particle Physics CERN, 1994, p. 291.

[28] Tai Sakuma. *Cutaway diagrams of CMS detector*. 2019. URL: `https://cds.cern.ch/record/2665537`. accessed: 07.11.2022.

[29] A Oh. "The CMS DAQ and run control system". In: *Journal of Physics: Conference Series* 110.9 (May 2008), p. 092020. DOI: `10.1088/1742-6596/110/9/092020`. URL: `https://dx.doi.org/10.1088/1742-6596/110/9/092020`.

[30] Remigius Mommsen et al. *The CMS event-builder system for LHC run 3 (2021-23)*. Tech. rep. Geneva: CERN, 2019. DOI: `10.1051/epjconf/201921401006`. URL: `https://cds.cern.ch/record/2649145`.

[31] *LHCb : Technical Proposal*. Geneva: CERN, 1998. URL: `https://cds.cern.ch/record/622031`.

[32] The LHCb Collaboration et al. "The LHCb Detector at the LHC". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005. URL: https://dx.doi.org/10.1088/1748-0221/3/08/S08005.

[33] J.-P. Dufey et al. "The LHCb trigger and data acquisition system". In: *1999 IEEE Conference on Real-Time Computer Applications in Nuclear Particle and Plasma Physics. 11th IEEE NPSS Real Time Conference. Conference Record (Cat. No.99EX295)*. 1999, pp. 49–53. DOI: 10.1109/RTCON.1999.842561.

[34] J Albrecht et al. "The upgrade of the LHCb trigger system". In: *Journal of Instrumentation* 9.10 (Oct. 2014), p. C10026. DOI: 10.1088/1748-0221/9/10/C10026. URL: https://dx.doi.org/10.1088/1748-0221/9/10/C10026.

[35] Hennessy, Karol. "The DAQ systems of the DUNE Prototypes at CERN/". In: *EPJ Web Conf.* 214 (2019), p. 09001. DOI: 10.1051/epjconf/201921409001. URL: https://doi.org/10.1051/epjconf/201921409001.

[36] *DUNE at LBNF - Fermilab*. 2022. URL: https://lbnf-dune.fnal.gov/. accessed: 07.11.2022.

[37] R. Acciarri et al. *Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE) Conceptual Design Report Volume 1: The LBNF and DUNE Projects*. 2016. DOI: 10.48550/ARXIV.1601.05471. URL: https://arxiv.org/abs/1601.05471.

[38] Kurt Biery et al. "artdaq: DAQ software development made simple". In: *Journal of Physics: Conference Series* 898.3 (Oct. 2017), p. 032013. DOI: 10.1088/1742-6596/898/3/032013. URL: https://dx.doi.org/10.1088/1742-6596/898/3/032013.

[39] F Gautheron et al. *COMPASS-II Proposal*. Tech. rep. SPSC-P-340, CERN-SPSC-2010-014. May 2010. URL: https://cds.cern.ch/record/1265628.

[40]   Jan Friedrich et al. "Measuring the Proton Radius in High-Energy Muon-Proton Scattering". In: *PoS* DIS2019 (2019), 222. 6 p. DOI: 10.22323/1.352.0222. URL: https://cds.cern.ch/record/2701397.

[41]   Ewa Lopienska. "The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022". In: (2022). General Photo. URL: https://cds.cern.ch/record/2800984.

[42]   Dipanwita Banerjee et al. "The North Experimental Area at the Cern Super Proton Synchrotron". In: (2021). URL: https://cds.cern.ch/record/2774716.

[43]   Subrt Ond et al. "The Continuously Running iFDAQ of the COMPASS Experiment". In: *EPJ Web Conf.* 214 (2019), p. 01032. DOI: 10.1051/epjconf/201921401032.

[44]   B. Ketzer et al. "GEM detectors for COMPASS". In: *IEEE Transactions on Nuclear Science* 48.4 (2001), pp. 1065–1069. DOI: 10.1109/23.958724.

[45]   Yuri N Kharzheev. "Scintillation Detectors in Modern High Energy Physics Experiments and Prospect of Their use in Future Experiments". In: *Journal of Lasers, Optics & Photonics* 2017 (2017), pp. 1–9.

[46]   Ix-B García Ferreira, J García Herrera, and L Villaseñor. "The Drift Chambers Handbook, introductory laboratory course (based on, and adapted from, A H Walenta's course notes)". In: *Journal of Physics: Conference Series* 18.1 (Jan. 2005), p. 346. DOI: 10.1088/1742-6596/18/1/010. URL: https://dx.doi.org/10.1088/1742-6596/18/1/010.

[47]   P. Abbon et al. "The COMPASS setup for physics with hadron beams". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 779 (2015), pp. 69–115. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2015.01.

035. URL: https : / / www . sciencedirect . com / science / article / pii / S0168900215000662.

[48]   G.K. Mallot. "The COMPASS spectrometer at CERN". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 518.1 (2004). Frontier Detectors for Frontier Physics: Proceedin, pp. 121–124. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2003.10.038.

[49]   Gregorio Landi and Giovanni E. Landi. "Silicon Micro-Strip Detectors". In: *Encyclopedia* 1.4 (2021), pp. 1076–1083. ISSN: 2673-8392. DOI: 10 . 3390 / encyclopedia1040082. URL: https://www.mdpi.com/2673-8392/1/4/82.

[50]   D Neyret et al. "New pixelized Micromegas detector for the COMPASS experiment". In: *Journal of Instrumentation* 4.12 (Dec. 2009), P12004–P12004. DOI: 10.1088/1748-0221/4/12/p12004. URL: https://doi.org/10.1088%2F1748-0221%2F4%2F12%2Fp12004.

[51]   Dong Jing et al. "Images of triple gas electron multiplier with pixel-pads". In: *Chinese Physics B* 18.10 (Oct. 2009), p. 4229. DOI: 10.1088/1674-1056/18/10/024. URL: https://dx.doi.org/10.1088/1674-1056/18/10/024.

[52]   Y. Giomataris. "Development and prospects of the new gaseous detector "Micromegas"". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 419.2 (1998), pp. 239–250. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(98)00865-1.

[53]   Fabio Sauli. *Principles of operation of multiwire proportional and drift chambers.* Tech. rep. CERN, Geneva, 1975 - 1976. Geneva, 1977. DOI: 10.5170/CERN-1977-009. URL: https://cds.cern.ch/record/117989.

[54] Malte Wilfert. "Final results on the spin dependent structure function g_1^d from COMPASS". In: *PoS* DIS2016 (2016), p. 228. DOI: `10.22323/1.265.0228`.

[55] C. Bernet et al. "The COMPASS trigger system for muon scattering". In: *Nucl. Instrum. Meth. A* 550 (2005), pp. 217–240. DOI: `10.1016/j.nima.2005.05.043`.

[56] *Electronic developments for COMPASS at Freiburg, CATCH*. URL: `https://twiki.cern.ch/twiki/pub/Compass/Detectors/FrontEndElectronics/catch-userguide.ps`. accessed: 24.10.2022.

[57] *iMUX/HGESICA module*. URL: `https://twiki.cern.ch/twiki/pub/Compass/Detectors/FrontEndElectronics/imux_manual.pdf`. accessed: 24.10.2022.

[58] *The GANDALF Module*. URL: `https://gandalf-framework.web.cern.ch/`. accessed: 24.10.2022.

[59] *S-Link — High Speed Interconnect*. URL: `http://hsi.web.cern.ch/HSI/s-link/`. accessed: 24.10.2022.

[60] Jean-Philippe Baud et al. "CASTOR status and evolution". In: (2003). DOI: `10.48550/ARXIV.CS/0305047`. URL: `https://arxiv.org/abs/cs/0305047`.

[61] B Adams et al. *A New QCD facility at the M2 beam line of the CERN SPS: COMPASS++/AMBER*. Tech. rep. Geneva: CERN, 2019. URL: `http://cds.cern.ch/record/2653603`.

[62] *COMPASS Setup for year 2022, CERN Internal Document*. 2022. URL: `https://wwwcompass.cern.ch/`. accessed: 29.11.2022.

[63] H. J. Hilke. *Time Projection Chambers*. Tech. rep. CERN-PH-EP-2010-047. CERN, Oct. 2010.

[64] Jan Friedrich. "Hadron research with AMBER at CERN". In: *Suplemento De La Revista Mexicana De Física* 3.3 (Aug. 2022), pp. 1–7. DOI: `10.31349/SuplRevMexFis.3.0308009`.

[65] Y. Bai et al. "Overview and future developments of the FPGA-based DAQ of COMPASS". In: *Journal of Instrumentation* 11 (Feb. 2016), pp. C02025–C02025. DOI: `10.1088/1748-0221/11/02/C02025`.

[66] B. Adams et al. *COMPASS++/AMBER: Proposal for Measurements at the M2 beam line of the CERN SPS Phase-1*. Tech. rep. CERN-SPSC-2019-022. SPSC-P-360. CERN, Oct. 2019.

[67] *Electromagnetic Shower schema - CERN Internal Document*. 2022. URL: `https://indico.cern.ch/event/318531/attachments/612850/843143/daniela_l5.pdf`. accessed: 07.11.2022.

[68] Christian W. Fabjan and Fabiola Gianotti. "Calorimetry for particle physics". In: *Rev. Mod. Phys.* 75 (4 Oct. 2003), pp. 1243–1286. DOI: `10.1103/RevModPhys.75.1243`. URL: `https://link.aps.org/doi/10.1103/RevModPhys.75.1243`.

[69] In: *Compton Scattering-Investigating the Structure of the Nucleon with Real Photons*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004. Chap. 3, pp. 9–36. ISBN: 978-3-540-40742-3. DOI: `10.1007/b12999`.

[70] Bruno Rossi. "High-Energy Particles". In: *American Journal of Physics* 21.3 (1953), pp. 236–236. DOI: `10.1119/1.1933408`.

[71] *Photomultiplier tube - Matsusada*. 2022. URL: `https://www.matsusada.com/application/ps/photomultiplier_tubes/`. accessed: 29.11.2022.

[72] *HAMAMATSU – About PMTs — Photomultiplier tubes (PMTs)*. 2021. URL: `https:%20//www.hamamatsu.com/eu/en/product/optical-sensors/pmt/about_pmts.html`. accessed: 23.11.2022.

[73] F. Binon et al. "Hodoscope multiphoton spectrometer GAMS-2000". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 248.1 (1986), pp. 86–102.

ISSN: 0168-9002. DOI: `https://doi.org/10.1016/0168-9002(86)90501-2`. URL: `https://www.sciencedirect.com/science/article/pii/0168900286905012`.

[74]   "Radiation hardness of lead glasses TF1 and TF101". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 345.1 (1994), pp. 210–212. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/0168-9002(94)90990-3`. URL: `https://www.sciencedirect.com/science/article/pii/0168900294909903`.

[75]   L. Dobrzynski. "R&D Proposal - Shashlik Calorimetry - A combined Shashlik + Preshower detector for LHC". In: CERN/DRDC 93-28 (1993).

[76]   *FEU84-3 Datasheet.* 1982. URL: `http://lampes-et-tubes.info/pm/FEU_84-3.pdf`. accessed: 13.11.2022.

[77]   M. V. Akopyan et al. *Stand for testing photomultipliers FEU 84-3*. USSR, 1988. URL: `http://inis.iaea.org/search/search.aspx?orig_q=RN:21028877`.

[78]   Thiemo Nagel. "Measurement of the Charged Pion Polarizability at COMPASS". PhD thesis. Technischen Universität München, 2012.

[79]   A Brunner et al. "A Cockcroft–Walton base for the FEU84-3 photomultiplier tube". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 414 (Sept. 1998), pp. 466–476. DOI: `10.1016/S0168-9002(98)00651-2`.

[80]   Alexander Mann, Igor Konorov, and Stephan Paul. "A Versatile Sampling ADC System for On-Detector Applications and the AdvancedTCA Crate Standard". In: *2007 15th IEEE-NPSS Real-Time Conference*. 2007, pp. 1–5. DOI: `10.1109/RTC.2007.4382804`.

[81] *ADS5271 - Eight-Channel, 12-Bit, 50-MSPS Analog-to-Digital Converter (ADC).* 2022. URL: https://www.ti.com/product/ADS5271. accessed: 07.11.2022.

[82] *HOTlink.* 1999. URL: http://hsi.web.cern.ch/components/serialisers/cypress/cy7b923.html. accessed: 07.11.2022.

[83] Xilinx. *Xilinx UG070 Virtex-4 FPGA User Guide, User Guide.* English. AMD-Xilinx. 406 pp.

[84] Alexander Mann, Igor Konorov, and Stephan Paul. "A Versatile Sampling ADC System for On-Detector Applications and the AdvancedTCA Crate Standard". In: *2007 15th IEEE-NPSS Real-Time Conference.* 2007, pp. 1–5. DOI: 10.1109/RTC.2007.4382804.

[85] Judit Freijedo et al. "Impact of Power Supply Voltage Variations on FPGA-Based Digital Systems Performance". In: *Journal of Low Power Electronics* 6 (Aug. 2010), pp. 339–349. DOI: 10.1166/jolpe.2010.1076.

[86] Andres Garcia G, Luis Gonzalez, and Reynaldo Felix. "Power consumption management on FPGAs". In: Mar. 2005, pp. 240–245. ISBN: 0-7695-2283-1. DOI: 10.1109/CONIEL.2005.60.

[87] *Virtex-4 FPGA Data Sheet: DC and Switching Characteristics.* 2009. URL: https://docs.xilinx.com/v/u/en-US/ds302. accessed: 04.10.2022.

[88] *Tips for succesfull power-up of today's high-performance FPGAs.* 2004. URL: https://www.ti.com/lit/an/slyt079/slyt079.pdf. accessed: 04.10.2022.

[89] Texas Instruments. *Understanding Linear Regulators.* Application Note SLVA769A. Texas Instruments, 2016. URL: https://www.ti.com/lit/an/slva769a/slva769a.pdf.

[90] *The ADC08D1520RB: Low-Power, 8-Bit, Dual 1.5 GSPS or Single 3.0 GSPS A/D Converter Reference Board.* 2007. URL: `https://www.ti.com/tool/ADC08D1520RB`. accessed: 04.10.2022.

[91] *Computadora Industrial Abierta Argentina - Alta Capacidad de Computo.* 2017. URL: `https://github.com/ciaa/CIAA_ACC`. accessed: 05.10.2022.

[92] Larry D. Smith & Eric Bogatin. *Principles of power integrity for PDN design – simplified : robust and cost effective design for high speed digital products.* Prentice Hall modern semiconductor design series. Prentice Hall signal integrity library. Prentice Hall, 2017. ISBN: 9780132735551.

[93] *Analog 120-Pin Probing Board Manual.* 2013. URL: `https://www.analog.com/media/en/technical-documentation/user-guides/analog_120_probing_rev.1.0.pdf`. accessed: 11.10.2022.

[94] Eric Bogatin. *Signal and Power Integrity - Simplified.* Prentice Hall, 2009. ISBN: 9780132349796.

[95] *Arria V FPGA Development Kit Board Description.* accessed: 29.03.2023. 2014. URL: `https://www.intel.com/content/dam/www/programmable/us/en/pdfs/literature/ug/ug_arria5_gx_dev_kit.pdf`.

[96] *SmartFusion2 Advanced Development Kit User Guide.* accessed: 29.03.2023. 2021. URL: `https://www.microsemi.com/document-portal/doc_view/144317-smartfusion2-advanced-development-kit-user-guide`.

[97] *ZCU102 Evaluation Board for the Zynq UltraScale+ MPSoC User Guide.* accessed: 29.03.2023. 2018. URL: `https://www.xilinx.com/support/documentation/boards_and_kits/zcu102/ug1182-zcu102-eval-bd.pdf`.

[98] Vladimir Frolov et al. "Data Acquisition System for the COMPASS++/AMBER Experiment". In: *IEEE Transactions on Nuclear Science* 68.8 (2021), pp. 1891–1898. DOI: `10.1109/TNS.2021.3093701`.

[99]   Chenren Xu et al. "The Case for FPGA-Based Edge Computing". In: *IEEE Transactions on Mobile Computing* 21.7 (2022), pp. 2610–2619. DOI: 10.1109/TMC.2020.3041781.

[100] INTEL. *AN 114: Board Design Guidelines for Intel Programmable Device Packages.* Tech. rep. INTEL, 2022.

[101] Xilinx. *UltraScale Architecture Memory Resources User Guide (UG573).* English. AMD-Xilinx. 139 pp.

[102] Xilinx. *UltraRAM: Breakthrough Embedded Memory Integration on UltraScale+ Devices (WP477).* English. AMD-Xilinx. 11 pp.

[103] *mouser XCZU4EG-1SFVC784E price.* 2022. URL: https://www.mouser.it/ProductDetail/Xilinx/XCZU4EG-1SFVC784E?qs=rrS6PyfT74c7fmt2bCEueQ%3D%3D. accessed: 29.10.2022.

[104] *Trenz TE0813-01-4BE11-A price.* 2022. URL: https : / / shop . trenz - electronic.de/en/TE0813-01-4BE11-A-MPSoC-Module-with-Xilinx-Zynq-UltraScale-ZU4EG-1E-2-GByte-DDR4-5.2-x-7.6-cm. accessed: 29.10.2022.

[105] Samtec. *High Speed Characterization Report - Razor beam LP Socket/Terminal Strip Series.* English. Samtec. 36 pp. URL: https://wiki.trenz-electronic.de/download/attachments/36245987/hsc-report-sma_st5-ss5-04mm_web.pdf?api=v2.

[106] Xilinx. *Zynq UltraScale+ MPSoC Data Sheet: DC and AC Switching Characteristics (DS925).* English. AMD-Xilinx. June 2022. 109 pp.

[107] *Copper SFP vs Fiber SFP: Which One is Better?* Aug. 2018. URL: https://www.fiber-optic-solutions.com/copper-sfp-vs-fiber-sfp.html. accessed: 31.10.2022.

[108] Marvell. *Marvell Alaska 88E1512 – Datasheet*. English. MArvell. 166 pp. URL: `https://www.marvell.com/content/dam/marvell/en/public-collateral/phys - transceivers / marvell - phys - transceivers - alaska - 88e151x - datasheet.pdf`.

[109] *Standard NIM instrumentation system*. Revision of the NIM document : AEC Report TID-20893 (Rev. 4) dated July 1974. Washington, DC: United States Department of Energy, 1990. URL: `https://cds.cern.ch/record/2026631`.

[110] "IEEE Standard Test Access Port and Boundary-Scan Architecture". In: *IEEE Std 1149.1-1990* (1990), pp. 1–139. DOI: `10.1109/IEEESTD.1990.114395`.

[111] Erik O. Hammerstad. "Equations for Microstrip Circuit Design". In: *1975 5th European Microwave Conference*. 1975, pp. 268–272. DOI: `10.1109/EUMA.1975.332206`.

[112] IPC. *IPC2141A - Design Guide for High-Speed Controlled Impedance Circuit Boards*. English. IPC. Mar. 2004.

[113] *EN5335QI 3A PowerSoC Datasheet*. Mar. 2019. URL: `https://www.intel.com/content/www/us/en/content-details/632773/intel-enpirion-power-solutions-en5335qi-3a-powersoc-datasheet.html`. accessed: 31.10.2022.

[114] *TEBF0808 Getting Started*. June 2021. URL: `https://wiki.trenz-electronic.de/display/PD/TEBF0808+Getting+Started`. accessed: 31.10.2022.

[115] Xilinx. *JESD204C LogiCORE IP Product Guide (PG242)*. English. AMD-Xilinx. Feb. 2023. 120 pp.

[116] Xilinx. *IBERT for UltraScale GTH Transceivers v1.4 Product Guide (PG173)*. English. AMD-Xilinx. Feb. 2021. 28 pp.

[117] *ZedBoard*. 2022. URL: `https://www.avnet.com/wps/portal/us/products/avnet-boards/avnet-board-families/zedboard/`. accessed: 14.11.2022.

REFERENCES

[118]   Kasun S. Mannatunga et al. "Design for Portability of Reconfigurable Virtual Instrumentation". In: *2019 X Southern Conference on Programmable Logic (SPL)*. 2019, pp. 45–52. DOI: 10.1109/SPL.2019.8714446.

[119]   *AMBA AXI and ACE Protocol Specification – Version H.c.* 2021. URL: https://developer.arm.com/documentation/ihi0022/latest/. accessed: 13.11.2022.

[120]   Xilinx. *Mixed-Mode Clock Management (DS737)*. English. AMD-Xilinx. June 2009. 5 pp.

[121]   *FreeRTOS*. 2022. URL: https://www.freertos.org/. accessed: 15.07.2022.

[122]   *Lightweight IP stack*. 2021. URL: https://www.nongnu.org/lwip/2_1_x/index.html. accessed: 13.11.2022.

[123]   Rodrigo A. Melo et al. "Study of the Data Exchange Between Programmable Logic and Processor System of Zynq-7000 Devices". In: *2019 X Southern Conference on Programmable Logic (SPL)*. 2019, pp. 3–8. DOI: 10.1109/SPL.2019.8714328.

[124]   Werner Florian et al. "An Open-Source Hardware/Software Architecture for Remote Control of SoC-FPGA Based Systems". In: *Applications in Electronics Pervading Industry, Environment and Society*. Ed. by Sergio Saponara and Alessandro De Gloria. Cham: Springer International Publishing, 2022, pp. 69–75. ISBN: 978-3-030-95498-7.

[125]   S Huber et al. "Upgrade of the COMPASS calorimetric trigger". In: *Journal of Instrumentation* 8.02 (Feb. 2013), p. C02038. DOI: 10.1088/1748-0221/8/02/C02038.

[126]   *An Interface for Texas Instruments Analog-to-Digital Converters with Serial LVDS Outputs*. Apr. 2008. URL: https://docs.xilinx.com/v/u/en-US/xapp1208-bitslip-logic. accessed: 13.11.2022.

[127] DA Huffman. "A method for the construction of minimum-redundancy codes". In: *Proceedings of the IRE* 40.9 (1952), pp. 1098–1101.

[128] J.V. Bennett et al. "Precision timing measurement of phototube pulses using a flash analog-to-digital converter". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 622.1 (2010), pp. 225–230. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/j.nima.2010.06.216`. URL: `https://www.sciencedirect.com/science/article/pii/S0168900210014063`.

[129] Maria Liz Crespo et al. "Remote Laboratory for E-Learning of Systems on Chip and Their Applications to Nuclear and Scientific Instrumentation". In: *Electronics* 10.18 (2021). ISSN: 2079-9292. DOI: `10.3390/electronics10182191`.

[130] *LXPLUS*. 2022. URL: `https://abpcomputing.web.cern.ch/computing_resources/lxplus/`. accessed: 10.08.2022.

[131] *EOS*. 2022. URL: `https://eos-web.web.cern.ch/eos-web/`. accessed: 10.08.2022.

[132] Kasun Sameera Mannatunga et al. "Data Analysis and Filter Optimization for Pulse-Amplitude Measurement: A Case Study on High-Resolution X-ray Spectroscopy". In: *Sensors* 22.13 (2022). ISSN: 1424-8220. DOI: `10.3390/s22134776`. URL: `https://www.mdpi.com/1424-8220/22/13/4776`.

[133] V. Jordanov and G.F. Knoll. "Digital pulse processor using moving average technique". In: *IEEE Transactions on Nuclear Science* 40.4 (1993), pp. 764–769. DOI: `10.1109/23.256658`.

[134] Valentin T. Jordanov and Glenn F. Knoll. "Digital synthesis of pulse shapes in real time for high resolution radiation spectroscopy". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers,*

Detectors and Associated Equipment 345.2 (1994), pp. 337–345. ISSN: 0168-9002. DOI: `10.1016/0168-9002(94)91011-1`.

[135] Valentin T. Jordanov et al. "Digital techniques for real-time pulse shaping in radiation measurements". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 353.1 (1994), pp. 261–264. ISSN: 0168-9002. DOI: `10.1016/0168-9002(94)91652-7`.

[136] Zbigniew Guzik and Tomasz Krakowski. "Algorithms for digital γ-ray spectroscopy". In: *Nukleonika* (2013).

[137] Salar Sajedi et al. "A novel non-linear recursive filter design for extracting high rate pulse features in nuclear medicine imaging and spectroscopy". In: *Medical Engineering & Physics* 35.6 (2013), pp. 754–764. ISSN: 1350-4533. DOI: `10.1016/j.medengphy.2012.08.003`.

[138] E. Gatti, A. Geraci, and G. Ripamonti. "Automatic synthesis of optimum filters with arbitrary constraints and noises: a new method". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 381.1 (1996), pp. 117–127. DOI: `10.1016/0168-9002(96)00653-5`.

[139] E. Gatti, A. Geraci, and G. Ripamonti. "Optimum filters from experimentally measured noise in high resolution nuclear spectroscopy". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 417.1 (1998), pp. 131–136. DOI: `10.1016/S0168-9002(98)00612-3`.

[140] S. Riboldi et al. "A new method for LMS synthesis of optimum finite impulse response (FIR) filters with arbitrary time and frequency constraints and noises".

In: *2002 IEEE Nuclear Science Symposium Conference Record.* Vol. 1. 2002, 198–202 vol.1. DOI: `10.1109/NSSMIC.2002.1239298`.

[141] E. Gatti et al. "Digital Penalized LMS method for filter synthesis with arbitrary constraints and noise". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 523.1 (2004), pp. 167–185. DOI: `10.1016/j.nima.2003.12.032`.

[142] S. Riboldi et al. "Optimum synthesis of FIR filters with arbitrary time and frequency constraints for energy and time estimations in case of pulse-correlated noise". In: *2007 IEEE Nuclear Science Symposium Conference Record.* Vol. 1. 2007. DOI: `10.1109/NSSMIC.2007.4436375`.

[143] A. Abba et al. "Experimental implementation of LMS synthesis of optimum FIR filters with arbitrary time and frequency constraints and noises". In: *2011 IEEE Nuclear Science Symposium Conference Record.* 2011, pp. 862–865. DOI: `10.1109/NSSMIC.2011.6154556`.

[144] Yuyan Huang, Hui Gong, and Jianmin Li. "Trapezoidal shaping algorithm based on digital penalized LMS method". In: *2016 IEEE Nuclear Science Symposium, Medical Imaging Conference and Room-Temperature Semiconductor Detector Workshop (NSS/MIC/RTSD).* 2016, pp. 1–3. DOI: `10.1109/NSSMIC.2016.8069678`.

[145] Florian Rettenmeier and Linus Maurer. "Design of optimum filters for signal processing with silicon drift detectors". In: *X-Ray Spectrometry* 50.6 (Feb. 2021), pp. 501–513. DOI: `10.1002/xrs.3227`. URL: `https://doi.org/10.1002/xrs.3227`.

[146] Kasun S. Mannatunga et al. "Design for Portability of Reconfigurable Virtual Instrumentation". In: *2019 X Southern Conference on Programmable Logic (SPL).* 2019, pp. 45–52. DOI: `10.1109/SPL.2019.8714446`.

[147]   E. Gatti and P. Rehak. "The concept of a solid-state drift chamber". In: *DPF Workshop on Collider Detectors: Present Capabilities and Future Possibilities*. 1983.

[148]   Emilio Gatti and Pavel Rehak. "Semiconductor drift chamber - An application of a novel charge transport scheme". In: *Nuclear Instruments and Methods in Physics Research* 225.3 (1984), pp. 608–614. ISSN: 0167-5087. DOI: `10.1016/0167-5087(84)90113-3`.

[149]   M. Sampietro, A. Geraci, and P. Fazzi A. & Lechner. "Advanced experimental application of a digital signal processor in high resolution x-ray spectroscopy". In: *Review Of Scientific Instruments* 66 (1995), pp. 5381–5382. DOI: `10.1063/1.1146056`.

[150]   A Cerdeira Estrada et al. "Readout electronics in large detector matrix for soft X-ray in medical applications". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 409.1 (1998), pp. 497–500. ISSN: 0168-9002. DOI: `10.1016/S0168-9002(97)01302-8`.

[151]   C. Cottini et al. "Minimum noise pre-amplifier for fast ionization chambers". In: *Il Nuovo Cimento* 3.2 (Feb. 1956), pp. 473–483. DOI: `10.1007/bf02745432`. URL: `https://doi.org/10.1007/bf02745432`.

[152]   Giuseppe Bertuccio and Stefano Caccia. "Progress in ultra-low-noise ASICs for radiation detectors". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 579.1 (2007). Proceedings of the 11th Symposium on Radiation Measurements and Applications, pp. 243–246. DOI: `10.1016/j.nima.2007.04.042`.

[153] Giuseppe Bertuccio et al. "A CMOS Charge Sensitive Amplifier with sub-electron equivalent noise charge". In: *IEEE Symposium on Nuclear Science (NSS/MIC)*. Nov. 2014, pp. 1–3. DOI: 10.1109/NSSMIC.2014.7431123.

[154] Pavel Rehak et al. "ASIC for SDD-based X-ray spectrometers". In: *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*. 2009, pp. 1088–1095. DOI: 10.1109/NSSMIC.2009.5402415.

[155] Alexandre Rachevski et al. "X-ray spectroscopic performance of a matrix of silicon drift diodes". In: *Nuclear Inst. and Methods in Physics Research, A* 718 (2013), pp. 353–355. DOI: 10.1016/j.nima.2012.10.017.

[156] G. Bertuccio et al. "A Silicon Drift Detector – CMOS Front-End System for High Resolution X-Ray Spectroscopy up to Room Temperature". In: *Journal of Instrumentation* 10 (2015). DOI: 10.1088/1748-0221/10/01/P01002.

[157] Martin Berger, Qiao Yang, and Andreas Maier. "X-ray Imaging". In: *Medical Imaging Systems: An Introductory Guide*. Cham: Springer International Publishing, 2018, pp. 119–145. DOI: 10.1007/978-3-319-96520-8_7.

[158] Shinya Yamada et al. "Poisson vs. Gaussian statistics for sparse X-ray data: Application to the soft X-ray spectrometer". In: *Publications of the Astronomical Society of Japan* 71.4 (July 2019). DOI: 10.1093/pasj/psz053.

[159] Glenn F Knoll. *Radiation detection and measurement*. John Wiley & Sons, 2010. Chap. 17.

[160] Wojciech Walewski, Patryk Nowak Vel Nowakowski, and Dariusz Makowski. "Giga-sample Pulse Acquisition and Digital Processing for Photomultiplier Detectors". In: *Journal of Fusion Energy* 41 (June 2022). DOI: 10.1007/s10894-022-00320-0.

[161] W. Skulski and M. Momayezi. "Particle identification in CsI(Tl) using digital pulse shape analysis". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 458.3 (2001), pp. 759–771. ISSN: 0168-9002. DOI: `10.1016/S0168-9002(00)00938-4`.

[162] Andres Cicuttin et al. "A programmable System-on-Chip based digital pulse processing for high resolution X-ray spectroscopy". In: *2016 International Conference on Advances in Electrical, Electronic and Systems Engineering (ICAEES)*. 2016, pp. 520–525. DOI: `10.1109/ICAEES.2016.7888100`.

[163] V. Radeka. "Trapezoidal Filtering of Signals from Large Germanium Detectors at High Rates". In: *IEEE Transactions on Nuclear Science* 19.1 (1972), pp. 412–428. DOI: `10.1109/TNS.1972.4326542`.

[164] C. de Cesare et al. "An FPGA-based algorithm to correct the instability of high-resolution and high-flux X-ray spectroscopic imaging detectors". In: *Journal of Instrumentation* 13.08 (Aug. 2018), P08022–P08022. DOI: `10.1088/1748-0221/13/08/p08022`. URL: `https://doi.org/10.1088/1748-0221/13/08/p08022`.

[165] K.P. Burnham and D.R. Anderson. "Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach". In: Springer New York, 2007. ISBN: 9780387224565. URL: `https://books.google.ca/books?id=IWUKBwAAQBAJ`.

[166] G. Hölzer et al. *X-Ray Transition Energies Database -NIST Standard Reference Database*. URL: `https://physics.nist.gov/cgi-bin/XrayTrans/` (visited on 2022).

[167] S O Flyckt and Carole Marmonier. *Photomultiplier tubes: principles and applications; 2nd ed.* Brive: Photonis, 2002. URL: https://cds.cern.ch/record/712713.

[168] R. Storn and K. Price. "Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces". In: *Journal of Global Optimization* 11.4 (1997), pp. 341–359. DOI: 10.1023/A:1008202821328.

[169] Jack E. Kemmerly William Hayt and Steven M. Durbin. *Engineering Circuit Analysis.* McGraw-Hill, 2002.

[170] Xilinx. *FIR Compiler v7.2.* Product Guide (PG149). Xilinx, Inc. 2016. URL: https://www.xilinx.com/support/documentation/ip_documentation/fir_compiler/v7_2/pg149-fir-compiler.pdf.

[171] P.W. Wong. "Quantization and roundoff noises in fixed-point FIR digital filters". In: *IEEE Transactions on Signal Processing* 39.7 (1991), pp. 1552–1563. DOI: 10.1109/78.134394.

[172] Andres Cicuttin et al. "A Simplified Correlation Index for Fast Real-Time Pulse Shape Recognition". In: *Sensors* 22.20 (2022). ISSN: 1424-8220. DOI: 10.3390/s22207697. URL: https://www.mdpi.com/1424-8220/22/20/7697.

[173] S. Carrato et al. "A scalable High Voltage Power Supply System with system on chip control for Micro Pattern Gaseous Detectors". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 963 (2020), p. 163763. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2020.163763. URL: https://www.sciencedirect.com/science/article/pii/S0168900220303016.

[174] *UDMA.* 2022. URL: https://gitlab.com/ictp-mlab/udma. accessed: 25.07.2022.

[175] K. Chen et al. "Evaluation of commercial ADC radiation tolerance for accelerator experiments". In: *Journal of Instrumentation* 10.08 (Aug. 2015), P08009. DOI: 10.1088/1748-0221/10/08/P08009. URL: https://dx.doi.org/10.1088/1748-0221/10/08/P08009.

[BSD-3] *BSD 3-Clause "New" or "Revised" License.* Version 3. Berkeley Software Distrubution. URL: https://spdx.org/licenses/BSD-3-Clause.html.

[GPL] *GNU General Public License.* Version 3. Free Software Foundation, June 29, 2007. URL: http://www.gnu.org/licenses/gpl.html.

[176] *quasar OPC UA servers.* 2023. URL: https://quasar.docs.cern.ch/. accessed: 15.05.2023.