APTM and GRID signal processing and Errant beam detection at ESS

Agnimesh Ghosh
ag6604gh-s@student.lu.se

Department of Electrical and Information Technology
Lund University

Supervisor: Anders J Johansson, Cyrille Thomas, Kaj Rosengren

Examiner: Fredrik Rusek

September 28, 2020

Printed in Sweden
Tryckeriet i E-huset, Lund

Abstract

The **European Spallation Source (ESS)** will be the world's most powerful next-generation neutron source. It will deliver a high-energy proton beam to the target with a high degree of control. With many components operating at their engineering limits, instrumentation must detect errant beam conditions and suppress proton beam production. One of the beam diagnostic instruments installed will be Aperture Monitor (APTM). This system consists of thermocouples and metallic blades fitted along the edges along a rectangular footprint. The Aperture Monitor determines the total beam current outside of the desired footprint. Apart from the Aperture Monitor, another beam diagnostic instrument is a multi-wire profiling monitor called Grid. This system consists of horizontal and vertical wires, and it will measure the amount of current generated by the bombardment of high energy protons. The high energy proton, while interacting with the silicon carbide wires in the profiling monitor, generates electrical current. This system measures the current value at the rate of 1 mega sample/second. The goal of this thesis aims to design a signal processing system with the aid of FPGA that will characterize the beam properties and detect any beam condition outside its defined operation range by analyzing the current signal in the shortest time possible.

Acknowledgement

I want to dedicate this section to the people who have helped me by supporting throughout the entire Master's Thesis project. I would like to express my appreciation to my supervisors Cyrille Thomas and Kaj Rosengren at ESS, who has helped me to understand and develop my thesis project. I would also like to thank my supervisor from the LTH, Docent Anders J Johansson, who has given me helpful feedback during this work. A special thanks to the rest of the Beam Diagnostic group, who have been extremely helpful during these challenging times. Thanks to Thomas Shea, who has been keeping eye on me during the project, making sure I had everything to continue my thesis. I would also like to thank Henrique Silva and Clement Derrez who have helped me with the hardware setup for the project in the Beam Instrumentation lab. I would also like to thank Mehdi Mohammednezhad, who helped me clear my doubts regarding the project whenever I needed it. I would also like to thank the entire ESS team, who managed to continue the work during this COVID-19 pandemic while maintaining WHO guidelines at workplaces in ESS campus. I appreciate all of us, continuing their work both during "normal" times and during COVID-19!

Last but not the least, I would like to thank my family and friends for never-ending love and support throughout my education and my father, because of whom I am here today, may his soul rest in peace.

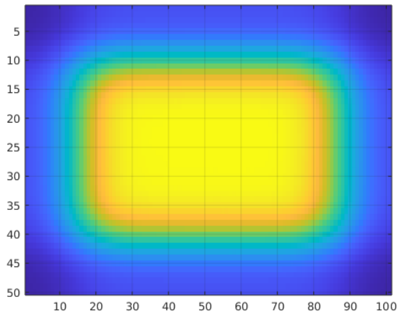
Popular Science Summary

ESS is a high power neutron source, which uses a 5 MW proton beam to generate neutrons. These highly energetic protons are sent to collide with the target. The interaction between the protons and the target's atoms (tungsten) produces energetic neutrons. This process is called Spallation. These neutrons are then directed to different experimental stations, where research will be done on different materials.

The ESS LINear ACcelerator (LINAC) uses a magnet system that sweeps these high energy protons transversely across the defined footprint of target. To protect the machine from an errant beam, a suite of interceptive beam diagnostic instruments will be used at different places along the LINAC. Among these beam diagnostic instruments, Aperture Monitor (APTM) and Multi-Wire Harp Profiling Monitor (Grid) will be used to monitor the beam properties to ensure the nominal operation of the LINAC.

The APTM employs nickel (Ni) plates at its perimeter and is designed to measure the fraction of the beam which goes through a defined aperture, whereas the Grid is fitted with silicon-carbide (SiC) wires along its vertical and horizontal plane and it measures the projected horizontal and vertical profiles of the beam.

When a beam enters or exits a material, secondary electrons are generated as some of the beam's kinetic energy is used to excite the valence shell electrons of the material. This process is called Secondary Electron Emission and the number of the secondary electrons excited per incident particle is called Secondary Emission Yield. These excited electrons then collectively generate current in the material. The Secondary Emission Yield is dependent upon the material upon which the incident particles are hit.
Both of the beam diagnostic instruments (APTM and Grid) discussed in this thesis, use this principle to generate current from the high energy proton beam. At ESS, the Beam Delivery System incorporates a raster



Grid Beam Profile

system that sweeps the proton pulse beamlet in a transverse pattern across the target surface. This system lowers the time-averaged beam intensity while containing the beam in a defined footprint. The figure shown above describes the Beam intensity over a beam pulse length (2.8 mS), generated by the 101 vertically aligned and 50 horizontally aligned wires.

In this master thesis, I have extracted various beam properties from the current signal generated due to the interaction of the APTM blades and the Grid wires with the high energy proton beam by means of signal processing. These beam properties are then characterized to detect the errant beam conditions. To detect an errant beam, different signal processing algorithms have been explored on a FPGA (Field-programmable gate array) based board in order to achieve minimum computation latency. To protect the target from an errant beam production, two algorithms (one for APTM and one for Grid) have been proposed to detect errant beam with minimum latency.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

ESS, the European Spallation Source, will be a world-class user facility, where researchers from academia as well as from the industry will investigate their scientific queries using neutron beams [1]. Neutron methods provide insights about the molecular building blocks of matter not available by other means. They are used for both basic and applied research. Neutron scattering, as a tool for the investigation of materials at a sub-atomic level, was pioneered in the North American sub-continent in the 1950s [1]. The neutron itself had been discovered in 1932 in Cambridge by James Chadwick. The first moderately intense sources of neutron beams were extracted from the early research reactors, constructed in several national laboratories in the USA as well as in Canada. It was on these research installations that the early instrumental techniques using neutrons were developed in order to begin to unravel the atomic structures of relatively simple materials and, uniquely, the atomic dynamics of these same materials. For this work, Cliff Shull and Bert Brockhouse were awarded the Nobel prize in physics in 1994 [1].

As the new technology emerged over the following decades after that, there was a huge increase of interest in neutron production due to its high usability in the nuclear power industry. The early sources of these accelerator-driven neutron sources were based on electron linear accelerators, which had a significant background problem caused by the generated intense gamma radiation bursts. To overcome this issue that affects electron sources, proton-driven neutron sources were introduced since the spallation reaction employed in proton machines generates significantly less heat per useful neutron [1].

Europe furnished with its rich network of neutron sources, and innovative instrumentation solutions resulted in an international project comprising of 20 European countries designing a neutron source capable of delivering its scientific goals. The need for such a facility resulted in ESS. The ESS design compromised a 5 MW proton linear accelerator delivering a 2 to 3-millisecond long pulse to a single target station surrounded by a suite of 20 to 25 neutron instruments [1]. As being an accelerator device, the accelerator at ESS is thus a critical component of the ESS facility. The accelerator creates protons at the ion source, accelerates them to an appropriate energy, and steers them onto the target to create neutrons via the spallation process.

Beam Instrumentation will provide detailed information on the condition and location of the proton beam. ESS will have a dense array of standard beam instrumentation to acquire such information for safe and reliable operation of the accelerator. To achieve a neutron production run of several months between major maintenance periods, the beam delivery and target systems must function reliably within the nominal conditions. With many components operating near their engineering limits, instrumentation must detect errant beam conditions before damage occurs, and the Machine Protection System (MPS) must suppress proton beam production.

In order to achieve this, a suite of beam on target diagnostic instruments have been implemented. Among these suites of instruments, which will provide measurement of beam characteristics in a drift space a few meters from the target, The Aperture Monitor (APTM) and the Multi-Wire Harp Profiling Monitor (Grid) are used to extract such information. The APTM is designed to measure the fraction of beam that goes through the defined aperture and covers the range of time from intra-pulse at micro-second sampling time to many pulses over seconds. Whereas the Grid measures the projected horizontal and vertical profiles of the beam.

1.1 Background and Motivation

The ESS will be delivering a high-power beam to the spallation target, where the nominal beam is defined, and departure from these nominal values causes an errant beam condition. Therefore, a beam in errant condition may bring damaging risk to the target. Preventing any risk condition to occur during the phases of tuning and neutron production is critical and is a mandatory task to be completed before the commissioning of the accelerator. The beam diagnostic instruments provide a diverse and somewhat redundant measurement of some of the beam parameters and ensure the support for beam tuning and beam operation together with machine protection. Essentially, these systems are part of the diagnostics suite that will be used to protect the machine from errant beam condition.

This master thesis discusses two of these aforementioned beam diagnostic instruments called APTM and Grid. The APTM system consists of metallic blades and thermocouples and is designed to measure the beam current through a defined aperture. This system will be deployed at four sites along the accelerator region. Whereas the Grid system will consist of one multi-wire grid assembly located in the target monolith. This assembly system will measure the horizontal and vertical projections of the beam current density, such that a change of 20% with respect to nominal peak density can be accurately determined. These diagnostic systems are interfaced to the beam interlock system so that beam can be interrupted within the pulse if current density exceeds a programmable threshold.

1.2 Purpose

The goal of this master thesis is to develop the signal processing for the beam on the target instruments called APTM and Grid. The signal processing aims at characterizing the beam properties and detect any beam condition outside its defined operation range.

A model of this APTM and Grid module had been proposed by Cyrille Thomas, prior to this master thesis. The design virtually emulates the current generation due to the interaction between the wires and the neutron beamlets. A brief study of the beam dynamics and beam on the target raster system has been done at first for the characterization of the signal acquired by the APTM and the Grid. The beam parameters are extracted by various signal processing means, and the conditions for the errant beam conditions are found. Finally, the deterministic conditions for the errant beam are optimized to be implemented on a FPGA based acquisition card with eight inputs. These acquisition cards are called Advanced Mezzanine Card (AMC), which will be mounted in a MicroTCA (MTCA) system. Since there are multiple AMCs used in this project, A proof of concept for the communication protocol between these AMCs have been discussed in this thesis.

This master's thesis is divided into six chapters. This introduction section is the first chapter of the thesis. Whereas in Chapter 2, the beam physics and some relevant technical design of ESS, which will provide a safe, reliable, high-power proton beam to the neutron-generating spallation target have been discussed. The third chapter describes the various beam parameters and how to extract them from the measured current signal. In the fourth chapter, various algorithms that can be applicable in a FPGA are explored as well as the data transfer protocol between AMC boards. The fifth chapter discusses the results from the various applied algorithms and their applicability in real-time systems at ESS. The last chapter summarizes the overall thesis and discusses its future scope.

Chapter 2

Background Theory

The ESS accelerator will facilitate a high power 5 MW linear beam to the spallation target, producing 2.86 ms long pulses of low-emittance 2 GeV proton beam at a rate of 14 Hz. At full power, it will be the most powerful linear proton accelerator ever built. Figure 2.1 describes the deliverance of beam from accelerator to target region.

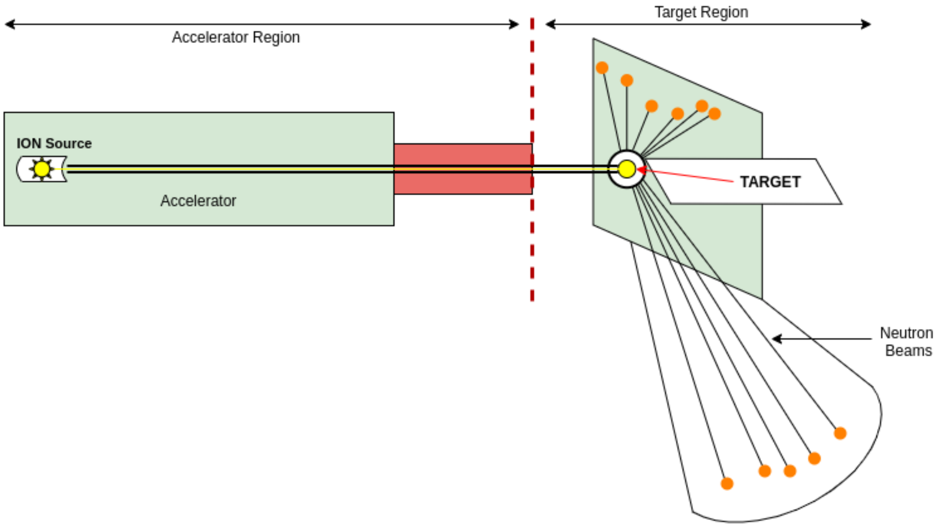


Figure 2.1: Illustration of Beam Transport

In the accelerator region as shown in figure 2.3, the beam follows a sequence of sections viz. the ion source, the normal conducting linac , the super conducting linac and the beam transport sections .

The Target region consists of a rotating wheel (as shown in Figure 2.2) of 2.6 meter in diameter. The wheel consists of hundreds of heavy metal tungsten bricks encased in a disk of stainless steel shielding and rotates at $23\frac{1}{3}$ RPMs.

It also incorporates a powerful helium-based cooling system able to dissipate the heat generated by the world's most energetic proton beam hitting the target.

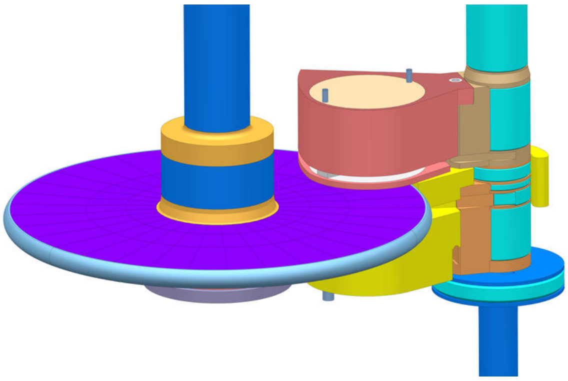


Figure 2.2: Model of ESS Target Wheel [1]

The Accelerator region is separated from the Target region by a 2 m thick neutron shield wall (NSW), and both of them have different requirements. In the accelerator region, the design is driven by the hands-on maintenance requirement. This defines a region with not so activated materials, which can be possible to do a hands-on maintenance. This is achieved with a very low beam loss in the accelerator of less than 1 W/m. Whereas downstream the NSW, the high-speed protons kick out the neutrons in a process known as spallation, producing a high radiation level and material activation. The target station requires to absorb all particles generated during the interaction of the proton beam with the target and to guide the moderated neutrons down to the neutron beamlines towards the neutron instruments.

2.1 How It Works

ESS is an accelerator driven neutron source. The accelerator is thus a critical component of the ESS facility and role of it is straightforward. The accelerator creates protons, accelerates them to an appropriate energy and steers them onto target to create neutrons via the spallation process. A proton of kinetic energy of 1-3 GeV is optimal for practical target, keeping the radiation shielding at reasonable limits. At ESS, the proton energy of 2.0 GeV requires an average macro-pulse current of 62.5 mA with a 4% duty cycle to reach a beam power of 5 MW. This current is consistent with the need of high reliability.

The Figure 2.3 describes the main components of the LINAC beam transportation model at ESS. The model is divided into different sub-parts for understanding.

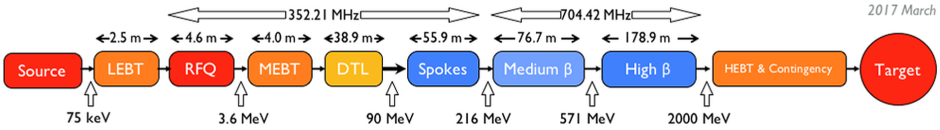


Figure 2.3: Block diagram of the ESS linear accelerator (LINAC) [1]

2.1.1 Ion Source (IS)

The proton beam is produced by heating hydrogen (H_2) gas with the help of a varying electromagnetic fields at high frequency. The heating method dissociates H_2 molecules, producing a plasma of protons and electrons. The protons are extracted by means of a 75 keV electric field, which produces the first beam introduced in the first transport section, the LEBT. The source is capable of delivering proton current of at least 80 mA.

2.1.2 Low Energy Beam Transport (LEBT)

After being emitted from the Ion source, the proton beam is then transported through the Low energy beam transport (LEBT) section under vacuum. The LEBT has dual purpose of bringing the beam from the source to the sharp focus required at the RFQ entrance and providing a mechanism for chopping the beam. The source produces a proton pulse of 6 ms, and the chopper reduces the pulse duration further to roughly match the required 2.86 ms pulse.
The focusing in LEBT is done by two solenoid magnets, whereas the chopper consists of a pair of electrical deflection plates between two solenoids. The chopper will be abort the beam production in approximately 100 nano-seconds.

2.1.3 Radio Frequency Quadrupole (RFQ)

The radio frequency quadrupole (RFQ) is a part of the low energy normal conducting linac. The beam from the LEBT is tansported to the RFQ section for bunching and acceleration. The energy that accelerates the proton beam is provided by the radio frequency (RF) system that converts the AC power from the electrical Grid to the appropriate frequencies. In RFQ, the proton beams are bunched and accelerated up to 3.6 MeV.

2.1.4 Medium Energy Beam Transport (MEBT)

The medium energy beam transport (MEBT) has a comprehensive set of beam instrumentation devices to match and steer the beam from RFQ into the drift tube linac (DTL). The design choices of the MEBT results in further acceleration of the beam. The beam is matched to the normal conducting drift tube linac (DTL) through this section.

2.1.5 Drift Tube Linac (DTL)

The drift tube linac (DTL) accelerates the protons from 3.6 MeV to 90 MeV in separate stages. The beam is equi-partitioned and steered with magnets in alternate stages. Leaving the DTL, the beam enters the superconducting portion of the linac.

2.1.6 Superconducting linac

The acceleration of the protons from 90 MeV to the final energy of 2 GeV is achieved with 3 stages of superconducting cavities. The acceleration is accomplished via superconducting radio frequency cavities immersed in liquid helium.

2.1.7 High Energy Beam Transport (HEBT)

From the superconducting portion of the linac, the beam is transported through the high energy beam transport (HEBT) section to the target. This section contains around 100 meters for potential energy upgrades for the accelerator. After the upgrade space, the beam is brought achromatically to the surface and the level of the target by two vertical bends, each composed of two dipole magnets in series.
At target level, the raster magnet system expands the beam to the desired dimensions on the beam entrance window (BEW) of the target wheel. The protons reach 96% of the speed of light before they hit the rotating target wheel.

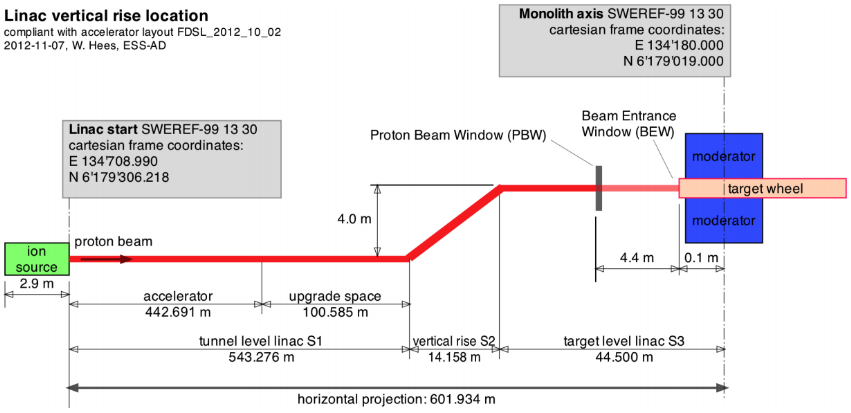


Figure 2.4: End Co-ordinates and overall dimensions of the accelerator [1]

2.2 Beam Instrumentation

The beam instrumentation is a part of beam diagnostics section and an essential part of the accelerator for safe and reliable operation. It will provide detailed critical information on the condition and location of the proton beam. At ESS, a dense array of standard beam instrumentation will be used at beam transport areas. Most of the instrumentation have been developed based on experience at similar facilities.

The first instrumentation in the beam transportation system occurs at the Ion Source (IS) and the Low Energy Beam Transport (LEBT) line (shown in figure 2.5). Various beam instruments will be deployed at the later linac stages to monitor beam properties. To ensure the interfacing between the linac and the target system, beam on target instruments will also be deployed at the target monolith (as shown in figure 2.6). These suites of instruments also should help to abort the pulse within 10 micro-second after any kind of errant beam is detected.

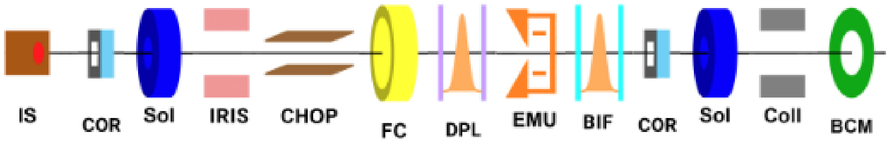


Figure 2.5: Schematic of Beam Instrumentation in LEBT [2].
IS: Ion Source, COR: Corrector Magnet, Sol: Solenoid, IRIS: multiblade Iris , CHOP: Beam chopper, FC: Faraday Cup, DPL: Doppler, EMU: Emittance Measurement Unit, BIF: Beam Induced Florescence monitor, Coll: Collimator, BCM: Beam Current Monitor

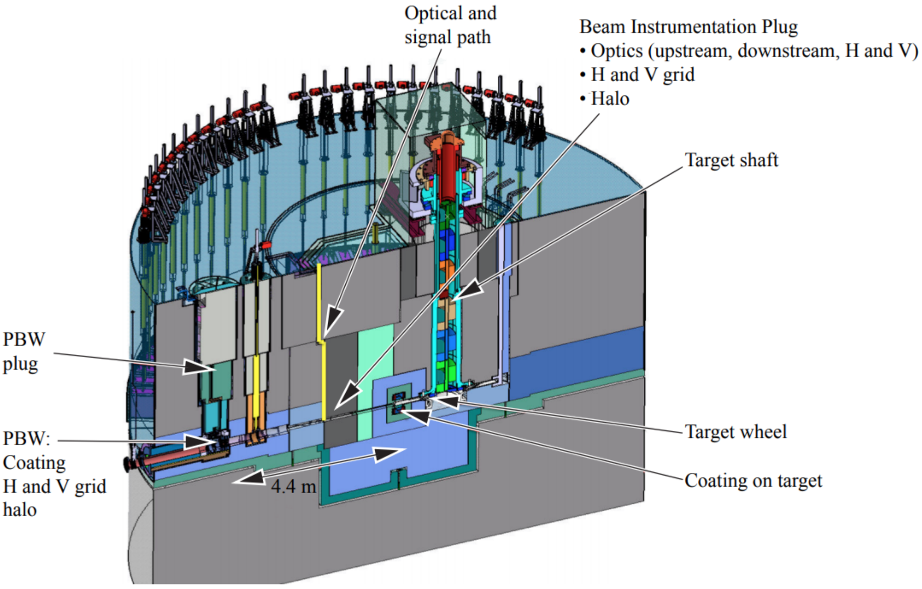


Figure 2.6: Location of Beam on Target instrumentation electronics [1]

The focus of this master thesis is on part of the beam on target instrumentation system. This system will be deployed to measure the beam properties at the Proton Beam Instrumentation Plug. The aim of this thesis demonstrate a robust way to detect errant beam conditions, using the suite of instruments designed to that aim. In the instrumentation plug, a multi-wire harp profiling monitor (Grid) will measure the horizontal and vertical beam profile. In the Proton Beam Window, upstream to the instrumentation plug, an instrument protecting the beamline element, called APerTure Monitor (APTM) is installed. Another APTM is also installed in the proton beam instrumentation plug. In addition, two imaging system perform an image of the proton beam traversing the proton beam window and the Beam Entrance Window (BEW) before the target (figure 2.7) .

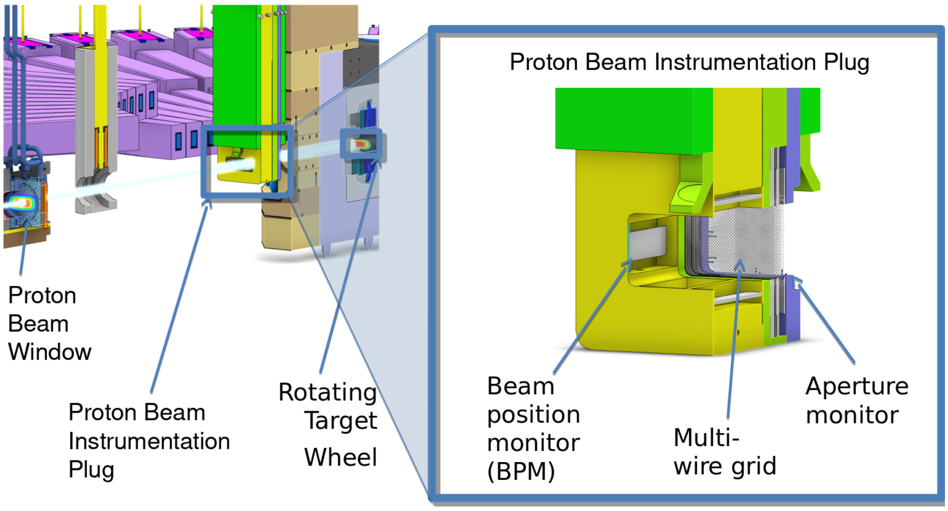

Proton Beam Instrumentation Plug
Proton Beam Window
Proton Beam Instrumentation Plug
Rotating Target Wheel
Beam position monitor (BPM)
Multi-wire grid
Aperture monitor

Figure 2.7: Beam on Target instrumentation electronics at PBIP [3]

The beam profile measurements also rely on the Beam Current Monitor (BCM) for normalisation. Synchronisation of the instruments is critical: it permits to trigger the instruments, and to define a unique time for all events.

2.2.1 Beam Current Monitor

The Beam Current Monitor system (BCM) measures the beam current versus time at various locations along the LINAC. With a precision of ±1%, the BCMs will provide the beam current waveform over a beam pulse, the charge per pulse. Real-time data of the average beam current and the cumulative charge is also sent to the beam on target instrumentation systems and to the neutron instruments. This data will be used to normalize the beam- on-target current density measurement, and will be included in the neutron instruments' event stream. The BCM output is sampled at a rate of 88 MHz by a fast ADC and stored in local memory, so that the data can be retrieved later and examined upon a user request or machine failure.

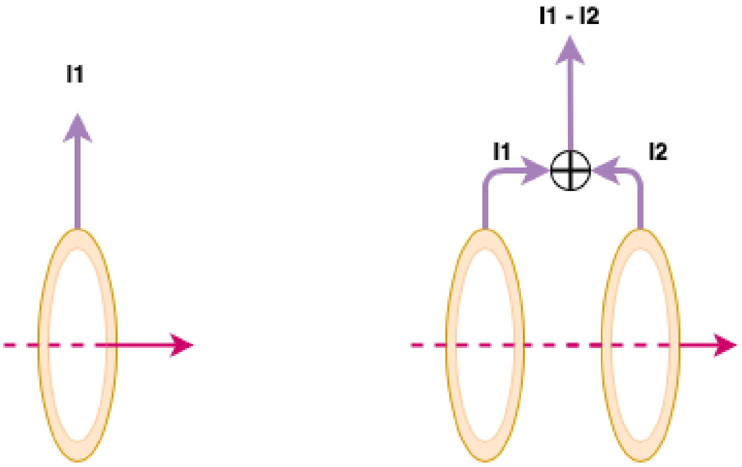


Figure 2.8: Beam current measurements due to the beam passing through the coils Left: Absolute Right: Differential [1]

Fast active current transformers are used as BCM to measure the beam current up to a few gigahertz bandwidth. In addition to the absolute current measurement, some of the BCM will be used for differential current measurement. These BCMs will also be connected to the MPS so that the beam can be aborted if either of the absolute or differential current value exceeds user-defined thresholds.

2.2.2 Beam Position Monitor

The Beam position monitor (BPM) system is used to measure the beam position, the beam phase, and an approximate beam current. This system provides the horizontal and the vertical beam position and beam phase waveform throughout the beam pulse. In the accelerator-to-target (A2T) line, the beam should be centered in the magnets, so BPMs will be located on or near these magnets.

The BPM electronics consist of a fast analog front-end. The signals are acquired by the sensitive analog electronics and a digital section where these signals are digitized and processed using an FPGA. The BPM electronics have a large dynamic range and a good signal-to-noise ratio since it'll be used from low power diagnostic mode of the BPMs to fully operational linac.

2.2.3 Grid, The multi-wire harp profiling monitor

The Grid is designed to measure the one dimensional projected profile of the beam on the orthogonal axis. The Grid will employ thin Silicon-Carbide (SiC) wires of high heat resistivity. The design assembly (as shown in Figure 2.9) is based on a harp of wires and will measure the beam induced current.

The planes of wires (horizontal and vertical) are separated by additional planes (as shown in figure 2.12) to provide electric fields. The generated electric fields are designed to prevent electrons present in the vicinity of the wires to be absorbed. The Grid will be installed in a replaceable part in the PBIP assembly, which is already equipped with cooling systems. The wire material is chosen after considering the radiation tolerance and the material properties, and the wires are spaced consistently. The Grid will be operational at vacuum.

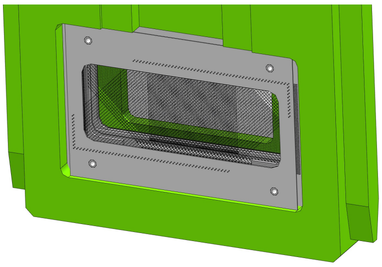


Figure 2.9: CAD image of ESS Multi-Wire Harp Profiling Monitor [4]

The evaluation of the current is calculated by the Sternglass expression [10] (eq.2.1). The Sternglass theory describes the secondary emission yield when a beam enters or exits a material. The number of secondary electrons created is dependent upon the condition of the surface of the material.

$$Y = \frac{P\,d_s}{E^*}\frac{dE}{dx} \tag{2.1}$$

where Y is the secondary electron emission yield, P is the probability of electron escaping, $d_s \approx 1\text{nm}$ is the average depth from which secondaries arise, E^* is the average kinetic energy of the secondary electrons, $E^* \approx 25\text{eV}$, and dE/dx is the stopping power of the target.

The current in the wires is defined as the rate of flow of these secondary electron with respect to time and is calculated by [5],

$$I_{wire} = Y \times \delta I_b \tag{2.2}$$

where δI_b is the proton beam current interacting with the wire. The current in the wires is measured by means of a digitizer, which will be discussed in later part of this chapter.

2.2.4 APTM

The APTM is an instrument designed to ensure that 99.9% of the proton beam current is going through a defined aperture. The APTM is then used to protect the part of the machine that should never interact with the proton beam, but may survive a small dose. They are based on the detection of current intercepting metallic plates (the dark grey area, as shown in Figure 2.10) and thermocouples surrounding the aperture. The protons intercepted by the metallic plates produces current that is read with a fast electronics system as the requirement is to detect errant beam within a pulse. The K-type thermocouples are to read the temperature increase due to the beam energy deposition in the thermocouple.

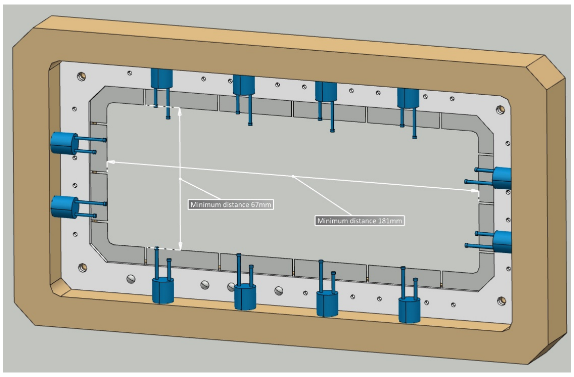


Figure 2.10: APTM [4]

There are four APTMs to be installed for the accelerator at different locations viz. at the PBIP, the PBW, one in the neutron shield wall and in the tuning dump line. The current in the blades are calculated by [5],

$$I_{blades} = Y \times N \times \frac{q_e}{dt} \tag{2.3}$$

where Y is the secondary electron yield as calculated with the Eq. 2.1 for the blade material, N is the number of protons interacting with the metallic blades and q_e/dt is the current produced due to a single electron. The mechanical design of the system takes care of the electrical insulation for signal acquisition from the blades as well as the heat load and cooling of the assembly.

2.3 Electronics Design

The electronics design and the hardware implementation (on FPGA) for the APTM and Grid instrument used in this thesis is a prototype version of the final implementation. The design will be verified at Japan Proton Accelerator Research Complex (J-PARC) which is a high intensity proton accelerator facility, before deploying in larger scale at ESS.

2.3.1 Signal Acquisition

The current signals measured from the wires are in the nano to the micro-ampere range. At ESS 16/20- bit floating ammeter card is selected to use for reading out the current signals. The current values are sampled at 1 Msamples/s by FPGA Mezzanine Card (CAEN-ELS: FMC-Pico-1M4 [11]). In this thesis, these FMC-modules will be placed on Struck SIS8160 [12] boards, which in addition have an FPGA for data filtering and analysis.
For the APTM, it has eight channels of current measurement, which is handled by two FMCs. In addition, we need a timing receiver for synchronization and a CPU board for EPICS [13] integration. As an ESS standardized component, the temperature sensors will be handled by EtherCat Beckhoff devices [14].

For the Grid, a higher number of signals are required to handle the approximately 160 signals (80 horizontal and 80 vertical signals), which will be managed by additional MTCA slots. Each AMC card gives 8 channels, so with a design following from the aperture monitors, two 12-slot chassis will support this configuration. For the detection of errant beam conditions on either APTM or Grid require a high-speed, low latency communication protocol. The architecture for the errant beam detection and for the triggering, the BIS will be better discussed in later chapters in this thesis.

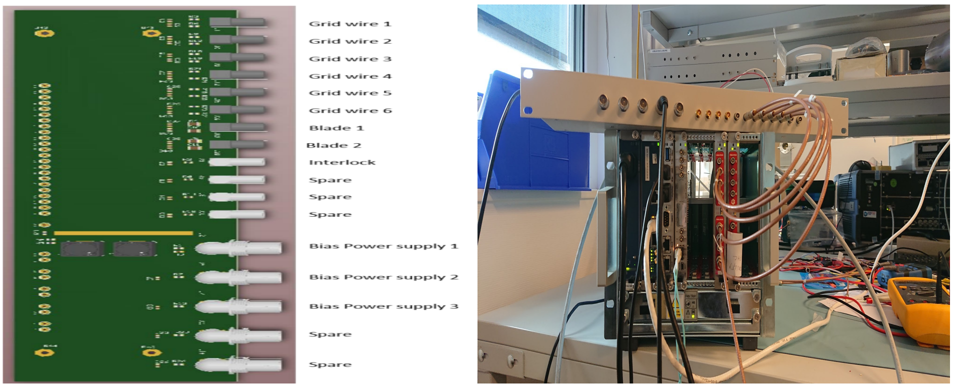
2.3.2 Calibration

The current signal from the FMC-Pico modules is 20-bit fixed-point values since the use of floating-point in hardware increases the complexity of the design, which in turn incurs undesired latency. These values use Fixed-point data Qm, n format [15] for internal precision. The precision can go from -524288 to 524288, depending upon the values of m and n.

However, the values are converted into nano to micro-ampere range by multiplying the 20-bit values with a coefficient which are unique for each channel and stored on an EEPROM on the FMC-Pico. These calibration coefficients are used for plotting/reading purposes.

2.3.3 Bias of APTM and Grid

In signal processing, there is always an advantage of a clean signal. In a real-time environment, a noisy signal is expected from the monitor blades and the Grid wires.



(a) Hardware connectors for the Bias

(b) Bias connected to the MTCA crate

Figure 2.11: APTM- Grid Bias Panel at ESS

When the beam interacts with the wires, there are some electrons emitted interacts with the other blades, producing increased current in the neighbouring blades in the APTM as well as there is cross talk between between the wire read outs. These factors leads to noise in the channel. To get rid of this noise, a voltage potential was set to either of the aperture monitor blades and the Grid itself or to a separate structure. The level and polarity of this bias potential is dependent on the magnitude of noise and is controlled and monitored over EPICS.

2.3.4 Processing And Analysis

A design of a prototype of the APTM and Grid has been done outside of this thesis and the prototype will be tested at the Japan proton accelerator research complex (J-PARC). The prototype design were to test various aspects of the conceptual design:

- Signal Intensity in Grid wires and APTM blades.
- Acquisition of the current signal and temperature in real environment.
- Effect of acquired signals due to cross-talk between APTM and Grid.
- APTM temperature increase under defined current signal and perturbation from the real and radioactive environment.

Figure 2.12 describes a detailed photo of the head of the assembly manufactured at ESS. The green coloured ceramic is made in U shape to let the beam to be intercepted from one side without intercepting any other materials than the Grid wires which are again connected to 2 external wires. Also there are two APTM blades are connected at the extremities of the ceramic. The Grid is upstream to the APTM, whereas both of the instruments are independent and can be exposed to the beam.

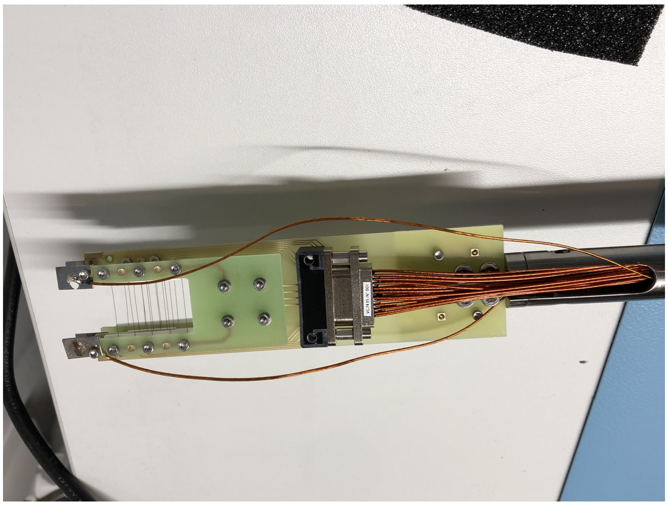


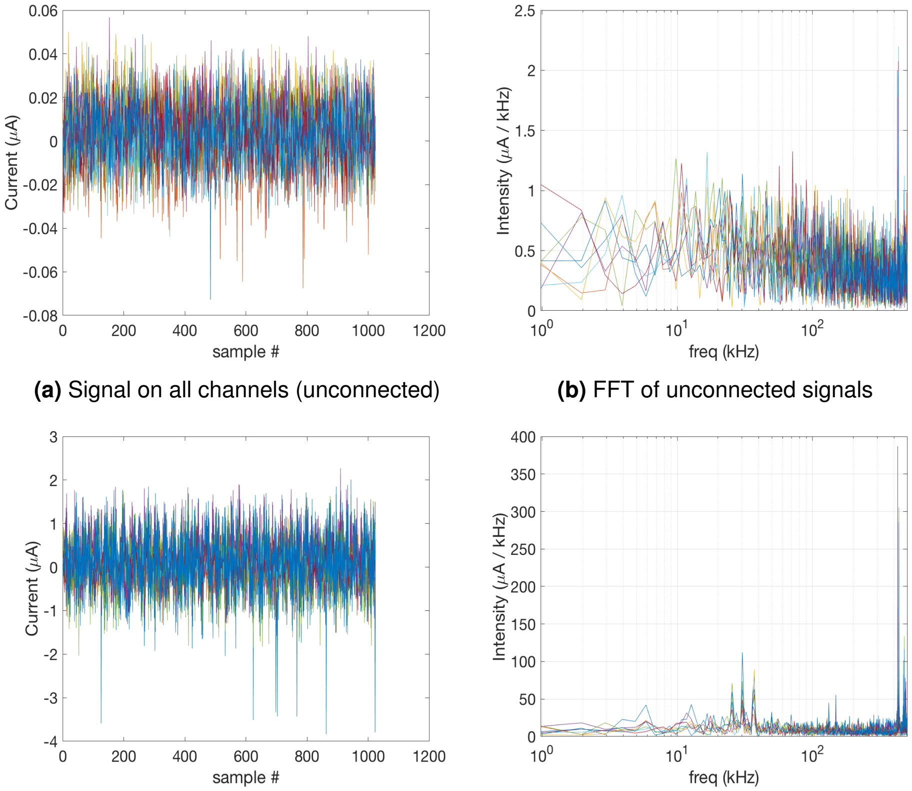
Figure 2.12: APTM-Grid Prototype

2.3.4.1 Scanning the APTM

The APTM is mounted on an actuator, that drives horizontally the prototype head to the beam region. The beam position can be steered too. Once the beam position is assured to be at the centre of the U-shaped instrument, the actuator can be scanned to map the temperature and current read in the thermocouple plates and Ni blades, as function of the distance to the beam (Figure 2.12).

2.3.4.2 Scanning the Grid wires

The results from the tests in J-PARC, Japan, will be acquired later on. In this thesis, a current signal generator (Keithley 6221 PV) is used to test signal acquisition. The current is read out by the pico digitizer. White noise is observable on all channels when it's not connected. However, some noise around 10 kHz frequency is noticed on the channels when it's connected via long haul cables (as shown in figure 2.13d). The noise is likely to be generated by a bad connection, either from the cables or from the connection board. It is under investigation. However, the noise level, on average less than 1μA, is acceptable for the experiment to be run at J-PARC, where the peak current is expected to be on the 100μA range.



(a) Signal on all channels (unconnected)

(b) FFT of unconnected signals

(c) Signal on all channels (connected)

(d) FFT of connected signals

Figure 2.13: Current signals acquired by the FMC

2.4 Requirements for APTM and Grid

In this thesis, the beam physics is explored (discussed in the next chapter) according to the requirements for the errant beam detection in Grid and APTM systems. The Grid and APTM systems provide a diverse and redundant measurement of some of the beam parameters that are measured by analysing the current values from the wires. These requirements [4] specify the nominal beam parameters to be delivered to the target in safe condition for any beam, and the extreme condition that leads to damage to the machine.

Parameter	**Nominal value**	**Threshold value**
Beam peak current density (mA/cm^2)	1.325	< 1.6
Beamlet size area $\sigma_x \times \sigma_y$ (mm^2)	53	> 32
Raster Horizontal and Vertical frequency (kHz)	39 and 29	> 10
Beam horizontal center position accuracy (mm)	±4	-
Beam vertical center position accuracy (mm)	±2	-
Beam top and bottom edge vertical displacement away from the nominal edge (mm)	-	< 5
Beam left and right edge horizontal displacement away from the nominal edge (mm)	-	< 13

Table 2.1: Requirements for the beam on target properties [9]

Chapter 3

Beam Diagnostics

The analysis of the beam profile from a high-power proton beam has been done previously in many places around the world. In general, to observe the characteristics of the high-intensity beam projected towards the spallation target, a reliable profile monitor system is necessary. Different kinds of measurements are done by a suite of instruments to monitor the beam's properties in similar research facilities. At Spallation Neutron Source (SNS, Unites States), the GRID is called a HARP and observes in the same way the integrated current in the wires to obtain the beam profile [16] and it is position many meters upstream to the target [17]. In addition, a similar APTM, called halo monitor, is located in from of the target and measures the beam loss by means of thermocouples only. Whereas at J-PARC RCS (Rapid Cycle Synchrotron, Japan), the Grid called Multi Wires Profile Monitor is also measuring the profile by integrating the current in the wires [18]. CERN, among other instruments, employs fast movable Wire-Scanners to obtain the beam profile by correlating the wire position with the detected signal [19].

Like other spallation sources, ESS also employs a suite of instruments to measure beam properties on target. Among these beam diagnostic instruments, Aperture Monitors (APTM) and Muti-Wire harp profiling monitor (Grid) are considered for this master thesis. In this chapter, the beam physics regarding those beam diagnostic instruments (APTM and Grid) is presented. The signal processing algorithm (developed in this thesis) to extract the beam properties from the signals acquired by those beam diagnostic instruments is also discussed later in this chapter. Prior to this thesis, a simulation model of the beam instrumentation suite, showing the performance of the instrumentation for the ESS beam on target characteristics, including the beam rastering system [5].

Nominally the ESS will operate with an average peak current (I_{peak}) of 62.5 mA and a pulse length of 2.86 ms at 2.0 GeV power. The long pulse duration leaves ample time for implementing a two-dimensional raster system that illuminates a defined footprint area by sweeping a slightly magnified beam ("beamlet") [20].

3.1 Beam Rastering

At ESS, the Lissajous like linear raster system will be provided by a fast 2D raster magnet system. The system will typically run at 10's kHz but the raster pattern here typically dictated by the ratio of the sweep frequencies (f_x, f_y), the waveform shape and relative phase.

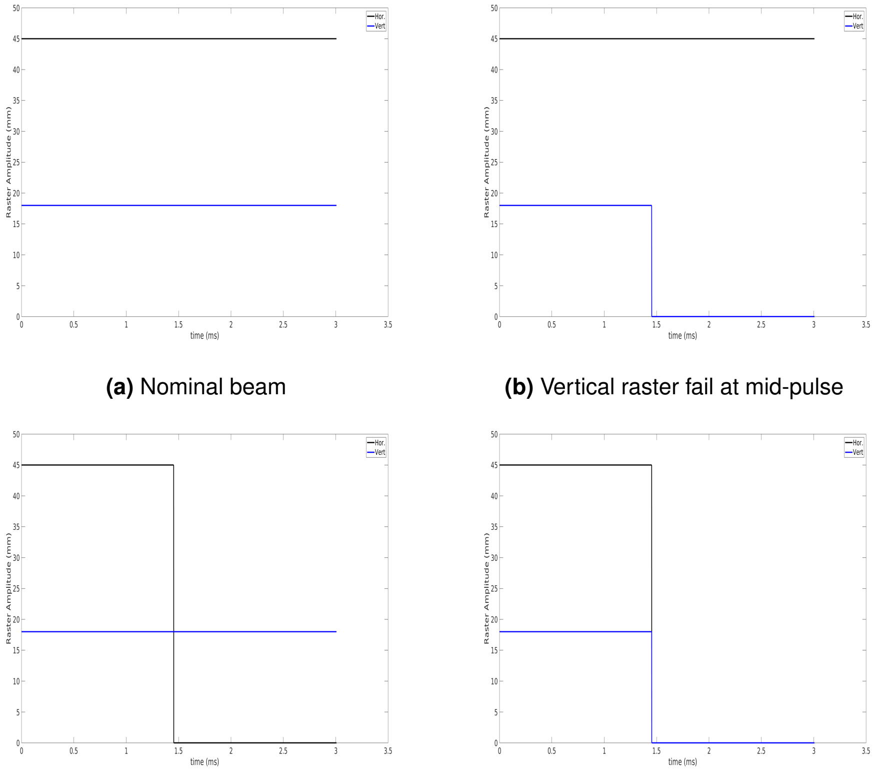
The raster function is modeled as shown below in the Eq. 3.1 ([5]) The waveform shape of the raster beam is modelled by applying a sawtooth function ($y = sin^{-1}(sinx)$) on centre lattice orbit position (cx_0, cy_0) along beam X and Y-axis respectively.

$$cx_0 = \frac{cx_r + 2ax}{\pi} sin^{-1}(sin2\pi f_x t) \tag{3.1}$$

$$cy_0 = \frac{cy_r + 2ay}{\pi} sin^{-1}(sin2\pi f_y t) \tag{3.2}$$

where f_x, f_y corresponds to the raster frequency in x and y-axis, cx_r, cy_r are the beamlet centers on target as guided by the beam lattice without raster system, with respect to the beam axis reference, ax, ay are the raster amplitude along X and Y -axis respectively.

In nominal conditions of the beam, the raster amplitude remains constant throughout the beam pulse length. Any disruptions in raster amplitude, or any other parameter like the beamlet position or the raster frequencies, leads to a production of errant beam at the target region.



(a) Nominal beam

(b) Vertical raster fail at mid-pulse

(c) Horizontal raster fail at mid-pulse

(d) Complete raster fail at mid-pulse

Figure 3.1: Raster Amplitudes of different modes of raster failures[5]

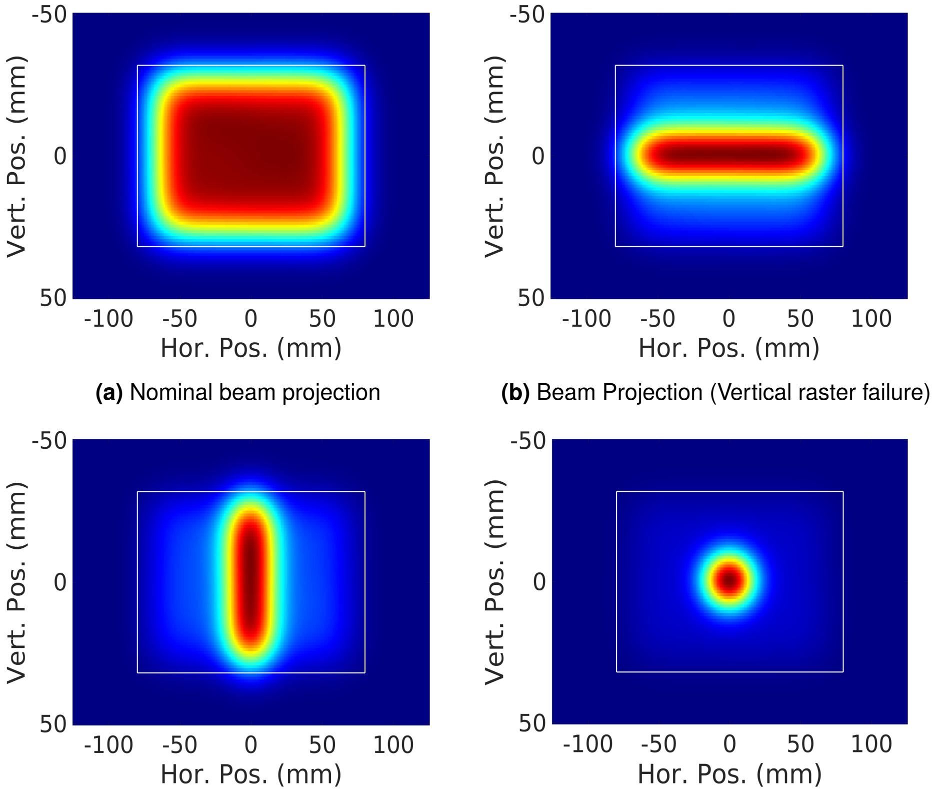
The figure 3.1 shows different possible raster failures at mid-pulse. In these examples, the failure manifests by having no deflection amplitude in on or both planes.

The model of the beam raster is composed of a 2-dimensional Gaussian beam current distribution, which centre moves as described in Eqs. 3.1

The expression is given by [5]:

$$I_m(x, y, t) = I_{0,t} \exp(-\frac{(x - cx_0(t))^2}{2\sigma_x^2} - \frac{(y - cy_0(t))^2}{2\sigma_y^2}) \tag{3.3}$$

where I_{im} is the amplitude of the current at time t in the (x,y) plane; $I_{0,t} = \frac{I_peak}{2\pi\sigma_x\sigma_y}$ is the normalised current peak density of the beam at time t; I_{peak} is the beam peak current; $\sigma_{x,y}$ the beam sizes in horizontal and vertical planes respectively.



(a) Nominal beam projection

(b) Beam Projection (Vertical raster failure)

(c) Beam Projection (Horizontal raster failure)

(d) Beam Projection (Complete raster failure)

Figure 3.2: Beam projections of different modes of raster, as simulated in Matlab environment [5]

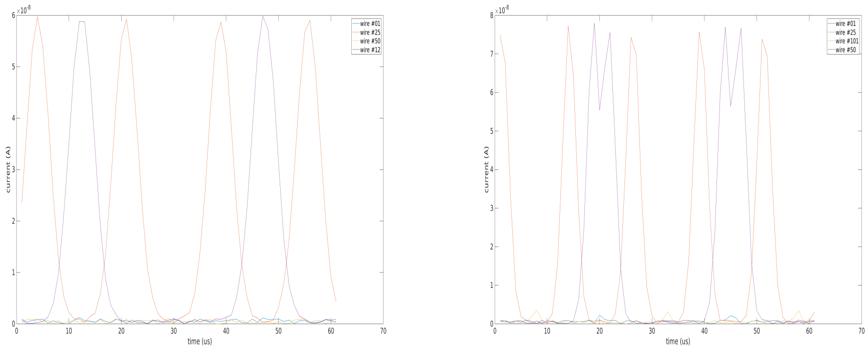
The figures 3.2 shows the result of the simulation model (In this model, 101 vertical wires (PV) and 50 horizontal wires (PH) are taken into consideration), where the produced image (shown in figure 3.2a) is the integrated current distribution for a nominal ESS beam.

The model produces an image of the beam current distribution at given sampling time. Here the pico-8 1µs sampling time. For each sampled imaged, the Grid wires current can be produced by integrating the proton charge over the wire projected area, dividing by the sampling time, and then applying Eq. 2.2 to get the current in the wire. The expression is shown in the Eqs. 3.4:

$$I_{wire,y} = \int I_m(x, y, t = ti)dx \tag{3.4}$$

$$I_{wire,x} = \int I_m(x, y, t = ti)dy \tag{3.5}$$

where $I_{wire,y}$ and $I_{wire,x}$ are defined as the current generated by a vertically and horizontally oriented Grid wires at position y and x respectively.



(a) Current by Horizontal wires

(b) Current by Vertical wires

Figure 3.3: Current generated on Grid wires, as simulated in Matlab environment (A)

The current generated by both of the horizontal and vertical Grid wires can be seen in the figure 3.3. An additive white gaussian noise of nano-range ($SNR = 10^3 : 1$) was also added as pico noise to the signal. The current generated at different wire position have different frequencies because of the rastering beam.

The frequency of the peaks, observed in horizontally oriented Grid wires (figure 3.3a) differs. It can be observed in the figure that the wires that are positioned centrally (wire #25) has higher frequency than the wires that are positioned at the edge (wire #01 or wire #50). Similarly for the vertically oriented Grid wires (as shown in figure 3.3b), the wires positioned centrally (wire #50) has two peaks consecutively than the wires that are positioned further away from it (wire #25 or wire #100).
These current signals generated by the Grid wires are the input to our system, as well as the current monitored by the BCM system (I_{peak}). The various beam properties is evaluated from those current signals by means of signal processing. The beam properties along with the algorithm for their evaluation is discussed in the next part of this thesis.

3.2 Beam Current Density from the Grid

The beam current density is one of the most crucial beam properties to be measured for the evaluation of errant beam conditions. The beam current density can be defined as the current generated through a unit area of defined cross section.
To avoid any kind of structural damage in the proton beam window, the requirement of the Grid system is to maintain the peak current density averaged over one pulse (ρ_I) less than 1.45 mA/cm^2, and also the peak current density for any sampled time should be less than 32 mA/cm^2.

The current density as measured with the Grid, $D_w(t)$, can be evaluated by the vector multiplication of the current values that are generated by the Grid wires and is given by,

$$D_w(t) = PH^T(t) \cdot PV(t) \tag{3.6}$$

Where the $D_w(t)$ is the vector product of the current values at any time instant t and $PV(t)$ & $PH(t)$ are the vectors representing the current from vertically and horizontally aligned Grid wires.

$$PV(t) = \begin{bmatrix} v_{1,t} & v_{2,t} & ... & v_{n,t} \end{bmatrix}, PH^T(t) = \begin{bmatrix} h_{1,t} \\ h_{2,t} \\ ... \\ h_{m,t} \end{bmatrix} \tag{3.7}$$

As mentioned before both $PV(t)$ and $PH(t)$ are time dependent vectors and are of n and m dimensions respectively, representing the peak current measured in each vertical and horizontal wires of the Grid.

For the calculation of the current density at sample time t , the matrix D_w is normalised to the proton beam current, I_{peak}, which again an input to the system and is delivered by the BCMs.

$$D_{w,norm}(t) = \frac{D_w(t)}{\sum_{i=o}^{m} \sum_{j=o}^{n} D_w(t) dx dy} I_{peak}(t) \tag{3.8}$$

where dx, dy are the wire spacing between the horizontally oriented and the vertically oriented wires respectively. The max current density (ρ_I) at time instant t is then evaluated by finding the maximum value over the elements of the normalised matrix $D_{w,norm}$.

$$\rho_I = max(D_{w,norm}(t)) \tag{3.9}$$

As the raster system sweeps the beam across the target footprint (160×64 mm^2) over time, the integrated current density over time decreases. The figure 3.4 shows the max current density for the cumulative sum of $D_{w,norm}(t)$ over time. The peak power rapidly decreases and converges towards $I_{peak}/4/(ax+\sigma_x)/(ay+\sigma_Y)$. After 100µs, the max current density is within a factor 2 of the converging value.

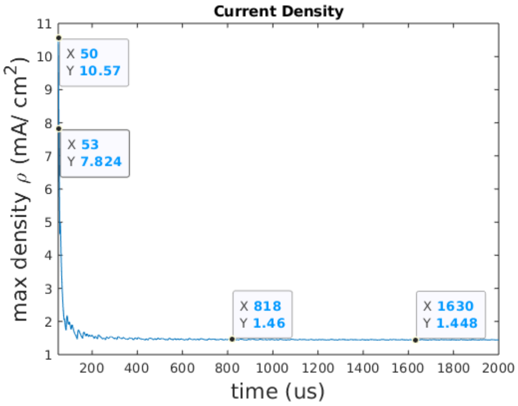


Figure 3.4: Max Current Density evolution during one pulse

In this master thesis, this behaviour of beam current density is used to detect errant beam. To perform this with minimum latency, three separate time windows have been chosen $n = 1\mu s$, $2\mu s$ and $5\mu s$. The beam current densities from those three windows have been compared and if there is any abnormality in the decreasing trend of the beam current density, a trigger for errant beam detection is raised. The hardware implementation portion of this current density evaluation will be discussed in the later chapters of this thesis.

3.3 Beamlet and Beam Size

The beam size in both axis is defined by the r.m.s width of the beam pulse distribution. Let's call these quantities $\sigma_{x,p}$ and $\sigma_{y,p}$. Let's recall the beamlet sizes σ_x and σ_y to be the r.m.s beam size as measured on $D_{w,norm}(t)$ at the sample time t.

One can measure the beamlet or the beam size along the horizontal and vertical axis by means of the two dimensional moment analysis of the normalised vector product $\int_{t=0}^{t=t_i} D_{w,norm}(t)dt$ (as defined in equation 3.8). For the beamlet size, t_i is one sample time, for the beam size, $t_i = t_{pulse}$, the pulse duration.

The equations for the beam sizes are given by,

$$\sigma_x = \frac{(Mu_{20} + Mu_{02}) + \sqrt{(Mu_{20} + Mu_{02})^2 - \{4(Mu_{20} + Mu_{02})Mu_{11}^2}}{2} \tag{3.10}$$

$$\sigma_y = \frac{(Mu_{20} + Mu_{02}) - \sqrt{(Mu_{20} + Mu_{02})^2 - \{4(Mu_{20} + Mu_{02})Mu_{11}^2}}{2} \tag{3.11}$$

Where Mu_{20} , Mu_{02} are the second order and Mu_{11} is the first order central moments. These are defined as,

$$Mu_{20} = \frac{\sum (X - \sum Mu_{10})^2 D_{w,norm}}{M_{00}} \tag{3.12}$$

$$Mu_{02} = \frac{\sum (Y - \sum Mu_{01})^2 D_{w,norm}}{M_{00}} \tag{3.13}$$

$$Mu_{11} = \frac{\sum (X - \sigma_x)(Y - \sigma_y) D_{w,norm}}{M_{00}} \tag{3.14}$$

Here X Y are initialization matrices and M_{00} is the zero order central moment. They are defined as,

$$X = \begin{bmatrix} 1 & 2 & ... & m \\ 1 & 2 & ... & m \\ ... & ... & ... & ... \\ ... & ... & ... & ... \\ 1 & 2 & ... & m \end{bmatrix} \tag{3.15}$$

$$Y = \begin{bmatrix} 1 & 1 & ... & 1 \\ 2 & 2 & ... & 2 \\ ... & ... & ... & ... \\ ... & ... & ... & ... \\ n & n & ... & n \end{bmatrix} \tag{3.16}$$

$$M_{00} = \sum_{i=o}^{m} \sum_{j=o}^{n} D_{w,norm} \tag{3.17}$$

In here, m and n are dimensions of the $D_{w,norm}$ matrix.

The max current density is inversely proportional to the beam(let) sizes $\sigma_{x,p}$ and $\sigma_{y,p}$. There are two requirements on the beam properties to be delivered to the target. The first one is that the pulse current density is less than $1.45mA/cm^2$. The other one is that the beamlet size is such that $\sigma_{xy} > S_{thr} = 30cm^2$. Where σ_{xy} is defined as $\sigma_x\sigma_y$.

The measurement of beamlet max current density verifies the second requirement at any time. For a Gaussian beamlet distribution, the max current density is express as

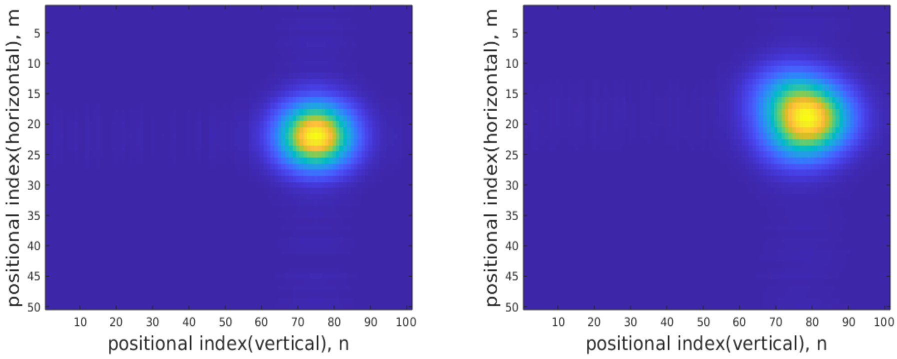
$$\rho = \frac{I_{peak}}{2\pi\sigma_x\sigma_y} \tag{3.18}$$

hence, the second condition is verified for any value of $2\pi\rho/I_{peak} < 1/S_{thr}$.

The first requirement is ensured by the raster system with the correct amplitude and with un-correlated frequencies in both axis. This translates to a converging max current density (as the integrated charge is swept over a larger area over time, the current density decreases), and thus measurement of the max current density with $t_i = 1, 2, 5$µs should always be sorted as

$$max(D_{w,norm,t_1}) > max(D_{w,norm,t_2}) > max(D_{w,norm,t_5}) \tag{3.19}$$

The equation 3.19 states that the maximum value of matrix $D_{w,norm,t_1}$ over its elements should be greater than that of the maximum value of $D_{w,norm,t_2}$ matrix over its elements, which again should be greater than the maximum of $D_{w,norm,t_1}$ matrix over its elements. When this relation is broken, an errant beam is detected.

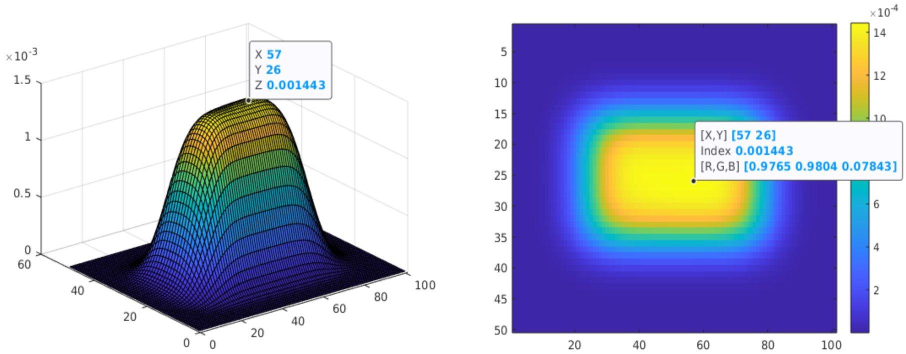


(a) Integrated beamlet profile ($1\mu s$)

(b) Integrated beamlet profile ($5\mu s$)

Figure 3.5: Beamlet generated on Grid wires, simulated in Matlab environment (A)

3.4 Beam Position



(a) Surface plot of the beam profile

(b) 2-D projection of the beam profile

Figure 3.6: Centroid of the Beam, as simulated in Matlab environment (A)

The beam position is defined as the centre of mass of the beam pulse current density. The beam position is given by the first order moment of the normalised matrix $D_{w,norm,t}$. First order moment calculation is defined as,

$$Mu_{10} = \frac{X \cdot D_{w,norm,t}}{M_{00}} \tag{3.20}$$

$$Mu_{01} = \frac{Y \cdot D_{w,norm,t}}{M_{00}} \tag{3.21}$$

$$M_{00} = \sum_{i=o}^{m} \sum_{j=o}^{n} D_{w,norm,t} \tag{3.22}$$

where Mu_{10}, Mu_{01} are the first order and Mu_{00} is the zeroth order moment of the $D_{w,norm,t}$ matrix.

The beam position coordinate is then given by: $(c_x, c_y) = (Mu_{10}, Mu_{01})$

The accuracy of the method depends mainly on the noise in the signal. This is not discussed in this thesis.

3.5 Raster Frequency

The beam on target must be rastered. Without raster, the beam will damage the target and compromise operations. In order to ensure the beam to be rastered, the wire signal can be processed in order to measure the raster frequency.

Failure of raster system results in failing to spray the beam across the footprint which in turn also increases the beam current density. Hence, maintaining a stable raster frequency on both horizontal and vertical axis is essential for a nominal beam production. The measurement of the beam raster frequency is not critical as a failure of the raster system will be detected by the peak density not decreasing. However, detecting and measuring it will help to diagnose the errant beam condition found.

Fast fourier transform (FFT) is used to determine the raster frequency on both horizontal and vertical direction. FFT computes discrete Fourier transform (DFT) of a signal, the time series signal is converted to a frequency domain signal in Fourier analysis. FFT reduces the computation complexity in DFT, which is defined as

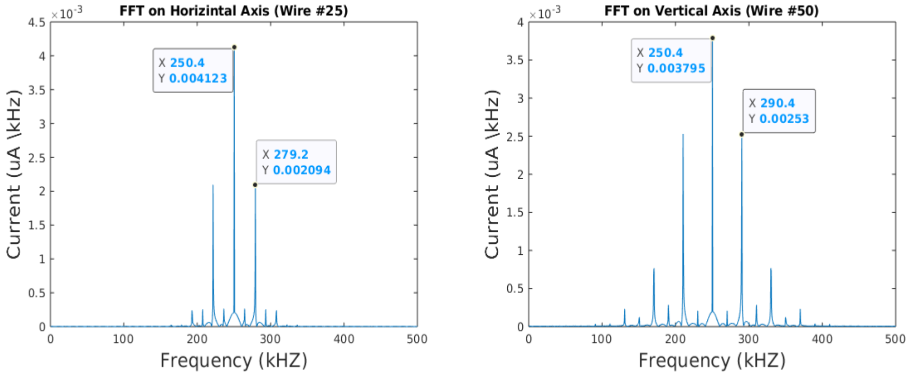
$$X_k = \sum_{n=0}^{N-1} x_n \exp \frac{-i2\pi kn}{N} \tag{3.23}$$

where $k = 0, 1, ..., N - 1$ and $\exp \frac{-i2\pi kn}{N}$ is a primitive Nth root of 1.

To compare speed computation and accuracy of the raster frequencies, FFT with $N = 1024, 512, 128$ were performed. The table 3.1 shows the comparison between the different calculation of raster frequencies due to different N-points. It also can be seen that higher N-point FFT has more accuracy in estimating raster frequencies from the current signals but higher N-point FFT also includes more computational complexity.

N-point (FFT)	**1024**	**512**	**128**
Lowest frequency detected (kHz)	0.98	1.95	7.81
Highest frequency detected (kHz)	500	500	500
Raster frequency (f_x, kHz)	40	40	39.1
Raster frequency (f_y, kHz)	28.8	28.3	27.4

Table 3.1: FFT comparison of Raster frequencies



(a) FFT of the current from horizontal Grid wires **(b)** FFT of the current from horizontal Grid wires

Figure 3.7: FFT of the Current Signals, as simulated in Matlab environment (A)

The figure 3.7 shows $N = 1024$ point fast fourier transform of the current signals of both horizontal and vertically oriented Grid wires. The raster frequency can be derived from this magnitude-frequency plot the given FFT and from the second harmonics of the FFT signal the raster frequency can be computed as,

$$f_y = (279.2 - 250.4) = 28.8 \text{ kHz}$$
$$f_x = (290.4 - 250.4) = 40 \text{ kHz}$$

The calculated raster frequencies f_x and f_y differs about ±0.2% from the pre-defined values. But the calculation complexity $Nlog(N)$ is still higher from implementation point of view. The fast fourier transform is highly susceptible to noise, a noise signal in the range of 10 kHz is also visible (as shown in figure 2.13). This makes the FFT inaccurate in estimating the raster frequencies in real time environment. Also the failure of raster system can also be detected by the gradual increase in beam current density over time.

Since the raster failure can be detected by monitoring the beam current density, it is unnecessary to implement a complex FFT algorithm in real time environment where minimal latency in computation time is a requirement for the errant beam detection system. Hence, this property of beam is not essential to be implemented on hardware to detect an errant beam.

Chapter 4
Method

This chapter explores the hardware implementation of the algorithms to extract the deterministic beam properties for determining an errant beam. Among all the deterministic beam properties, the evaluation of maximum beam current density is the most necessary beam property to detect an errant beam production. The selected algorithm to extract the peak current density from the beam also requires multiple AMC boards to evaluate current sample data from multiple GRID wires, which will be discussed later in this chapter (section 4.3).

In the previous chapter of this thesis, the calculation of the beam current density (ρ_I) along with the beam sizes (σ_x, σ_y) and beam centroids (X_{centre}, Y_{centre}) have been discussed in detail. Matlab implementation of those algorithms has been done on the software level to check the validity of the calculation. Still, when it comes to the implementation on the hardware level, the algorithms have to be modified to a certain level to meet the minimum latency requirements. The hardware implementation of the discussed algorithm to detect errant beam is done by means of electronics. An AMC card containing Xilinx Kintex Ultrascale FPGA is used to achieve this. Various algorithms were implemented to achieve minimum latency for computation of the maximum beam current density. The algorithms will be discussed in the next section of this chapter (section 4.2).

The algorithm that is developed in this master thesis for the detection of errant beam condition in the hardware level also has to comply with the existing FPGA framework at ESS. The ESS FPGA Framework aims at re-usability, configurability of the AMC cards to be used in different projects at ESS. The FPGA framework is divided into three major structural hierarchies.

- **sis8160_essffw_top:** This is the top-level structural hierarchy of the framework. It connects the AMC card peripherals (onboard DDR memory, FMC modules) to the *framework_core*.
- **framework_core:** This is the next structural hierarchy in the FPGA framework. It holds most of the registers to be used across the framework to connect between different modules. All the registers in the framework is connected via AXI communication protocol.
- **custom_logic_wrapper:** This structural hierarchy contains the registers for the user application and the user application logic itself. The custom logic registers

can also be accessed through the AXI communication protocol.

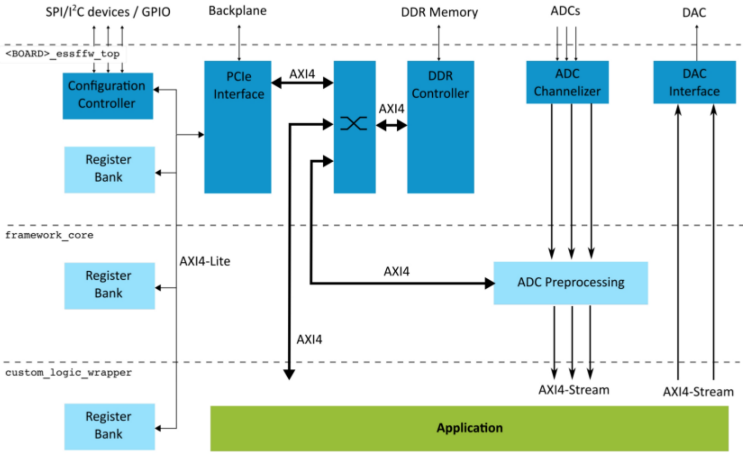


Figure 4.1: ESS FPGA Framework [6]

The proposed algorithm (discussed in the next section) will be implemented in the application part of the *custom_logic_wrapper* (as shown in figure 4.1).

4.1 Algorithm Overview (APTM)

This section describes the hardware implementation of a much simpler algorithm for the APTM instrument. This part of the machine protection system detects if the current in the blade is larger than a threshold value. If the current read by the APTM blades exceeds a pre-defined threshold, the algorithm should trigger a signal to stop the production of the beam. Higher current on a small surface translates to higher current density, which should be avoided to protect the machine.

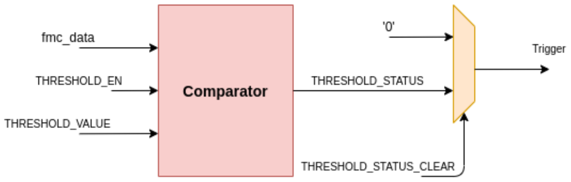


Figure 4.2: Hardware implementation of the APTM algorithm

Figure 4.2 shows the proposed logic for the hardware implementation of the algorithm. The *fmc_data* is the current data from the APTM blades, where *threshold_value* is the user-defined threshold for each channel. Two more control signals (*threshold_en* & *threshold_status_clear*) are used in this algorithm to enable the logic and to clear the threshold_status bit in the algorithm. The threshold value signal (THRESHOLD_VALUE) as well as the control signals (THRESHOLD_EN, THRESHOLD_STATUS and THRESHOLD_STATUS_CLEAR) are read from/ written to the Register Bank of *custom_logic* and can be accessed by the system CPU.

4.2 Algorithm Overview (Grid)

This section describes the hardware implementation of the algorithm to evaluate the maximum beam current density (as discussed in section 3.2) for the Grid instrument. To minimize the complexity of hardware implementation in FPGA, high-level synthesis (HLS) tool from Xilinx (Vivado HLS) is used to test the proposed algorithm.

To meet the latency requirements for the evaluation of maximum current density, the algorithm must evaluate every sample at every 1µS or at every 100 cycles (considering the FPGA is running at 100 MHz). Three algorithms were tested in HLS to achieve this. Since the HLS tool from Xilinx uses an automated design process, which interprets the algorithmic description of a desired function and implements that behaviour in hardware. **Pragma** (directives) are used to manipulate the HLS compiler for the generation of hardware.

4.2.1 Algorithm I

The first tested algorithm to evaluate the current density requires multiplication of two vectors containing current values from Grid wires (PV and PH respectively). In this thesis, 101 vertical wires (PV) and 50 horizontal wires (PH) are taken into consideration. Since all the current values from the Grid wires are sampled at 1 MSample/s. At each time instant ($t = 1\mu$S), we obtain a single dimension vector from both the horizontal and vertically oriented Grid wires.

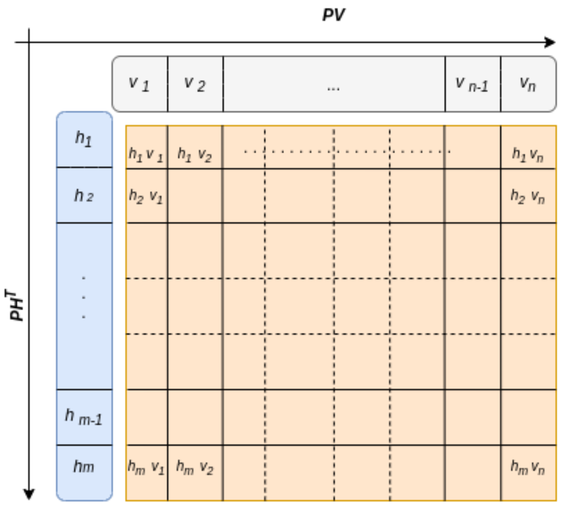


Figure 4.3: Algorithm I

The figure (4.3) shows calculation of the matrix multiplication defined by,

$$D_w = PH^T \cdot PV \tag{4.1}$$

In hardware implementation, normalisation of the entire D_w matrix is unnecessary, since the objective is to find the maximum value in this matrix.

For finding the maximum value in the calculated matrix D_w, a running buffer to search each elements of the matrix and replace the highest value with newly found element is implemented in this case. A summing algorithm is also implemented to find the summation of the matrix. The peak current density is thus found by dividing the maximum value by the summation of the matrix (over the elements of the matrix of dimensions m, n) and is given by,

$$\rho_{I,max} = \frac{max(D_w)}{\sum_{i=0}^{m}\sum_{j=0}^{n} D_w(m,n)} \tag{4.2}$$

The calculated $\rho_{I,max}$ is the peak current density for 1μS window. To detect the errant beam production, we need to confirm the decreasing trend of peak current density (as shown in figure 3.4). This behaviour is achieved by creating three separate buffers for the D_w matrix viz. D_{2w}, D_{5w}. Hence the previous 4 matrices are stored and the values are used for the evaluation of separate windows.

The D_w buffer is designed to evaluate at each $t = 1\mu$S. Whereas the D_{2w} buffer evaluates the average of the new D_w and the D_{2w} which is the previous D_w and stores it in D_{2w}. Similarly the D_{5w} buffer evaluates at each 1µS instant.

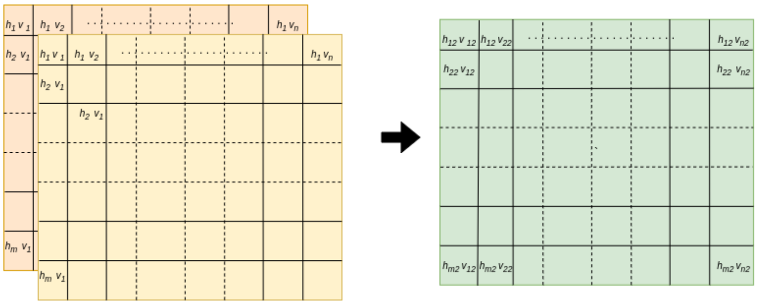


Figure 4.4: Calculation of 2μS window

The figure 4.4 shows the calculation of the 2μS buffer. Each element of the matrix D_{2w} is calculated by summing the corresponding elements of three D_w matrices. This can be given by,

$$h_{12}v_{12} = (h_1v_1)_t + (h_1v_1)_{t-1}$$
$$h_{22}v_{22} = (h_2v_1)_t + (h_2v_1)_{t-1}$$

Where $h_{12}v_{12}$ is the first element of the D_{2w} matrix and $(h_1v_1)_t$, $(h_1v_1)_{t-1}$ are the first element of the D_w matrices from current (t) and previous time $(t-1)$ instance respectively.

Similarly the 5μS buffer is calculated. The peak beam current density from each buffer is evaluated (as per equation 4.2) and compared. The current density condition for a nominal beam production is given by,

$$\rho_{I,max(1\mu S)} > \rho_{I,max(2\mu S)} > \rho_{I,max(5\mu S)} \tag{4.3}$$

Where $\rho_{I,max(2\mu S)}$, $\rho_{I,max(5\mu S)}$ are the peak current densities of 2μS and 5μS respectively. This algorithm evaluates at each 1µS, though it requires a startup time to have the past 4 matrices to be loaded to start the evaluation but once it has all the values for the previous matrices, the algorithm can run continuously for every sample. Apart from high hardware utilization, this algorithm also incurs a high latency (discussed in the next chapter) which enabled us to find for an alternative solution for the evaluation of peak current density.

4.2.2 Algorithm II

An alternative approach to calculate the peak current density is discussed in this algorithm. This method solves both the high hardware utilization and high latency problem we experienced in the native algorithm (Algorithm I). In this algorithm, instead of implementing three separate buffers to evaluate the $\rho_{I,max(1\mu S)}$, $\rho_{I,max(2\mu S)}$ and $\rho_{I,max(5\mu S)}$ respectively. A single buffer (Acc_X, Acc_Y) for each wire orientation(vertical and horizontal) is introduced. These buffers acts as Accumulators and are defined as,

$$Acc_X = \sum_{t=0}^{Ti} PH(t)$$

$$Acc_Y = \sum_{t=0}^{Ti} PV(t)$$

where $Ti = 5\mu$S.

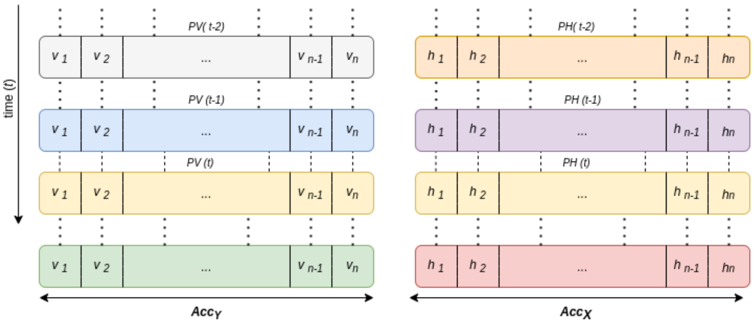


Figure 4.5: Accumulators for Algorithm II

The figure 4.5 shows the visual description of the accumulators, which add the current data over time t and integrate until $t = Ti$.

The complexity to evaluate peak current density could not be solved by the HLS tool since the tool is efficient when it comes to matrix multiplication however the tool can not optimize the evaluation of the maximum and the summation of the D_w matrix. Hence, two separate processes are introduced to solve the problem in this algorithm.

- **row process:** In this, a single element of Acc_X is multiplied with all the elements of Acc_Y (as shown in figure 4.6). The entire process results to a single row of D_w matrix and the process is repeated for every element in Acc_X. This process also computes the summation and maximum of each row simultaneously.

$$Row(1, j) = Acc_X(1).Acc_Y(j)$$
$$Row(2, j) = Acc_X(2).Acc_Y(j)$$

and so on.

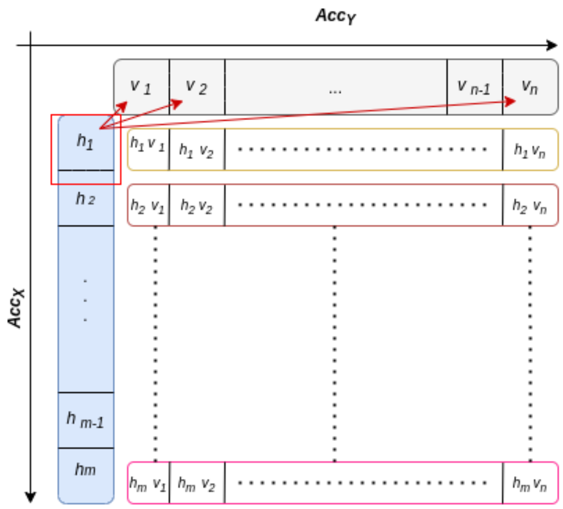


Figure 4.6: Algorthm II

- **column process:** In this process, the maximum value from each row is compared to obtain the overall maximum value of the D_w matrix and the sum from each row is also added to get the sum of D_w over its elements ($\sum D_w$). The values are then divided to get the peak current density (as shown in equation 4.2).

This algorithm inherently calculates the peak current density at each time instant due to the accumulators (Acc_X, Acc_Y). An internal timer is used to get the peak current density values at $t = 1\mu$S, 2μS and 5μS. All the peak current densities are checked again to match the criteria described in equation 4.3. Any deviation from this condition should point to an errant beam. This algorithm despite of having a nominal hardware utilization and latency (discussed in the next chapter) can only be evaluated at each 5µS, this enabled us to explore further more into algorithms for the evaluation of peak current density.

4.2.3 Algorithm III

Another approach to evaluate the peak current density with the help of the acquisition-concentrator AMC card configuration is explored in this thesis (section 4.3). In this algorithm, the integration of the current values is done on the acquisition board. Each of the acquisition board contains three separate buffers for the evaluation of 1μS, 2μS and 5μS window respectively.

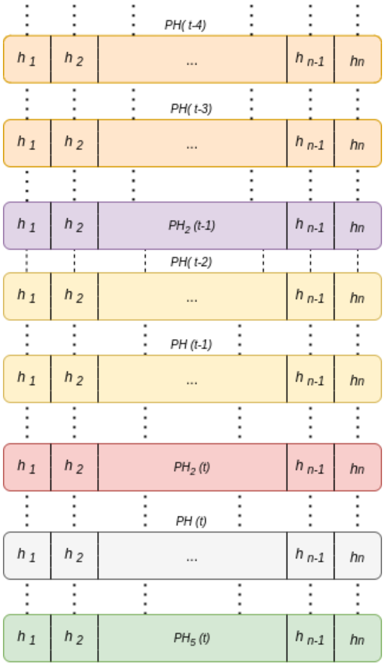


Figure 4.7: 1μS, 2μS, 5μS buffer for Horizontal Grid wires

The figure 4.7 shows the three separate windows PH, PH_2 and PH_5 for the integration of the current values from the horizontally oriented Grid wires at t = 1μS, 2μS, 5μS respectively.

The rest of the calculation for the sum and the maximum of the D_w matrix is done on the concentrator board for each window in parallel. Once all the integrated current values are acquired at the concentrator board, this method also implies three separate instances to evaluate peak beam current for each window (viz. 1us, 2us and 5us) continuously over time.

After the evaluation of current densities ($\rho_{I,max(1\mu S)}$, $\rho_{I,max(2\mu S)}$ and $\rho_{I,max(5\mu S)}$) for each window, the decreasing trend (according to equation 4.3) is checked for the detection of errant beam.

4.3 Data Transfer Overview

The necessity of high speed data transfers between the cards in this master thesis project is inevitable. Since in one AMC card, there will be a total of 8 channel inputs for the Grid module for acquiring current signals. To do a multiple signal acquisition (more than 8 inputs) for evaluating the current density of the incoming beam, multiple AMC boards are needed.

To achieve this multi-board implementation in real environment, a communication protocol is needed between boards. An acquisition - concentrator model was proposed for this communication between boards. The acquisition board acts as a signal collector from the Grid wires whereas the concentrator board acts as a master board to collect all the data from the acquisition boards and do the necessary computation to detect errant beam.

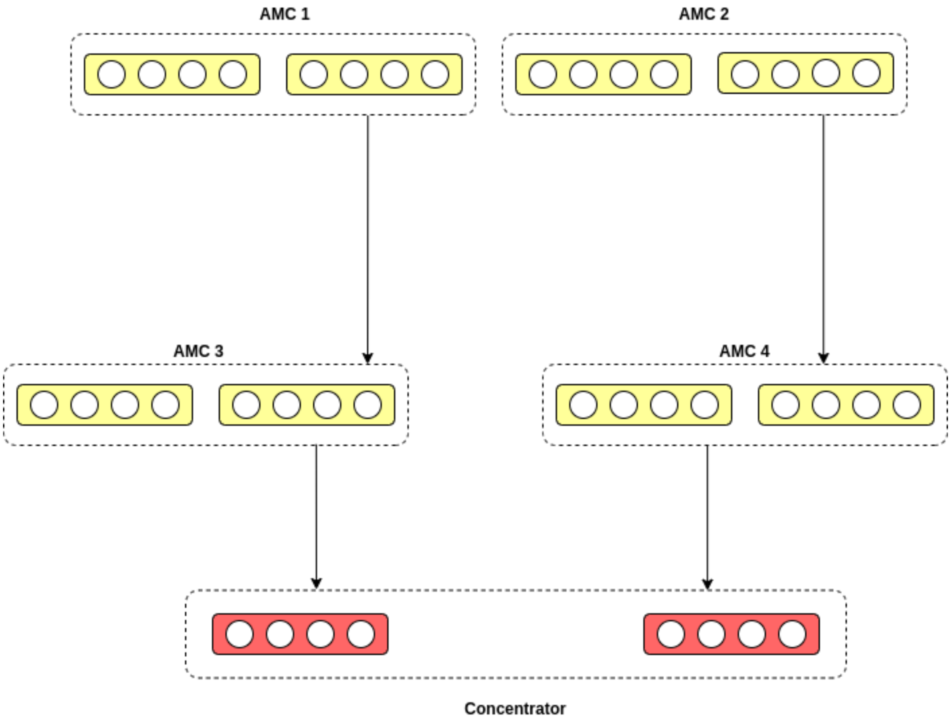


Figure 4.8: Acquisition- Concentrator Board Configuration

The MTCA crate (Schroff MTCA.4) can have multiple modes of communication protocols between the AMC boards viz. PCie, P2P, Gigabit ethernet, Optical Fibre etc. The Gigabit Ethernet and PCie protocols need a host computer to transfer data. These communication protocols can not connect two AMCs in the absence of a host computer, and the host computer can interrupt the data transfer to complete its system processes. But the P2P communication link uses the gigabit transceiver pins of the FPGAs to establish a non interrupting high-speed communication protocol between two AMC cards. Hence the point-to-point (P2P) communication link is considered as the most optimal solution since we have a timing constraint of minimum latency.

In point-to-point (P2P) communication protocol, data from device A (in here AMC1) is to be transferred to device B (in here the concentrator) without any switching medium in between and with minimal complexity. The point to point link between boards in the Micro-TCA (MTCA) crate has been described in the figure 4.8. Since the MTCA crate has limitations in P2P links between boards, 5 among 7 AMC slots can be used for P2P communication link. In this configuration, normally signals from 32 + 8 (concentrator board) channels/ wires can be acquired and sent to the concentrator board for computation.

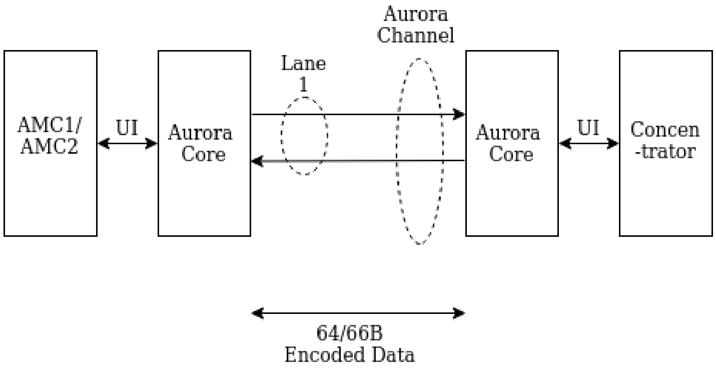


Figure 4.9: Xilinx 64/66B Aurora core configuration

To enable the implementation of high-speed serial communication protocol, Aurora IP from Xilinx is used. Xilinx boards offer a lightweight, low-cost, high-speed Aurora communication protocol for P2P connections with a minimum speed of 500 Mbps over a single lane. The figure 4.9 describes the aurora core configuration between the collector (AMC1 to AMC4) and the Concentrator board. To achieve the lowest latency for acquiring the signals from the acquisition boards, one 64/66B Aurora core with a single lane will be used for each of the AMC boards (AMC1, AMC2, AMC3, and AMC4).

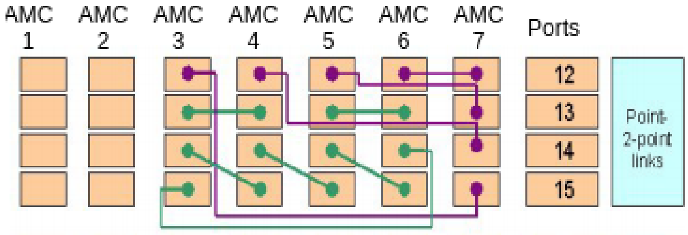


Figure 4.10: Backplane topology of MTCA.4 crate [7]

Figure 4.10 describes the back-plane topology of the Schroff 7 slot MTCA.4 crate. The crate is used for this prototype project where the concentrator board (AMC 7) has a single point-to-point (P2P) connection to every other acquisition board in the crate (AMC 3-6) as per the figure 4.10. One P2P link can transfer the acquisition data from 2 channels (current values from two-wire) at a time over a single Aurora lane. So, for the proposed communication protocol, one can acquire current signals from only 32 (8 × 4) + 8 (concentrator board) channels by using 4 Aurora IP(s) in total and do the required computation for detecting errant beam. But for the full implementation of the system, A crate with ten slots will be used where the sampled data needs to be transported across cards to connect all AMC boards to the concentrator (AMC7).

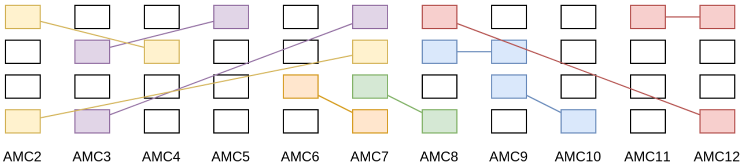


Figure 4.11: P2P connection of 12-slot crate [8]

Figure 4.11 shows an example of the P2P connection between AMC boards within the 9U crate (crate consisting of 12 AMC slots, out of which 11 AMC slots have P2P connection between them and one slot is used for the system CPU). It shows the complex connections between 10 boards (one of the card is used for the timing synchronisation) with shortest distances, though the longest path above is AMC11 and AMC12, which needs to cross 3 AMCs before arriving at the concentrator board (AMC7).

In Aurora IP configuration, the user can transmit data through two different interfaces (either framing interface or streaming interface). The framing interface requires a separate Framer, which will pack each data into 32-bit frames, including assigning additional bits like destination address, end of frame (EOF), start of frame (SOF) etc. The addition of these extra bits for framing incurs undesired latency, which doesn't meet our latency requirements in theory. Whereas streams are implemented in Aurora IP configuration are continuous unending frame. Since the FMC pico data is sampled at a high frequency (1 Msample/s), to avoid unnecessary latency streaming option is preferred over framing.

However, this streaming interface still has it's drawbacks in BER (bit error rate). This should be considered in this case since the data is transmitted through copper wires in the backplane of the crate. These backplane signals can be susceptible to cross-talks or any other kind of electrical interference in a real-time environment even though 64/66B core has lower BER compared to 8/10B core (theoretically).

4.3.1 Addressing Mode

Addressing the FMC-pico data is absolutely necessary for the evaluation of errant beam condition as the algorithm discussed earlier in this chapter requires multiplication of acquisition data from specific channels. To facilitate this, the acquisition data (current values) from each channel has to be identified for correct multiplication and evaluation. In the data streaming configuration of the Aurora IP, the packing and addressing of data has to be done by the user. Since the acquisition data from the FMC(s) will be coming from multiple channels at different positions, the user needs to include some addressing mode to the current data itself. The following proposal for addressing the data is an efficient method to meet the latency requirements since the addressing of the gird wires are embedded in the current data from the wires itself.

In the current hardware design, the FMC-pico(s) acquire 20-bit current values from the wire at a rate of 1 MSample/s. These 20-bit values are converted to 32-bit values while storing them to registers on the AMC board for further evaluation. This is being done by assigning 0's to the upper 12 bits of the 32-bit register data. In this thesis, the upper 12 bits are used to address the FMC pico data, as shown in figure 4.12.

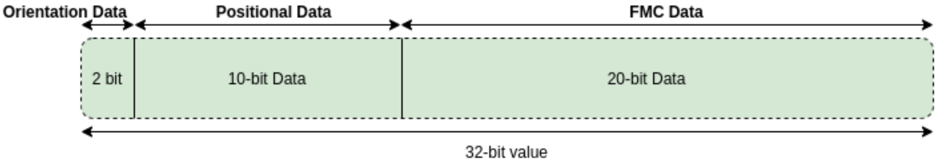


Figure 4.12: Addressing scheme of FMC-pico data

The 32-bit FMC pico data is divided into three sections in this case (figure 4.12). The lower 20-bit values are assigned as the pico data(s), and the two most significant bits (MSB) are called Orientation Data. The Orientation data can be used to address the horizontal or vertical orientation of the Grid wires. 2- bits are chosen for this to minimize the bit error rate during transmission. The 10-bit information between the Orientation data and the FMC pico data is used as the positional data of the Grid wires in a given orientation. In this given configuration, one can address 1024 no of wires either in vertical or in horizontal direction.

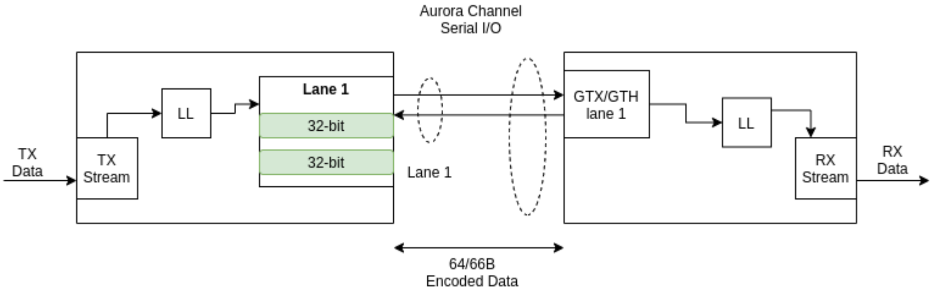


Figure 4.13: Aurora IP configuration for Grid

Figure 4.13 describes the Aurora IP core configuration of this master thesis. Each IP core is fed with transmission/ receiving data (TX/RX Data) through streaming user interface. The interface moves data from the application to the channel as well as performs TX flow controls. Each transceiver is also driven by an instance of lane logic (LL), which handles the encoding and decoding of control characters and also performs error detection.

In this master thesis, due to the limitation in P2P link connection between the boards (in MTCA.4 crate), only five out of seven AMC boards can be used. Four of them will be used as the acquisition boards (AMC3-7 as shown in figure 4.10). However, data from 8 channels can also be acquired at the concentrator board, but then the user needs to form a finite state machine to wait for the data from acquisition boards to arrive before starting the computation to detect an errant beam.
The latency to transmit data from 8 channels (one acquisition board) will be discussed in the next chapter in this thesis. This estimation is just to give the reader an idea of what the transfer protocol could look like and to give out the future prospects of data transmission between the AMC cards. Detailed evaluation of this data transmission is beyond the scope of this master thesis.

Chapter 5

Results

This chapter discusses about the hardware implementation of the proposed algorithms, discussed in the previous chapter. Each of the algorithms are tested and compared in terms of accuracy, hardware utilization, estimated latency to detect the condition for errant beam production.

5.1 Accuracy

The accuracy of the calculated beam parameters are crucial to detect the errant beam. The Matlab model as well as the derived C++/ HLS model uses floating point data type for calculating the beam parameters. To simplify the calculations on hardware level, fixed points instead of floating points are used for hardware implementation. The conversion from a floating point to fixed point calculation incurs error to the beam parameter calculation.

Beam Parameters	**MATLAB model**	**HLS/C++ model (floating point)**	**HLS/C++ model (64 bit fixed point)**	**HLS/C++ model (optimized)**
Current Density (A)	0.0011	0.00114568	0.00114568	0.0011571
σ_x	14.5722	14.5722	14.5722	14.5834
σ_y	6.9748	6.97477	6.97477	6.98012
Centre X	52.1973	52.1973	52.1973	52.2758
Centre Y	27.3983	27.3983	27.3983	27.4388

Table 5.1: Beam Parameters for 1µS buffer

Beam Parameters	**MATLAB model**	**HLS/C++ model (floating point)**	**HLS/C++ model (64 bit fixed point)**	**HLS/C++ model (optimized)**
Current Density (A)	0.00094031	0.000940313	0.000940317	0.00094533
σ_x	14.6448	14.6448	14.6448	14.6559
σ_y	7.0160	7.01603	7.01603	7.02135
Centre X	51.1746	51.1746	51.1746	51.2506
Centre Y	26.7489	26.7489	26.7489	26.7879

Table 5.2: Beam Parameters for 2μS buffer

Beam Parameters	**MATLAB model**	**HLS/C++ model (floating point)**	**HLS/C++ model (64 bit fixed point)**	**HLS/C++ model (optimized)**
Current Density (A)	0.00074330	0.000743302	0.000743304	0.000746489
σ_x	14.4949	14.4949	14.4949	14.506
σ_y	7.1678	7.1678	7.1678	7.17333
Centre X	50.7079	50.7079	50.7079	52.2758
Centre Y	25.9975	25.9975	25.9975	26.0357

Table 5.3: Beam Parameters for 5μS buffer

Beam Parameters	**Error (%)**, 1μS window	**Error (%)**, 2μS window	**Error (%)**, 5μS window
Current Density (A)	0.9	0.5	0.4
σ_x	0.07	0.07	0.07
σ_y	0.07	0.07	0.07
Centre X	0.15	0.148	0.149
Centre Y	0.147	0.145	0.146

Table 5.4: Accuracy of beam parameters

As it can be seen in tables 5.1, 5.2 and 5.3, the beam parameters lose precision when floating point calculation is changed to fixed point calculation. The error(%) due to this fixed-point calculation is shown in table 5.4 for each of the windows (1µS, 2µS and 5µS respectively). To optimize the hardware utilization of the algorithms, the data types of each variable are further optimized which also further reduced the precision of the beam parameters to a permissible limit. The optimized version of the HLS/ C++ model also optimizes the number of multipliers (DSP48E) since in this model a single multiplier is needed for the evaluation of a single element in D_w matrix.

5.2 Hardware Utilization

The resource utilization of FPGA in digital design is one of the important design parameter since most of the time the synthesis and place-n-route of an implemented algorithm is handled by an automated tool (in this master thesis Xilinx tools have been used).
The implementation of these proposed algorithms will be placed in the *custom logic* part of the *ess fpga framework* (as shown in figure 4.1). The framework itself occupies a significant portion of the hardware resources (as shown in table 5.5), which leaves a specific amount for user applications. As the puzzle of place and route gets more complicated as the design gets closer to the area limit, the user must optimize the custom logic design for hardware resources to meet the timing enclosure.

Resource	**Utilization**	**Available**	**Utilization (%)**
LUT	102418	242400	42.25
LUTRAM	9431	112800	8.36
FF	123877	484800	25.55
BRAM	246.5	600	41.08
DSP	52	1920	2.71
IO	321	520	61.73
GT	4	20	20.00
BUFG	19	480	3.96
MMCM	4	10	40.00
PLL	6	20	30.00
PCIe	1	3	33.33

Table 5.5: HW Utilization (ESS FPGA Framework)

5.2.1 Algorithm (APTM)

As discussed in section 4.1, the algorithm used to detect errant beam in this case is simple. The algorithm uses a comparator logic to compare between two sets of value. The value can be read from/ written to the registers externally through direct memory access (DMA), hence the pre-defined threshold values are re-configurable. These threshold values are different for each channel/ APTM blades. Should the current value from the FMC increases beyond the pre-defined threshold, the algorithm assigns '1' to the *THRESHOLD_STATUS* register (as shown in figure 4.2).

Resource	**Utilization**	**Available**	**Utilization (%)**
LUT	121147	242400	49.98
LUTRAM	10204	112800	9.05
FF	141632	484800	29.21
BRAM	338.50	600	56.42
DSP	73	1920	3.80
IO	321	520	61.73
GT	4	20	20.00
BUFG	19	480	3.96
MMCM	4	10	40.00
PLL	6	20	30.00
PCIe	1	3	33.33

Table 5.6: HW Utilization (APTM)

This algorithm is already implemented in the *custom_logic* part of the *ess_fpga_framework*. As it can be seen in table 5.6, there is an overall increase in hardware utilization due to the initialization of separate registers and LUTs for the proposed comparator logic in the *custom_logic* region of the framework.

5.2.2 Algorithm I

The native algorithm (Algorithm I) to evaluate the peak current density of the beam implements a near straightforward approach to the multiplication of the vectors. For hardware implementation, each of the current vectors (*PH* and *PV*) were declared as one dimensional row/cloumn vectors of sizes 1×50 and 1×101 in the test bench respectively.
Since implementation of this algorithm, initialize three rolling windows (being evaluated every 1µS) requires a lot of computational power per 1µS window. The algorithm also needed to store four of the previously calculated matrices to make the 2µS, 5µS windows which required it to store $50 \times 101 \times 64$ bits per matrix. This sums up to be total of 157.8 kBytes of data which can not be stored in registers. Hence block RAMs (BRAMs) were

initialized to store these vectors (PH and PV) as well as the multiplication result D_w. The vertical input vector (PV) are unrolled with a factor of 10 to unroll loops to create multiple independent multipliers (DSP48E) for the vector multiplication. The algorithm synthesizes $(50 \times (100/10) + 1) \times 2 = 1002$ no. of multipliers (since each element of the matrix uses 2 multipliers for 64-bit multiplication) as shown in table 5.7.

Utilization Estimates (summary)					
Name	BRAM18K	DSP48E	FF	LUT	URAM
DSP	-	-	-	-	-
Expression	-	-	0	34	-
FIFO	0	-	35	272	-
Instance	57	1002	214491	176766	0
Memory	66	-	0	0	0
Multiplexer	-	-	-	54	-
Register	-	-	6	-	-
Total	123	1002	214532	177126	0
Available	1200	1920	484800	242400	0
Utilization (%)	10	52	44	73	0

Table 5.7: HW Utilization (Algorithm I)

In addition to this, a significant amount of flip flops (FF) and look-up tables (LUT) are synthesized due to declaration of temporary variables in the design. This custom logic paired with the fpga framework reaches the area limit for the FPGA. Such high hardware utilization is not favourable, manual modifications of the code did not help the HLS tool minimize the HW utilization as the complexity was too much. Hence, this algorithm is discarded for hardware implementation.

5.2.3 Algorithm II

This algorithm is a much simplified version of the native algorithm (Algorithm I). In this algorithm. accumulating buffers (Acc_X & Acc_Y) are used to minimize the hardware utilization as it replaces the need of storing intermediate values of D_w matrix. This accumulators also stores values for only 3 different time instances as it resets to zeros at every 5µS, and also it can be evaluated every 5µS.

Apart from this, to reduce the number of multipliers, the number of bits are reduced from 32-bit fixed point to 20-bit fixed point. The reduction of fixed point bits of the input current vectors results in a single multiplier for a single element multiplication between the input vectors ($PV(i) \times PH(j)$).

Unlike the native algorithm (Algorithm I), this algorithm did not store a huge amount of data to perform the peak current density evaluation. Hence in this case, the input current vectors (*PH* and *PV*) were also forced to generate registers (FF) for each elements instead of generating BRAMs. Along with this, to speed up the evaluation one of the accumulators (Acc_Y) were completely unrolled for the row process, which generated 101 no. of multiplier(DSP48E) instances (as shown in figure 4.6) to parallelize multiplication operation.

Utilization Estimates (summary)					
Name	BRAM18K	DSP48E	FF	LUT	URAM
DSP	-	-	-	-	-
Expression	-	-	0	4936	-
FIFO	0	-	-	-	-
Instance	0	101	27680	25808	-
Memory	-	-	-	-	-
Multiplexer	-	-	-	1360	-
Register	0	-	12266	128	-
Total	0	101	39946	32232	0
Available	1200	1920	484800	242400	0
Utilization (%)	0	5	8	13	0

Table 5.8: HW Utilization (Algorithm II)

This algorithm has a decent amount of hardware utilization which can be integrated to the fpga framework. The lower HW utilization also helps with easy development with the algorithm in future.

5.2.4 Algorithm III

In this algorithm, the hardware implementation part is almost as same as Algorithm II. But in this case three separate matrix multiplications are created for three separate windows (1μS, 2μS and 5μS respectively). For these three matrix multiplication instances, the row processing unit were completely unrolled with a factor of 10, which generated thrice as many multipliers (DSP48E) as Algorithm II.

The column process in this case is pipelined since there are only one instance generated for use.

Utilization Estimates (summary)					
Name	BRAM18K	DSP48E	FF	LUT	URAM
DSP	-	-	-	-	-
Expression	-	-	0	32	-
FIFO	0	-	-	-	-
Instance	0	303	34131	47361	-
Memory	-	-	-	-	-
Multiplexer	-	-	-	273	-
Register	0	-	10005	32	-
Total	0	303	44136	47698	0
Available	1200	1920	484800	242400	0
Utilization (%)	0	15	9	19	0

Table 5.9: HW Utilization (Algorithm III)

As it can be seen in table 5.9, the number of multiplier is 3 times as that of Algorithm II. The rest of the HW utilization increased obviously due to multiple instance generation in hardware level. The HW utilization definitely increases with the inclusion of the fpga framework and is greater than the HW utilization of Algorithm II, but the utilization is much less than Algorithm I.

The selection between Algorithm II and Algorithm III depends upon the latency in hardware level between the systems. This algorithm (Algorithm III) has the downside since it needs to transfer three times the amount of samples from the acquisition cards to the concentrator card then the Algorithm II. The user has a trade-off between hardware utilization and latency of the system (discussed in the next part of this chapter). If the latency of the system improves for Algorithm III than the Algorithm II on a level that oversees the increase in HW utilization of Algorithm III, then Algorithm III will be chosen; otherwise, Algorithm II will be the best option for every metrics of hardware design.

5.3 Timing

This master thesis deals with the detection of an errant beam within a specific time limit. The beam production must be suppressed before the next production of the beam pulse to save the target from an errant beam by the proposed design. Hence the algorithm to detect errant beam must have minimum latency (less than t_{pulse}). In timing, the latency of the design is defined as the time required for the design to generate a flag from an input. The latency of different algorithms varies due to different hardware implementation of the same logic.

5.3.1 Algorithm I

In Algorithm I, the latency is much high (as shown in table 5.10) than desired (which is around 1µS since the current values are being sampled at 1 MHz), considering the FPGA is running about at 100 MHz and the next beam pulse will be generated at $t < t_{pulse}$. This algorithm executes (50×101 operations) for each sample, every 1µS. In addition to this, it also executes 5050 additions for 2µS window and another 15150 ($50 \times 101 \times 3$) additions for the 5µS window. This, along with the memory access time to readout / write the values from/ to the memory (BRAM), incurs in high latency of the system.

Latency summary					
Latency (cycles)		Latency (absolute)		Interval (cycles)	
max	min	max	min	max	min
125818	125818	1.258 mS	1.258 mS	101012	101012

Table 5.10: Latency (Algorithm I)

Many manual modifications are made by re-organizing the algorithm, but the latency of the system could not be improved to a suitable limit (close to 1µS). Due to this high latency, this algorithm can not be selected for the evaluation of the peak current density and hence the detection of errant beam.

5.3.2 Algorithm II

The Algorithm II describes an efficient and much simpler way to evaluate the peak current density. The parallelization of multiplication units in row process speeds up the evaluation of peak current density in this case. Along with this, all of the input values as well as the intermediate values are stored in registers, hence reading out/ writing to data has almost no latency due to memory access.

Latency summary					
Latency (cycles)		Latency (absolute)		Interval (cycles)	
max	min	max	min	max	min
136	136	1.381 µS	1.381 µS	50	50

Table 5.11: Latency (Algorithm II)

The latency of this algorithm is much lower than the native algorithm (Algorithm I) and can be selected as the final algorithm due to its low hardware utilization for the evaluation of peak current density. However, the algorithm itself puts a limitation as it can be evaluated every 5µS.

5.3.3 Algorithm III

The Algorithm III implements a similar algorithm to evaluate the peak current density as Algorithm II, but this method also synthesizes thrice the number of multipliers than

Algorithm II. As it can be seen in the table 5.12, the latency of Algorithm III is almost similar to Algorithm II as both of the algorithm varies by a little amount. However, this algorithm has a slightly less latency due to the parallelization of multiplication in the row process for different windows. Whereas in Algorithm II, the same instance is shared over different windows.

Latency summary					
Latency (cycles)		Latency (absolute)		Interval (cycles)	
max	min	max	min	max	min
135	135	1.35 µS	1.35 µS	135	135

Table 5.12: Latency (Algorithm III)

This algorithm also reduces the latency but with a cost of hardware utilization. It also has an increased number of samples to be transferred before it can start the computation. The latency for the evaluation of peak current density in this case also didn't improve much as compared to the increase in HW resources. But this algorithm evaluates the peak current density for each window (1μS, 2μS and 5μS respectively) every 1µS. Hence, the Algorithm III is chosen over Algorithm II for the evaluation of peak current density on the hardware level.

5.3.4 Aurora 64/66B Core

This section discusses the proof of concept about the latency due to data transfer methodology between multiple AMC cards. To evaluate the peak current density of the beam, the 20-bit current data from the FMC pico are needed to be transferred from the Acquisition boards to the Concentrator board.
The section 4.3 discusses about the data transfer topology using P2P links, available in the micro-TCA crate. For the ease of implementation, Aurora IP is used to implement the high-speed serial data communication between boards. In this thesis, a loop test was performed to estimate the latency of the proposed system.

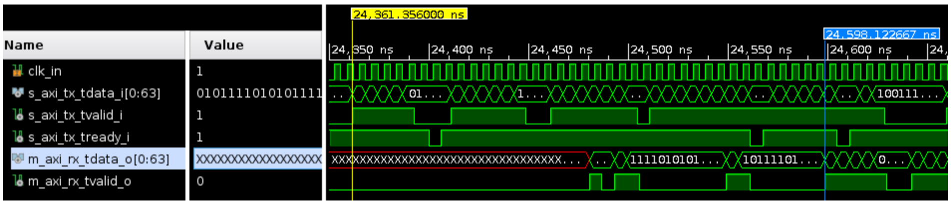


Figure 5.1: Latency (Aurora Core)

As it can be seen in figure 5.1, 64 -bit data is sent back-to-back to the TX port (*s_axi_tx_tdata* signal) of the Aurora IP from testbench and is read out at the RX port (*m_axi_rx_tdata* signal) of the IP. The streaming interface of Aurora IP core uses AXI4-Stream for data transmission. In this loop test, both of the *s_axi_tx_tvalid* and *s_axi_tx_tready* signals are asserted to transmit data. When the *s_axi_tx_tvalid* signal deasserted by the user through testbench no data is transferred through the IP core, as the core only samples data when both of *s_axi_tx_tvalid* and *s_axi_tx_tready* signals are asserted.

This loop test takes about 23 clock cycles to get the data flowing from TX port to RX port of the Aurora IP. Hence, to transmit the data from all the 8 channels 26 (23 + 3) clock cycles would be needed. This assumption is based on that the acquisition cards are running at 100 MHz, but the user can increase the frequency to reduce the latency a bit more.

5.3.5 Latency Estimation

In conclusion, the latency for the prototype project to evaluate the maximum current density of the beam would be about 161 clock cycles (26 clock cycles for data transfer + 135 clock cycles for the evaluation of peak current density) approximately. This is equivalent to 1.6µS (considering the FPGA is running at 100 MHz), which is an acceptable latency for the system in a real-time environment.

Chapter 6

Discussion

This master thesis project aimed to develop a hardware-based design of the signal processing for the beam on the target instruments, called APTM and Grid. A signal processing algorithm is developed in this thesis, characterizing the various beam properties from a set of vectors that are read out current values from the metallic plates/ wires due to the interaction between accelerated protons and metal surface. To achieve a neutron production safely, the beam delivery and target systems must function reliably within the nominal conditions. Any deviation in the beam delivery system will result in an errant beam situation. The proposed algorithm in this thesis also provides a robust characterization of the ESS beam on target and the detection of any errant condition.

However, a direct implementation of the signal processing algorithm was not possible on the hardware level. Certain timing constraints had to be met to detect the errant beam before the beam damages the target. Among the other beam properties, the maximum beam current density or peak current density is chosen to detect an errant beam where it is evaluated for three separate windows (1µS, 2µS and 5µS respectively) over time. Any deviation from the converging behaviour of the maximum current density for those separate windows points towards errant beam condition.

Three hardware optimised algorithms were explored, out of which one was selected to implement as the selected algorithm has shown promising results for the evaluation of peak current density. The selection of the algorithms were based on the accuracy, hardware utilization, and the latency of the proposed system. The proposed algorithm in this master's thesis also requires multiple AMC boards in a real-time environment. Hence, an acquisition-concentrator board connection model, along with the data transfer topology is proposed. The acquisition cards are used to sample data for each window simultaneously and send the values to the concentrator board, whereas the concentrator board computes the maximum current density once all the samples from each wire for each window have arrived. The evaluation of maximum current density at the concentrator board has an estimated latency of about 1.6 µS, which is promising for the final system setup fulfilling the requirements.

6.1 Future Work

This master thesis proposes the algorithm to be implemented on the hardware level (Algorithm III). This gives us an idea about the next part of the Grid system, which can be pursued in the future. The data transfers between two AMC cards needed to be tested in a real-time environment before the implementation of the algorithm on hardware as the algorithm (Algorithm III) requires to implement the 1μS, 2μS, and 5μS averaging window on the acquisition cards and the algorithm to evaluate peak current density on the concentrator card. It is also necessary to check the latency in real hardware from input to physical output.

The final prototype of the system can be expanded to the final hardware, where the data is shared between two or more 12-slots crates. An investigation needed to be done for the data transfer configuration that can result in minimum latency. The final implementation of the full-scale electronics configuration (with at least 50 and 100 wires on horizontal and vertical orientation respectively) will be deployed in real-time at ESS.

References

[1] List of Authors: https://docdb01.esss.lu.se/cgi-bin/public/DocDB/ListAuthors. *ESS Technical Design Report*. ESS ERIC, Lund, SE, 2013. http://eval.esss.lu.se/cgi-bin/public/DocDB/ShowDocument?docid=27.

[2] Schematic of beam instrumentation in lebt. "https://confluence.esss.lu.se/display/ALSV/next+LEBT".

[3] Cyrille Thomas. Beam on target instrumentation system (presentation).

[4] Anders J Johansson C. Thomas, E. Donegani and Mattias Wilborgsson. Preliminary Design for the APTM and GRID. (ESS-0492447):30, 2019.

[5] C.A. Thomas. Beam on target instrumentation model. https://gitlab.esss.lu.se/beam-diagnostics/beam_on_target_instrumentation_model.

[6] C. Amstutz, M. Donna, A. J. Johansson, and M. Mohammednezhad. The ess fpga framework and its application on the ess llrf system, 2018.

[7] nvent Schroff. *User Manual MTCA.4 Cube*, Jan 2016. "https://schroff.nvent.com/wcsstore/ExtendedSitesCatalogAssetStore/Attachment/SchroffAttachments/Documents/63972-347.pdf".

[8] nvent Schroff. *User Manual MTCA.4 Shelf*, Jan 2016. "https://schroff.nvent.com/wcsstore/ExtendedSitesCatalogAssetStore/Attachment/SchroffAttachments/Documents/63972-328.pdf".

[9] Kristoffer Sjögreen. Beam on target requirements, apr 2020.

[10] E. J. Sternglass. Theory of secondary electron emission by high-speed ions. *Physical Review*, 108(1):1–12, 1957.

[11] CAEnels. Fmc-pico-1m4 4-channel 20-bit 1 msps fmc floating ammeter. "https://www.caenels.com/products/fmc-pico-1m4/".

[12] Struck Innovative Systeme. Sis8160 dual hpc fmc carrier amc for digitizer fmcs. "https://struck.de/sis8160.html".

[13] EPICS. Experimental physics and industrial control system. https://epics-controls.org/.

[14] Beckhoff.com. *BECKHOFF New Automation Technology*, 2020. "https://www.beckhoff.com/EL3314-0010/".

[15] ARM Limited. *ARM Developer Suite AXD and armsd Debuggers Guide*, 2006 [1999]. https://developer.arm.com/documentation/dui0066/d/axd/axd-facilities/data-formatting/q-format.

[16] W. Blokland. High Beam Intensity Harp Studies and Developments at SNS. In *Proc. 6th International Particle Accelerator Conference (IPAC'15), Richmond, VA, USA, May 3-8, 2015*, number 6 in International Particle Accelerator Conference, pages 17–20, Geneva, Switzerland, June 2015. JACoW. https://doi.org/10.18429/JACoW-IPAC2015-MOAB1.

[17] T. J. Shea W. Blokland, T. McManamy. Sns target imaging system software and analysis. In *Proc. of the 14th Beam Instrumentation Workshop (BIW 10), Santa Fe, New Mexico, May 2-6 2010*, pages 93–95.

[18] Kiyomi Ikezaki Tomoyuki Kawasaki Hidetaka Kinoshita Atsushi Akutsu Masaaki Nishikawa Shin-ichiro Meigo, Motoki Ooi and Shinpei Fukuta. Beam profile monitor at the 1 mw spallation neutron source. *IBIC*, pages 1–7, 2015.

[19] J.L. Sirvent et al. Performance Assessment of Pre-Series Fast Beam Wire Scanner Prototypes for the Upgrade of the CERN LHC Injector Complex. In *Proc. of International Beam Instrumentation Conference (IBIC'17), Grand Rapids, MI, USA, 20-24 August 2017*, number 6 in International Beam Instrumentation Conference, pages 338–341, Geneva, Switzerland, Mar. 2018. JACoW. https://doi.org/10.18429/JACoW-IBIC2017-WEPCC03.

[20] Heine Thomsen, Anne Ivalu Holm, and Søren Møller. A Linear Beam Raster System for the European Spallation Source? In *4th International Particle Accelerator Conference*, page MOPEA005, 2013.

Appendix A

Code

```
 %% part of the Raster beam code to extract the beam
     properties from the current signals from grid wires

%%
close all;
%%
nC = 1e-9 ;
PH = aa{2} ;  %current from the horizontal grid wires

[nyh nxh] = size(PH)

xh = 1:nxh ;

figure(20)

plot(xh,sum(PH,1)*dt*1e-3/nC,'ks--',  ...
     'MarkerSize',12,'LineWidth',2)

xlabel('Hor. position (mm)','FontSize',fts)
ylabel('charge / pulse (nC)','FontSize',fts)

set(gca,'FontSize',fts)

set(gcf,'PaperSize',[ww0 hh0],'PaperPosition',[0 0 ww0 hh0
   ],'Units','centimeters')
fnam = [prefix 'Vert-Grid_beam_raster' suffix]
saveas(gcf,fnam,'png')

%%
PV = aa{3} ;  %current from the vertical grid wires

[nyv nxv] = size(PV)

xv = 1:nxv ;
```

```

figure(30)

plot(xv,sum(PV,1)*dt*1e-3/nC,'ks--', ...
     'MarkerSize',12,'LineWidth',2)

xlabel('Vert. position (mm)','FontSize',fts)
ylabel('charge / pulse (nC)','FontSize',fts)

set(gca,'FontSize',fts)

set(gcf,'PaperSize',[ww0 hh0],'PaperPosition',[0 0 ww0 hh0],'Units','centimeters')
fnam = [prefix 'Hor-Grid_beam_raster' suffix]
saveas(gcf,fnam,'png')

%%
id = [10 25];

figure(50)
clf
subplot(3,1,1)
plot(t,ax,...
     t,ay,...
 'MarkerSize',12,'LineWidth',3)
xlabel('time (ms)','FontSize',fts-6)
ylabel('Raster Amplitude (mm)','FontSize',fts-6)
set(gca,'FontSize',fts-6)

subplot(3,1,2)
plot(t,(PH(:,id(1))/muA),'-' ,...
 'MarkerSize',12,'LineWidth',3)
xlabel('time (ms)','FontSize',fts-6)
ylabel('Peak Current (\muA)','FontSize',fts-6)
legend(['Horizontal wire #' num2str(id(1))])
set(gca,'FontSize',fts-6)

subplot(3,1,3)
plot(t,(PH(:,id(2))/muA),'-' ,...
 'MarkerSize',12,'LineWidth',3)
xlabel('time (ms)','FontSize',fts-6)
ylabel('Peak Current (\muA)','FontSize',fts-6)
legend(['Horizontal wire #' num2str(id(2))])
set(gca,'FontSize',fts-6)

set(gcf,'PaperSize',2*[ww0 hh0],'PaperPosition',[0 0 2*ww0 2*hh0],'Units','centimeters')
```

```

fnam = [prefix 'Raw-Hor-Grid_beam_raster' suffix]
saveas(gcf,fnam,'png')
%%

f0 =1/mean(diff(t*1e-3)) ;
kHz = 1e3

nu = (1:length(t))/length(t) * f0 ;
nuc = ((1:length(t))/(length(t)) -0.5) * f0 ;

xV = 1:2:100;

FFTPV = (fft(PV-mean(PV(:)),[],1)) ;


figure(1)
imagesc(nuc/kHz,xV, (abs(fftshift(FFTPV.',2))))

xlim([0 100])

%%
xV = 1:2:200;

FFTPH = (fft(PH-mean(PH(:)),[],1)) ;


figure(10)
imagesc(nuc/kHz,xV, (abs(fftshift(FFTPH.',2))))

xlim([0 100])

%%
figure(1)
plot(t,PV(:,25))

%%
Y = PV(:,25) / max(PV(:,25)) ;
[AA,BB] = peakdet(Y,0.5)

(mean(diff(AA(:,1)))/kHz)



%%
figure(1)
hold all
```

```
plot(nu/kHz,unwrap(angle(fftshift(fft(PV(:,30))))))
grid on
hold off
% xlim([0 50])

%%
id1 = [20 45];

figure(51)
clf
subplot(3,1,1)
plot(t,ax,...
      t,ay,...
 'MarkerSize',12,'LineWidth',3)
xlabel('time (ms)','FontSize',fts)
ylabel('Raster Amplitude (mm)','FontSize',fts-6)

legend('Hor.','Vert')
set(gca,'FontSize',fts-6)


subplot(3,1,2)
plot(t,(PV(:,id1(1))/muA),'-',...
 'MarkerSize',12,'LineWidth',3)
xlabel('time (ms)','FontSize',fts-6)
ylabel('Peak Current (\muA)','FontSize',fts-6)
legend(['Vertical wire #' num2str(id1(1))])
set(gca,'FontSize',fts-6)

subplot(3,1,3)
plot(t,(PV(:,id(2))/muA),'-',...
 'MarkerSize',12,'LineWidth',3)
xlabel('time (ms)','FontSize',fts-6)
ylabel('Peak Current (\muA)','FontSize',fts-6)

legend(['Vertical wire #' num2str(id1(2))])

set(gca,'FontSize',fts-6)

set(gcf,'PaperSize',2*[ww0 hh0],'PaperPosition',[0 0 2*ww0 2*hh0],'Units','centimeters')

fnam = [prefix 'Raw-Vert-Grid_beam_raster' suffix]
saveas(gcf,fnam,'png')


%%
```

```

idxx =1:33 ;
m2 = zeros(1,2500) ;
MY = zeros(1,2500) ;

vx = 2 * ax0 * Fx * dt
%%
wid = 51;

for jj = 1:2500

    xxx = jj+idxx ;
    yyy=PV(xxx,wid).';

%       figure(20)
%       plot(idxx,yyy)
%       drawnow
%
    mmm = Moments(xxx,yyy,3);
    m2(jj)=mmm{3} ;

    MY(jj) = sum(yyy) ./ mmm{3} / vx ;

end
%
%%
%

Dw = zeros(10, 50,101) ;
dxdy = dx * dy ;

for jj=45:2991

    Dw(1,:,:) = PV(jj,:) .* PH(jj,:).' ;
    Dw(2,:,:) = PV(jj+1,:) .* PH(jj+1,:).' ;
    Dw(3,:,:) = PV(jj+2,:) .* PH(jj+2,:).' ;
    Dw(4,:,:) = PV(jj+3,:) .* PH(jj+3,:).' ;
    Dw(5,:,:) = PV(jj+4,:) .* PH(jj+4,:).' ;
    Dw(6,:,:) = PV(jj+5,:) .* PH(jj+5,:).' ;
    Dw(7,:,:) = PV(jj+6,:) .* PH(jj+6,:).' ;
    Dw(8,:,:) = PV(jj+7,:) .* PH(jj+7,:).' ;
    Dw(9,:,:) = PV(jj+8,:) .* PH(jj+8,:).' ;
    Dw(10,:,:) = PV(jj+9,:) .* PH(jj+9,:).' ;

    Dw00 = squeeze(sum(Dw(1,:,:),1) ) ;
    Dw1_m = Dw00 ;
    Dw1_m = Dw1_m ./ sum(Dw1_m(:)) * I_p_peak / dxdy / 4 ;

    Dw00 = squeeze(sum(Dw(1:2,:,:),1) );
    Dw2_m = Dw00 ;
    Dw2_m = Dw2_m ./ sum(Dw2_m(:)) * I_p_peak / dxdy / 4  ;

    Dw00 = squeeze(sum(Dw(1:5,:,:),1) );
    Dw5_m = Dw00 ;
    Dw5_m = Dw5_m ./ sum(Dw5_m(:)) * I_p_peak / dxdy / 4  ;

    trend3 = 1 - max(Dw2_m(:))./max(Dw1_m(:)) ;
    trend5 = 1 - max(Dw5_m(:))./max(Dw1_m(:)) ;

      figure(1)
      subplot(2,2,1)
      imagesc(Dw1_m)
      title(jj)
      colorbar
      drawnow
      subplot(2,2,2)
      imagesc(Dw3_m)
      colorbar
      drawnow
      subplot(2,2,3)
      imagesc(Dw5_m)
      colorbar
      drawnow
      subplot(2,2,4)
      bar([3 5],[trend3 trend5])
      ylabel('max(I1) - max(Ii)')

      drawnow





end
%%
figure(1)
clf
subplot(2,1,1)
plot(1:2500,m2 * vx,'k-','LineWidth',2)
xlabel('time (\mus)')
ylabel('r.m.s width')
set(gca,'FontSize',22)
subplot(2,1,2)
plot(1:2500,MY/mA/A0 * 2,'r','LineWidth',2)
```

```
xlabel('time (\mus)')
ylabel('Current Density (mA/cm^2) ')
set(gca,'FontSize',22)

drawnow

%%
ms = 1e-3;
% jj = 100

for jj = 1:50
     Dw = PV(jj,:) .* PH(jj,:).'   ;
end
     Dw = Dw ./ sum(Dw(:)) * I_p_peak    ;

im100 =  aa{4}   ;

gg = fit_2d_gaussian(im100) ;

ggw = fit_2d_gaussian(Dw)   ;

wspace = 2
sigma_xw = ggw(5) * wspace
sigma_yw = ggw(6) * wspace

sigma_xi = gg(5)
sigma_yi = gg(6)

rho_im = max(im100(:) * qe / dt/ms /dx/dy) /ms
rho_w = max(Dw(:) /4 /dx/dy) /ms


ttw = {['total current from wires image: ' num2str(sum(Dw(:))/ms) ' mA'] ...
         ['\sigma_x = ' num2str(sigma_xw,3)  ' mm -  \sigma_y = ' num2str(sigma_yw,3) ' mm']  ...
         ['max density : \rho = ' num2str(rho_w,3) ' mA/cm^2'] ...
         };

tti = {['total current from beamlet image: ' num2str(sum(im100(:)) * qe  / dt / ms /ms  ,3) ' mA'] ...
         ['\sigma_x = ' num2str(sigma_xi,3) ' mm -  \sigma_y = ' num2str(sigma_yi,3) ' mm' ] ...
         ['max density : \rho = ' num2str(rho_im,3) ' mA/cm^2'] ...
         };
```

```


figure(110)
imagesc(Dw/dx/dy/4)
colorbar
title(ttw)


figure(111)
imagesc(im100 * qe / dt/ms /dx/dy )
colorbar
title(tti)
```