Characterization of the Front-End Electronics for the Pixel-Strip Module of the Phase 2 CMS Tracker

Christian Leefmans
School of Applied and Engineering Physics
Cornell University, Ithaca, NY 14853

In the mid-2020s, the Large Hadron Collider (LHC) will be upgraded to the High-Luminosity LHC (HL-LHC). The HL-LHC will generate up to 3000 fb^{-1} by 2035 and produce a maximum instantaneous luminosity of $5-7\times10^{34}\ cm^{-2}s^{-1}$. To cope with the high luminosity and radiation present in the HL-LHC environment, the CMS Collaboration will upgrade its tracking detector. To deal with an unprecedented number of events per bunch crossing (pileup), CMS will produce input to the L1 trigger on the tracker for the first time with innovative pixel-strip (PS) and strip-strip (2S) modules. This paper reviews the prototype of the pixel electronics for the PS module, the Macro Pixel ASIC-Light (MPA-Light), before introducing measurements of the device's timewalk and peaking time.

I. THE STANDARD MODEL

For the past half century, the Standard Model has been the cornerstone of particle physics research, transforming our understanding of matter and its interactions. With the aid of particle accelerators such as the Tevatron and the LHC, physicists have discovered a set of 17 fundamental particles consistent with the Standard Model, which constitute all known visible matter, mediate the electromagnetic, strong, and weak interactions, and enable the mechanism by which particles acquire mass. All of these particles can be broadly classified as either fermions or bosons [12],[8].

Fermions are particles that obey the Pauli exclusion principle, and they are characterized by one-half integer spins [9]. Under this brand, the fundamental fermions are divided into six quarks and six leptons, which are further divided into three generations. All everyday matter is composed of Generation I quarks − the up and the down − and the electron, which is a Generation I lepton.

Quarks are unique in that they experience the strong interaction, mediated by the force carrying boson, the gluon [12]. Unlike gravity or the electromagnetic interaction, the strong interaction becomes greater as the distance between quarks increases, and at a certain distance the force between the quarks becomes so strong that it is energetically favorable to generate new quarks, each bound to a member of the newly separated pair. This is the phenomenon of quark confinement − quarks are only found in multi-quark particles, called hadrons. The only known exception is the massive top quark, but the top quark is so unstable that it decays almost immediately.

In contrast, leptons − the electron, muon, tau, and their associated neutrinos − only exist individually. The muon and tau are like heavier cousins of the electron, and, like the quarks as well, the charged leptons exhibit a pattern of becoming more massive with each successive generation. On the other hand, the neutrino mass hierarchy is still being measured. The neutrino mass is so small

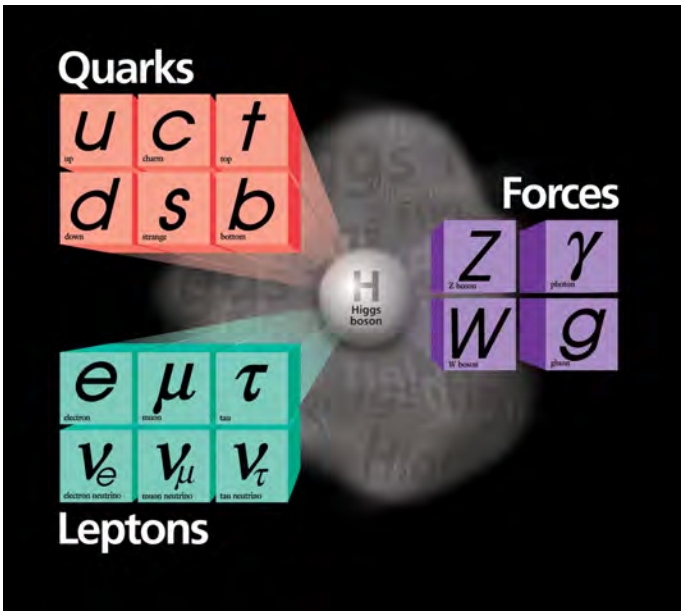

FIG. 1: The Standard Model: The current Standard Model includes 17 particles. There are six quarks, six leptons, and four force carrying particles. The Higgs Boson, discovered in 2012, is associated with the mechanism by which particles acquire mass.

that these particles were once thought to be massless, but the discovery of neutrino oscillations − neutrinos changing from one type to another − indicated that neutrinos indeed have mass [11]. The mysterious nature of neutrinos is part of what makes them such an active area of research.

In the present Standard Model, four of the bosons − the gauge bosons − are associated with three of the fundamental interactions. Besides the gluon, associated with the strong interaction, is the photon, arbiter of the electromagnetic interaction, and the W and the Z bosons, which together mediate the weak interaction. Unlike

fermions, bosons do not obey the Pauli exclusion principle, and they possess whole-integer spins [9].

The culminating achievement of the current Standard Model, however, is the Higgs boson, discovered in 2012. The Higgs is the only boson associated with a scalar field, and through the Higgs mechanism, it provides an explanation for how particles acquire mass [9].

Despite the progress accompanying the discovery of the Higgs, many exciting questions in particle physics remain unanswered, and for this reason physicists continue to build new and better particle accelerators, to generate higher energy collisions and more interesting events. With more powerful accelerators, physicists hope to shed light on dark matter, prove or disprove the existence of supersymmetry, continue the search for a graviton, and answer the question of whether neutrinos are their own antiparticle. With the inception of the High Luminosity LHC (HL-LHC) in the mid-2020s, physicists hope that the unprecedented luminosity will provide answers to these questions and many more.

II. THE LHC AND HL-LHC

As a proton-proton collider, the LHC merges complexity with precision. Throughout its 27 km circumference, the LHC operates in an incredibly well-controlled manner, maintaining the orbit of the proton beams while consistently producing high energy events in the LHCs various detectors [6]. At the heart of the LHCs operation is an intricate system of radiofrequency (RF) cavities, magnets, cryogenic systems, and vacuums that accelerate the protons to the correct energy and maintain the beam in its curved trajectory.

The oscillating electric fields in the RF cavities serve two primary functions. Firstly, the cavities accelerate the proton beams. Currently each beam achieves an energy of 6.5 TeV in order to produce 13 TeV head-on collisions. Secondly, the electric fields help to separate the particle beams into bunches, each consisting of about 10^{11} protons, by slowing down particles traveling too quickly and speeding up others traveling too slowly. The 400 MHz RF cavities separate the bunches with a 25 ns periodicity [10].

The $\sim$ 9300 magnets in the LHC are tasked with everything from making gross and fine adjustments to the beam path to compressing and manipulating the beam in preparation for collisions [2],[1]. The LHC currently uses an 8.33 T magnetic field in order to guide the beam around the accelerator [10].

In order for the RF cavities and superconducting dipole magnets of the LHC to function properly, the accelerator needs massive and sophisticated cryogenic and vacuum systems. To keep the dipole magnets at a superconducting temperature of 1.9K, the magnets are immersed in superfluid helium, and around these magnets a vacuum system acts as an insulator, preventing outside heat from reaching the magnets by conduction. The RF cavities are also held at superconducting temperatures in order to minimize power losses and heat dissipation, and inside the beam pipe another vacuum system creates a pressure as low as 10^{-10}-10^{-11} mbar. This is meant to minimize unwanted interactions with extraneous particles in the beam pipe [10],[18].

To date the LHC has generated $\sim$20 fb^{-1} of 13 TeV collisions, with an instantaneous luminosity reaching 1.2×10^{34} $cm^{-2}s^{-1}$. By the end of its run in 2023, it is expected to achieve 300 fb^{-1} while increasing the center of mass energy of its collisions to 14 TeV. The HL-LHC will operate under similar principles as the LHC, but the higher luminosity will require revamping the current system. By the end of its run in 2035, the HL-LHC is projected to produce an additional 2700 fb^{-1} of 14 TeV collisions, while increasing the maximum instantaneous luminosity to 5-7$\times 10^{34}$ $cm^{-2}s^{-1}$. Such high luminosity necessitates new magnets and RF cavities and upgrades to both the cryogenic and vacuum systems. For example, keeping the beams in their paths will require magnets capable of producing magnetic fields between 11 and 12 T [14].

In addition, the increased luminosity will increase the mean number of collisions per bunch crossing (pileup) from 23 to between 140 and 200 [15]. This unprecedented number of events requires the experiments in the LHC to update their detectors so as not to lose signs of new and interesting physics. For this reason, CMS is preparing the Phase 2 upgrade to its detector to occur in tandem with the construction of the HL-LHC.

III. CMS

At a basic level, CMS is based on the Lorentz force. The 3.8 T superconducting solenoid of the CMS detector bends particles as they pass through the detector. In addition to the solenoid, CMS consists of four layers of sub-detectors − the muon detector, hadron calorimeter (HCAL), electromagnetic calorimeter (ECAL), and tracking detector. The muon detector and outer barrel of the HCAL sit outside the solenoid, while the rest of the detector sits within [13].

A. The Muon Detector

As its name implies, the muon detector is exclusively meant to detect muons. Unlike many other particles, muons are able to pass through the rest of the subdetectors, enabling physicists to conclude with high probability that most hits in the muon detector result from muons. The muon detector consists of three different types of detectors − resistive plate chambers (RPC), cathode strip chambers (CSC), and drift tubes (DT) − interleaved by the three layers of the solenoids iron return yoke. The CSCs and DTs are found in the barrel region and endcap of the muon detector, respectively,

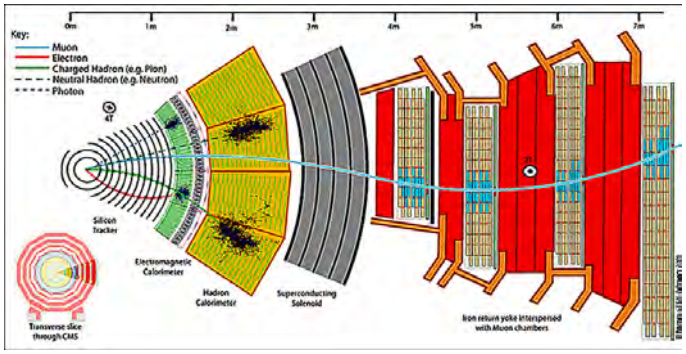

FIG. 2: The CMS Detector: The CMS detector can be broken down into five layers: the muon detector, HCAL, solenoid, ECAL, and tracking detector. Each sub-detector plays an important role in reconstructing the energy and momenta of particles from an event.

and the RPCs are found in both. Because of the magnetic field′s return path through an iron return yoke, the muons will bend oppositely in the muon detector than they did within the solenoid, and measuring the curvature of their trajectory through the muon detector provides a second measurement of the muons momentum − the first measurement coming in the tracking detector [13].

B. The Hadron Calorimeter

The HCAL is optimized to measure the energy of incident hadrons. It does this by utilizing a brass absorber, a plastic scintillator, and the interactions between the hadrons and the nuclei of the absorber. The absorber and scintillator are arranged in layers. When the hadrons interact with the nuclei of the absorbers, they create other particles, which excite electrons in the tiles of scintillator. When these electrons fall back to their original energy levels they emit light, which is then read out via optical fibers. The secondary particles pass on to another layer of absorber, where they can interact with the absorber to produce more particles. The repetition of this process gives rise to HCAL showers, and by taking the optical sum of the light generated by a single shower, one can measure the energy of the incident hadron. The HCAL is built with layers inside and outside the solenoid to ensure that no energy from the HCAL showers is lost [13].

C. The Electromagnetic Calorimeter

The ECAL, lying directly within the HCAL, also relies on a scintillator − lead tungstate − to measure the energy of charged particles, primarily photons and electrons. In measuring these particles, the ECAL relies on the principle of secondary emission. When the photons or electrons strike the scintillator nuclei, they generate photons, electrons, and positrons, which in turn can generate more photons, electrons, and positrons, resulting in ECAL showers. Just as with the scintillator in the HCAL, when excited electrons in the lead tungstate crystals return to their original state, they emit a photon which is detected and recorded. In addition to these crystals, the ECAL also consists of the ECAL preshower in the endcap regions, which uses a combination of lead and silicon detectors to distinguish low energy photon pairs − resulting from pion decays −from single high energy photons [13].

Inside the ECAL is the tracking detector, the innermost layer of the CMS detector. This layer consists entirely of silicon detectors, which measure the trajectories of charged particles emerging from the beam pipe. As the results presented in this paper are primarily concerned with the HL-LHC upgrade to the outer tracker, a separate section will be dedicated to the current CMS tracking detector and the plans to improve it.

IV. TRACKING DETECTOR

The primary purpose of the tracking detector is to measure the charge and trajectory − and consequently momentum − of charged particles as they bend in the 3.8 T CMS magnetic field. To accomplish this, the CMS tracking detector is built from multiple layers of silicon detectors, which register hits as charged particles pass through. Reconstruction algorithms then use the readout from these detectors to retrace particles′ paths through the tracker and to identify the primary and secondary vertices of events. An overview of reconstruction algorithms used to generate tracks is available in [5].

The current CMS tracking detector contains a pixel tracker and a strip tracker. There are a total of 66 million pixels (1440 modules), covering an area of ~ 1 m^2 and 9.3 million strips (15148 modules), covering an area of ~ 198 m^2. Each tracker consists of multiple layers of detectors [5].

As can be seen from Figure 3, the pixel and strip trackers are arranged in barrel regions and endcap regions. All pixels, whether in the barrel or the endcap are 100x150um, and in the barrel region the 150um side lies parallel to the z-axis − along the beam line. In contrast, the sizes and shapes of the strips varies throughout the strip layers, with the pitch varying as well [5].

The pixel tracker, or vertex detector, contains three layers in the barrel region and two endcap layers on either side of the barrel region. To permit 3D measurements of particle position, all pixel layers consist of pixel modules arranged back-to-back as so-called stereo modules [5]. The hits from these stereo modules are correlated to provide improved resolution in the z-direction in the barrel region and the r-direction in the endcap region.

Improved resolution in the z-direction is especially important in the vertex detector, as it permits secondary vertex localization of particle decay events.

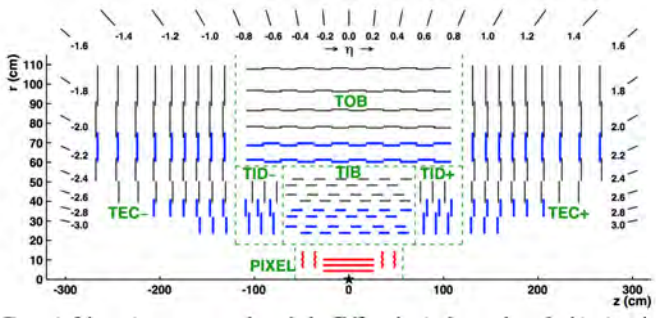

FIG. 3: The Tracking Detector: The current CMS tracking detector is separated into a pixel tracker and a strip tracker. These trackers are organized into barrel and endcap regions, with the thick, colored lines representing stereo modules, and the thin, black lines representing single-layered modules.

The strip tracker is subdivided into four different segments: the tracker inner barrel (TIB), tracker inner disks (TID), tracker outer barrel (TOB), and tracker endcaps (TEC). Together the TIB and TOB form 10 layers of strip modules in the barrel region, the TID adds three layers of modules to each endcap, and the TEC contributes a nine more layers to each endcap.

In the strip tracker, stereo strip modules are present in the inner two layers of the TIB and TOB, the inner two rings of the TID and TEC. The fifth innermost ring of the TEC also has stereo modules. All other layers and rings in the strip tracker consist of single strip modules [5].

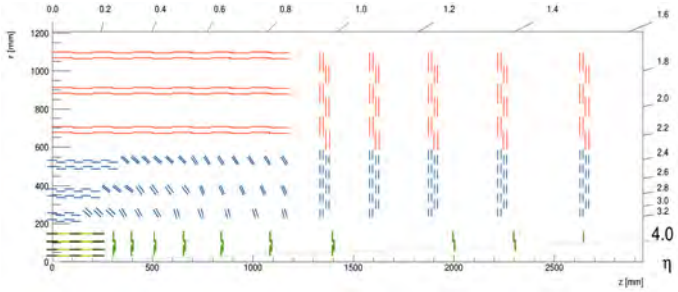

FIG. 4: The Phase 2 Tracking Detector: The Phase 2 tracking detector will include several innovations. In addition to extending the pixel detector farther into the end-cap region, the outer tracker will consist of new PS and 2S modules, which will provide input to the L1 trigger.

In the Phase 2 upgrade, the CMS tracker will receive major modifications in order to handle the increased luminosity and increased pileup of the HL-LHC. In the HL-LHC, each event will produce ∼200,000 hits in the tracking detector. In order to correctly identify high p_T tracks, this volume of data requires CMS to use the tracking information as an input to the detector's L1 trigger – rather than just muon detector and calorimeter information as in the present detector. Furthermore, the Phase 2 detector will reduce multiple scattering and be radiation harder. To decrease multiple scattering, the tracker upgrade will retain the evaporative CO_2 cooling system introduced in the Phase 1 upgrade while decreasing the number of cables by placing DC-CD converters within the tracking detector. Both the Phase 2 sensors and electronics will be radiation hard enough to withstand the projected pileup of 140-200. [15],[7].

In order to produce trigger input on the tracking detector, CMS will introduce special modules in the outer tracker. A layout of the Phase 2 upgrade to the CMS tracker is shown in Figure 4. The vertex detector will be constructed of single layers high granularity pixel detectors – the pixel detectors will extend farther into the end-cap regions, covering an angle up to 4 η – while the outer tracker will consist of two types of "p_T modules": pixel-strip (PS) and strip-strip (2S) modules. The PS modules will comprise a strip sensor mounted behind a pixel sensors, while the 2S modules will include two strip sensors, one mounted behind the other. In each case, the modules will be separated by only a few mm [15].

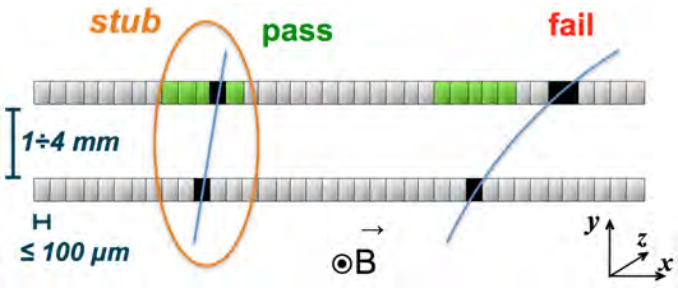

FIG. 5: Stub Finding: The outer tracker will produce information for the L1 trigger via a process called stub finding. This process involves correlating a hit on the front sensor of a p_T module with a hit in a narrow region on the sensor behind it.

The PS and 2S modules will provide triggering information to the back-end electronics through an innovative process called stub finding, which is illustrated in Figure 5. As high-momentum particles bend less in the CMS magnetic field, these modules will identify potential high-p_T candidates by correlating the locations of hits on the front and back modules. Upon receiving a hit on the front sensor, the module′s electronics will search a narrow region of strips – optimized to select 2-3 GeV tracks – on the back module. If the electronics register a hit on a strip within this window, the two hits will be correlated into a stub, or a potential high-p_T candidate. Hits arriving outside this window will also be saved, but they will not generate stubs for triggering [4]. Each bunch crossing (40 MHz), stubs will be sent to the track trigger, which will quickly use the stubs to find tracks. The tracks from the track trigger will then be used by the L1 trigger.

The information presented in this paper is concerned primarily with the development of the PS modules, and

in particular the first prototype of the pixel module's ASIC. Before introducing this technology, however, I will briefly review basic principles of silicon detectors.

V. SILICON DETECTORS

A. The Silicon Sensor

In a simplified description, the silicon sensor can be treated as a single, reverse-biased diode [3]. A diode can be formed in multiple ways, but in general it consists of adjacent regions of n- and p-type silicon. n-type silicon is formed by doping silicon with pentavalent donor impurities, while p-type silicon is doped with trivalent acceptor impurities. The device discussed in this paper uses n^+-on-p pixels [4], but it should be noted that in the current tracking detector consists of n^+-on-n pixels and p^+-on-n strips [19]. Doping in these manners introduces donor and acceptor energy levels near the conduction and valence bands of the silicon, which greatly increase the conductivity of the silicon by increasing the number of mobile electrons or holes [17].

Currents in semiconductors result from two different mechanisms: drift and diffusion. The drift current consists in the motion of electrons and holes in an electric field, while the diffusion current consists in the aggregate motion of these carriers due to concentration gradients. When n- and p-type silicon are made form a junction, an energy barrier arises between the valence and conduction bands of the two regions, such that the equilibrium Fermi level is constant throughout the silicon. This energy barrier equivalently manifests as a voltage difference between the two regions, and the region where the electric field accompanying this voltage difference appears is known as the depletion region. As can be seen in Figure 6, in a reverse-biased p-n junction, an external bias increases the energy barrier between the n- and p-type material, widening the depletion region. This larger energy barrier tends to prevent the diffusion of carriers between the two regions, and therefore the diffusion current in a reverse-biased diode is typically quite small [17].

While the diffusion current in a reverse-bias diode is negligible, a reverse-bias saturation current arises in reverse-bias diodes due to a combination of the drift and diffusion mechanisms. Minority carriers located within a diffusion length of the depletion region can diffuse into that regions electric field, causing the carriers to drift to the other side. Furthermore, carriers generated within the depletion region are carried away by the electric field before they have a chance to recombine [17].

As it turns out, because of the relative heights of the energy barrier, it is easier to thermally generate electron-hole pairs in the depletion region than in the surrounding material. Since the pairs are separated before recombination can occur, it happens that as the depletion region widens, the reverse-bias saturation current increases. This continues until the silicon is fully depleted

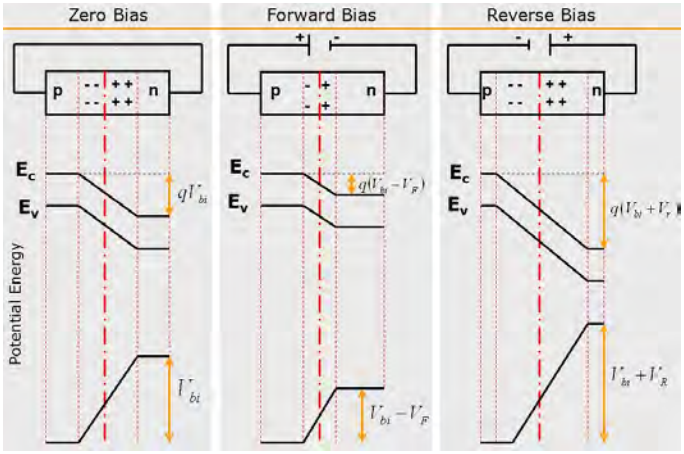

FIG. 6: The Energy Barrier in p-n Junctions: Application of bias to a p-n junction changes the energy barrier between the two regions. Forward bias lowers the barrier, facilitating diffusion, while reverse bias raises the barrier, inhibiting diffusion.

– at which point the reverse-bias saturation current remains constant.

Despite the increased reverse-bias saturation current resulting from a fully depleted p-n junction, silicon sensors must be fully depleted. When a charged particle passes through a silicon sensor, it generates electron-hole pairs. These carriers then drift toward the the electrodes of the device, and the magnitude of the charge measured is proportional to the energy of the incident particle.

In the case of a partially depleted sensor, some of the electron hole pairs generated by the passing particle do not necessarily reach the electrode, and therefore part of the signal is lost. Therefore, in order to properly detect particles, the sensor must be fully depleted [3].

B. The Analog Front-End

The silicon sensor is the first element of the analog front end (FE) of a silicon detector. In a typical analog FE, the signal from the silicon sensor generally passes through a preamplifier, a pulse shaper, and finally a ADC or discriminator before being transmitted to the next layer of electronics. A high-level schematic of an analog FE is shown in Figure 7. In order to receive a measure of the charge from the incident particle, the resulting current pulse must be integrated, and in principle this can be done at either the sensor, preamplifier, or ADC stage of the FE [16].

Each stage of the analog FE contributes to produce a useful, storable electronic signal. Because the signals from the silicon sensors quite small, the preamplifier amplifies this signal before it reaches the shaper. The shaper then has the job of making the signal wide enough for the ADC or discriminator while keeping it narrow enough so that two signals do not overlap – a phenomenon known

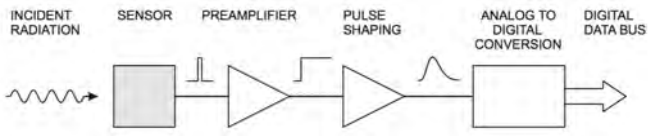

FIG. 7: The Analog FE: This high-level schematic shows a general outline of the typical analog FE. A signal is received by a sensor and passes through a preamplifier, shaper, and ADC. In the MPA-Light, the device used for the measurements in this paper, the ADC is replaced by a discriminator; however, this discriminator achieves the same effect of processing the analog signal for the digital electronics.

as pileup. The ADC or discriminator then produces a signal that can be processed by the digital electronics [16].

C. Timewalk and Peaking Time

Two phenomena associated with the analog FE are important to the results presented in this paper: timewalk and peaking time.

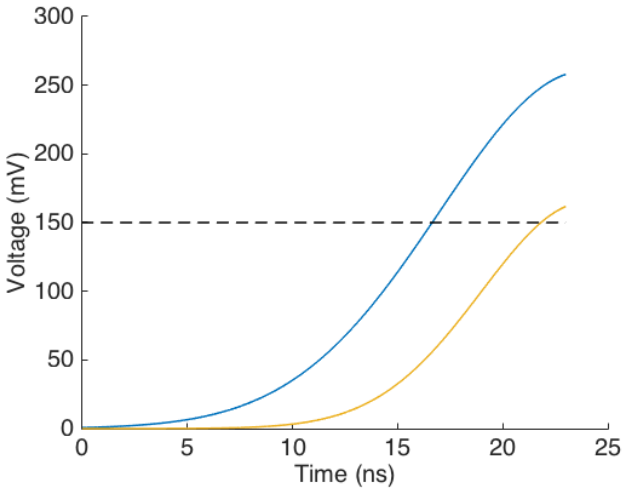

FIG. 8: Timewalk: When two signals of different amplitudes begin at the same time, the larger signal will cross a threshold first, causing a disparity in the measured arrival times of the two signals. This phenomenon is known as timewalk.

An example of timewalk is shown in Figure 16. The two signals in this diagram begin at the same time, but because one signal has a greater amplitude than the other, it crosses threshold first. According to the electronics, the larger signal arrives earlier than the smaller one, and, in this way, timewalk creates problems when trying to extract timing information about a signal. In the context of the LHC, if the timewalk is too large, it might mean that a large-amplitude signal is assigned to one bunch crossing, while a small-amplitude signal is assigned to the next one. This would make reconstructing any information about the particles nearly impossible.

Peaking time is simply the time it takes for a signal to reach its maximum value. It is desirable that the peaking time be roughly the same for all signal amplitudes, as this timing shift can then be treated approximately as a constant when performing track reconstruction.

D. Noise

Both the timewalk and peaking time are influenced by noise, and an illustration of how this occurs is shown in Figure 9. In addition to affecting pulse amplitude, noise can influence timing information by changing when the signal crosses threshold. The dominant types of noise in silicon detectors are Johnson noise and shot noise [3],[17],[16].

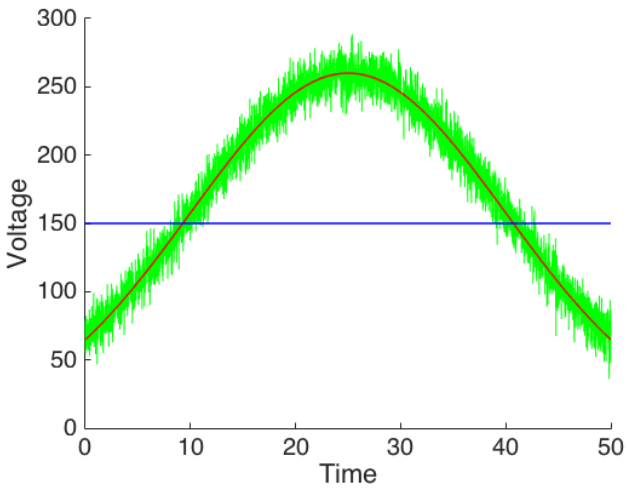

FIG. 9: The Impact of Noise on Timewalk and Peaking Time: Noise influences timewalk and peaking time measurements by influencing both amplitude and timing information. The red curve represents the signal, while the green represents the signal plus noise. Here the blue line is the threshold.

Johnson noise is caused by the thermal excitation of carriers, which causes slight variations in signal current. Shot noise results from the finite probability that a carrier will overcome the energy barrier in the depletion region. Each carrier has some probability of overcoming the barrier, and at random times certain carriers do. These carriers contribute to the signal as noise.

Due to the random nature of Johnson noise and shot noise, they produce a Gaussian distribution. This distribution will be illustrated while discussing the calibration of the MPA-Light device.

VI. MPA-LIGHT AND MAPSA

As discussed above, the Phase 2 upgrade to the outer tracker will contain PS modules. In order to record hits and correlate stubs, the pixels and the strips will be associated with two separate types of ASICs − the Macro Pixel ASIC (MPA) and the Short Strip ASIC (SSA). The SSAs process hit information from the strip sensors and transmit this data to the MPAs. The MPAs use this information, as well as hit information from the pixel sensors, to produce stubs. In the tracker, the stubs will be transmitted to the track trigger at 40 MHz, and after receiving the Level 1 (L1) trigger, the MPA will send an entire event to the back end electronics.

As preparation for the Phase 2 upgrade, developers at CERN produced the MPA-Light, a prototype of the MPA, in order to gain a better understanding of the ASICs design. The MPA-Light was then integrated into the Macro Pixel Subassembly (MaPSA) for testing. This assembly consists of six MPAs, pixel detectors, with a periphery capable of being either bump bonded or wire bonded to a test system. The system used for the measurements in this paper utilizes the wire bonding connectivity.

The MPA-Light is quite similar to is successor, but with several key differences. Each one of the final MPAs will be connected to an array of 120x16 pixels, while each MPA-Light is connected to an array of 16x3 pixels. However, the size of these pixels − 100x1446um − is the same for each design. Furthermore, the nominal clock speed of the MPA-Light is 160MHz (320MHz in the MPA), and the I/O power supply in the MPA-Light has a nominal value 2.5V (1.2V in the MPA). The following sections provide a brief overview of MPA-Light operation, while detailed information about the MPA and MPA-Light can be found in the MPA-Light User Manual, published by Dr. Davide Ceresa in 2015 [4].

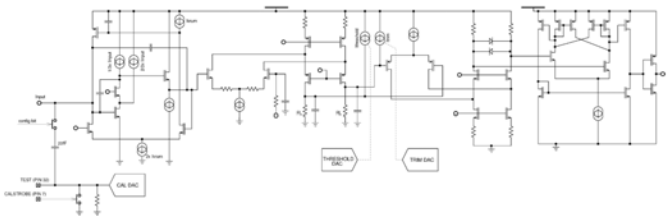

FIG. 10: The MPA-Light Analog FE: Like the analog FEs described earlier, the MPA-Light FE contains a preamplifier, shaper, and discriminator. Also shown is the connection between the device's test capacitance and the FE. More information about this circuit is available in the MPA-Light User Manual.

A. MPA-Light Design

Each MPA-Light is divided into two components: the pixel matrix and the periphery. The pixel matrix contains analog FEs as well as logic to preprocess and digitize the output of the FE. The periphery receives and processes hit information from the pixels, and it holds the control logic used to mediate data acquisition and serial configuration and readout.

The pixel matrix possesses 48 analog FEs, one for each pixel. Like the FEs discussed above, each pixel matrix FE contains a preamplifier, shaper, and discriminator. The preamplifier amplifies the signal, the shaper modifies the pulse width and shape, and the discriminator determines whether the pulse is above a particular threshold. A circuit diagram of the analog FE is shown in Figure 10. If the output of the shaper is large enough to cross threshold, the output of the discriminator is sent on to the pixel logic.

The pixel logic preprocesses hit information for the periphery logic and stores asynchronous hit information. Each pixel logic module services two pixels, although some logic, such as asynchronous hit counters and the configuration registers, are repeated for each pixel. A module containing the pixel clustering logic is common to both pixels, as is a buffer that buffers the output of this logic. The idea of clustering will be explained in the context of chip configuration.

The periphery serves the important functions of processing and temporarily storing information, communicating with the back end electronics, and generating signals to control operation of the MPA-Light. It possesses distinct analog bias, periphery logic, and control logic modules. The analog bias produces bias signals for the analog FEs, while the periphery logic is responsible for processing and storage. The control logic handles several important signals, including ones that mediate serial configuration and readout of the chip. The most important signal in this context is the Shutter signal, which controls when the chip acquires data.

Shutter must be LOW to configure the MPA-Light, and data can only be acquired when Shutter is HIGH. When Shutter returns LOW after data acquisition, the MPA-Light initiates data readout. Shutter is also responsible for initiating several other signals during data acquisition. Before collecting data, however, the chip must be properly calibrated and configured [4].

B. Calibration

Before operation, each pixel must be calibrated to compensate for the discrepancies between pixels arising in the manufacturing process. These variations result in slightly different noise levels in each pixel, and in order to use a single discriminator threshold for a whole MPA, the noise levels must be more or less equalized for each pixel. This is accomplished through a combination of the analog FE,

the analog bias system, and the pixel and periphery configuration bits.

Each pixel possesses an analog FE containing a 20 fF test capacitance. This test capacitance is used for calibration, although it can also be used in lieu of an external signal while acquiring data. During calibration, a calibration enable (CEL or CER) signal establishes a connection between an 8-bit DAC in the analog bias system with the pixels 20 fF capacitor. Another signal, the CALSTROBE signal, mediates when the calibration pulse will be sent and how long the pulse will last. The voltage applied to the test capacitor is set by the user via an 8-bit signal called CALDAC during configuration of the periphery [4].

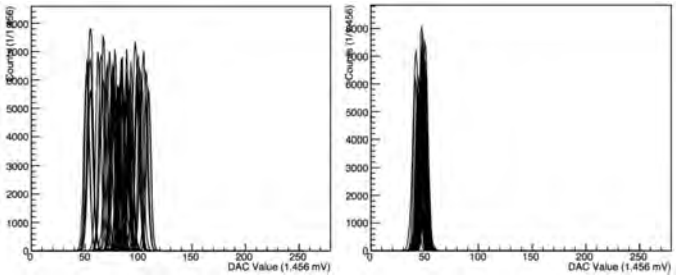

FIG. 11: Calibration Curves: The response from each pixel on an MPA-Light prior to calibration (left) and after calibration (right). During calibration, a 5-bit value, TRIMDAC is set for each pixel. These values equalize the response from each pixel.

Plots from before and after calibration are shown in Figure 11. The test capacitor injects its charge into the analog FE of the enabled pixel, where it is compared, one value at a time, to all 256 thresholds of the 8-bit threshold DAC. Since the FE is only sensitive to transitions across threshold rising above or falling below threshold, depending on the selected polarity the injected pulse only registers hits in a particular range of threshold values.

As can be seen from the plots in Figure 11, at very low thresholds the charge injected by the test capacitance is always greater than the threshold value. At these levels the FE registers no hits. Similarly, at large threshold values, the injected charge is always below the threshold value, and again, no hits are recorded. It is only in proximity of the voltage set by CALDAC that threshold transitions occur due to the presence of noise. Because of the statistically random nature of this noise, the plots of hits vs. Threshold DAC appear approximately as Gaussians. The width of these Gaussians increases with increasing noise, and Gaussian approximation improves markedly near the top of the curves, as the tails deviate from a Gaussian shape. This deviation could result from any number of sources, including an incommensurate amplification of noise due to negative feedback in the FE preamplifier or additional noise introduced by the bump bond between the pixels and the MPA-Light.

After iterating over all threshold values, a 5-bit TRIMDAC value is set for each pixel. This signal offsets the threshold values in the FE to create a near uniform response from each pixel. This effect is also seen in Figure 11 by the overlap of the post-calibration curves [4].

C. Configuration

Both the pixel logic and the periphery logic must be configured before performing data acquisition with an MPA-Light. In addition to controlling parts of the calibration, the configuration establishes important information such as the operating mode of the chip and whether synchronous and asynchronous readouts are performed. Configuration data is loaded serially.

Each pixel logic module is configured with a 20-bit word. As each module is shared between two pixels, some of these bits carry the same information for the different pixels. For example, each word contains two 5-bit TRIMDAC values, one for the left pixel and one for the right. The pixel configuration word also includes separate pixel mask, asynchronous readout enable, and calibration enable bits for each pixel. The pixel mask is AND-ed with the output of the FE to filter the signal. The remaining bits specify the signal polarity, the width of pixel clusters, and whether synchronous readout is enabled for both pixels.

The periphery is configured 32-bit word. Bits 0-15 configure the periphery logic, while bits 16-31 specify the CALDAC and TRIMDAC values discussed above. The first 16 bits determine the chips operating mode, the retiming information, the correlation offset for the first eight columns and last eight columns of pixels, and the cluster width information for strips. Strip data must be input via four pins on the periphery, and note that MaPSA does not include any strip sensors.

A cluster is a group of consecutive pixels that all receive a hit during the same clock cycle, or adopting the parlance of the particle accelerator, the same bunch crossing. The cluster width signals are used to determine the maximum cluster sizes that will be used to generate stubs. For example, the pixel clustering logic − part of the pixel logic − will find the centroid − via a process known as centroid extraction − of all clusters with cluster widths smaller than the value set by the cluster width signal. In the case when the cluster width is 00, no centroid extraction occurs and no clusters are disregarded.

The correlation offset bits are used to match the hits received by the strip pins on the periphery with those registered by the pixels, and the retiming bits are used to correlate pixel data and strip data in time.

For synchronous acquisition, the MPA-Light can operate in four different modes. These modes processes hit information differently, but each generates a 72-bit word for each bunch crossing in which a hit is received. The four operating modes are stub finding mode, in which the periphery logic uses stub data received in the periphery to generate stubs; stub emulator mode, in which the

OR of each pixel column produces a 16-bit word that is used to simulate strip data; centroid extraction mode, in which the MPA-Light stores up to eight centroids per bunch crossing; and no processing mode, in which each MPA-Light simply stores hit and bunch crossing information. In our setup, we operate exclusively in no processing mode, as we only have pixel detectors, and the laser signal used generate hits is well focused and positioned [4].

D. Acquisition

Once the MPA-Lights are configured, they are prepared for data acquisition. In general, the two acquisition modes are synchronous and asynchronous, but as described above, there are multiple operating modes for synchronous acquisition. In order for data acquisition to occur, the shutter signal must be HIGH. When the shutter goes HIGH, it initiates the sample clock, a 40MHz clock derived from the 160MHz system clock, which serves the FE and periphery logic.

During asynchronous readout, this sampling clock is ignored, as the number of hits during the shutter opening is simply stored using a 16-bit ripple counter. However, during synchronous acquisition, each period of the clock defines a bunch crossing, starting with the first clock cycle after the shutter opens. A signal diagram illustrating synchronous acquisition is shown in Figure 12.

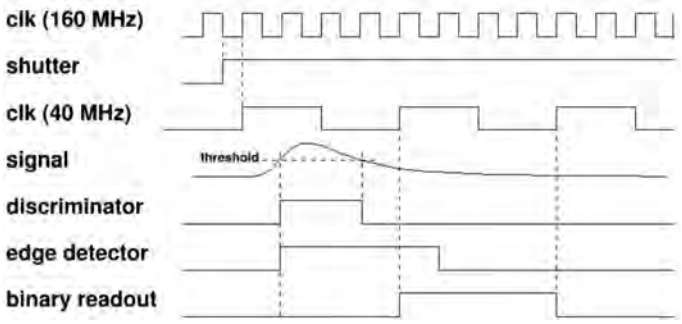

FIG. 12: Synchronous Readout: When the Shutter signal opens, it generates a Sample Clock and acquisition begins. When a hit crosses threshold, Discriminator and Edge Detector signals are generated, and a Binary Readout signal is produced on the next rising edge of the Sample Clock.

Synchronous acquisition occurs through a binary readout. Upon receiving a pulse from the discriminator, the pixel generates a one-clock-cycle-width pulse. The polarity selected in configuration determines whether this signal is generated by the rising or falling edge of the discriminator pulse. The pixel logic then performs any centroid extraction and removes any clusters exceeding the maximum cluster width before transmitting the hit information to the periphery logic, where it is encoded into 72-bit words according to the operating mode selected. The periphery can store up to 96 words per shutter opening, and any other words are simply lost. Note that memory writes occur only when hits are received.

In addition to initiating the sample clock, the shutter signal also generates two more signals: a Clear signal, which clears the counters in the pixel logic and the memory in the periphery; and an internal shutter signal (IntShutter), which is essentially the inverted Clear signal. Similarly to the pixel masks, the IntShutter signal is AND-ed with the output of the discriminator in the pixel logic to filter the hit information. The Clear signal is set for a nominal value of 1.5ns at the beginning of data acquisition. This signal resets the 16-bit ripple counter and all registers except that containing the asynchronous readout header [4].

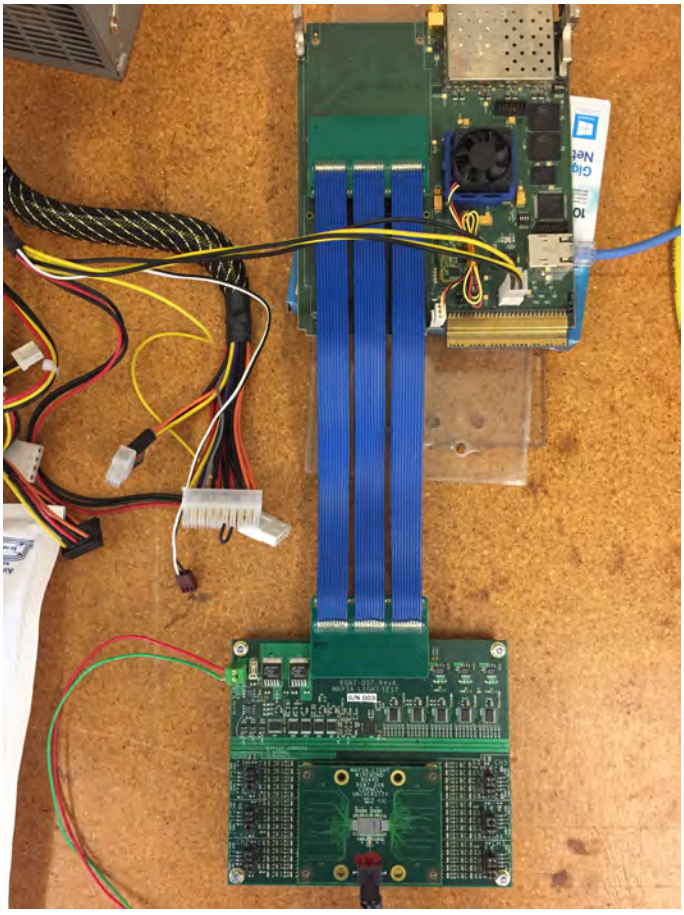

FIG. 13: The MaPSA Setup: Toward the bottom of the picture, the MaPSA − located under an aluminum cover − is mounted to a carrier board, which in turn is connected to a testboard. This testboard transmits information back to the GLIB, where it is sent via an Ethernet connection to the lab computer.

E. Readout

When the Shutter signal returns LOW, the data is read out to the back end electronics. There are dedicated pins for the synchronous and asynchronous readouts, but both readouts are done serially.

The counters in the pixel logic, which record asynchronous hit information, are arranged as a shift register. During asynchronous readout, the information from these counters is shifted out on the falling edge of the 160MHz readout clock. Meanwhile, a 32-bit header is loaded after the final pixel information in the shift register. This header is programmable and its readout enables the user to verify that no information is lost during the asynchronous readout.

The synchronous readout is very similar, but there is no synchronous header. During synchronous readout all 96 words from the periphery memory are transmitted to the back end electronics [4].

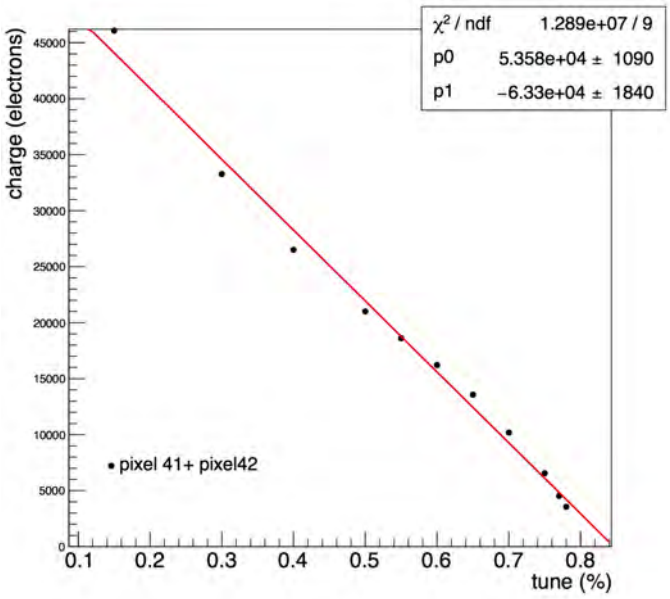

FIG. 14: Correlation Between Tune and Charge: The laser tune and signal charge for 300μm silicon were correlated using a Phase 2 pixel. By applying a best-fit line to this data, we produced the rough conversion between laser tune and signal charge that is used to analyze our measurements.

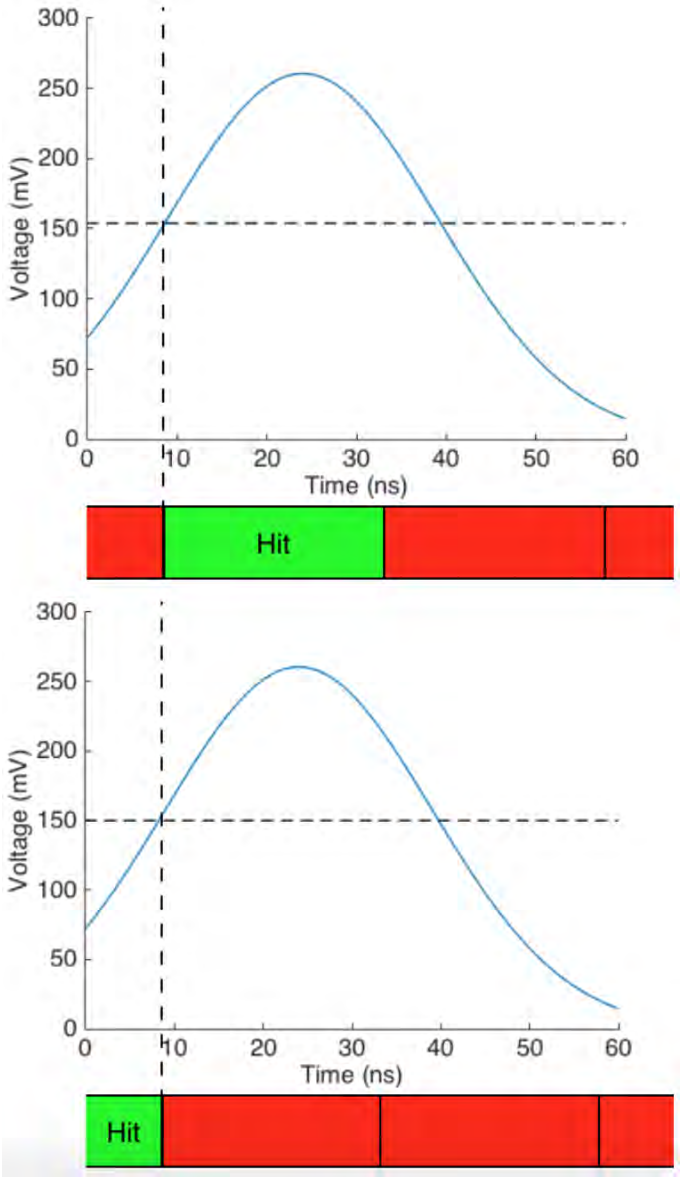

FIG. 15: Making Measurements with Bunch Crossings: Adjusting the laser delay with our oscilloscope, we positioned the pulse at the edge of two bunch crossings. Then by adjusting the discriminator threshold, and seeing when a majority of hits crossed from one bunch crossing to another, we were able to correlate a point on the pulse with a particular laser delay. This technique was used extensively for reconstructing pulse profiles, and a variant was used to perform timewalk measurements.

VII. MEASUREMENTS

In order to gain a better understanding of the MPA-Light, we used the setup shown in Figure 13, in which the MaPSA is wirebonded to a carrier board. This carrier board is mounted on a testboard, which is then connected to a GLIB. The GLIB transmits information from the MaPSA to our lab computer via an Ethernet connection.

The GLIB is powered by a 12V power supply, while the testboard is powered with 3V power supply. An additional power supply is used to reverse bias the pixel sensors bonded to the MaPSA. During data acquisition it is always set at a nominal value of 90 V. This bias volt-

age fully depletes the silicon sensor, reducing the effects of Johnson noise and shot noise on our data. In a effort to reduce noise further, we also inserted a low pass filter between the bias voltage power supply and the pixel sensor.

To test our silicon detector, we use a 900 nm laser, calibrated to reliably provide the signal equivalent of a MIP in 300μm silicon – 20,000-25,000 electrons, or 3.2-4 fC – with a laser tune of 50%. After obtaining the correlation between several laser tunes and signal charge for 300μm silicon, we produced a best-fit line to provide an approximate conversion for our measurements. This correlation between tune and charge is shown in Figure 14.

In order to limit the amount of ambient light reaching the sensors, the MaPSA is encased under a reflective aluminum surface, which contains a slit through which only three of its six pixels are visible. Using stepper motors, we are able to scan this laser in steps of 4.8 nm in the x-, y-, and z-directions, and this facilitates focusing the laser as well as studying the transition region in which the laser generates hits simultaneously on two pixels. The laser is considered to be focused when the transition region is only 15-20 nm wide – that is, when there is only a 15-20 nm region between seeing the laser exclusively on one pixel and seeing the laser exclusively on the adjacent pixel.

Because the MaPSA only produces digital information regarding pixel hits, it was necessary to use roundabout techniques to measure the pulse shape. Ultimately, we were able to glean this information from our data by looking at hits arriving at the transition between two bunch crossings. An illustration of how this is accomplished is shown in Figure 15.

In our effort to understand the MPA-Light, we sought to confirm the device′s timewalk and peaking time specifications. Note that all of the following measurements were performed at room temperature.

A. Timewalk

From simulations, the specified timewalk of the MPA-Light is listed as < 14 ns for a threshold of 0.5 fC with signals between 0.75 fC and 12 fC. This value was verified in a measurement made by the developer, who found the timewalk to be 15±1.6ns for thresholds of 0.5 fC and 1 fC and with signals between 0.5 fC and 9 fC. Unfortunately, from the best-fit approximation above, the maximum signal attainable with our current setup is ~ 8.6 fC, meaning we could verify neither the specification nor developer′s measurement exactly. Furthermore, because the developer worked with a bare chip, while our MaPSA included a pixel sensor, our baseline noise level was notably higher, making a threshold as low as 0.5 fC inadvisable for our system.

We therefore determined to measure the timewalk over a range of signals we might reasonably expect to see in the final tracker – signals close to that produced by a MIP. The results of this measurement are shown in Figure 16 With a threshold of ~ 0.9 fC and signals between ~ 1.5 fC and ~ 8.6 fC, we measured a timewalk of ~ 8 ns, well within the specified value.

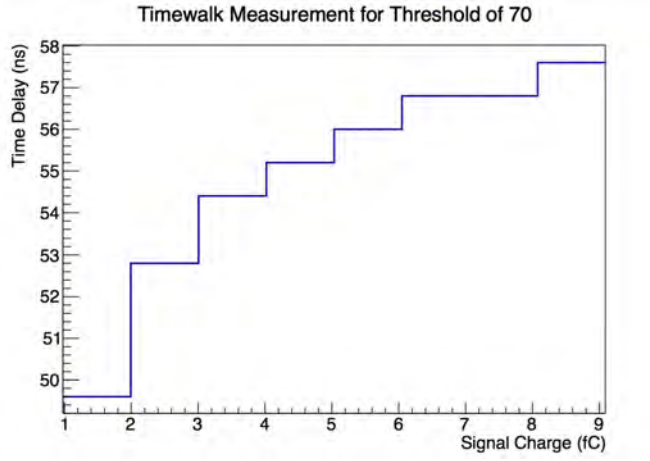

FIG. 16: Timewalk: The timewalk was measured by comparing the relative arrival times of eight signals between 1.5 fC and 8.6 fC for a threshold of 0.9 fC. This measurement showed the timewalk to be ~ 8 ns, well below the specified value.

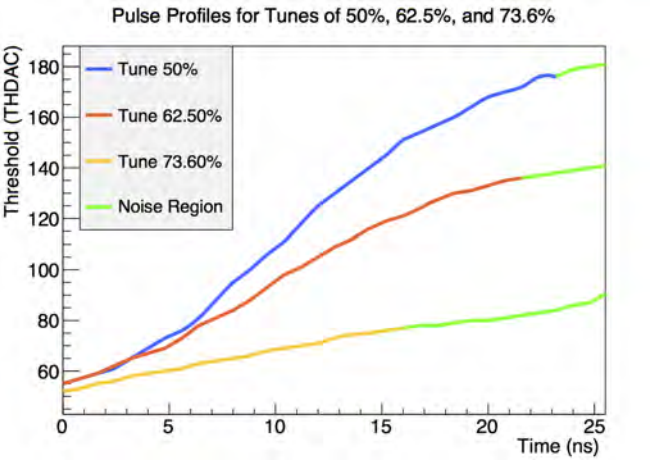

FIG. 17: Peaking Time: For the 3.5 fC, 2.3 fC, and 1.1 fC signals, the peaking times were measured to be ∼23.2 ns, ∼21.6 ns, and ∼16 ns, respectively. The green portions of the curves are not included in this calculation, and they represent regions where we saw hits with $< 100\%$ efficiency. These regions are associated primarily with noise.

B. Peaking Time

The peaking time for the MPA-Light is listed as $< 24ns$. In order to corroborate this value, we profiled the

rising edges of 3.5 fC, 2.3 fC, and 1.1 fC signals. This was accomplished using the method described in Figure 15, starting from the peak of the signal and descending until the signal was indistinguishably enveloped in noise. In this way we determined the peaking times of these signals to be $\sim$ 23.2 ns for the 3.5 fC signal, $\sim$ 21.6 fC for the 2.3 fC signal, and $\sim$ 16 fC for the 1.1 fC signal. These results are shown in Figure 17.

These results indicate that the peaking times of the 3.5 fC ($\sim$ 1 MIP) and 2.3 fC signals are reasonably close to the specified value, while seemingly implying an inability of the pixel's analog FE to properly shape such low-amplitude pulses. However, it is also highly probable that the method used to profile the pulse shapes is inadequate for low signal charges, as is evidenced by the unexpected linearity of the 1.1 fC pulse. Regardless, such low-amplitude signals are not likely to be present in the environment of the HL-LHC, making the agreement between the 3.5 fC and 2.3 fC pulses and the specification much more significant.

VIII. CONCLUSION

In general, our timewalk and peaking time measurements verify the specifications for the MPA-Light by demonstrating a timewalk well below 14ns and peaking times that remained reasonably close to 24ns for two different signals. In the future, the attenuator of our laser might be adjusted to permit a timewalk measurement over an even broader range of signals, and the peaking time might be measured for a higher amplitude signal, say a signal equivalent to two MIPs. Nonetheless, our findings indicate that, with regards to timewalk and peaking time, the MPA-Light electronics are ready to be implemented in the actual MPA.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Anadi Canepa, and Basil Schneider for their advice and assistance in understanding the material covered in this document. I would also like to thank Gino Bolla, Caterina Vernieri, Irene Zoi, and Lorenzo Uplegger for their contributions to this project and all those involved with the SIST Program for providing me the opportunity to work at FNAL, with a special thanks to Elliot McCrory, Sandra Charles, Judy Nunez, Charley Orozco, and Mayling Wong-Squires.

[1] Magnetic Multipoles: Taking a closer look at LHC. http://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.magnetic_multipoles.
[2] Ben Berger. Large hadron collider. *Symmetry Magazine*, 2006. http://www.symmetrymagazine.org/article/august-2006/deconstruction-large-hadron-collider.
[3] Francesca Cenna. Signal formation in innovative silicon detectors. 2012-2013.
[4] Davide Ceresa. Mpa-light user manual. *Journal of Instrumentation*, 2015.
[5] The CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the cms tracker. *Journal of Instrumentation*, 9(10):1–8, 2014.
[6] The Large Hadron Collider. http://home.cern/topics/large-hadron-collider.
[7] D. Contardo. Technical proposal for the phase-ii upgrade of the cms detector. 2015.
[8] Science: Worldwide Particle Physics Discoveries. http://www.fnal.gov/pub/science/particle-physics-101/worldwide-discoveries.html.
[9] Glenn Elert. The hyperphysics textbook: The standard model. http://physics.info/standard/.
[10] Engineering. http://home.cern/about/engineering/.
[11] Science: Neutrinos. http://www.fnal.gov/pub/science/particle-physics/experiments/neutrinos.html.
[12] The Particle Adventure: The Fundamentals of Matter and Force. http://particleadventure.org/how-does-the-higgs-boson-get-its-mass.html.
[13] How CMS Detects Particles. http://cms.web.cern.ch/news/how-cms-detects-particlesod.
[14] The HL-LHC Project. http://hilumilhc.web.cern.ch/about/hl-lhc-project.
[15] Giacomo Sguazzoni. Overview of the tracker system. Presentation, 2016.
[16] Helmuth Spieler. Analog and digital electronics for detectors. 2003.
[17] Ben G. Streetman. *Solid State Electronic Devices*. Prentice Hall, Englewood Cliffs, New Jersey, 1990.
[18] The LHC Cryogenic System. https://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/components/doc/TheLHCcryogenicsystem.pdf.
[19] Viktor Veszpremi. Operation and performance of the cms tracker. *Journal of Instrumentation*, pages 1–2, 2014.

APPENDIX

During initial testing with the MaPSA, it was noted that some hit data was lost during synchronous acquisition. This resulted in an "inefficiency window", in which signals arriving during that period of time were lost. It was ultimately determined that this inefficiency was due to an internal RESET signal, which reset the edge detector signal – used in generating the binary readout – at the beginning of each bunch crossing [1]. As any hits arriving while RESET=1 are irretrievably lost, developers at CERN plan to remove this signal from the final MPA design. This appendix will briefly discuss measurements associated with this inefficiency window.

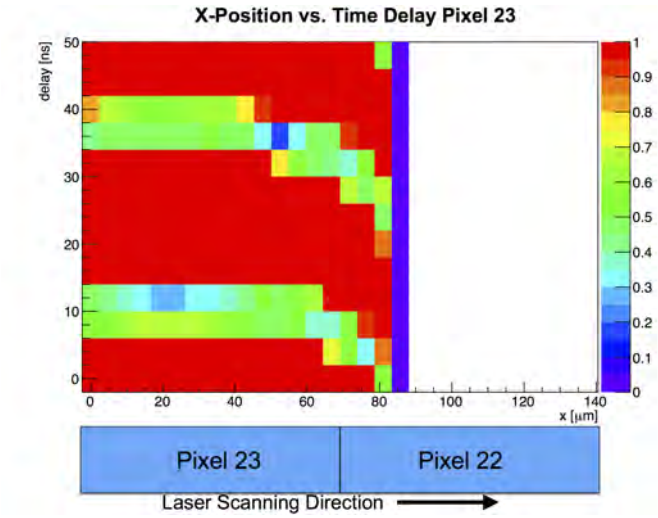

FIG. 1: Position vs. Laser Delay: Scanning across two pixels for different laser delays exposes an inefficiency in the synchronous readout of the MPA-Light. This inefficiency exhibits 25 ns periodicity, and was determined to arise from the device's internal RESET signal.

The inefficiency window was originally identified while scanning across the transition region between two pixels for different laser delays. The result of such a measurement is shown in Figure 1, where the non-red bands represent the regions of inefficiency.

There are two noteworthy features in this plot. Firstly, in accordance with the 40 MHz frequency of the RESET signal, the inefficiency window exhibits a 25 ns periodicity. Secondly, as the laser moves into the transition region between the two pixels, the inefficiency window appears to "drift", occurring at earlier time delays.

This second phenomenon is actually due to timewalk. As the laser moves into the transition region, the photons from its pulses are shared by the two pixels, resulting in smaller signals than arise when the laser is directly positioned over a single pixel. Because of timewalk, the smaller signals in the transition region arrive later, entering later bunch crossings with an earlier time delay. This is exactly the behavior seen in the inefficiency windows: The smaller signals in the transition region coincide with the RESET signal with a smaller time delay than the signals not in the transition region.

If taken as a timewalk measurement, this plot shows a timewalk of approximately ∼12 ns between signals in the center of a pixel and those in the center of the transition region. This is still within the specification of < 14 ns.

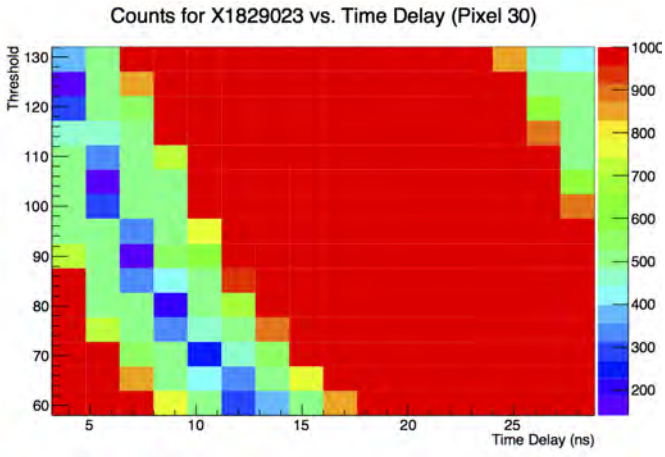

FIG. 2: Width of the Inefficiency Window: A higher granularity measurement reveals the width of the inefficiency window to be ∼5.4 ns.

In order to understand the width of this inefficiency window, we performed the measurement shown in Figure 2, in which we decreased the step size in our laser delay to produce a finer granularity measurement. This plot demonstrates the width of the inefficiency window to be ~ 5 ns.

[1] Davide Ceresa. Personal Communication, 2016.