Characterization of the H2M Monolithic CMOS Sensor

Rafael Ballabriga[a], Eric Buschmann[b], Michael Campbell[a], Raimon Casanova Mohr[c], Dominik Dannheim[a], Jona Dilg[d], Ana Dorda[a], Ono Feyens[d], Finn King[d], Philipp Gadow[a], Ingrid-Maria Gregor[d], Karsten Hansen[d], Yajun He[d], Lennart Huth[d], Iraklis Kremastiotis[a], Stephan Lachnit[d], Corentin Lemoine[a,e], Stefano Maffessanti[d], Larissa Mendes[d], Younes Otarid[a], Christian Reckleben[d], Sébastien Rettie[a], Manuel Alejandro del Rio Viera[d], Sara Ruiz Daza[d,*], Judith Schlaadt[d], Adriana Simancas[d], Walter Snoeys[a], Simon Spannagel[d], Tomas Vanat[d], Anastasiia Velyka[d], Gianpiero Vignola[d], Håkan Wennlöf[d]

[a] *CERN, Esplanade des Particules 1, CH-1211 Geneva 23, Switzerland*
[b] *Brookhaven National Laboratory (BNL), New York 11973-5000, Upton, USA*
[c] *Institut de Física d'Altes Energies (IFAE), Edifici CN, UAB campus, 08193 Bellaterra (Barcelona), Spain*
[d] *Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany*
[e] *IPHC, Université de Strasbourg, 23 rue du Loess, Strasbourg, France*

Abstract

The H2M (Hybrid-to-Monolithic) is a monolithic pixel sensor manufactured in a modified 65 nm CMOS imaging process with a small collection electrode. Its design addresses the challenges of porting an existing hybrid pixel detector architecture into a monolithic chip, using a digital-on-top design methodology, and developing a compact digital cell library. Each square pixel integrates an analog front-end and digital pulse processing with an 8-bit counter within a 35 µm pitch.

This contribution presents the performance of H2M based on laboratory and test beam measurements, including a comparison with analog front-end simulations in terms of gain and noise. A particular emphasis is placed on backside thinning in order to reduce material budget, down to a total chip thickness of 21 µm for which no degradation in MIP detection performance is observed. For all investigated samples, a MIP detection efficiency above 99 % is achieved below a threshold of approximately 205 electrons. At this threshold, the fake-hit rate corresponds to a matrix occupancy of fewer than one pixel per the 500 ns frame.

Measurements reveal a non-uniform in-pixel response, attributed to the formation of local potential wells in regions with low electric field. A simulation flow combining technology computer-aided design, Monte Carlo, and circuit simulations is used to investigate and describe this behavior, and is applied to develop mitigation strategies for future chip submissions with similar features.

Keywords: Solid state detectors, Silicon sensors, Charged particle detection, MAPS, CMOS imaging process, Test beam, Simulation, TCAD

Contents

*Corresponding author
Email address: sara.ruiz.daza@desy.de (Sara Ruiz Daza)

1. Introduction

Monolithic active pixel sensors (MAPS) with a small collection electrode are a promising sensor type for future high-energy physics (HEP) experiments. This includes tracking or vertexing applications at a lepton collider like CLIC [1] or FCC-ee [2]. MAPS come with the benefit of integrating the readout electronics and sensitive volume in the same die, which inherently reduces the material budget but also removes the necessity of the costly bump bonding process. A small collection electrode layout results in a small input capacitance to the first amplifier stage (order of fF), which reduces power consumption for a given signal charge [3]. The TPSCo 65 nm ISC (Image Sensor CMOS) process [4] is a promising technology in that regard. It allows for a higher density of circuit elements than larger feature size processes, and technology demonstrators such as APTS [5] or DPTS [6] have qualified it for applications in HEP.

The hybrid-to-monolithic (H2M) test chip is a MAPS produced in this 65 nm ISC process. It explores complex per-pixel pulse processing functionality for time and amplitude measurements, at the cost of a pitch of 35 µm, which is 40 % larger than other prototypes developed for High Energy Physics in this process. The architecture of the chip is ported from a Timepix-like hybrid pixel detector. This porting process is interesting to explore, as the large set of hybrid chips for HEP applications may speed up the development of future monolithic sensors. The design applied a digital-on-top workflow, which is a contemporary strategy for the design of complex integrated circuitry. The digital logic is synthesized around the analog building blocks such that certain constraints (e.g. on timing) are fulfilled. The design effort included the composition of a compact digital cell library comprising a set of standard circuit elements with a footprint reduced by about 25 % compared to the standard cells.

Preliminary results from the characterization and simulation of H2M are presented in [7, 8], respectively. Section 2 of this article describes the H2M chip design, its features, and the data acquisition system utilised for the presented measurements. Laboratory characterization and calibration are presented in Section 3, while the methodology for test-beam measurements and the resulting performance in the detection of minimum-ionizing particles (MIPs) are detailed in Section 4 and Section 5. Section 6 summarizes the employed simulation procedure, which has proven to be a valuable tool to understand the results of the measurements. Conclusions and an outlook to further plans are provided in Section 7.

2. The H2M Chip

The H2M chip matrix consists of 64×16 square pixels with a pitch of 35 µm, resulting in a total sensitive area of 2.24×0.56 mm^2. The total chip size is 3×1.5 mm^2, including the periphery, differential-to-CMOS receivers/drivers, and I/O connectors at the bottom of the pixel matrix. Each pixel features a collection electrode, an analog front-end, and digital logic capable of 8-bit pulse processing. The following section describes the sensor design, including the analog front-end and digital logic, as well as the chip periphery and readout architecture.

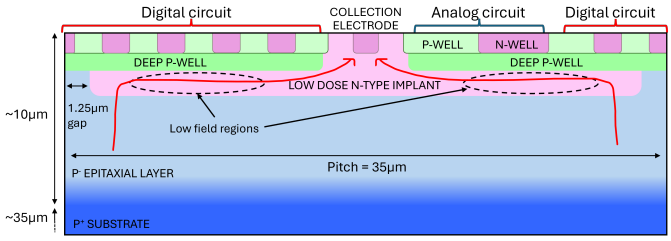


Figure 1: H2M cross section sketch of one pixel, with drift paths of electrons indicated by red arrows. The approximately 5 µm metal interconnect layers are not shown in the schematic.

2.1. Sensor Design

The sensor is manufactured in a modified TPSCo 65 nm ISC process [4], and has been thinned on wafer level to a total thickness of approximately 50 µm. This includes around 5 µm of metal interconnect layers, a 10 µm p-type epitaxial layer, and a 35 µm p-type substrate. Due to the comparatively high doping concentration of the substrate, the lifetime of charge carriers is significantly reduced in this region, leading to rapid recombination before they can contribute to signal formation. Therefore, mainly the epitaxial layer serves as an active volume for the signal formation.

In order to enhance the depletion within the epitaxial layer, a low-dose n-type implant is introduced, as illustrated in Figure 1. This implant has a 2.5 µm gap at the pixel boundaries, forming a vertical pn-junction that introduces a lateral component of the electric field. As a result, the charge collection efficiency and timing performance of the sensor are improved. This design is known as *modified with gap* layout [4].

Each pixel has a small n-type collection electrode and integrates full CMOS circuitry using the quadruple well technology [9]. This enables the implementation of PMOS transistors in n-wells and NMOS transistors in p-wells, shielded by a deep p-well to prevent charge collection by these n-wells. The positions of these n-wells and p-wells within the deep p-well are sketched in Figure 2.

The sensor is reverse-biased from the front side via a contact located outside the pixel matrix, connecting the p-well and p-substrate to the same potential. The limitations of this reverse bias are discussed in Section 3.1. Since there is no junction isolation at the location of the gap in the deep n-type implant between the p-well and the p-substrate, current flows through the deep implant gap when the p-well is biased differently from the p-substrate. This prevents the sensor from operating with different p-well and p-substrate bias voltages.

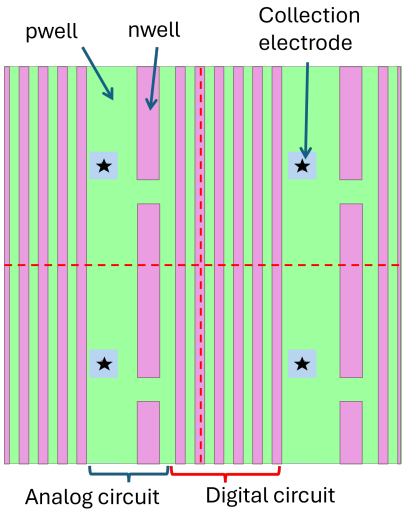


Figure 2: Simplified layout of the n-well and p-well positions within the deep p-well in four pixels. The locations of the analog front-end and digital logic are shown. The collection electrode, placed outside the deep p-well, is also indicated. The pixel cell boundary is marked with a dashed red line.

In this work, the impact of backside thinning (reduction of the substrate thickness) on chip performance will be discussed through laboratory and test beam measurements. In particular, results on samples thinned on single-die level by backside grinding to between 21 µm and 30 µm will be presented.

2.2. Analog Front-End

The in-pixel front-end consists of a collection electrode with a charge-sensitive amplifier (CSA) with Krummenacher feedback [10], followed by a continuous-time threshold comparator.

The global threshold is defined by `dac_vthr`, an 8-bit digital-to-analog converter (DAC) located in the chip periphery. Local fine-tuning of the threshold is achieved through an additional 4-bit DAC (`tuning_dac`) implemented within each pixel. The step size, and thus the dynamic range, of this fine-tuning is determined by another 8-bit DAC in the periphery (`dac_itrim`), which defines the current used in the tuning DACs.

The front-end also offers test pulse injection through a dedicated capacitor connected to the CSA input. Test pulses are enabled via the `tp_enable` DAC, and their amplitude is controlled by a global 8-bit DAC (`dac_vtpulse`). This amplitude determines the voltage step applied to the injection capacitor, enabling the injection of a controlled amount of charge into the pixel. Individual pixels can be masked by disabling their digital logic and setting the analog front-end into a low-power mode. A schematic of the pixel analog front-end is shown in Figure 3.

The inverting amplifier amplifies and converts the negative output of the charge collected by electrons into a positive output. The gain simulations are presented in Section Section 2.4. The linear discharge rate of the CSA

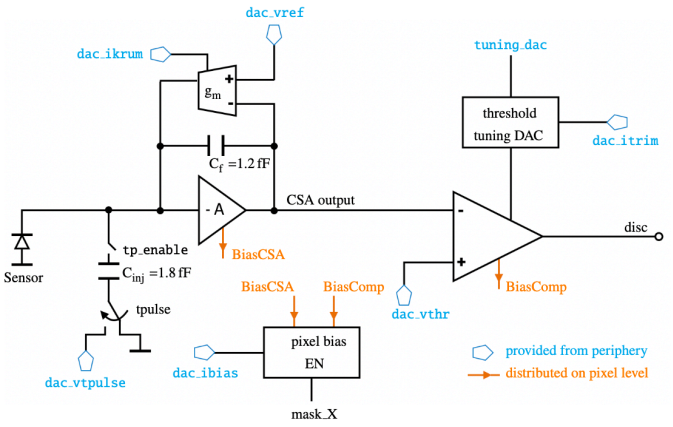


Figure 3: Schematic of the pixel analog front-end.

is controlled by a feedback current (*ikrum*), which is set by an 8-bit DAC in the chip periphery. This feedback current defines how fast the amplifier resets after a hit. It compensates for leakage current, and enables energy measurements, as the time the signal remains above the threshold is proportional to the input charge. Figure 4 (top) illustrates a triangular signal shape for two different *ikrum* settings. A higher *ikrum* results in a steeper return-to-baseline slope (faster signal decay).

Two additional 8-bit DACs are included in the analog periphery, `dac_ibias` and `dac_vref`, which set the CSA current bias and the Krummenacher reference voltage, respectively.

2.3. Digital Logic & Acquisition Modes

The H2M test chip can operate in four non-simultaneous acquisition modes: Time-over-Threshold (ToT), Time-of-Arrival (ToA), photon counting, and triggered. For this purpose, each pixel includes a configurable 8-bit counter, which is used differently in each acquisition mode. A schematic layout of the four acquisition modes is shown in Figure 4. In ToT, ToA, and photon counting modes, a shutter signal controls the data acquisition, while in triggered mode, it is controlled by a strobe signal.

ToA records the time at which the signal exceeds the threshold during charge collection, while ToT measures how long the signal stays above that threshold. More specifically, the ToA is defined as the number of clock cycles from the moment the signal crosses the threshold until the shutter closes, and the ToT as the number of clock cycles between the rising and falling edges of the signal crossing the threshold within a single shutter. If multiple signals cross the threshold within the same shutter window, their respective ToT values are accumulated for each pixel. The acquisition clock runs at 100 MHz, corresponding to 10 ns per clock cycle, and setting a lower limit on the time/energy resolution in ToA/ToT modes of $10\,\text{ns}/\sqrt{12} \approx 2.9\,\text{ns}$. A time walk correction cannot be applied since a simultaneous ToT and ToA measurement is not possible with the chip.

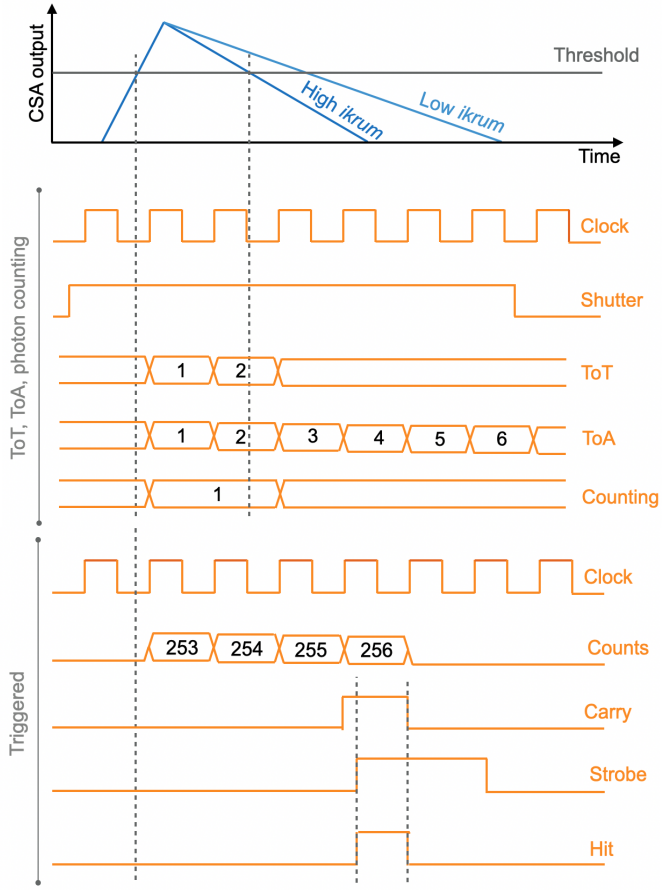


Figure 4: Schematic layout of the four acquisition modes. On top, the CSA output signal is represented with a triangular shape, whose return-to-baseline slope depends on the *ikrum* setting. In the middle, for the high *ikrum*, the signal is compared to a reference voltage threshold, and the corresponding measured counts in ToT, ToA, and photon counting modes are indicated. At the bottom, a readout in triggered mode is depicted, using as an example, a configurable preset of 253 counts and a strobe duration of 2 clock cycles.

In the photon counting mode, each pixel registers the number of times that the signal crosses the threshold within a shutter window.

In triggered mode, an external trigger signal is used to validate hits before readout. The in-pixel 8-bit counter is preset with a value corresponding to the expected fixed delay of the trigger latency. When a pixel fires, the counter starts incrementing from this preset value and generates a carry signal once it reaches its maximum. This carry signal lasts for one clock cycle. A strobe signal, provided externally after the trigger signal arrives, is distributed to the pixels using the same line as the shutter in other acquisition modes. This signal defines an adjustable coincidence window, accommodating for the H2M time walk. If the carry and strobe signals coincide, a binary readout is issued. In this way, shorter shutter durations can be utilized in test beam measurements compared to ToT and ToA acquisition modes (see Section 4), reducing the probability of noise hits and allowing the operation of the chip at lower thresholds.

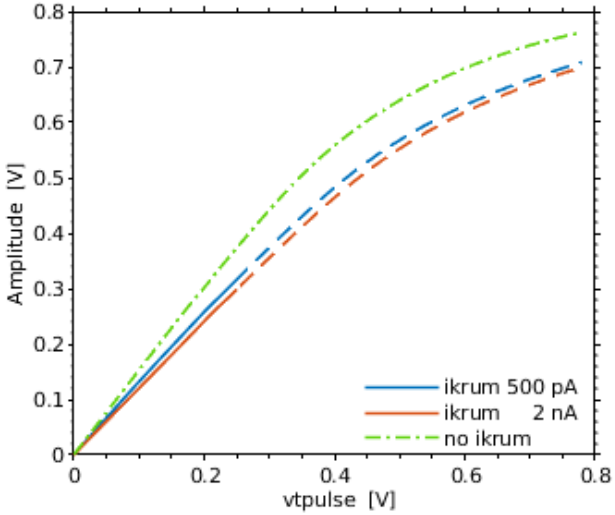


Figure 5: Simulated amplitude response of the CSA as a function of the injected signal `vtpulse`, for three different *ikrum* settings.

2.4. Analog Front-End Simulations

Analog front-end simulations have been performed to validate the circuit performance of the final layout, including all the parasitic elements of the circuit. The sensor is emulated with a 2 fF capacitance.

Figure 5 shows the signal amplitude at the output of the CSA as a function of the injected test pulse `vtpulse` for two different *ikrum* configurations. Because of the faster discharging of the feedback capacitance C_f for larger Krummenacher bias currents, the gain is reduced from 1.269 V/V to 1.198 V/V [11]. The values are extracted from the linear range, indicated in solid lines. Recalculating both gains using the injection capacitance yields a charge sensitivity of 113 mV/ke$^-$ and 107 mV/ke$^-$, respectively. The maximum gain is theoretically achievable for *ikrum* = 0 (no discharging of C_f), included in Figure 5. It is given by the ratio $C_{inj}/C_f = 1.5$ when the CSA is operated as a voltage amplifier.

The simulated ToT behavior is shown in Figure 6 as a function of the injected test pulse `vtpulse` for the same *ikrum* settings as in Figure 5. For small *ikrum* values, the CSA pulse width is longer, and with it, the ToT is also increased, resulting in a larger ToT gain. The nonlinear region for small input signals is suppressed by setting the comparator threshold to 70 mV (corresponding to 620 electrons using the above-mentioned gain) above the baseline. Below this threshold, C_f discharges exponentially with a non-constant `vtpulse`-dependent discharging current. Whereas for `vtpulse` values above 70 mV, a constant current of $0.5 \cdot ikrum$ [11] leads to a linear decrease of the CSA output signal. In this region, the CSA output can be approximated as a triangular waveform over time, with a steep rising edge and a slow falling edge, giving:

$$\mathrm{ToT} \simeq 2\ C_f \cdot \frac{\mathrm{Amplitude} - \mathrm{vthr}}{ikrum}. \qquad (1)$$

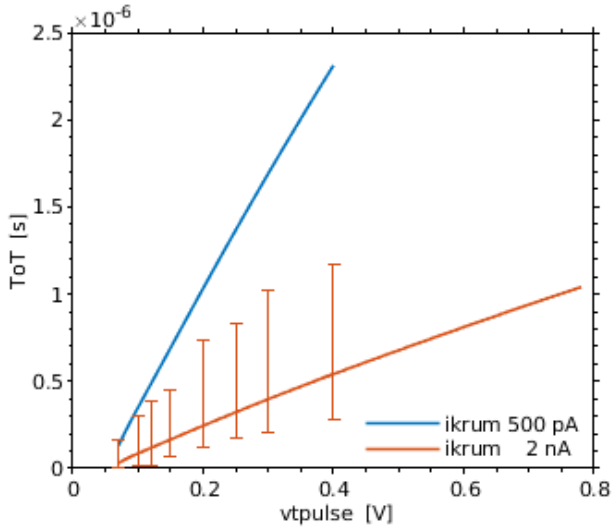


Figure 6: Simulated ToT response to injected signal `vtpulse`. The deviations (error bars) are primarily caused by variances in *ikrum*.

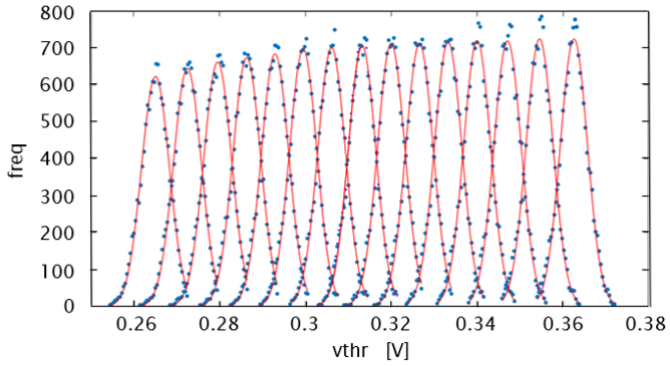
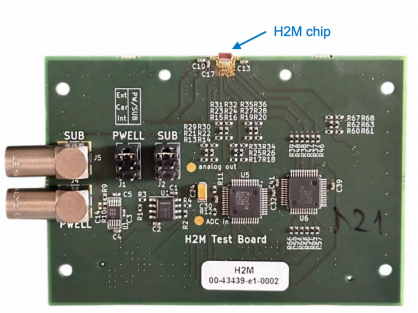


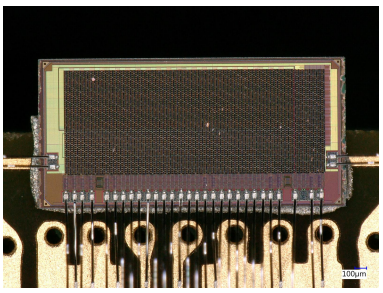
Figure 7: Simulated occupancy as a function of threshold, `vthr`, for the different `tuning_dac` settings. The curves are the fitted Gaussian functions.

The error bars 6 indicate the ToT variance arising from process statistics in the simulation, with the minimum and maximum deviations shown. The main source of these deviations is transistor mismatch affecting the *ikrum* current distribution from the periphery across the pixel matrix. The error bars are asymmetric because *ikrum* appears in the denominator of Equation (1).

Figure 7 shows the result of a transient noise [12] analysis while scanning the threshold. The simulation is repeated for each of the 15 DAC settings of the threshold `tuning_dac`. The RMS of the Gaussian-like distributions represents the CSA noise. It is the same for all of the 15 distributions $\sigma \approx 3.2\ \mathrm{mV_{rms}}$, which corresponds to approximately $28\ \mathrm{e^-_{rms}}$. The mean of the distributions is shifted by about 6 mV per DAC step. This value can be adjusted by varying the periphery-generated DAC bias current `itrim`.



(a) Picture of an H2M chip mounted on a chip board.



(b) Zoomed-in picture of an H2M chip mounted on a chip board.

Figure 8: Picture of an H2M chip glued and wire-bonded at the edge of a chip board.

2.5. Configuration & Readout

The chip configuration is performed via a slow control interface based on the Medipix4/Timepix4 protocol [13, 14], featuring a serial communication bus with data and clock inputs and outputs. Configuration signals are distributed either globally (to control settings such as acquisition start/stop, shutter, acquisition mode selection, and trigger latency), or locally to individual pixels (enabling features like test pulse injection, pixel masking, and tuning DAC).

While the clock frequency of the slow control is 40 MHz, the readout is performed using a 25 MHz clock. During readout, the 8-bit pixel counters are connected as a shift register and their contents are sequentially shifted from the top to the bottom of each column across the entire matrix, without zero suppression.

2.5.1 Caribou DAQ System

The characterization of the H2M chip is performed using the Caribou data acquisition (DAQ) system [15]. This is a modular and flexible platform developed for the prototyping and testing of detectors. It combines a System-on-Chip (SoC) running an embedded Linux Operating System (OS), a user-configurable FPGA firmware, and the *Peary* software to provide a streamlined solution for detector control, configuration, and data readout. Caribou's hardware, centered around the Control and Readout (CaR) board, offers interfaces, programmable power supplies, and versatile connections to easily integrate a wide range of detector prototypes.

For this work, the H2M device was integrated into the Caribou system through a three-step approach. First, a custom printed circuit board (PCB), referred to as the chip board, was developed to host the H2M chip and incorporate the necessary passive components for signal conditioning of all control and readout lines. To allow for backside illumination and reduction of the material budget, the chip is mounted at the edge of the board, with approximately half of its area extending beyond the edge without mechanical support. A picture of an H2M chip mounted on a chip board is shown in Figure 8. Second, dedicated FPGA firmware was implemented to provide the device's slow control interface and handle its custom readout algorithms. Finally, a Peary software interface was developed to manage power-up, configuration, control, and data acquisition in a unified and automated manner. A comprehensive set of test procedures was then established to validate the system's correct operation and to assess the performance of the H2M device under both laboratory and test-beam conditions, as detailed below.

3. Laboratory Characterization

This section presents the electrical and functional characterization of the H2M chip, including threshold and ToT calibration using radioactive sources and test pulses. Unless explicitly stated otherwise, the discussed results correspond to the H2M-3 sample, which has a total thickness of 50 µm. However, all the samples presented in the paper underwent the same measurements and yielded similar results, as will be demonstrated.

3.1. Current-Voltage Characteristic

A current-voltage (IV) measurement has been performed, where the full chip p-well and substrate bias currents are recorded as a function of their bias voltages (the sensor bias voltage). The configured chip is kept in the dark, at room temperature, and the p-well and substrate bias voltages are increased simultaneously in steps of 0.02 V. After each step, a delay of 1.5 s is applied to allow for stabilization, followed by 20 measurements from which the mean and RMS are calculated.

Figure 9 shows the IV measurement for samples with different total chip thicknesses. Significantly higher currents are observed in samples with thicknesses below 25 µm, and further systematic studies of thin samples are currently ongoing to have a better understanding of this behavior. All samples show a punch-through current at absolute bias voltages above 4.8 V. At this point, the change in current between two consecutive points exceeds 0.2 µA, and the current begins to increase exponentially. An upper limit of −4.2 V is therefore defined as a safe operational sensor bias voltage for laboratory and test beam measurements.

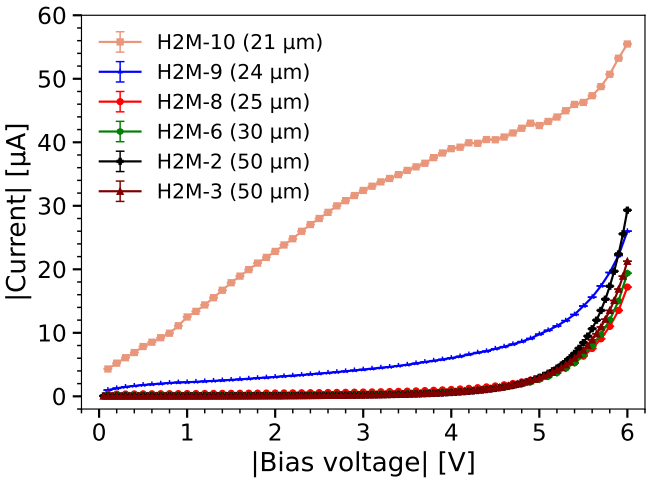


Figure 9: IV measurements for samples with different total chip thicknesses.

3.2. Threshold Equalisation

Variations in the characteristics of circuit elements, and current and voltage supplies over the pixel matrix cause pixel-to-pixel variations of the threshold [16, 17]. If unaccounted for, these variations limit the minimum achievable threshold for noise-free operation. Hence, the analog front-end allows adjustment of the threshold for each individual pixel (`tuning_dac`), as detailed in Section 2.2. To find the optimal `tuning_dac` setting for each pixel, a noise-based equalisation procedure, similar to the one described in [18], is applied. The advantage of this method, compared to those based on calibration-pulse injection, is its insensitivity with respect to gain variations.

The first step of the equalisation procedure is a scan of the parameter `dac_vthr` in units of the DAC (THL DAC), referred to as a threshold scan. This is performed while operating the chip in the dark, configured in counting mode, and acquiring data for a fixed duration at each threshold, without test-pulse injection. Since applying a masking pattern to avoid instabilities has not been found to improve the results, the procedure is applied to the entire matrix at once. Plotting the number of acquired counts as a function of the threshold for each pixel reveals the baseline position of the pixels with respect to the threshold (see top of Figure 10). When the baseline and the threshold are close, the baseline noise is sufficient to toggle the discriminator. Hence, an excess of counts indicates the relative position of the baseline.

This threshold scan is repeated for all possible settings of `tuning_dac`. The baseline of each pixel is determined from a Gaussian fit to the above-mentioned plot, as shown at the bottom of Figure 10. A linear fit is used to describe the relation between the relative baseline position of each pixel and the tuning DAC setting, in order to find the setting closest to a trimming target. This trimming target is chosen such that it is reachable by the largest possible number of pixels. In particular, a pixel is masked when the

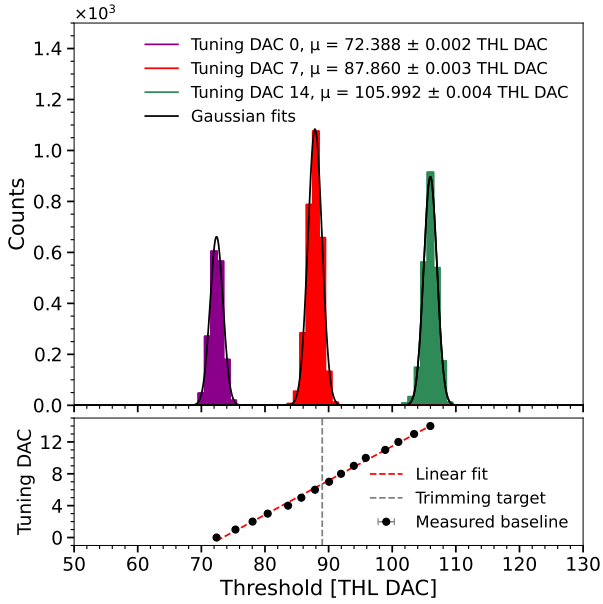


Figure 10: On the top: Occupancy as a function of the threshold for a pixel operated at three different tuning DACs. The mean of the Gaussian fit defines the measured baseline position. On the bottom: the tuning DAC setting as a function of its baseline position. The sensor is biased at −1.2 V.

Gaussian fit does not converge or if the trimming target cannot be reached with any tuning DAC setting.

The distribution of the tuning DAC settings across the matrix, as well as the masked pixels, is shown in Figure 12. The uniform distribution, which makes full use of the available tuning range and results in fewer than 0.3 % of pixels being masked, indicates that no systematic effects are observed that would lead to localised variations of the pixel baselines or discriminator thresholds.

Using the equalised matrix, with the pixels configured as in Figure 12, a new threshold scan in counting mode is performed. As before, the count distribution for each pixel is fitted with a Gaussian, where the mean represents its baseline position and the width indicates its single-pixel noise. Figure 13 shows the baseline distributions of the entire matrix when all pixels are set to the lowest and highest tuning DAC value, as well as after equalisation. The mean $\mu = (94.604 \pm 0.018)$ THL DAC of the equalised distribution corresponds to the chip baseline position, and the width $\sigma = (0.564 \pm 0.015)$ THL DAC represents the threshold dispersion between pixels. Additionally, The single-pixel noise, shown in Figure 11, is measured as the RMS of the curve obtained in the threshold scan (see Figure 10 top), and has a mean of (1.386 ± 0.004) THL DAC after equalisation. This threshold dispersion and single-pixel noise can be converted into electrons after the threshold calibration described in Section 3.3. It yields to a threshold dispersion of approximately 17 electrons and a mean single-pixel noise of 45 electrons RMS when the sensor

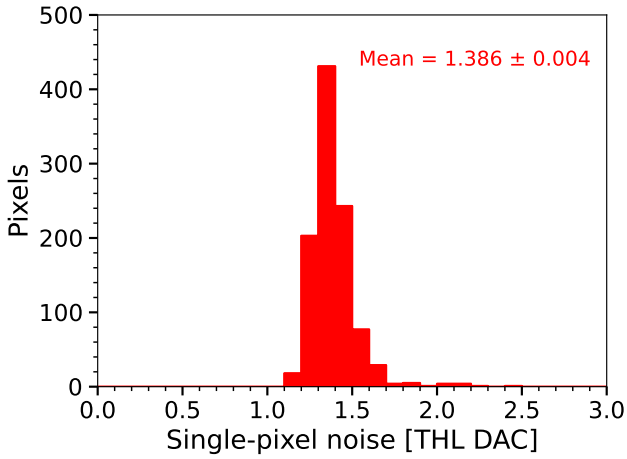


Figure 11: Single-pixel noise distribution of the pixel matrix.

is biased at −1.2 V. The measured single-pixel noise is slightly higher than the simulated (see Section 2.4). Although the presented simulations rely on several assumptions, the differences are still under investigation. In particular, the outliers in the single-pixel noise distribution (Figure 11), together with ongoing measurements, suggest that Random Telegraph Noise [19] is a likely contributing factor.

Figure 14 shows the impact of the sensor bias voltages on the threshold dispersion and single-pixel noise for samples with different total chip thicknesses. No significant differences are observed between samples. The combination of noise and threshold dispersion defines the rate of fake hits at a certain threshold level. Towards lower bias voltages (< 2 V), the detector capacitance increases and noise dominates the fake hit rate. At higher bias voltages, threshold dispersion among pixels becomes more relevant due to the slowdown of the NMOS transistors [20], but the single-pixel noise still dominates the fake hit rate. To overcome the increased threshold dispersion at −4.2 V, a lower value for the `dac_itrim` is used at the cost of masking additional pixels. This results in a smaller step size in the tuning process, covering a smaller threshold dynamic range. Consequently, 5 extra pixels (8 out of 1024) are masked compared to the other bias voltages. Operating the chip at different sensor bias voltages changes the tuning DAC setting by one or two DACs for approximately 20% of the pixels in the matrix.

3.3. Threshold Calibration

To determine the applied threshold in electrons, the chip is calibrated using an iron (^{55}Fe) radioactive source. As a cross-check of the measurements, a 75 µm thick titanium (^{48}Ti) layer is placed beneath the ^{55}Fe to produce X-ray fluorescence from ^{48}Ti. Assuming a mean energy of (3.66 ± 0.03) eV is required to produce an electron-hole pair in silicon [17], the characteristic peaks of ^{55}Fe and ^{48}Ti K_{α_1}

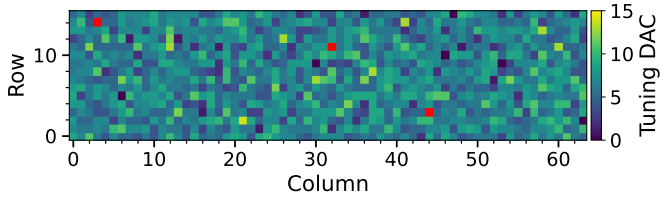


Figure 12: Tuning DAC setting map after equalisation. The three masked pixels are indicated in red, and the sensor is biased at −1.2 V.

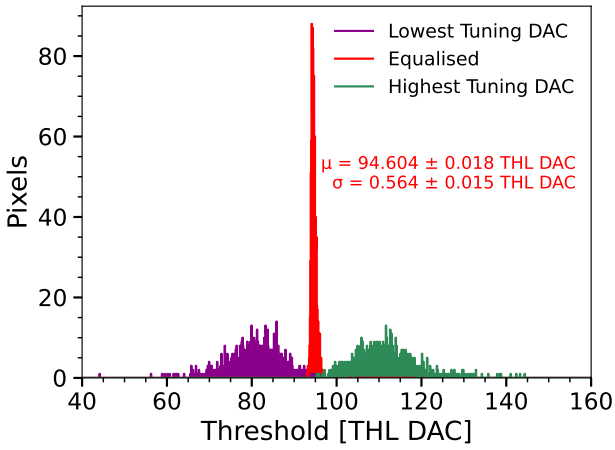


Figure 13: Baseline distribution for the equalised matrix, and the lowest and highest tuning DACs. The sensor is biased at −1.2 V.

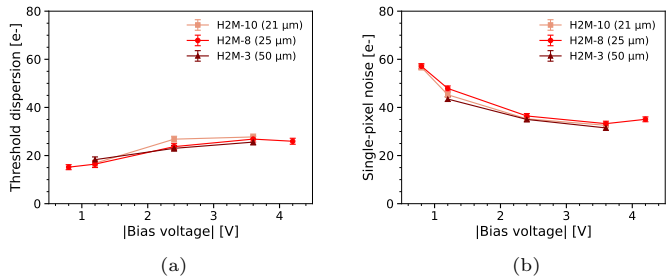


Figure 14: Threshold dispersion (a) and mean single-pixel noise (b) as a function of the bias voltage for samples with different total chip thicknesses.

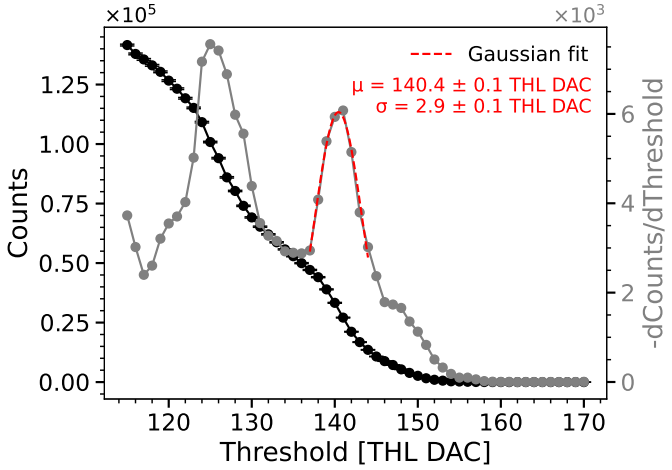


Figure 15: Occupancy (left axis) and its derivative (right axis) for the full matrix as a function of the threshold for the ^{55}Fe source.

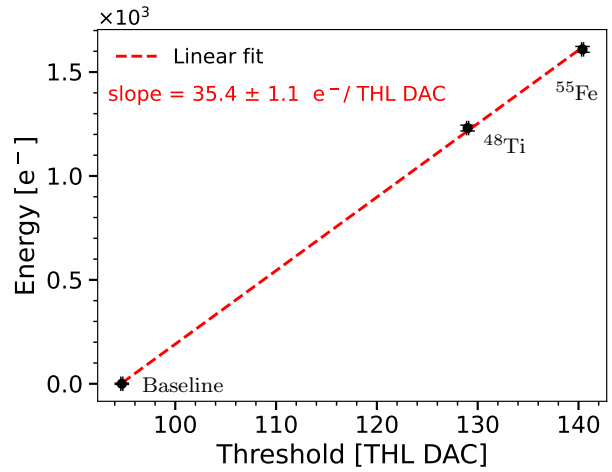


Figure 16: Energy calibration of the threshold using the determined baseline, and the ^{48}Ti K_{α_1} and ^{55}Fe K_{α_1} peaks.

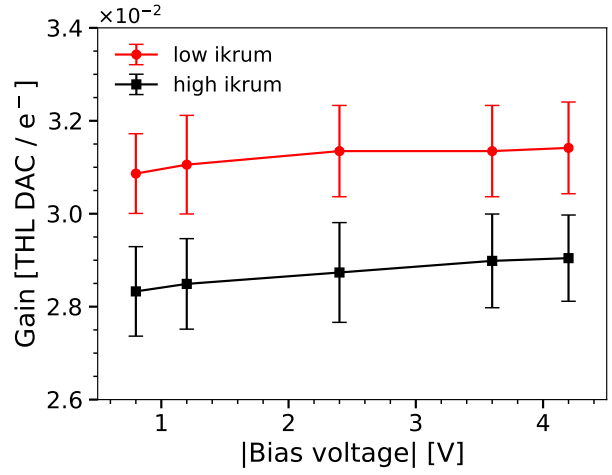


Figure 17: Gain as a function of the bias voltage for low (∼1.65 nA) and high (∼3.75 nA) feedback currents.

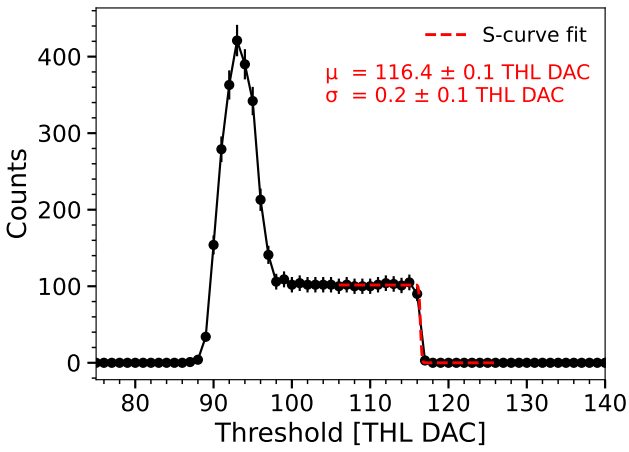


Figure 18: Occupancy for a single-pixel as a function of the threshold with 100 injected test pulses. The injection charge is approximately equivalent to a MIP.

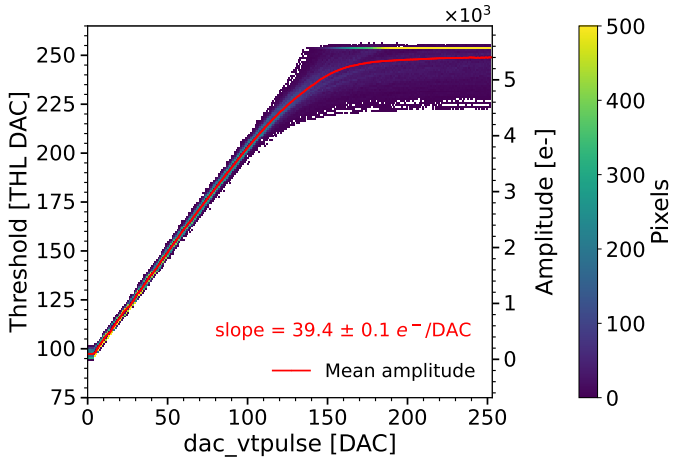


Figure 19: Uncalibrated (left axis) and calibrated (right axis) test pulse amplitudes of the full matrix. The mean amplitude is shown in red.

are expected to correspond to energies of (1609 ± 14) and (1232 ± 10) electrons [21], respectively. To identify these peaks, the threshold is scanned in counting mode.

Figure 15 shows the obtained occupancy and its derivative as a function of the threshold using the ^{55}Fe source. A significant drop in counts occurs at the threshold corresponding to the K_{α_1} energy. This threshold is determined by fitting a Gaussian function to the right visible peak of the occupancy derivative. The mean of the Gaussian fit $\mu = (140.4 \pm 0.1)$ THL DAC represents the energy of the K_{α_1} peak. At a threshold of approximately 148 THL DAC, a small peak is visible in the derivative, corresponding to the K_β energy.

Figure 16 shows the mean values of the Gaussian fit for the ^{55}Fe and ^{48}Ti K_{α_1} peaks as a function of the energy. It also includes the baseline of the chip as determined in Figure 13, which corresponds to a charge deposit of zero electrons. A linear function is fitted, obtaining a calibration factor of (35.4 ± 1.1) electrons per THL DAC. This value corresponds to 111 mV/e$^-$, and agrees with the analog front-end simulations presented in 2.4. The gain is the inverse of the calibration factor, and it is shown in Figure 17 for different bias voltages of the sensor and feedback currents. The *low ikrum* setting corresponds to 1.65 nA, while the *high ikrum* setting corresponds to 3.75 nA. Additionally, no significant differences are observed between samples with different thicknesses.

A second peak is visible at lower amplitudes of approximately 126 DACs in Figure 15. This effect, which is related to loss of signal height due to ballistic deficit [22], as explained in Section 5.1, results in two signal amplitudes for the K_{α_1} peak, depending on the position of the X-ray energy deposition within the pixel cell. Since the ballistic deficit does not strongly affect the position of the right peak, this is utilized for calibration. If spatial information is available, the calibration procedure could be improved by applying different calibration factors to different parts of the pixel cell.

3.4. ToT Calibration

The results of the threshold calibration can now be used to calibrate the ToT response of each pixel individually. To achieve this, test pulses are injected into the analog front-end and the corresponding response is characterised in terms of the amplitude (in threshold units) and ToT.

Similar to the procedure described in Section 3.3, the threshold is scanned in counting mode while injecting test pulses of a given positive amplitude above baseline. A pixel registers one count (i.e., a threshold crossing) per injected test pulse if the threshold lies between the baseline level and the pulse amplitude. If the threshold is lower than the baseline or higher than the pulse amplitude, the pulse is not detected (no threshold crossing), and no count is added. When the threshold is close to the baseline level, noise (or other perturbations such as undershoot in the return to baseline) can cause a higher number of threshold crossings, similar to what is shown in Figure 10.

Figure 18 shows the occupancy as a function of the threshold for 100 test pulses per threshold, using `dac_vtpulse` equals 20 DACs (approximate most probable value for a MIP, about 700 electrons), injected into a single pixel. Above the Gaussian part (baseline), an occupancy of 100 is expected until the maximum pulse height is reached, after which the occupancy drops to zero. In this region, an S-curve function is fitted, with the mean $\mu = (118.5 \pm 0.1)$ THL DAC representing the test pulse amplitude and the width $\sigma = (0.2 \pm 0.1)$ THL DAC reflecting the pixel noise and the amplitude variations across the 100 injected pulses. This procedure is repeated for each pixel and all possible `dac_vtpulse` values, as shown in Figure 19. A linear regression in the region from `dac_vtpulse`

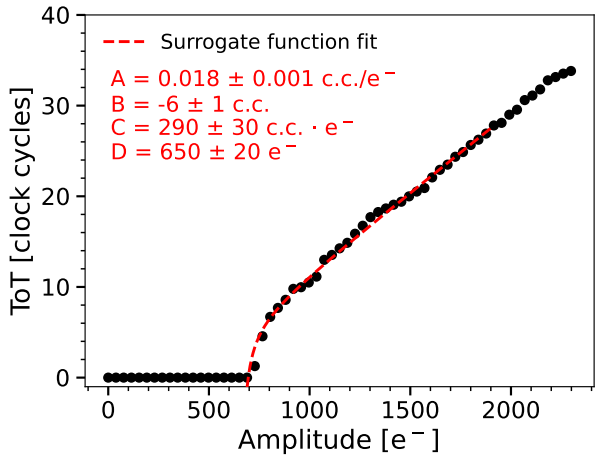


Figure 20: Calibrated ToT response for a single pixel.

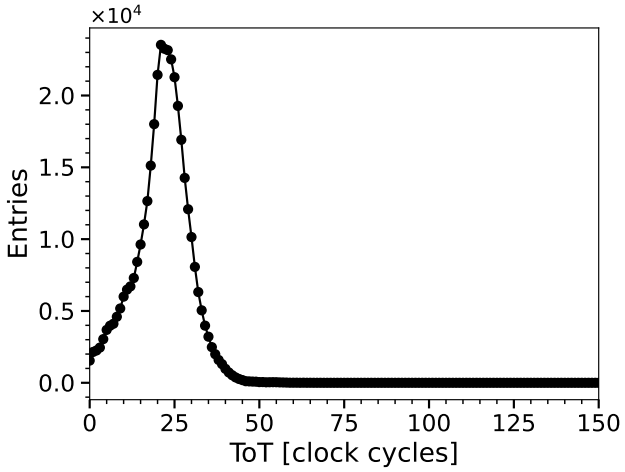


Figure 22: ToT spectrum of the ^{55}Fe source before calibration.

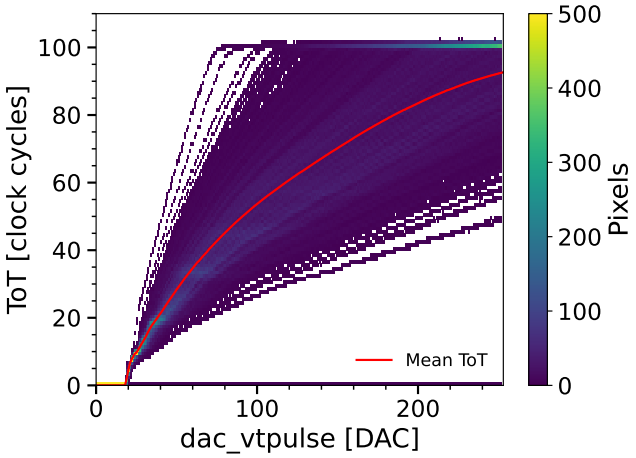


Figure 21: ToT response for the full matrix as a function of the amplitude of the injected test pulses. The threshold is approximately 600 electrons. The mean ToT is shown in red.

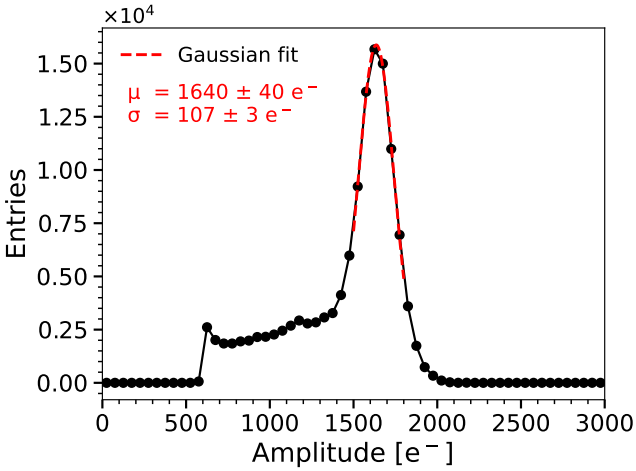


Figure 23: Amplitude spectrum of the ^{55}Fe source after calibration.

between 25 and 100 DAC is used to extract the conversion factor between electrons and DACs. All pixels have a similar linear gain, with maximum measured amplitude reaching the 8-bit `dac_vthr` limit.

The ToT value for each test pulse amplitude is calculated from the average of 100 test pulse measurements at a fixed threshold. This procedure is repeated for each pixel and all possible `dac_vtpulse` values, as shown in Figure 21 for a threshold of approximately 600 electrons. At higher test pulse amplitudes, far from the expected MIP value, the measured ToT values vary between pixels, and the maximum mean value is reached at 1 µs, coinciding with the acquisition frame duration. This large threshold of 600 electrons is chosen in this measurement for consistency with the ToT spectrum of the ^{55}Fe source measurements presented below.

Next, the ToT and amplitude measurements can be combined. Figure 20 shows the ToT as a function of the test pulse amplitude in electrons for a single pixel. A surrogate function

$$\mathrm{ToT} = \mathrm{Ax} + \mathrm{B} + \frac{\mathrm{C}}{\mathrm{x} - \mathrm{D}} \tag{2}$$

is used to fit each pixels data, where x is the test pulse amplitude, A modulates the linear component at higher amplitudes, B adjusts the ToT offset, C represents the curvature part at low amplitudes, and D indicates the offset in amplitude [23]. The inverse of the surrogate function allows the conversion of measured ToT in clock cycles to signal amplitude in electrons.

To verify the calibration procedure, a ToT spectrum of ^{55}Fe is recorded in ToT mode, as shown in Figure 22. The threshold of approximately 600 electrons is the same as used during the calibration. This spectrum is then converted into electrons, as shown in Figure 23, after applying the threshold and ToT calibration. The peak is de-

scribed by a Gaussian with $\mu = 1640 \pm 40$ electrons corresponding to the resolved K_{α_1} amplitude with an accuracy smaller than 5 %. The dominant systematic uncertainty arises from the threshold calibration uncertainty.

The ToT calibration is applied to the test beam reconstructed data acquired in ToT mode down to a threshold of 300 electrons. For lower thresholds, the surrogate function fits (Equation (2)) do not converge —particularly the parameter C, which models the curvature at low thresholds.

4. Methodology for Test Beam Measurements

This section describes the experimental setup and analysis procedure for the test beam measurements, where the H2M performance in MIP detection is studied.

4.1. Experimental Setup

The data have been recorded at the DESY II [24] and CERN SPS [25] Test Beam Facilities employing the ADENIUM [26] and CLICdp Timepix3 [27] beam telescopes, respectively. Each beam telescope comprises three upstream and three downstream reference planes, with the device under test (DUT) in between. These planes are arranged and oriented to optimise the pointing spatial resolution of the beam telescope at the DUT position. With the used configurations and particle beams (∼4.8 GeV electrons at DESY II, and 120 GeV charged pions at CERN SPS), the track resolution at the DUT position for the ADENIUM telescope is found to be ∼3.8 µm, and ∼1.5 µm for the CLICdp Timepix3 telescope, determined using the *GBL Track Resolution Calculator* [28].

At DESY II, the Telepix2 HV-CMOS chip [29] is placed downstream the ADENIUM telescope as a region of interest (ROI) trigger and timing layer, providing a reference time with a resolution below 4 ns. When H2M is operated in ToT and ToA modes, the shutter closes 100 ns after receiving the trigger signal. It opens again once the readout is completed. This results in a readout frequency of approximately 100 Hz, given the selected energy and small trigger area. The use of Telepix2 as a trigger also enables measurements in triggered mode. In this acquisition mode, the H2M shutter is opened for 500 ns, once the external trigger arrives, accounting for the timewalk from H2M. The preset of the pixel counter compensates for the trigger latency of about 2 µs. The binary readout is issued only for hits that are validated by the external trigger.

The AIDA Trigger Logic Unit [30] ensures synchronization among all devices mentioned, and its signal processing dominates the trigger latency. The EUDAQ2 [31] data acquisition framework is used for configuring and reading out all these devices.

At the CERN SPS, the Timepix3 telescope planes provide a timestamp for each track with a resolution of approximately 1 ns. The H2M is read out continuously with a fixed shutter duration of 300 µs in ToT mode and 2.56 µs in ToA mode. With a readout time of 500 µs between shutters, the resulting readout frequency can go up to approximately 2 kHz. Outside the SPS spill, the shutter remains closed, and no data are acquired. The SPIDR DAQ system [32] is used for the readout of the Timepix3 reference planes, and EUDAQ2 for configuration and readout of the H2M chip.

4.2. Reconstruction and Analysis

The Corryvreckan framework [33] is used for the reconstruction and analysis of the test-beam data.

The duration of each H2M frame defines the event time window, and the ADENIUM beam telescope provides matching data if its trigger signal has a timestamp within this window. For the Timepix3 beam telescope, all hits with a timestamp within this window are considered.

The cluster position of recorded adjacent pixel hits is reconstructed using the center-of-gravity algorithm. The number of hits in a cluster determines the cluster size. For events with a cluster size of two in data acquired in ToT mode, η-correction [34] is applied to account for nonlinear charge sharing within the pixel cell.

The trajectory of the particle is reconstructed using the General Broken Lines (GBL) track model [35]. This is performed requesting a hit in each of the six telescope planes and a track χ^2 per degree of freedom smaller than 5. Then, the reconstructed tracks are associated with a cluster on H2M if the projected track position on H2M is within a distance of 52.5 µm (1.5 pixel pitches) of the cluster center.

The hit detection efficiency is defined as the ratio of tracks with an associated cluster on H2M to the total number of tracks going through H2M. To avoid sensor-edge effects, tracks that pass through the outermost columns and rows of the pixel matrix are excluded from the calculation. Tracks crossing a masked pixel or any of its neighboring pixels are also excluded. These masked pixels are identified in the equalization procedure described in Section 3.2 and are disabled during the data taking. The total percentage of masked pixels remains below 0.5 % for all datasets.

The fake-hit rate is also measured in the test beam environment in the absence of beam. For this, the shutter is opened for 100 µs, and the number of hits above threshold is counted.

5. Results

The performance of H2M in test beam measurements is discussed in the following. Unless explicitly stated, the results correspond to the H2M-3 (50 µm) sample, with data recorded at DESY II using the ADENIUM telescope.

5.1. Efficiency

Figure 24 shows the efficiency and fake-hit rate as a function of the threshold for different sensor bias voltages. While the efficiency does not depend significantly on the

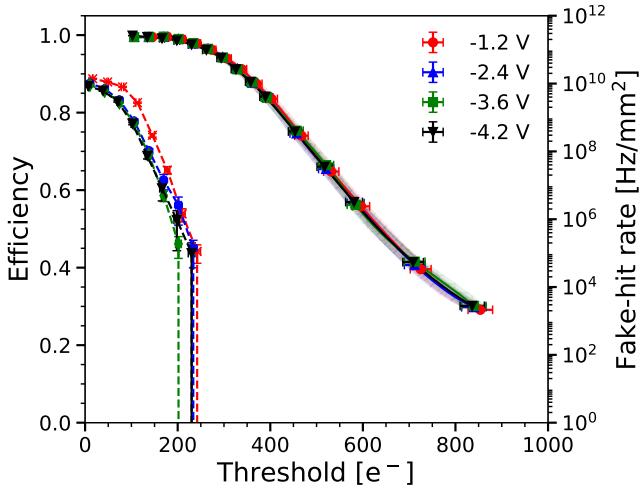


Figure 24: Efficiency (left axis) in solid lines and fake-hit rate (right axis) in dashed lines as a function of the threshold for different sensor bias voltages, obtained in triggered mode.

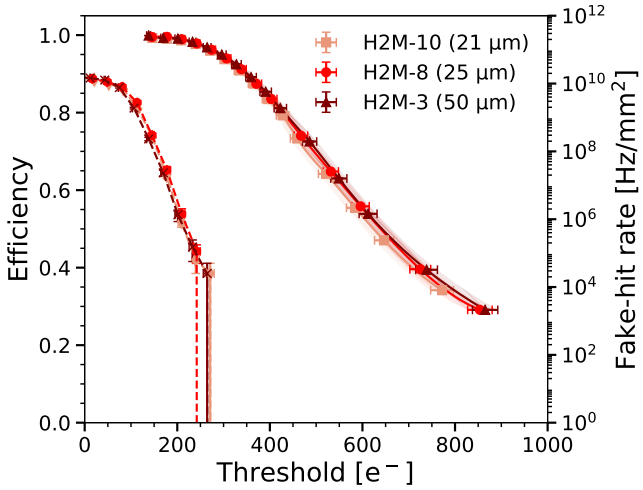


Figure 25: Efficiency (left axis) in solid lines and fake-hit rate (right axis) in dashed lines as a function of the threshold for samples with different total chip thicknesses, obtained in triggered mode. The sensors are biased at −1.2 V.

sensor bias voltage, the fake-hit rate rises at low bias voltages due to the increase in sensor capacitance (see Figure 14). An efficiency of 99.6 % is achieved at a threshold of 144 electrons when the sensor is biased at −3.6 V. The fake-hit rate at this point is around 1.5×10^7 Hz/mm^2 (i.e., 18.4 kHz per pixel), corresponding to a matrix occupancy of approximately nine pixels hit (less than 1% of the active pixels) per 500 ns frame. Longer frame durations employed in ToT and ToA acquisition modes result in higher matrix occupancies, limiting the lowest achievable thresholds and, thus, the chip efficiency.

Table 1 summarizes the highest threshold at which efficiencies above 99 % and 99.5 % are achieved for different bias voltages. The uncertainty in the threshold calibration factor dominates the uncertainty. The maximum threshold allowing for 99 % efficiency is larger than the result from most sensors manufactured in the same process, as comparison with e.g. the studies on APTS [5], where many pitches, designs and bias voltages are investigated, shows. A reason for that might be the larger pitch, which reduces the relative size of the region where charge sharing occurs, as the studies on APTS suggest. In the meantime, the fake-hit rate of H2M is comparably high, due to the threshold dispersion and single-pixel noise discussed in Section 3.2. At higher thresholds, the measured efficiency is comparably low. The reason for this is investigated in the following sections.

Figure 25 shows the efficiency as a function of the thresholds for three chips thinned to different total thicknesses obtained in triggered mode. No deterioration in efficiency or fake-hit rate is observed between samples for thresholds below 450 electrons. In particular, all samples achieve an efficiency above 99 % at a threshold of approximately 200 electrons. At this point, the measured fake-hit rate is around 4×10^5 Hz/mm^2, corresponding to a matrix occupancy of fewer than one pixel hit per 500 ns frame.

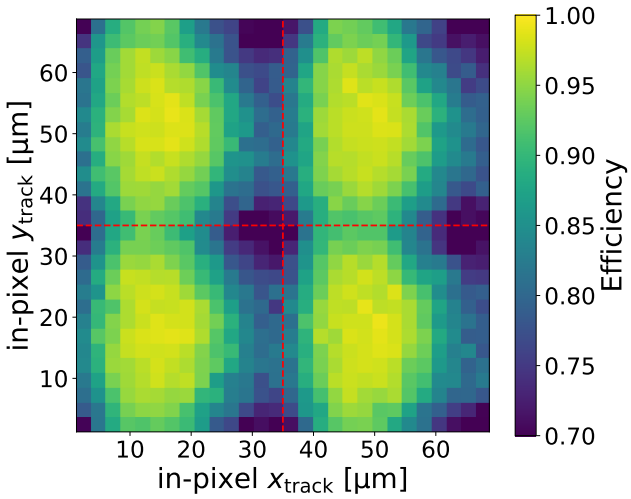
The short shutter durations achieved in triggered mode allowed operation of the chip at low thresholds. However, measurements in ToT and ToA mode have been performed with longer shutters (see Section 4.1), which increases the probability of fake hits and thus requires a higher operational threshold of around 200 electrons. The results achieved at these higher thresholds are compatible with Table 1, where an efficiency of 99 % is measured.

Table 1: Highest threshold at which efficiencies above 99 % and 99.5 % are achieved for different bias voltages in triggered mode.

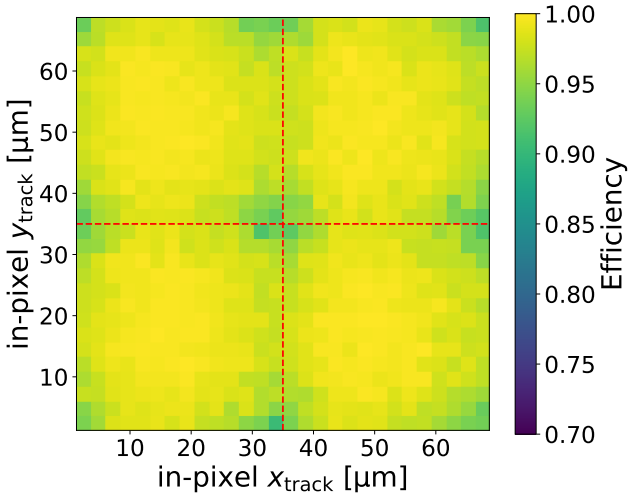
	−1.2 V	−2.4 V	−3.6 V	−4.2 V
99 %	(205 ± 6)	(208 ± 6)	(208 ± 6)	(208 ± 6)
99.5 %	(179 ± 6)	(178 ± 6)	(182 ± 6)	(183 ± 6)

5.1.1. Non-uniform In-pixel Response

At higher thresholds, a non-uniformity in the in-pixel response is observed. This is shown in Figure 26a, where the right part of the pixel cell shows a lower efficiency than the rest of the pixel. The efficiency drop location coincides with the position of the 4 µm wide n-well of the analog front-end shown in Figure 2. This behavior is attributed to local potential wells of very shallow amplitude at the interface between the low-dose n-implant and the deep p-well [7, 8]. The very low electric field where some of these local potential wells are located, allows them to slow down charge collection, as visible in the in-pixel representation of the mean arrival time discussed in Section 5.5 (Figure 35). This affects not only the charge carriers generated near the interface, but also those originating from the full epitaxial layer due to the drift path of charge carriers (see Figure 1). In contrast to the n-well of the analog front-end, the n-wells of the digital logic are thin, as sketched in Figure 3,



(a) High *ikrum*, −1.2 V, 330 e$^-$.



(b) Low *ikrum*, −3.6 V, 224 e$^-$.

Figure 26: Efficiency maps projected onto four pixels for different feedback currents, sensor bias voltages, and thresholds. The pixel cell boundary is marked with a dashed red line.

mitigating this effect.

This slow charge collection leads to a mismatch between the charge collection time and the fast response of the CSA. As a result, the CSA begins to reset before the full charge is collected, resulting in a reduced signal amplitude (ballistic deficit) and, consequently, a drop in the hit efficiency over a large part of the pixel cell.

In Figure 26a, the sensor is operated at low bias voltage, −1.2 V, with a high threshold of 330 electrons and high *ikrum* to aggravate this effect. Increasing the bias voltage enhances the electric field within the sensor, and reducing the feedback current effectively increases the integration time of the CSA. Both factors contribute to reducing the ballistic deficit, and improving the hit detection efficiency. Figure 26b shows a homogeneous in-pixel efficiency map at low *ikrum*, a threshold of 224 electrons, and a sensor bias of −3.6 V.

To fully understand this effect, detailed simulations of the sensor and front-end have been performed, which are presented in Section 6.

5.2. Cluster Size

The relatively large pixel pitch, thin epitaxial layer, and the *modified with n-gap* sensor layout minimize charge sharing between neighboring pixels. Lowering the threshold increases the probability of inducing a signal in adjacent pixels, making two-pixel clusters more frequent, as shown in Figure 27. For a threshold of 177 electrons, the average cluster size is (1.19 ± 0.01), with 18% of events with two-pixel clusters. Figure 28 shows the mean cluster size as a function of the threshold. At lower bias voltages, the electric field is reduced, leading to slower charge collection. This leaves more time for charge carrier diffusion, resulting in a slight increase in cluster size. Towards higher detection thresholds (above 700 electrons), the mean cluster size increases due to the detection of delta-rays, which are more likely to deposit large amounts of energy.

Although the right side of the pixel cell experiences slower charge collection as discussed in Section 5.1, this does not impact the charge sharing. As a result, the in-pixel cluster-size maps remain symmetric, as shown in Figure 29. Near the collection electrode, charge carriers are collected by the nearest pixel. A slight increase in cluster size is observed only at the boundaries of the pixel cell.

Figure 30 shows the mean cluster size as a function of the threshold for different total chip thicknesses. Similar to the hit detection efficiency results, there is no effect from thinning on the cluster size.

5.3. Spatial Resolution

Figure 31 shows the spatial residuals between the track impact position and the H2M hit position, for a threshold of 224 electrons. At higher thresholds (>400 electrons), the residual distributions in the column direction become asymmetric due to reduced efficiency in the right part of the pixel cell (see Figure 26). The RMS is calculated over

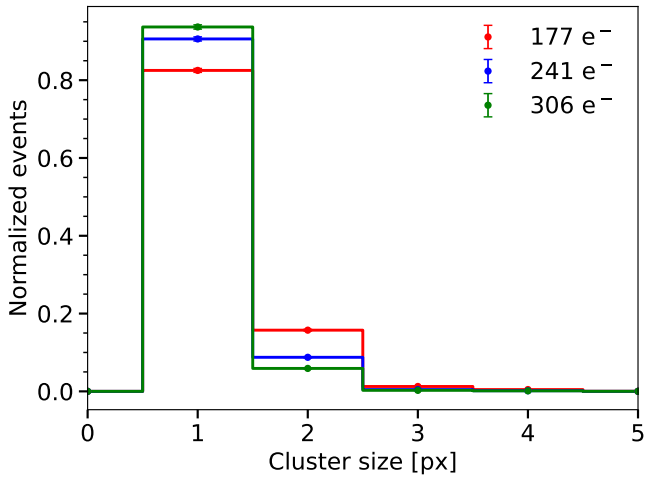


Figure 27: Cluster size distributions for different thresholds. The sensor is biased at −1.2 V.

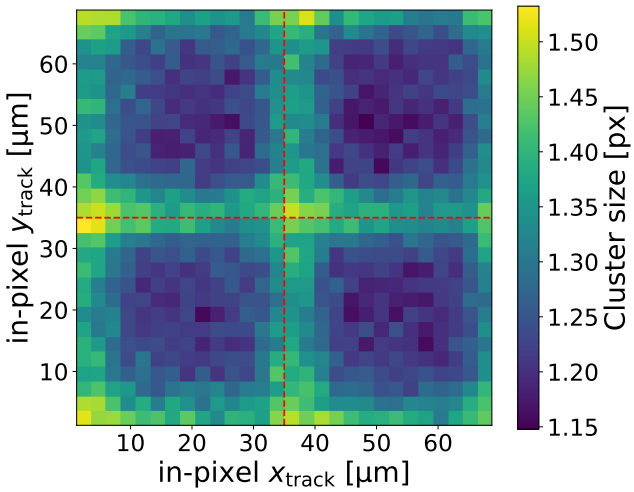


Figure 29: Cluster size map projected onto four pixels. The sensor is biased at −1.2 V and the threshold is 224 electrons. The pixel cell boundary is marked with a dashed red line.

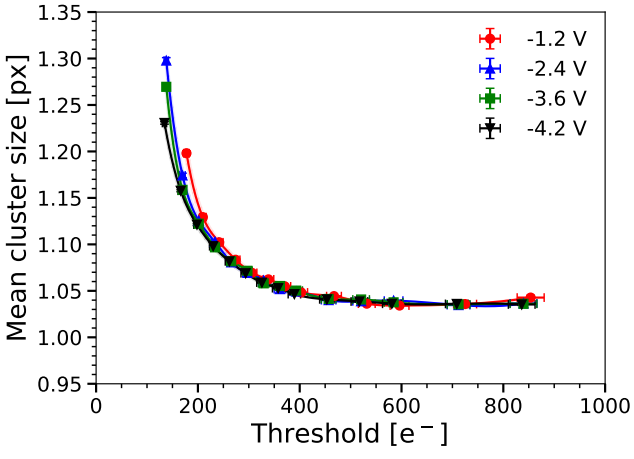


Figure 28: Cluster size as a function of the hit detection threshold for different sensor bias voltages.

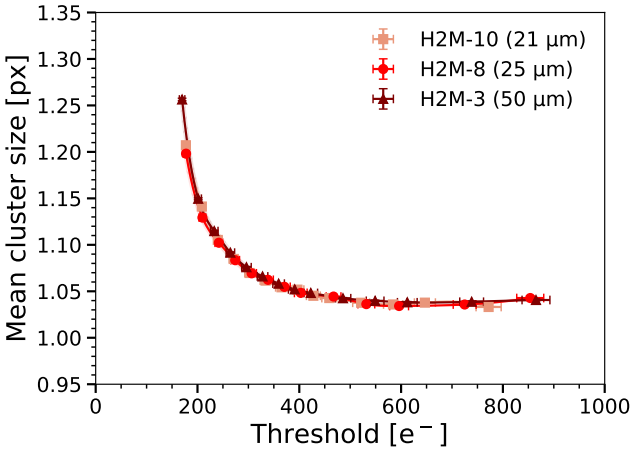


Figure 30: Cluster size as a function of the threshold for samples with different total chip thicknesses. The sensor is biased at −1.2 V.

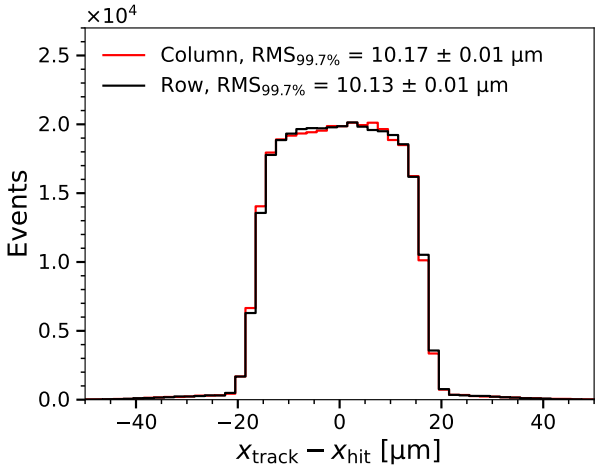


Figure 31: Spatial residuals in column and row direction recorded in ToT mode using the CLICdp Timepix3 beam telescope. The sensor is biased at −3.6 V and the threshold is 224 electrons.

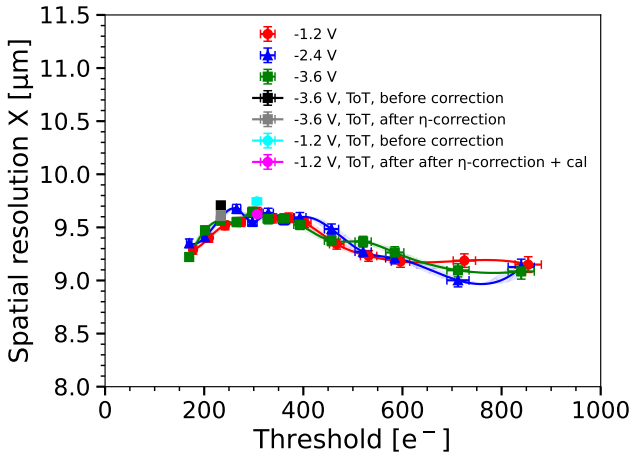


Figure 32: Spatial resolution as a function of the threshold for different sensor bias voltages in triggered and ToT modes.

the central 99.7 % of the shown distribution ($RMS_{99.7\%}$). By subtracting the tracking resolution of the beam telescope from this value, the intrinsic spatial resolution of H2M can be determined.

Figure 32 presents the H2M spatial resolution in the row direction as a function of the threshold for different sensor bias voltages in triggered mode. A few additional points measured in ToT mode, for which calibration and η-correction have been applied, are also included. The pixel-to-pixel variations and nonlinear charge sharing inside the pixel cell lead to a deterioration of the reconstructed cluster position when these corrections are not applied. Due to the fact that a majority of the interactions have a cluster size of one, no significant difference is found between the spatial resolution in ToT mode (with 8 bits) and triggered mode (binary readout), and a spatial resolution close to the pitch $/\sqrt{12} \sim 10.1\,\mu m$ is measured. Moreover, no significant differences between sensor bias voltages are observed.

The increase in charge sharing at a threshold below 240 electrons leads to an improvement of the spatial resolution. In particular, a spatial resolution of $(9.3 \pm 0.1)\,\mu m$ is measured at a threshold of (178 ± 5) electrons. No significant dependence on the sensor bias voltages is observed. Between approximately 240 and 400 electrons, the spatial resolution remains constant at about 9.7 µm, corresponding to a range dominated by single-cluster events. Above 400 electrons, the above-mentioned asymmetric residuals (caused by the lower efficiency) result in a reduced $\sigma_{RMS,3\sigma}$, and therefore in an improvement of the spatial resolution.

5.4. Signal Distribution

The seed pixel is defined as the pixel with the highest signal within a cluster. Figure 33 shows the seed pixel charge distribution fitted using a convolution of a Landau and Gaussian function.

An MPV of (680 ± 20) electrons has been measured, which is higher than measured in other prototypes from the same submission [5]. The discrepancy could be attributed to the ToT calibration procedure, which assumes a uniform in-pixel response of the sensor. During calibration, test pulses are injected and their amplitudes and corresponding ToT values are measured to obtain a single calibration function per pixel (see Section 3.4). However, when a charged particle passes through the sensor, the non-uniform in-pixel response, which leads to slower charge collection and smaller amplitudes, also results in larger ToT values in the right part of the pixel. This is shown in the in-pixel seed-pixel ToT map of Figure 34. This occurs because, for lower signal amplitudes, the feedback current of the CSA is not fully saturated, effectively reducing the *ikrum*, slowing the pulse discharge, and increasing the measured ToT value. As a consequence, the MPV of the seed-pixel charge distribution shifts towards larger values.

5.5. Time Resolution

The time residuals are defined as the difference between the trigger timestamp ($t_{trigger}$) and the H2M hit timestamp (t_{hit}), measured when the chip operates in ToA mode. For clusters with more than one pixel, t_{hit} is assigned to the earliest pixel timestamp within the cluster.

Figure 35 shows the in-pixel ToA projected onto four pixels. Under the position of the n-wells of the analog front-end, there is a slower charge collection, leading to lower efficiency due to the ballistic deficit, as explained in Section 5.1. In contrast to the efficiency pattern, the ToA asymmetry remains visible even at the lowest thresholds and high sensor bias voltages.

This non-uniform in-pixel response results in non-Gaussian time residuals. An example is shown in Figure 36

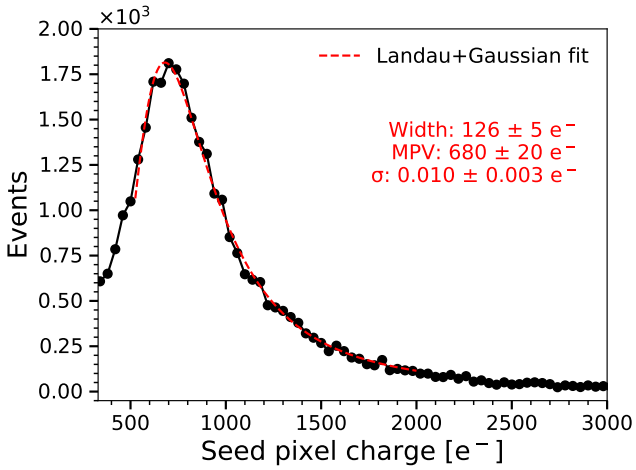


Figure 33: Seed-pixel charge distribution after calibration. The sensor is biased at −1.2 V, and the threshold is approximately 300 electrons. The scale (width) and Most Probable Value (MPV) of the Landau distribution, and the standard deviation of the Gaussian component (σ) in the Landau-Gaussian convolution fit, are shown.

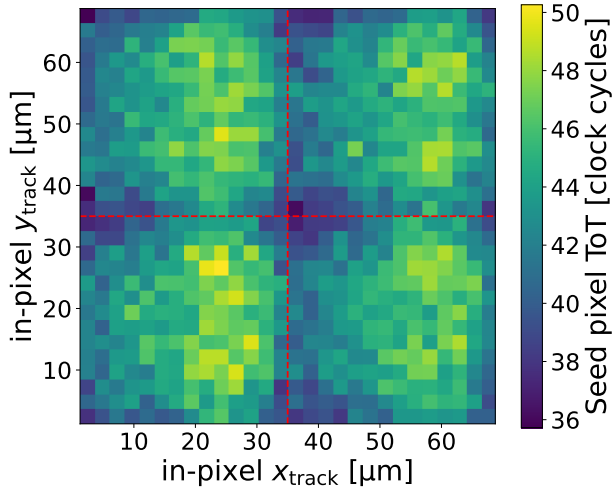


Figure 34: Mean seed pixel ToT map projected onto four pixels. The sensor is biased at −1.2 V, the threshold is 300 electrons, and a low *ikrum* is set. The pixel cell boundary is marked with a dashed red line.

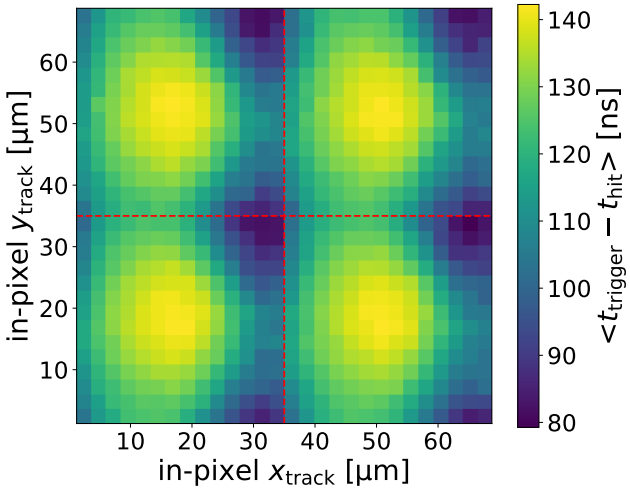


Figure 35: ToA map projected onto four pixels. The sensor is biased at −3.6 V, the threshold is 224 electrons, and a low *ikrum* is set. The pixel cell boundary is marked with a dashed red line.

for a threshold of 224 electrons and a sensor bias voltage of −3.6 V, where a long left-side tail, more pronounced than what would be expected from time walk, is visible. The truncated RMS ($\sigma_{t_{trigger}-t_{hit}}$), computed within the histogram range shown, is (28.4 ± 0.2) ns. Since the time resolution of the reference detector is negligible compared to that of the DUT resolution, the measured $\sigma_{t_{trigger}-t_{hit}}$ is attributed directly to the DUT time resolution.

Figure 37 shows the $\sigma_{t_{trigger}-t_{hit}}$ as a function of the in-pixel position. Regions with slower charge collection exhibit worse time resolution, whereas a faster and more uniform charge collection is observed near the pixel center. The opening in the n-well hosting the analog front-end (see Figure 2) affects the time resolution locally, and an improvement in the time resolution is visible in that region.

Figure 38 shows the time resolution of H2M as a function of the threshold, for two different total chip thicknesses and sensor bias voltages. While no significant difference is observed between samples, increasing the sensor bias voltages improves the time resolution by enhancing the charge collection speed.

6. Simulation

To gain a better understanding of the sensor and its behaviour when combined with the circuit of the H2M, simulations were performed. A simulation flow combining technology computer-aided design (TCAD), Monte Carlo and circuit simulations is employed.

This simulation procedure focuses on qualitatively reproducing the behaviour of the chip by using generic configuration such as typical transistors or capacitors without

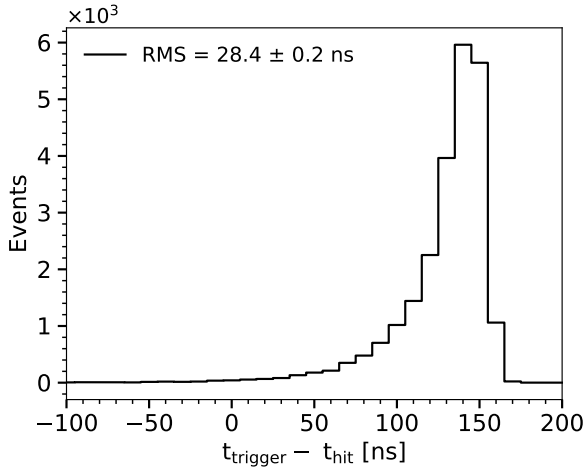


Figure 36: Time residuals distribution. The sensor is biased at −3.6 V, the threshold is 224 electrons, and a low *ikrum* is set.

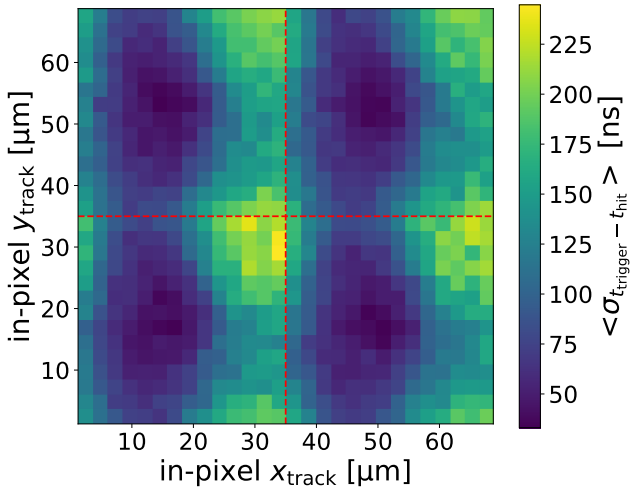


Figure 37: Time residuals RMS map projected onto four pixels. The sensor is biased at −3.6 V, the threshold is 224 electrons and a low *ikrum* is set. The pixel cell boundary is marked with a dashed red line.

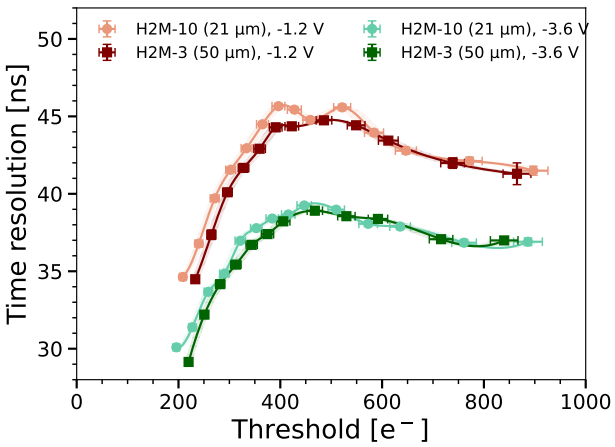


Figure 38: Time resolution as a function of the threshold for two different total chip thicknesses and sensor bias voltages.

considering variations due to e.g. manufacturing. Though the method could in theory be extended and fitted with measurement results, results presented here are only qualitative and obtained independently from measurement in order to demonstrate root cause and understand general trends.

6.1. TCAD Sensor Simulations

The first step of the simulation consists of computing the electric field inside the sensor with a TCAD simulation. The Sentaurus [36] framework from Synopsys is used to first build a 3D finite element model of the sensor and solve Poisson and continuity equations on this mesh. Considering the goal of understanding the asymmetric response, the TCAD simulation includes not only the collection electrode, epitaxial layer, p-well and deep p-well but also the n-wells of the circuitry inside the deep p-well, as it represents the main asymmetric component of the layout. The layout of these n-wells used in the simulation is shown in Figure 2; it is slightly simplified compared to the chip layout to eliminate small features of the design. The n-wells hosting the in-pixel circuit are biased at 1.2 V while the p-well and substrate are biased at −1.2 V and the collection electrode at 0.8 V.

To reproduce the asymmetric features of the sensor design, the boundary conditions of the TCAD simulation need to be carefully set. Ideally, to model the real device included in a large matrix, a pixel with periodic boundary conditions should be simulated, but to achieve numeric convergence, a 2 by 2 pixel matrix composed of a central pixel surrounded by parts of its 8 neighbors was simulated with mirror boundary conditions instead and later cropped to the central pixel. A fine mesh was used in the central pixel, especially close to the collection electrode and in the deep n-type implant, to allow accurate electric field simulation in these sensitive regions.

Transient TCAD simulations in [8] demonstrated that including the n-wells from the in-pixel circuitry leads to a significant slow down of the charge collection due to low amplitude potential wells in the deep n-type implants underneath the n-wells as can be inferred from the electrostatic potential. This effect appears in combination with the large pixel pitch creating a region of very low lateral electric field and the specific layout of the n-wells. It will be shown later in Section 6.5 that minor layout modifications can have a significant impact on charge collection slowdown.

The electric field and doping concentration (for computing mobility and lifetime) are extracted from the output of the TCAD simulation to be used in a Monte Carlo simulation.

6.2. Monte Carlo Sensor Simulations

Monte Carlo simulation of the H2M is executed with the Allpix Squared framework [37]. Different configurations are simulated, either using a uniform charge deposition in order to ignore stochastic effects and emphasize the intrinsic response of the sensor or using Geant4 [38, 39, 40] to simulate realistic energy deposition corresponding to a perpendicularly incident 5 GeV electron beam or X-rays produced by a ^{55}Fe radioactive source. The charge carrier mobility is computed using the extended Canali model [41] and recombination with the combined Shockley-Read-Hall-Auger model [42, 43, 44]. To increase computation speed, the signal induced on the collection electrode is not computed by applying the Shockley-Ramo theorem [45, 46] but simply by registering the time of arrival of electrons at the collection electrode which is a reasonable assumption for small collection electrode sensors.

Several studies [47, 48] demonstrated that Allpix Squared can accurately simulate silicon sensors. A dedicated comparison of transient TCAD and Allpix Squared simulations of the H2M showed that Allpix Squared reproduces the charge-collection slow down induced by the n-well layout [8].

Figure 39 and 40 show respectively the average collected charge and the average collection time (10% to 90%) of the seed signal as a function of the particle impinging position in the pixel. For this simulation, the sensor is biased at $-1.2\,\mathrm{V}$ and $63\,\mathrm{e^-\,\mu m^{-1}}$ are deposited uniformly along lines orthogonal to the sensor surface. Deposition steps are small enough to consider the deposition as perfectly uniform in depth, and such injection is performed with a step of 0.1 µm on both x and y axes before averaging it over bins of 1 µm in Figures 39 and 40. Except in the pixel edge where charge sharing occurs, the collected charge after the 500 ns simulation varies between $640\,\mathrm{e^-}$ and $700\,\mathrm{e^-}$, with the visible pattern corresponding to the part of the charge deposited within the circuit n-wells (and thus not collected by the collection electrode). On the other hand, the collection time exhibits a clear asymmetric pattern, with the right part of the pixel where the large analog n-well is located having a significantly slower rise time.

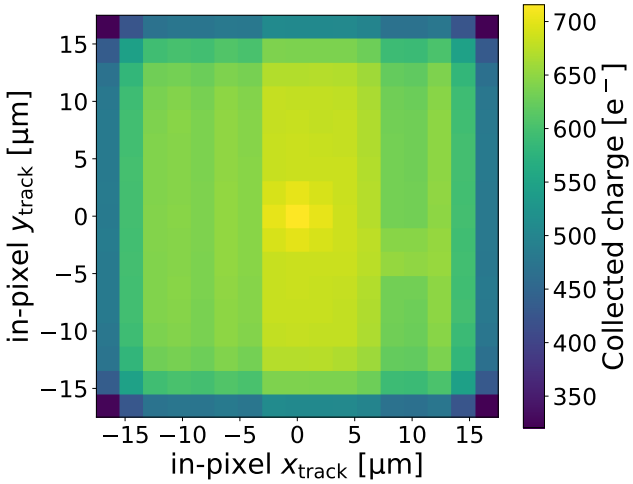


Figure 39: Simulated seed pixel average collected charge map projected onto one pixel for a uniform $63\,\mathrm{e^-\,\mu m^{-1}}$ deposition. The sensor is biased at $-1.2\,\mathrm{V}$ and an integration time of 500 ns is used.

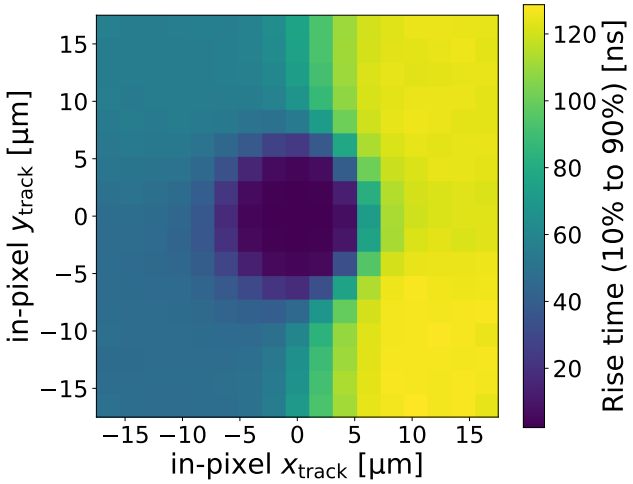


Figure 40: Simulated seed pixel average charge collection time (10% to 90%) map projected onto one pixel for a uniform $63\,\mathrm{e^-\,\mu m^{-1}}$ deposition. The sensor is biased at $-1.2\,\mathrm{V}$.

The combination of both information suggests that the observed asymmetric efficiency pattern might not be caused by charge loss in the sensor but rather by the slower collection leading to partial reset of the amplifier output before full charge collection, i.e. the ballistic deficit.

6.3. Analog Front-End Simulations

Ballistic deficit is an effect intrinsically related to the front-end circuit and to the shape of the signal induced by charge carrier motion. Considering the non-linearity of the H2M front-end, a simulation of the circuit was executed in a dedicated electronic circuit simulator, Spectre [49].

In order to simulate a comparable number of events as in measurements in a reasonable time, only the charge sensitive amplifier with Krummenacher feedback is simulated, without including noise. The signal at the input of the front-end is extracted from the Allpix Squared simulation with a script and injected in the form of a current source in parallel to a 3 fF capacitance representing the sensor capacitance plus parasitic capacitance [5].

By simulating the analog front-end for two inputs corresponding to an ^{55}Fe simulation in Allpix Squared, one from the 'fast' left part of the pixel and one from the 'slow' right part, [8] showed that even though both events collect the same number of charges, the amplitude at the output of the CSA is significantly smaller for the 'slow' event than the 'fast' one. This simulation confirms that ballistic deficit is a relevant effect in this chip. Additionally, the two different collection times identified in Figure 40 will lead to two different magnitudes of ballistic deficit, thus causing a double peak in the iron spectrum. This double peak is indeed present in the measured ^{55}Fe spectrum of Figure 15 but also in the simulated one in Figure 41. As expected by the use of a typical simulation without any adjustments, more in-depth comparison of the simulated and measured spectra yields significant differences both in the peak amplitude and overall shape of the spectrum.

The ballistic deficit impacts mainly the amplitude of the signal, but since the total integrated charge is quite uniform (see Figure 39), it has no large effect on the time over threshold (for thresholds low enough compared to the input charge). Therefore only a single peak is observed in the iron spectrum acquired in TOT mode (Figure 23).

6.4. H2M Efficiency Simulation

The H2M efficiency in a test beam can be estimated by combining TCAD, Monte-Carlo and analog circuit simulations. The fit of the right K_α peak (at ≈ 600 mV) of the iron spectrum from Figure 41 is used for calibration of the simulated amplitude into electrons. Signals are smeared with a noise corresponding to the measured one in Section 3.2 (45 e^- single-pixel noise and 17 e^- for threshold dispersion) and a threshold of 330 e^- is used. A smearing of 3 µm on the position of the particle is also applied to account for the telescope resolution. The obtained efficiency map is shown in Figure 42 and the average efficiency is

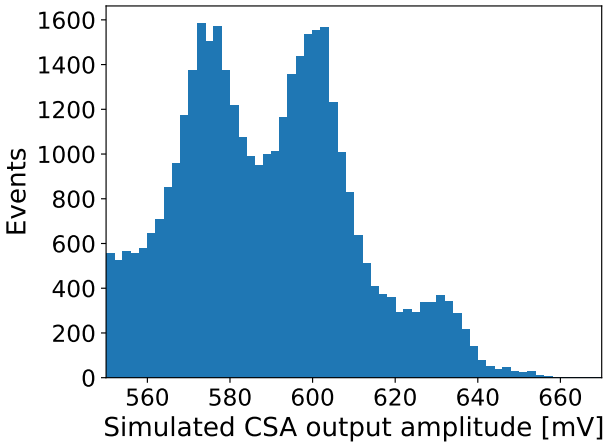


Figure 41: Simulated single pixel amplitude spectrum at the output of the CSA for a ^{55}Fe source. The sensor is biased at −1.2 V and the front-end with high *ikrum*.

projected along each axis and compared to measurements. The efficiency pattern reproduces approximately the measured one from Figure 26a. The simulated average efficiency is 86% while the measured one is 88%. The simulated efficiency pattern and average efficiency agree with the measurements within 3%.

6.5. Optimization of Circuit Layout

The process described above, which leads to slower charge collection, is expected to be dependent on the layout of the in-pixel electronics. A design optimization was thus performed in simulation to determine if a more uniform charge collection can be achieved while maintaining the pitch and functionalities. This simulation employed the same simulation workflow, and the only change is the position and shape of the n-wells in the analog circuitry, as shown in Figure 43. The area of n-wells is preserved compared to the H2M but no redesign of the analog front-end to fit with this new n-well shape was attempted and simulation use the same circuit as before.

Shape and position are chosen such that the junction between the analog circuit n-well and the p-well in the direction of the collection electrode is located in a region of high electric field (as close as possible to the collection electrode). Other variations of the layout not following this condition were also simulated and produced as expected worse results compared to the one presented here.

The resulting signal rise time is shown in Figure 44. Despite the analog n-well being even wider than in the original H2M, the proposed layout features a three times lower maximal rise time: 42 ns to be compared to the 129 ns from Figure 40.

The simulated performance of this new n-well layout can be estimated with the same method as in Section 6.4, the results are presented in Figure 45. As expected, significant improvement can be observed, the efficiency is more

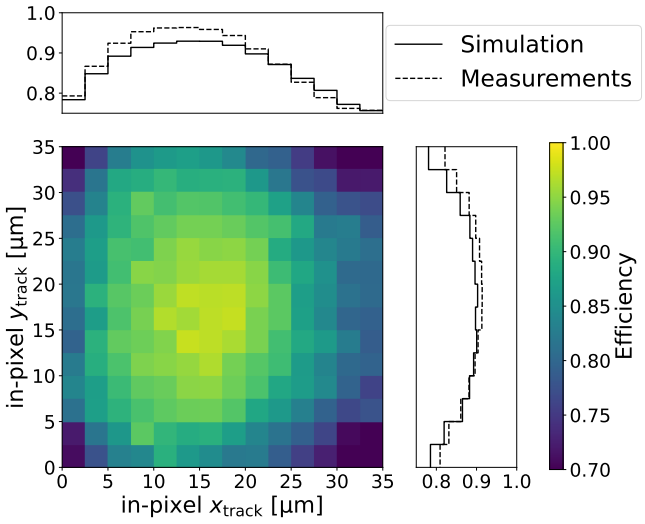


Figure 42: Simulated efficiency maps of the H2M for one pixel and its average projected on each axis. The sensor is biased at −1.2 V, and a high *ikrum* is used. The projected average is compared to measured data from Figure 26a.

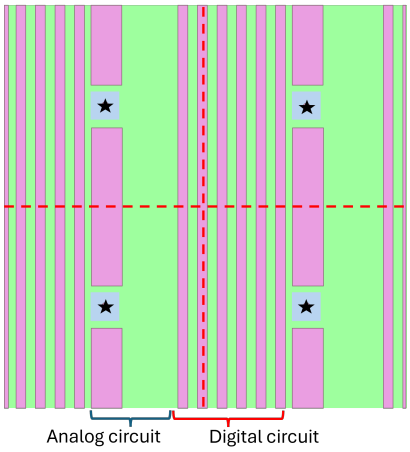


Figure 43: Schematic of the n-wells and p-wells positions within the deep p-well in four pixels for the proposed layout discussed in Section 6.5. The position of analog front-end, and digital logic is also indicated.

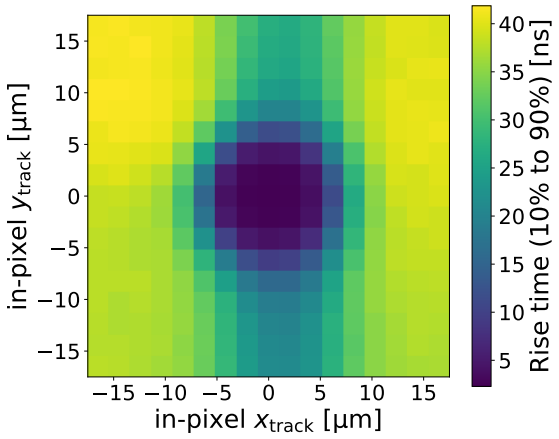


Figure 44: Simulated seed pixel average charge collection time (10% to 90%) map projected onto one pixel for a uniform 63 e^- µm^{-1} deposition. The n-well layout presented in 6.5 is used and the sensor is biased at −1.2 V.

uniform than in Figure 42 and the average simulated efficiency reaches 95%, to be compared to the 86% simulated with the H2M layout.

7. Conclusion & Outlook

The H2M chip has proven to be fully functional in all four data acquisition modes. All relevant circuit elements work in the foreseen way, and their performance matches specifications and expectations from analog front-end simulations. This shows that the strategy behind the project — porting the architecture of a hybrid pixel detector readout chip into a MAPS, and applying the digital-on-top workflow — works, so that the development of future MAPS may profit from the gained experience. It also validates the performance of the building blocks designed to be a part of the new, compact digital cell library.

The characterization of the H2M chip has shown that the chip can be thinned down to a total thickness of 21 µm, without significant effect on the key performance characteristics. A MIP detection efficiency above 99 %, is reached for thresholds below 205 electrons, which is among the best results from prototypes produced in the same process, and attributed to the large pixel pitch. However, single-pixel noise and threshold dispersion are found to be on the order of 40 and 25 electrons, respectively, and depend on the bias voltage. This will limit the lowest attainable operation threshold, depending on limits on the fake-hit rate. The spatial hit resolution is (9.3 ± 0.1) µm at a threshold of 178 electrons, and limited by a combination of the pitch and the cluster size, as expected for sensors employing the modified with a gap layout. The temporal hit resolution

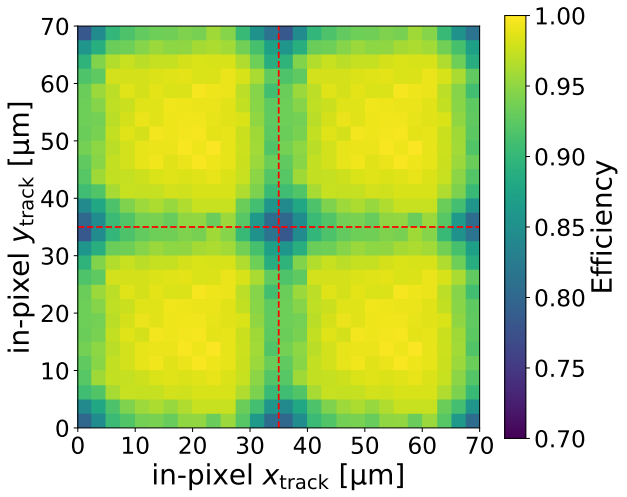


Figure 45: Simulated efficiency maps projected onto four pixels for the proposed modified n-well layout. The simulated sensor is biased at $-1.2\,\mathrm{V}$, high *ikrum* and a threshold of $330\,\mathrm{e}^-$ are used. The pixel cell boundary is marked with a dashed red line (all four pixels contain the same simulated data).

goes down to $(28.4 \pm 0.2)\,\mathrm{ns}$ at a threshold of 224 electrons, and is limited by the non-uniformity of the in-pixel response.

The cause for this non-uniformity is thought to be local potential wells slowing down the charge collection in regions of very low electric field, caused by the layout of n-wells within the deep p-well. This leads to an asymmetric temporal response as a function of the in-pixel position. Due to the fast response of the CSA this reduces the pulse amplitude, which results in an accelerated loss of efficiency as a function of threshold. The presented simulation procedure shows to be capable of reproducing this effect qualitatively and is an indispensable tool to understand the underlying mechanism. The effect can be mitigated by reducing the feedback current of the CSA, or by re-positioning the large n-wells in a new design, as the simulations suggest.

Overall, the characterization of the H2M prototype led to a better understanding of MAPS produced in the 65 nm CMOS imaging process of Tower Partner Semiconductor Co. (TPSCo). This knowledge is fundamental to optimize the performance of future sensors — especially those with a comparably large pitch.

Acknowledgments

Some measurements leading to these results have been performed at the Test Beam Facility at DESY Hamburg (Germany), a member of the Helmholtz Association (HGF).

The developments presented in this contribution are performed in collaboration with the CERN EP R&D program on technologies for future experiments.

This project has received funding from the European Union's Horizon 2020 Research and Innovation program under GA no 101004761.

References

[1] E. Berger, M. Demarteau, J. Repond, L. Xia, H. Weerts, A. N. L. U. States), CLIC CDR - physics and detectors: CLIC conceptual design report (Feb 2012). doi:10.2172/1035023.

[2] A. Abada, M. Abbrescia, S. S. AbdusSalam, I. Abdyukhanov, J. Abelleira Fernandez, et al., FCC-ee: The Lepton Collider, Eur. Phys. J. Spec. Top. 228 (2019) 261–623. doi:10.1140/epjst/e2019-900045-4.

[3] W. Snoeys, Monolithic pixel detectors for high energy physics, Nucl. Instrum. Methods Phys. Res. A 731 (2013) 125–130, Pixel 2012. doi:10.1016/j.nima.2013.05.073.

[4] W. Snoeys, G. Aglieri Rinella, A. Andronic, M. Antonelli, R. Baccomi, et al., Optimization of a 65 nm CMOS imaging process for monolithic CMOS sensors for high energy physics, PoS Pixel2022 (2023) 083. doi:10.22323/1.420.0083.

[5] G. Aglieri Rinella, G. Alocco, M. Antonelli, R. Baccomi, S. M. Beole, et al., Characterization of analogue Monolithic Active Pixel Sensor test structures implemented in a 65 nm CMOS imaging process, Nucl. Instrum. Methods Phys. Res. A 1069 (2024) 169896. doi:10.1016/j.nima.2024.169896.

[6] G. Aglieri Rinella, A. Andronic, M. Antonelli, M. Aresti, R. Baccomi, et al., Digital pixel test structures implemented in a 65 nm CMOS process, Nucl. Instrum. Methods Phys. Res. A 1056 (2023) 168589. doi:10.1016/j.nima.2023.168589.

[7] S. Ruiz Daza, R. Ballabriga, E. Buschmann, M. Campbell, R. Casanova Mohr, et al., The H2M Monolithic Active Pixel Sensor — characterizing non-uniform in-pixel response in a 65 nm CMOS imaging technology, J. Instrum. 20 (06) (2025) C06037. doi:10.1088/1748-0221/20/06/C06037.

[8] C. Lemoine, R. Ballabriga, E. Buschmann, M. Campbell, R. Casanova Mohr, et al., Impact of the circuit layout on the charge collection in a monolithic pixel sensor, Journal of Instrumentation 20 (06) (2025) C06052. doi:10.1088/1748-0221/20/06/C06052.
URL https://dx.doi.org/10.1088/1748-0221/20/06/C06052

[9] J. A. Ballin, J. P. Crooks, P. D. Dauncey, A.-M. Magnan, Y. Mikami, et al., Monolithic Active Pixel Sensors (MAPS) in a Quadruple Well Technology for Nearly 100% Fill Factor and Full CMOS Pixels (2008). doi:10.3390/s8095336.

[10] F. Krummenacher, Pixel detectors with local intelligence: an IC designer point of view, Nucl. Instrum. Methods Phys. Res. A 305 (1991) 527–532. doi:10.1016/0168-9002(91)90152-G.

[11] L.Ratti, F.DeCanio, M.Manghisoni, V.Re, G.Traversi, A Front-End Channel in 65 nm CMOS for Pixel Detectors at the HL-LHC Experiment Upgrades, IEEE Trans. Nucl. Sci. 64 (2) (2017) 789–799. doi:10.1109/TNS.2016.2646908.

[12] S. V. Vaseghi, Transient Noise, Vieweg+Teubner Verlag, Wiesbaden, 1996, pp. 314–327. doi:10.1007/978-3-322-92773-6_12.
URL https://doi.org/10.1007/978-3-322-92773-6_12

[13] V. Sriskaran, J. Alozy, R. Ballabriga, M. Campbell, P. Christodoulou, et al., High-rate, high-resolution single photon X-ray imaging: Medipix4, a large 4-side buttable pixel readout chip with high granularity and spectroscopic capabilities (2024). doi:10.1088/1748-0221/19/02/P02024.

[14] X. Llopart, J. Alozy, R. Ballabriga, M. Campbell, R. Casanova, et al., Timepix4, a large area pixel detector readout chip which can be tiled on 4 sides providing sub-200 ps timestamp binning, J. Instrum. 17 (01) (2022) C01044. doi:10.1088/1748-0221/17/01/C01044.

[15] Y. Otarid, M. Benoit, E. Buschmann, H. Chen, D. Dannheim, et al., Caribou — a versatile data acquisition system for silicon pixel detector prototyping, J. Instrum. 20 (07) (2025) C07043. doi:10.1088/1748-0221/20/07/C07043.

[16] T. Heim, M. Garcia-Sciveres, Self-adjusting threshold mechanism for pixel detectors, Nucl. Instrum. Methods Phys. Res. A 867 (2017) 209–214. doi:10.1016/j.nima.2017.06.040.

[17] L. Rossi, P. Fischer, T. Rohe, N. Wermes, Particle Detectors — From Fundamentals to Applications, Springer Berlin, Heidelberg, 2006. doi:10.1007/3-540-28333-1.

[18] I. Kremastiotis, R. Ballabriga, K. Dort, N. Egidos, M. Munker, CLICTD: A monolithic HR-CMOS sensor chip for the CLIC silicon tracker, PoS Topical Workshop on Electronics for Particle Physics (TWEPP2019) (2020) 039. doi:10.22323/1.370.0039.

[19] F. M. Puglisi, A. Padovani, L. Larcher, P. Pavan, Random telegraph noise: Measurement, data analysis, and interpretation, in: 2017 IEEE 24th International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA), 2017, pp. 1–9. doi:10.1109/IPFA.2017.8060057.

[20] I. Kremastiotis, R. Ballabriga, M. Campbell, D. Dannheim, K. Dort, Design and Characterization of the CLICTD Pixelated Monolithic Sensor Chip, IEEE Trans. Nucl. Sci. (2020) 1–1doi:10.1109/TNS.2020.3019887.

[21] R. D. Deslattes, E. G. Kessler Jr., P. Indelicato, L. de Billy, E. Lindroth, et al., X-ray transition energies (version 1.2), accessed: 2025-05-13 (2005).
URL http://physics.nist.gov/XrayTrans

[22] B. Loo, F. Goulding, D. Gao, Ballistic deficits in pulse shaping amplifiers, IEEE Trans. Nucl. Sci. 35 (1) (1988) 114–118. doi:10.1109/23.12686.

[23] J. Jakubek, Precise energy calibration of pixel detector working in time-over-threshold mode, Nucl. Instrum. Methods Phys. Res. A 633 (2011) S262–S266, 11th International Workshop on Radiation Imaging Detectors (IWORID). doi:10.1016/j.nima.2010.06.183.

[24] R. Diener, J. Dreyling-Eschweiler, H. Ehrlichmann, I. Gregor, U. Kötz, et al., The DESY II test beam facility, Nucl. Instrum. Methods Phys. Res. A 922 (2019) 265–286. doi:10.1016/j.nima.2018.11.133.

[25] CERN, Secondary Beam Areas of the PS, SPS machines, accessed: 2025-05-13 (2019).
URL http://sba.web.cern.ch/sba/

[26] Y. Liu, C. Feng, I.-M. Gregor, A. Herkert, L. Huth, et al., ADENIUM — A demonstrator for a next-generation beam telescope at DESY, J. Instrum. 18 (06) (2023) P06025. doi:10.1088/1748-0221/18/06/p06025.

[27] D. Dannheim, K. Krüger, A. Levy, A. Nürnberg, E. Sicking (Eds.), Detector Technologies for CLIC, Vol. 1 of CERN Yellow Reports: Monographs, CERN, Geneva, 2019. doi:10.23731/CYRM-2019-001.

[28] S. Spannagel, H. Jansen, GBL Track Resolution Calculator v2.0 (2016). doi:10.5281/zenodo.48795.

[29] L. Huth, H. Augustin, L. Dittmann, S. Dittmeier, J. Hammerich, et al., TelePix2: Full scale fast region of interest trigger and timing for the EUDET-style telescopes at the DESY II test beam facility (2025). doi:10.1016/j.nima.2025.170720.

[30] P. Baesso, D. Cussans, J. Goldstein, The AIDA-2020 TLU: a flexible trigger logic unit for test beam facilities, J. Instrum. 14 (09) (2019) P09019–P09019. doi:10.1088/1748-0221/14/09/p09019.

[31] Y. Liu, M. Amjad, P. Baesso, D. Cussans, J. Dreyling-Eschweiler, et al., EUDAQ2—A flexible data acquisition software framework for common test beams, J. Instrum. 14 (10) (2019) P10033–P10033. doi:10.1088/1748-0221/14/10/p10033.

[32] J. Visser, M. v. Beuzekom, H. Boterenbrood, B. v. d. Heijden, J. Muñoz, et al., SPIDR: a read-out system for Medipix3 & Timepix3, J. Instrum. 10 (12) (2015) C12028. doi:10.1088/1748-0221/10/12/C12028.

[33] D. Dannheim, K. Dort, L. Huth, D. Hynds, I. Kremastiotis, et al., Corryvreckan: a modular 4D track reconstruction and analysis software for test beam data, J.Instrum. 16 (03) (2021) P03008. doi:10.1088/1748-0221/16/03/p03008.

[34] R. Turchetta, Spatial resolution of silicon microstrip detectors, Nucl. Instrum. Methods Phys. Res. A 335 (1) (1993) 44–58. doi:https://doi.org/10.1016/0168-9002(93)90255-G.

[35] V. Blobel, A new fast track-fit algorithm based on broken lines, Nucl. Instrum. Methods Phys. Res. A 566 (1) (2006) 14–17, TIME 2005: Proceedings of the 1st Workshop on Tracking in High Multiplicity Environments. doi:10.1016/j.nima.2006.05.156.

[36] Synopsys, Sentaurus, accessed: 2025-07-30 (2025).
URL https://www.synopsys.com/manufacturing/tcad.html

[37] S. Spannagel, K. Wolters, D. Hynds, N. Alipour Tehrani, M. Benoit, et al., Allpix2: A modular simulation framework for silicon detectors, Nucl. Instrum. Methods Phys. Res. A 901 (2018) 164–172. doi:10.1016/j.nima.2018.06.020.

[38] S. Agostinelli, J. Allison, K. a. Amako, J. Apostolakis, H. Araujo, et al., GEANT4–a simulation toolkit, Nucl. Instrum. Methods Phys. Res. A 506 (3) (2003) 250–303. doi:10.1016/S0168-9002(03)01368-8.

[39] J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. A. Dubois, et al., Geant4 developments and applications, IEEE Trans. Nucl. Sci. 53 (1) (2006) 270–278. doi:10.1109/TNS.2006.869826.

[40] J. Allison, K. Amako, J. Apostolakis, P. Arce, M. Asai, et al., Recent developments in GEANT4, Nucl. Instrum. Methods Phys. Res. A 835 (2016) 186–225. doi:10.1016/j.nima.2016.06.125.

[41] C. Canali, G. Majni, R. Minder, G. Ottaviani, Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature, IEEE Trans. Electron Devices 22 (11) (1975) 1045–1047. doi:10.1109/T-ED.1975.18267.

[42] W. Shockley, W. T. Read, Statistics of the recombinations of holes and electrons, Phys. Rev. 87 (5) (1952) 835. doi:10.1103/PhysRev.87.835.

[43] R. N. Hall, Recombination processes in semiconductors, Proceedings of the IEE - Part B: Electronic and Communication Engineering 106 (17S) (1959) 923–931. doi:10.1049/pi-b-2.1959.0171.

[44] J. G. Fossum, D. S. Lee, A physical model for the dependence of carrier lifetime on doping density in nondegenerate silicon, Solid-State Electron. 25 (8) (1982) 741–747. doi:10.1016/0038-1101(82)90203-9.

[45] S. Ramo, Currents induced by electron motion, Proc. IRE 27 (9) (2006) 584–585. doi:10.1109/JRPROC.1939.228757.

[46] W. Shockley, Currents to conductors induced by a moving point charge, J. Appl. Phys. 9 (10) (1938) 635–636. doi:10.1063/1.1710367.

[47] R. Ballabriga, J. Braach, E. Buschmann, M. Campbell, D. Dannheim, et al., Transient Monte Carlo simulations for the optimisation and characterisation of monolithic silicon sensors, Nucl. Instrum. Methods Phys. Res. A 1031 (2022) 166491. doi:10.1016/j.nima.2022.166491.

[48] H. Wennlöf, D. Dannheim, M. Del Rio Viera, K. Dort, D. Eckstein, Simulating monolithic active pixel sensors: A technology-independent approach using generic doping profiles, Nucl. Instrum. Methods Phys. Res. A 1073 (2025) 170227. doi:10.1016/j.nima.2025.170227.

[49] Cadence, Spectre, accessed: 2025-07-30 (2025).
URL https://www.cadence.com/en_US/home/tools/custom-ic-analog-rf-design/circuit-simulation.html