



Full Length Article

Test beam characterization of a digital silicon photomultiplier

Finn King¹*, Inge Diehl², Ono Feyens¹, Ingrid-Maria Gregor², Karsten Hansen²,
Stephan Lachnit², Frauke Poblitzki², Daniil Rastorguev³, Simon Spannagel²,
Tomas Vanat², Gianpiero Vignola²

Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany

ARTICLE INFO

Keywords:

Silicon sensors
Digital SiPMs
CMOS SPADs
Charged particle detection
MIP detection
Test beam

ABSTRACT

Conventional silicon photomultipliers (SiPMs) are well established as light detectors with single-photon-detection capability and used throughout high energy physics, medical, and commercial applications. The possibility to produce single photon avalanche diodes (SPADs) in commercial CMOS processes creates the opportunity to combine a matrix of SPADs and an application-specific integrated circuit in the same die. The potential of such digital SiPMs (dSiPMs) is still being explored, while it already is an established technology in certain applications, like light detection and ranging (LiDAR).

A prototype dSiPM, produced in the LFoundry 150-nm CMOS technology, was designed and tested at DESY. The dSiPM central part is a matrix of 32 by 32 pixels. Each pixel contains four SPADs, a digital front-end, and has an area of $69.6 \times 76 \mu\text{m}^2$. The chip has four time-to-digital converters and includes further circuitry for data serialization and data links.

This work focuses on the characterization of the prototype in an electron beam at the DESY II Test Beam facility, to study its capability as a tracking and timing detector for minimum ionizing particles (MIPs). The MIP detection efficiency is found to be dominated by the fill factor and on the order of 31%. The position of the impinging MIPs can be measured with a precision of about $20 \mu\text{m}$, and the time of the interaction can be measured with a precision better than 50 ps for about 85% of the detected events. In addition, laboratory studies on the breakdown voltage, dark count rate, and crosstalk probability, as well as the experimental methods required for the characterization of such a sensor type in a particle beam are presented.

1. Introduction

Silicon photomultipliers (SiPMs) contain an array of single-photon avalanche diodes (SPADs). SPADs are operated in Geiger-mode and sensitive to ionizing radiation and even single photons. One key feature of the Geiger-mode operation are fast-rising signal edges, facilitating single photon time resolutions on the order of 10 ps [1]. Structure, operation principle, and applications of SiPMs are discussed, for example, in [2–4].

An effect of the Geiger-mode operation is that the SPAD signal contains no information about the energy deposition by the interacting quanta. This inherent digital nature of the SPAD signal makes thinking of a SiPM as a digital device seem natural. As SPAD layouts in CMOS processes are available, CMOS circuitry can be embedded in the SiPM design, which allows profiting from digital signal processing. Introduced as digital SiPMs (dSiPMs) in [5], these devices come with several

advantages compared to analog SiPMs. The circuitry allows storing information about the position of a firing SPAD (order of $10 \mu\text{m}$ precision), the masking of noisy SPADs, and on-chip pattern recognition. Parallel connection of the whole SPAD array can be avoided, which reduces the input capacitance to the readout circuit, and thus improves signal rise times. Finally, it can ease system integration, as no separate device for signal digitization is required. One clear disadvantage is the quality of the SPADs, yielding generally higher dark count rates (DCRs) in dSiPMs [6,7], which will certainly improve once more sophisticated SPAD designs become available in CMOS processes. Another is the reduction of the fill factor, as a part of the surface area is needed for the CMOS circuitry, with increasing circuit complexity.

Possible applications have to find a trade-off between these advantages and disadvantages. Especially, the availability of position information is a key point to consider. One example is to use dSiPMs

* Corresponding author.

E-mail address: finn.king@desy.de (F. King).

¹ Also at Vrije Universiteit Brussel, Belgium.

² Also at Rheinische Friedrich-Wilhelms-Universität Bonn, Germany.

³ Also at Bergische Universität Wuppertal, Germany.

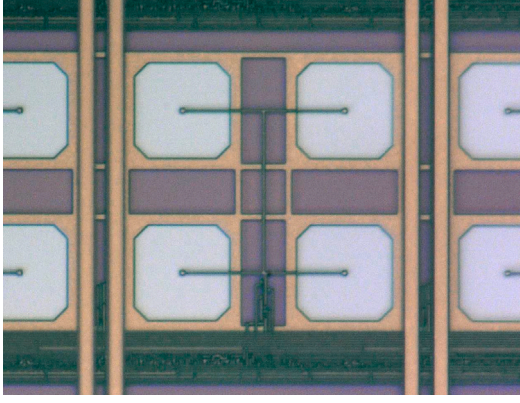


Fig. 1. Microscopic image of a pixel. The four-SPAD structure with shared electronics in the lower part of the pixel is visible. The total area of the pixel is $69.6 \times 76 \mu\text{m}^2$.

for the readout of scintillating-fiber bundles [8], where the available position information is used to identify individual fibers, potentially reducing complexity and cost of fiber-readout solutions. Another application would be 4D particle tracking [9], where the combination of position information on the order of $10 \mu\text{m}$ and time information on the order of 10 ps might be beneficial properties. Calculations of time resolution and efficiency of SPADs and SiPMs illuminated with photons or charged particles are discussed in [10]. The performance of SPAD-based sensors in terms of charged-particle detection is characterized in [11] (conventional SiPMs), and [12] (SPAD sensors produced in a CMOS technology). Another interesting concept in the context of charged-particle detection is the combination of two layers of such CMOS-based SPAD sensors to suppress noise or increase the active area, as presented in [7,13], respectively.

With this perspective, this work reports on the minimum ionizing particle (MIP) detection capabilities found for a new dSiPM prototype designed at DESY [14], and described in Section 2. Laboratory measurements, focusing on breakdown voltage, and crosstalk, are discussed in Section 3. Setup, analysis method and results of the MIP-detection studies are presented in Section 4, 5, and 6, respectively.

2. The DESY dSiPM

The DESY dSiPM prototype is fabricated in the LFoundry standard 150-nm CMOS technology. The basic building blocks of this dSiPM are pixels containing four SPADs with a shared quenching circuit, discriminator and 2-bit counter. A microscopic image of the pixel is shown in Fig. 1, while the readout scheme is shown in Fig. 2. The pixel area is $69.6 \times 76 \mu\text{m}^2$, where the area occupied by the SPADs amounts to 30%, determining the fill factor. A group of 16×16 pixels forms a quadrant with one 12-bit time-to-digital converter (TDC), generating a time stamp for the first hit (firing SPAD) in each quadrant and readout frame. The readout frames are defined by a 3 MHz clock, and the time and hit position information of each quadrant are buffered and read out in the following frame, if the chip is operated in 1-bit mode. In 2-bit mode, the in-pixel counter is used to store up to three hits per pixel and frame. Consequently, the data transmission takes two frames in this case. For the studies presented hereafter, the 1-bit mode is used. Individual pixels can be disabled (masked) using a transistor switch in series with the quenching transistor, which helps to reduce the DCR in the presence of noisy pixels. A temperature diode in the periphery of the chip allows for precise temperature monitoring during measurements and operation.

Further features of the DESY dSiPM, like a validation logic for topology-based event selection, external test circuits, and LED trigger pulse generation, are detailed in [14], along with a detailed report on the ASIC design and properties. One of the investigated properties is the

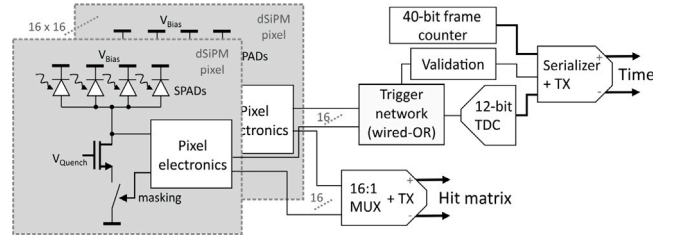


Fig. 2. Simplified readout scheme of a quadrant of the DESY dSiPM. From [15].

DCR as a function of temperature, overvoltage, and fraction of masked pixels. A mean DCR of $8.7 \text{ Hz } \mu\text{m}^{-2}$ is reported at room temperature and for an overvoltage of 1 V.

2.1. Data acquisition system

To facilitate data acquisition (DAQ) and save development time, the DESY dSiPM is integrated into the Caribou DAQ system, an open-source hardware, software (Peary) and firmware project dedicated to pixel detector development [16,17]. Its key component is a system-on-chip (SoC) board (AMD [18], Zynq 7000 SoC ZC706) containing an FPGA and a CPU. The CPU is running a fully-featured Yocto-based Linux distribution and Peary, while the FPGA executes custom written firmware blocks for configuration, control, and readout of the dSiPM. The SoC board is connected to a mezzanine board, the CaR board, that provides supplies like current and voltage sources, and a physical interface to the dSiPM. The dSiPM itself is placed on a custom chip board, connected to the CaR board, and hosting LVDS repeaters, for communication with the chip, and further interfacing circuits.

The Caribou system provides everything required for operation of the dSiPM, but not the high voltage required for biasing of SPADs, which requires an external voltage source. The provided supplies include power, current references, voltage references, LVDS links, a 3 MHz clock, a 408 MHz clock, and synchronous control signals generated in the FPGA. This way, the entire system is relatively compact, which eases physical integration at the test beam, in a laser setup, or for microscopy.

The Peary software provides an interface to control the supplies on the CaR board, and allows writing to registers in the FPGA for control of the firmware, and data transactions. It also provides a development environment to define sensor-specific powering, configuration and readout procedures. Following a certain structure, naming convention, and implementing key functional features, these procedures are intrinsically embedded in the EUDAQ data acquisition software framework [19,20], used with the test-beam instrumentation introduced in Section 4.

3. Laboratory characterization

Details on the laboratory characterization of the DESY dSiPM — with focus on digital properties — are reported in [14]. This section discusses breakdown voltage, and crosstalk measurements, to give additional insights in properties of the dSiPM and to enable the operation at comparable overvoltages.

3.1. Breakdown voltage measurements

The breakdown voltage V_{BD} marks the reverse bias voltage, at which SPADs start to exhibit avalanche multiplication leading to Geiger discharges. This is an important characteristic of a SiPM, since they are operated at an overvoltage V_{OV} above the breakdown voltage. As summarized in [21], the breakdown voltage can be extracted from a measurement of the bias current I_{bias} of a SiPM as a function of the reverse bias voltage V_{bias} (IV-curve).

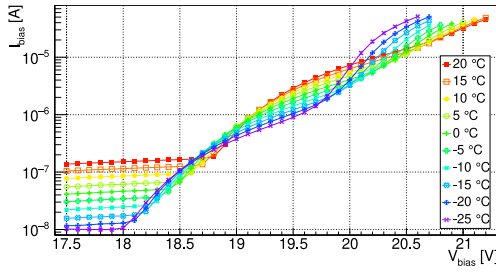


Fig. 3. IV-curves for one of the tested sensors, at temperatures between -25°C and 20°C . Each point represents the average of several measurements; the errors are within the marker.

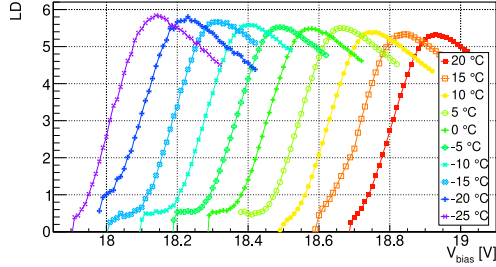


Fig. 4. Logarithmic derivatives of IV-curves, for one of the tested sensors, at temperatures between -25°C and 20°C .

Measurements of IV-curves are made inside a climate chamber (ACS [22], model DY200), assuring a temperature and humidity controlled environment and dark conditions. A source meter (Keithley [23], model 2410) is used for fine control of reverse bias and precise current measurement. Fig. 3 shows IV-curves for one of the tested sensors, at environmental temperatures between -25°C and 20°C . As can be seen, breakdown, marked by the sharp rise of the current, occurs between 18 V and 19 V and the current before breakdown is between 10 nA and 200 nA, both depending on temperature. It should be noted that this is a purely analog measurement and does not depend on the digital characteristics of the sensor but rather on SPAD design and manufacturing process.

Several methods to quantify V_{BD} are based on identifying the previously mentioned sharp rise in the IV-curve. Some of these methods are compared in [21]. Within this work, the breakdown voltage is defined as the maximum of the logarithmic derivative

$$LD = \frac{d}{dV_{bias}} \ln(I_{bias}), \quad (1)$$

which proved to provide a robust estimate, and is reported to be in good agreement with other methods. A set of IV-curves for the same temperatures, and with finer voltage steps, is taken around the breakdown voltage estimated from Fig. 3. The corresponding logarithmic derivative is shown in Fig. 4, and the estimated breakdown voltages are e.g. 18.14 V and 18.93 V, at -25°C , and 20°C , respectively. These measurements are performed for all investigated samples, so that consistent overvoltages for the following studies can be chosen in the relevant temperature range. It should be noted that the logarithmic derivative is calculated between points with a distance $dV_{bias} = 0.08$ V. This choice has a systematic impact on the measured breakdown voltage, but yields a robust estimate, and is the same for all measurements.

3.2. Crosstalk measurements

The Geiger discharge within a SPAD generates secondary photons, which can generate secondary electron-hole pairs in adjacent SPADs and thus trigger additional discharges. This effect is known as optical crosstalk, and it occurs with a probability proportional to the gain of

the SPAD [24]. As discussed in [2,21], one can differentiate prompt and delayed optical crosstalk, where the secondary charge carriers are generated in the depleted, and non-depleted region, respectively. The crosstalk probability can be reduced by means of trenches filled with optically dense material between SPADs [25].

While this effect is considered a nuisance to photon counting applications, it could be beneficial for MIP detection. This is because the affected SPAD cell might be in another pixel and hence indicate a particle interaction close to the boundary of the pixels, similar to charge sharing in silicon pixel or strip sensors. Crosstalk measurements on a single pixel level are discussed in the following, while the macroscopic effects of crosstalk are discussed in Section 6.2.

3.2.1. Single pixel crosstalk

As the DESY dSiPM provides information about the position of a firing pixel, the crosstalk probability can be measured for individual pixels using an approach similar to the one described in [26]. First, an isolated noisy pixel is chosen. In step 1, the isolated DCR of this pixel, and its eight neighbors, is measured successively by masking the entire matrix but the studied pixel. One should note, that this eliminates crosstalk contributions from neighboring pixels, as the masking mechanism prevents Geiger discharges in the masked pixels. In step two, the central pixel plus one of the neighbors are unmasked and the DCR measurement is repeated, again successively for all eight neighbors. The total recorded number of frames N and the number of frames N_c , N_n , where the central, or a neighboring pixel fires in step one, allow estimating the probabilities

$$p_c = \frac{N_c}{N}, \text{ and} \quad (2)$$

$$p_n = \frac{N_n}{N}. \quad (3)$$

From the number of frames where both pixels fire in step two, $N_{c \wedge n}$, one can derive

$$p_{c \wedge n} = \frac{N_{c \wedge n}}{N}. \quad (4)$$

This allows to derive the crosstalk probability p_x to a given neighbor

$$p_x = \frac{p_{c \wedge n} - p_c p_n}{p_c (1 - p_n)}, \quad (5)$$

assuming that p_c is significantly larger than p_n .

An example for one pixel, at an overvoltage of about 2 V, and a sensor temperature of about 0°C , is shown in Fig. 5. It can be seen that the probability of crosstalk is non-uniform, and largest for the pixel right of the central one. It is observed, that the total probability of crosstalk and the position of the pixel with the largest crosstalk, vary between several studied noisy pixels. The arrangement of the SPADs inside the pixel, see Fig. 1, suggests that a certain SPAD within the pixel is responsible for the increased DCR, and consequently crosstalk to proximate pixels is more likely. Assuming that a localized production defect determines the increased DCR, even the position of this defect within the noisy SPAD will impact the measured crosstalk probability.

To confirm this, the emission of visible light, generated during the Geiger discharge in noisy pixels, is measured in dark conditions. This method is based on [27], and performed on the same pixel, using a microscope (Zeiss [28], Axioscope), and a CCD camera (Jenoptik [29], model ProgRes SpeedXT core5). Fig. 6 shows microscopic pictures of the pixel matrix, super-imposed with the obtained image. The image is recorded with an exposure time of 50 s, at the maximum overvoltage, the lowest resistance of the quenching transistor, and corrected for dark noise from the CCD sensor to optimize the signal-to-noise ratio. A clear spot of emitted light is visible close to the top-right SPAD of the pixel marked in Fig. 5, confirming the hypothesis from above. One should note, that also some difference in crosstalk to the top and right neighbor is expected, as the placement of the in-pixel electronics on the bottom of a pixel results in a larger distance to the closest SPAD in the neighboring pixel. Further noisy pixels can be identified in the image,

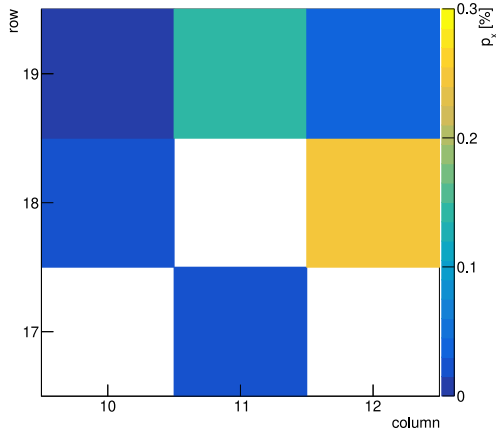


Fig. 5. Measurement of the probability of crosstalk between a noisy pixel (center), and its neighbors. Measured at an overvoltage of about 2 V and a temperature of about 0 °C. Count rates in the outermost pixels of the bottom row are too low for an estimate of the crosstalk probability.

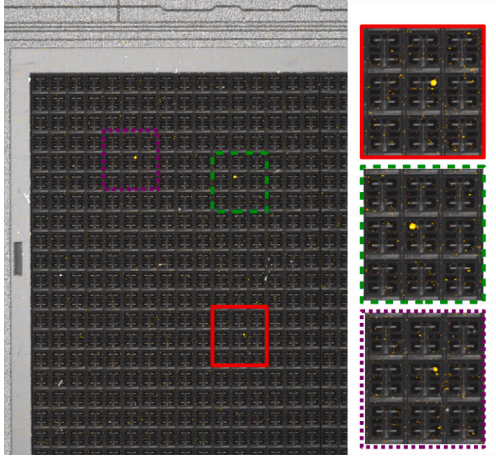


Fig. 6. Measurement of the light emitted in the avalanches process in noisy pixels. The image, in shades of yellow, is superimposed to a microscopic image of the matrix, displayed in grayscale. Close-ups of the region around the three noisiest pixels are shown on the right side of the image. The solid red box indicates the group of pixels shown in Fig. 5.

and the measurement of crosstalk to their neighbors yields consistent results.

The studies show that a crosstalk probability on percent level is to be expected for typical operation conditions, which is too small to have a significant effect on the spatial resolution. It also shows, that the DCR of noisy pixels is generally dominated by single noisy SPADs. Thus, a masking mechanism with higher granularity, that allows masking a SPAD instead of a pixel of 4 SPADs, would require only 25 % of the masked area to turn a noisy SPAD off. One should note that the shown crosstalk probability is not corrected for the occurrence of more than one dark count per frame (relevant at high DCR, or due to after-pulsing), and that delayed crosstalk events are not necessarily recorded in the same frame.

4. Test-beam setup

The presented test-beam measurements follow a simple concept. A collimated beam of electrons penetrates the device under test (DUT), in this case two DESY dSiPMs. A set of reference detectors provides trigger signals and tracking on a single electron level. This way, electron trajectories (tracks) can be reconstructed during (online) or after (offline)

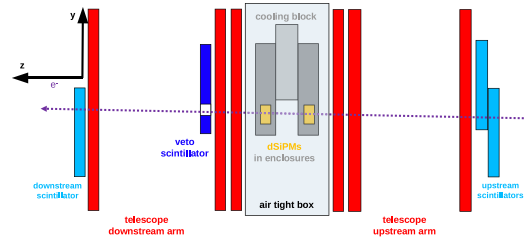


Fig. 7. Schematic image of the test-beam setup, showing the position of the DESY dSiPMs with respect to the reference detectors for trigger signals (scintillators) and tracking (telescope).

beam operation (run). This allows to reconstruct the position at which an electron hits the DUT and to characterize the DUT in terms of hit detection efficiency, and temporal as well as spatial hit resolution. The hit efficiency denotes the probability to detect MIPs. The spatial resolution refers to the uncertainty of the position measurement provided by the DUT, the temporal resolution is defined in the same way. The reconstruction of these properties is discussed in Section 5.

The test-beam measurements were conducted in two weeks, in March 2023, at the DESY II Test Team Facility [30]. An electron momentum of 4 GeV/c is selected as a trade-off between electron rate and track resolution, which is affected by multiple Coulomb scattering [31,32]. A schematic image of the test-beam setup is shown in Fig. 7. The set of reference detectors for particle tracking is the so-called DATURA beam telescope [33]. It consists of six MIMOSA 26 sensors [34] with an intrinsic spatial resolution of 3.24 μm . The sensors are arranged in two groups of three, with a spacing between about 20 mm and 300 mm. The trigger system consists of four scintillators coupled to photomultiplier tubes. A coincidence logic between three of the scintillators is used to suppress noise. A fourth scintillator features a hole of about 2 mm diameter, which is slightly smaller than the active area of the dSiPMs. It is placed downstream of the dSiPMs, as it adds no material in the area of interest, and used as a veto, to discard tracks outside the active area of the dSiPMs. The other three scintillators are placed outside the beam telescope, so that the additional material does not deteriorate the tracking resolution.

To facilitate synchronous operation of the dSiPMs, and the trigger and tracking system, an AIDA trigger logic unit (TLU) is used [35]. The TLU derives the trigger signal from the scintillator signals and distributes it to the other detectors in the setup. To ensure that the DAQ systems of the beam telescope and the dSiPMs are ready to accept a trigger, and not occupied processing a previous trigger signal, both devices provide a busy signal to the TLU, which vetoes the generation of trigger output signals. In addition, the TLU provides a clock and a run start signal to ensure synchronous operation of the detector systems. The firmware running on Caribou continuously buffers the data packets for each readout frame of the dSiPMs. As the trigger signals arrive with a certain delay, and asynchronously with respect to the 3 MHz frame clock, the corresponding hit might be registered in the current or previous readout frame. Thus, both frames are marked for storage, and combined for the later analysis. In some measurements, more than two frames are marked for storage, to have data for noise studies under the conditions at the test beam.

The chip boards with the dSiPMs are placed inside aluminum casings, which provide light shielding and thermal contact to a cooling element. This allows stabilizing the operational temperature at about 0 °C, limited by the power of the cooling unit and losses. The aluminum casings and cooling elements are inside a nitrogen flushed plastic casing, as shown in Fig. 8. The backplate of the casing is mounted on a set of position stages to allow movement orthogonal to the beam direction with micrometer precision. The chip boards, the aluminum casings, and the surrounding plastic casing feature holes to minimize the amount



Fig. 8. Test-beam setup used in March 2023. The transparent plastic on the left contains the dSiPMs in their aluminum casing, the CaR board, and tubes for the light blue coolant. The five aluminum casings on the left and right contain the MIMOSA 26 sensors.

of material affecting tracking precision. Thin layers of aluminum or Kapton foil were used to keep the casings light- and airtight.

One should note that the trigger area, given by the hole in one of the scintillators, is slightly smaller than the active area of the dSiPM. This is a compromise to maximize the fraction of triggers that actually hit the DUT. On the other hand, it requires the mechanical alignment of the DUT and the trigger system to be on the order of 0.1 mm. The mechanical alignment of the trigger system relative to the beam telescope is straight forward, as the trigger area is easily recognizable in the hit maps of the MIMOSA sensors. If the DUT were able to provide a trigger output itself, it could be aligned in the same way. Unfortunately, the high DCR does not permit this approach. Instead, an approach based on the evaluation of the electron scattering in the material around the DUT is applied, as presented in Sections 5 and 6.

5. Analysis procedure

The analysis of the test-beam data is performed with the Corryvreckan framework [36,37]. This software analyzes the raw data from the TLU, MIMOSA sensors, and DESY dSiPMs, allowing for the reconstruction of particle trajectories and observables quantifying the dSiPMs MIP-detection capabilities, within an event-based structure.

As a first reconstruction step, the raw data are decoded and the hit information is associated to events, which are defined by TLU trigger IDs. Next, clusters are reconstructed as groups of neighboring hits in the telescope planes and the dSiPMs (merging data from both readout frames associated with the given trigger ID). The cluster center is calculated as the mean position of all pixels in the cluster and estimates the point at which the electron traversed the corresponding detector. Electron tracks are reconstructed from the cluster centers in the telescope planes, applying the General-Broken-Lines track model [38,39], which allows for kink angles in the particle trajectory at scatterer positions (telescope planes, DUT, etc.). To maximize tracking precision, only tracks with an associated cluster on each telescope plane are analyzed, and tracks with $\chi^2/n_{\text{dof}} > 5$ are rejected, where n_{dof} denotes the degree of freedom of the track fit. Also, tracks outside a region of interest related to the trigger acceptance are discarded. Hits on the dSiPMs are associated to the reconstructed tracks when they are within a radius of 70 μm around the reconstructed intersection of the particle track and the corresponding dSiPM plane. Additionally, the time stamps of the dSiPM hits are requested to be within 230 μs around the time stamp of the corresponding TLU trigger, to account for the integration time of the MIMOSA 26 sensors.

The presented measurements and their analysis required some uncommon procedures for physical and software alignment, described in Section 5.1. The investigated observables, quantifying the dSiPMs MIP-detection performance, are introduced in Section 5.2.

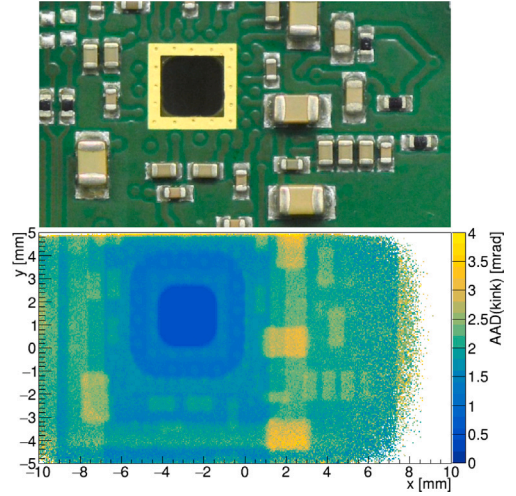


Fig. 9. The upper part shows a microscopic picture of the dSiPM chip board. The dSiPM is placed on the backside of the board, behind the hole in its center. The lower part shows the width of the scattering-angle distribution as a function of position in the DUT plane.

5.1. Physical and software alignment

The analysis of test-beam data relies on precise knowledge of the relative positions and orientations of the involved detectors. This is achieved using dedicated alignment modules within the Corryvreckan framework, which are based on minimizing the global χ^2 of all tracks, a standard procedure in testing particle-tracking detectors.

Given the small size of the hole in the trigger scintillator, defining the acceptance for electron tracks, mechanical alignment of the dSiPMs with respect to the reference detectors needs to be on the order of 0.1 mm. This needs to be done when setting up, before starting final measurements. Typically, this is achieved by using a trigger signal derived from the response of the DUT, and analyzing the position and shape of the corresponding acceptance on a close telescope plane. For the dSiPM, this approach is not applicable, as there are dark counts in the majority of the frames. Hence, the relation between the electron impact position and the presence of a trigger signal is not given.

Instead, measurements at an electron momentum of 2 GeV/c are recorded, using the coincidence of two scintillators without hole as a trigger input. This is to increase scattering effects and the investigated area, respectively, otherwise the setup remains unchanged. For the analysis of the acquired data, Corryvreckan is configured to perform so called material-budget imaging (MBI), using a module developed in the context of [40]. This approach assumes that the dominating amount of scattering material lies in the plane of a single DUT, between the upstream and downstream arm of the telescope (see Fig. 7). Straight lines are reconstructed in the upstream and downstream arm, separately. Pairs of tracks in both arms are defined, applying a matching cut of 200 μm on the intercept of the tracks with the DUT plane. This allows to derive a scattering angle (kink) between the upstream and downstream track.

The width of the distribution of scattering angles is related to the traversed material budget [31,41]. Calculating the average absolute deviation (AAD) of the kink angle as a function of position in the DUT plane yields the result shown in Fig. 9. It allows identifying the position of the dSiPM relative to the reference system by identifying the hole in the chip board and characteristic components, also shown in the figure. It also shows, that the two DUTs are well aligned with respect to each other.

Offline analysis of the data revealed movements of the dSiPM, relative to the reference system, on the scale of a few $\mu\text{m h}^{-1}$. This is assumed to be due to thermal expansion of mechanical components. To

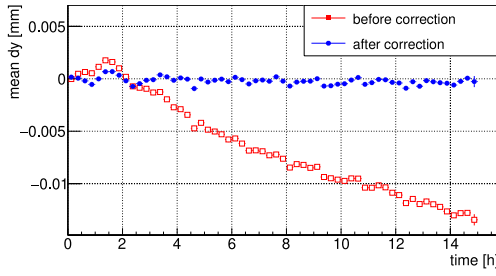


Fig. 10. Mean of the residual dy as a function of time, before and after applying the correction for thermal expansion described in the text.

correct for this effect, the alignment of the DUT is performed using a part of the data at the beginning of a run, where relative movement can be considered negligible. The entire data of the run is analyzed in a next step, using the previously found alignment parameter. The spatial residuals, defined in Section 5.2.2, are then analyzed as a function of time. A polynomial fit to the obtained data is then used to parameterize the time dependence of the alignment parameters, and used as correction in the resolution studies. Fig. 10 shows the effect of the described correction on the spatial residuals in y -direction.

5.2. Observables

To investigate the MIP-detection performance of the dSiPMs, clusters on the dSiPMs are associated with the reconstructed tracks, as described in the beginning of this section. The set of selection criteria is the same for all observables, unless mentioned otherwise.

5.2.1. Efficiency

The hit-detection efficiency of a tracking detector denotes the probability of identifying a signal, when a MIP penetrates the pixel matrix of a given sensor. In a test-beam experiment, it is reconstructed as the ratio

$$\epsilon_{\text{rec}} = \frac{N_{\text{hits}}}{N_{\text{tracks}}}, \quad (6)$$

where N_{tracks} denotes the number of tracks passing the selection criteria, and N_{hits} is the subset of tracks with an associated cluster on the DUT.

The relatively high DCR of the dSiPMs causes some additional corrections to be required, to prevent a bias to the quoted hit-detection efficiency. First, tracks that intercept a masked pixel, or one of its direct neighbors, are discarded. Pixels are masked when their hit rate is about two orders of magnitudes larger than the average. Still, the reconstructed efficiency ϵ_{rec} needs to be corrected for the non-negligible probability of associating fake hits with the reconstructed track. To account for the strong dependency of the DCR on operating conditions, the density of fake hits is estimated from the same data, defining clusters as fake when there is no track within a radius of 2 pixel-pitches around them. Normalizing to the area defined by the association cut yields the fake-hit probabilities for the two readout frames f_{F1} and f_{F2} . This allows to define a corrected efficiency

$$\epsilon = c_{\text{dt}} \cdot \frac{\epsilon_{\text{rec}} - (f_{F1} + f_{F2} - f_{F1}f_{F2})}{1 - (f_{F1} + f_{F2} - f_{F1}f_{F2})}, \quad (7)$$

where the factor $c_{\text{dt}} = 1.0342$ corrects for the dead time between the two readout frames.⁴

⁴ Each readout frame has a duration of 333 ns. However, due to the readout scheme described in [14], hits within a reset phase of 11 ns cannot be detected. It still happens, that the reference system issues trigger signals for particles interacting during that reset phase. The given correction factor corresponds to the ratio $333/(333 - 11)$.

The method to correct for the noise contamination was cross-validated by acquiring an additional 3rd frame, that does not contain hits due to particle interactions, as it is too late in time. Applying the described analysis procedure only on data from this third frame, the per-frame fake-hit contribution to the efficiency f_{F3} can be estimated directly. Replacing f_{F1} and f_{F2} in Eq. (7) with f_{F3} , allows applying the correction in a similar way. The results from these two methods are found to be compatible within the uncertainties.

5.2.2. Spatial residuals and resolution

The spatial resolution of a tracking detector is given by the error on its measurement of a MIPs impact position. This can be reconstructed from the residuals

$$dx = x_{\text{track}} - x_{\text{cluster}}, \quad (8)$$

where x_{cluster} is the reconstructed position of an associated cluster in a dSiPM, and x_{track} is the position at which the reconstructed particle track intercepts this dSiPM. Note that the coordinates are given in a Cartesian coordinate system, where the z -axis is orthogonal to the sensor plane, and the x - and y -axes are parallel to the column- and row-direction of the dSiPMs. The width of the distribution dx is determined by the uncertainties of x_{track} , and x_{cluster} , respectively denoted as $\sigma_{x,\text{track}}$, and $\sigma_{x,\text{cluster}}$. The tracking uncertainty $\sigma_{x,\text{track}}$ depends on the electron momentum, geometrical setup, and amount of material the electron needs to penetrate. It is estimated to be between $3.3\mu\text{m}$ and $3.7\mu\text{m}$ at the positions of the DUTs, using the General Broken Lines formalism [42] for the given setup and conditions.

The distribution of dx is determined by the shape of a signal and a background contribution, denoted $S(dx)$ and $B(dx)$, respectively. The sum of these parametrizations is fitted to the reconstructed distributions, to extract the spatial resolution without impact from background. The signal contribution is determined by the positions of the SPADs within the pixel cell (see Fig. 1), as they define the efficient part of the pixel. It is parameterized as the normalized sum of four error functions,

$$S(x) = N_S \cdot \frac{f_1 - f_2 + f_3 - f_4}{2}, \quad (9)$$

where

$$f_i(x) = \text{erf}\left(\frac{x \pm C \pm \frac{W}{2}}{\sqrt{2}\sigma}\right).$$

The parameters C , and W represent the position of the SPAD center, fixed to be at $17.4\mu\text{m}$, and the nominal active width of the SPAD, fixed at $19.93\mu\text{m}$, respectively. The free parameters, N_S and σ , represent a normalization factor and the width of the Gaussian smearing, which includes the tracking uncertainty and the gradual decline of the efficiency at the boundaries of the SPADs.

The background contribution, on the other hand, stems from erroneously associating a noise hit (dark count) with a track, that was not registered as a hit in any pixel. The noise contamination has been reduced through online or offline masking of noisy pixels, but it cannot be completely eliminated due to the typical properties of SPADs, especially at high overvoltages and temperatures. The shape of the background contribution is obtained integrating the acceptance of the radial cluster-association cut, and yields

$$B(x) = N_B \cdot 2 \cdot \left(\sqrt{|D^2 - x^2|}\right), \quad (10)$$

where D represents the radial cut, and N_B is a normalization factor and a free parameter of the fit.

Fitting the combination of the signal and the background distribution, allows their separation for an unbiased analysis of the dSiPMs spatial hit resolution.

5.2.3. Cluster size

The cluster size is defined as the number of pixels contributing to a cluster, and reconstructed for all clusters selected for the spatial residuals. It exceeds one for certain types of sensors and operation conditions, mainly because of charge sharing. Charge sharing describes the fact that the deposited charge might be collected in more than one readout electrode, due to the angular incidence of a particle or diffusion of the generated charge carriers. It is an interesting observable for tracking detectors, as charge sharing, in combination with a charge measurement, allows for an interpolation between the centers of pixels within a cluster, which can improve the spatial resolution significantly. For a digital SiPM, this observable is expected to be dominated by optical crosstalk.

5.2.4. Temporal resolution

One of the most interesting characteristics of SiPMs and SPADs is their fast response due to the small rise-time of the signals, as described in Section 1. Similar to the spatial resolution, the temporal resolution can be determined using a reference measurement, defining a time residual

$$\Delta t = t_{ref} - t_{cluster}. \quad (11)$$

Here, t_{ref} is the time stamp provided by a reference detector, and $t_{cluster}$ is the time stamp the investigated dSiPM provides for a given cluster. This cluster timestamp is given by the earliest (smallest) pixel timestamp in the cluster, considering the two selected readout frames. The reference measurement can be taken by the TLU, i.e. the trigger time stamp, marking the first of the coinciding scintillator signals, or by the second dSiPM. In the former case, the temporal resolution of the reference measurement dominates the residual, and allows estimating only an upper limit for the temporal resolution of the dSiPM. This is an expected limitation, and the reason the setup with two dSiPMs is used.

The analysis of the corresponding temporal residuals is challenging, as they have a non-Gaussian shape. First, there is a background originating from dark counts, determining the time stamp of the quadrant TDC. The shape and width of the signal part, however, is subject to four major effects. The first two, are due to the finite effective bin width of the TDC, and bin width fluctuations, characterized by the so-called integral nonlinearity (INL). They are investigated by means of a statistical code density method, as presented in [14]. The measurements were repeated for the investigated samples. The mean bin width is found to be 93.1 ps, yielding an effective TDC resolution of 26.9 ps for the two investigated samples. The maximum INL is found to be 41 ps, but can be corrected for, by taking the individually measured bin widths into account when calculating the time stamps. The third contribution, comes from variations of the propagation delay between the pixels and the quadrant TDC. The studies with a laser setup, also discussed in [14], reveal that this effect can be as large as $326 \text{ ps} \pm 86 \text{ ps}$. It is corrected for using the measured propagation delays for each pixel, which improves the quality of the fit shown below. Otherwise, the impact of this correction is small, since the sample-to-sample variations of the propagation delay are small, and the physical alignment of the two sensors in the test beam setup is sufficiently accurate. The last contribution, comes from an inhomogeneity of the temporal response over the area of the SPAD. Evidence for that is given in [15,43], and further discussed in Section 6.3. Measurements with focused light in [1] indicate a similar effect in analog SiPMs. This results in an exponential tail in the temporal response, which is taken into account in the analysis of the residuals.

Overall, this means that the temporal response function of a single dSiPM is characterized by three features, namely a Gaussian peak, an exponential tail due to the slow SPAD response, and a flat background. The shape of the temporal residual between the two dSiPMs is given by the convolution of two of these temporal response functions, and an

empirical fit model is used to characterize the different features. The used probability density function (PDF) is

$$f(\Delta t) = (1 - p_{\text{noise}}) \cdot f_{\text{signal}}(\Delta t) + p_{\text{noise}} \cdot u(\Delta t), \quad (12)$$

where $u(\Delta t)$ denotes the uniform distribution, which is a sufficiently accurate description of the noise contribution in the vicinity of the peak, and p_{noise} determines its relative amplitude with respect to the signal PDF $f_{\text{signal}}(\Delta t)$. The signal is the sum of four additional PDFs

$$\begin{aligned} f_{\text{signal}}(\Delta t) = & p_{\text{fast}}^2 \cdot f_{\text{fast,fast}}(\Delta t) \\ & + p_{\text{fast}} \cdot (1 - p_{\text{fast}}) \cdot (f_{\text{fast,slow}}(\Delta t) + f_{\text{slow,fast}}(\Delta t)) \\ & + (1 - p_{\text{fast}})^2 \cdot f_{\text{slow,slow}}(\Delta t), \end{aligned} \quad (13)$$

where p_{fast} denotes the proportion of a single dSiPMs fast signal to the entire signal contribution. The PDF $f_{\text{fast,fast}}$ describes the situation where both dSiPMs exhibit a fast response, and is modeled as a Gaussian distribution with the mean t_{μ} , and standard deviation of $\sigma_t \sqrt{2}$, such that σ_t describes the standard deviation of a single dSiPMs fast response. The case where both dSiPMs are slow is described by

$$f_{\text{slow,slow}}(\Delta t | \lambda) = \frac{\lambda}{2} \exp(-\lambda |\Delta t|). \quad (14)$$

Finally there is a case where one of the dSiPM signals is fast, and the other slow. The convolution of a Gaussian and an exponential distribution, proposed in [1], is used to model this contribution. The corresponding PDF reads

$$\begin{aligned} f_{\text{fast,slow}}(\Delta t | \sigma_t, \lambda, t_{\mu}, t_{\text{tail}}) = & \frac{\lambda}{2} \exp\left(\lambda(t_{\mu} + t_{\text{tail}} - \Delta t) + \frac{1}{2} \lambda^2 \sigma_t^2\right) \\ & \text{erfc}\left(\frac{t_{\mu} + t_{\text{tail}} - \Delta t + \lambda \sigma_t^2}{\sigma_t \sqrt{2}}\right). \end{aligned} \quad (15)$$

The PDF $f_{\text{slow,fast}}$ is defined similarly, but such that the exponential tail is on the opposite side.

An unbinned maximum likelihood estimator is used to fit the combined PDF to the temporal residuals. This approach is chosen, as it is independent of the binning of the histogram, and therefore more robust in the presence of bin-width fluctuations. Exemplary histograms and the results of the fitting procedure will be discussed in the next section, while more information on initial values and limits of the fit parameters are discussed in [43].

It should be noted, that the data selection for the analysis of the temporal resolution is slightly different with respect to the aforementioned observables. The first difference is related to the fact, that the dead time mentioned in Section 5.2.1, relates to hit detection, while there is some additional dead time while the TDCs are being reset. This leads to events with void time stamps, which are discarded for the analysis of timing properties. The second difference is that the track intercept is not required to be within the region of interest mentioned above. This increases the number of entries in the time-residual distribution, to compensate for the events discarded due to void time stamps.

6. Performance of the dSiPM

The performance of a tracking detector is characterized by its capability of detecting MIPs. The key benchmarks are the detection efficiency, the spatial, and the temporal resolution.

6.1. Hit-detection efficiency

One of the major differences in the use of SPADs for MIP detection, compared to their conventional use as photodetectors, lies in the energy deposited by a MIP. While a photon in the optical wavelength regime generates a single electron-hole pair, a MIP generates more than 50 electron-hole pairs per micrometer [44,45]. This gives raise to the assumption, that the hit-detection efficiency of the dSiPM is close to the nominal fill factor of 30%.

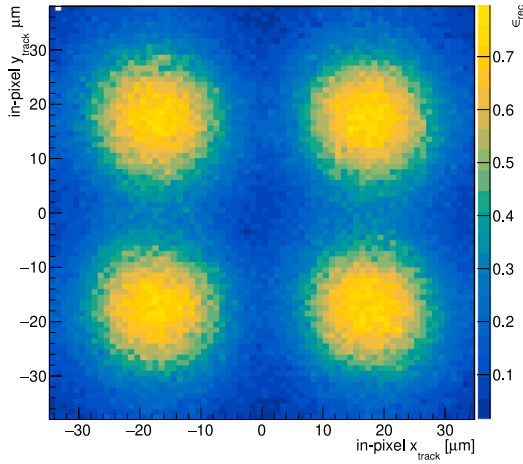


Fig. 11. Average hit-detection efficiency as a function of the in-pixel position, measured at an overvoltage of 2 V. The regions of maximum efficiency correspond to the position of the SPADs, shown in Fig. 1.

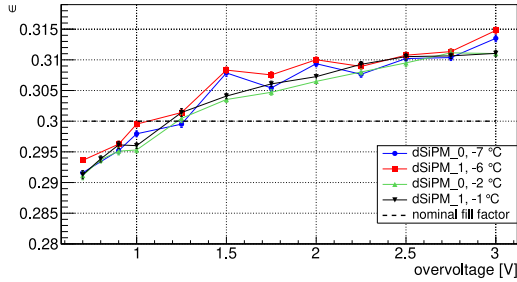


Fig. 12. Hit-detection efficiency of two different DESY dSiPMs as a function of overvoltage, for two temperature settings, which result in slightly different chip temperatures due to different cooling contacts. The expectation from the nominal fill factor is included.

This is confirmed in Fig. 11, where the reconstructed efficiency ϵ_{rec} is shown as a function of the in-pixel position, also discussed in [15]. The in-pixel positions correspond to the reconstructed track intercept with the dSiPM, overlaying all selected pixels to obtain sufficient statistics. Comparing Fig. 11 with Fig. 1, shows that the efficiency reaches a maximum when the track intercepts the center of the SPAD, and a minimum in-between the SPADs. One should note that the tracking uncertainty determines the resolution of the in-pixel efficiency image. It is assumed that the efficiency in the SPAD center is actually close to 100 %, and appears lower due to the smearing caused by this finite resolution.

Fig. 12 shows the hit-detection efficiency ϵ of the two installed dSiPMs (labeled dSiPM_0 and dSiPM_1) as a function of the overvoltage, and for two different temperature settings. The results include the corrections introduced in Eq. (7), and the error bars indicate statistical uncertainties only. All four measurements are found to agree within 0.5 %. The small systematic uncertainties are assumed to be due to the relative movement between the dSiPMs and the beam telescope, discussed in Section 5.1.

The hit-detection efficiency reaches a value of $\sim 31\%$ at 3 V. This is slightly above the nominal fill factor of 30 %, and significantly higher than the expected photon-detection efficiency, due to the large amount of electron-hole pairs created by a MIP, which is assumed to cause a significant avalanche probability even when they are created in an area with a vanishing electric field. Whether the slight increase of the hit-detection efficiency with rising overvoltage is due to an increasing avalanche trigger probability, as it is the case for single photons, or a growth of the active area cannot be resolved due to the limited track

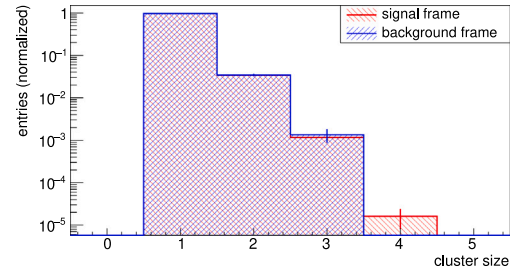


Fig. 13. Cluster size distribution for associated cluster reconstructed within a single frame, chosen to be in time (signal frame) or out-of-time (background frame) with respect to the particle interaction. Measured at an overvoltage of 2.0 V and a temperature of -2°C .

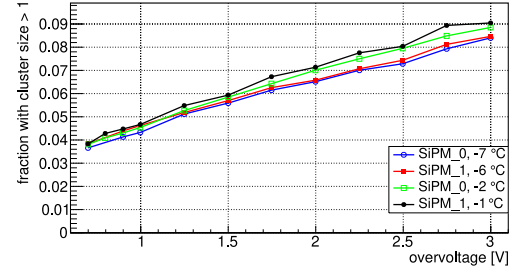


Fig. 14. Fraction of events with cluster size greater than one, for two different DESY dSiPMs as a function of overvoltage, for two temperature settings.

resolution with this setup. Operating at larger overvoltages was not attempted, as further increasing signal amplitudes could damage the in-pixel electronics. For lower overvoltages, the efficiency drops sharply, as the amplitude of the signals becomes too small to pass the threshold of the inverter used as discriminator.

6.2. Cluster size and spatial residuals

Due to the layout of the pixels and SPADs, which are relatively wide compared to the typical thickness of the absorption region, and separated by dedicated isolation structures [14], the cluster size for MIP interactions is expected to be close to one. This is confirmed by measurement, as shown in Fig. 13, where the size of associated clusters reconstructed within a single frame is shown. This allows to compare to a second frame, which is chosen to be late with respect to the electron interaction, and hence contains predominantly dark counts that are by coincidence associated with a track. The difference between the two distributions is insignificant. It is thus concluded, that the spatial distribution and amount of charge carriers generated in a MIP interaction comes with a vanishing probability of triggering an avalanche in more than one pixel.

Fig. 14 shows the percentage of events with a DUT cluster size greater than one as a function of the overvoltage, for different temperatures and samples. It shows a marginal dependence on temperature, but increases by more than a factor of two between the lowest and highest studied overvoltage. One should note that there are two possible reasons why the cluster size can exceed one. The first is crosstalk, the second is a random coincidence with noise in a neighboring pixel. Due the latter contribution, the numbers are larger than the single pixel measurements in Section 3 suggest. The increase in cluster size does not have a beneficial impact on the spatial resolution, as discussed below.

Fig. 15 shows the distribution of the spatial residual Δx , defined in Eq. (8), recorded at an overvoltage of 2 V and a temperature of -2°C . The distribution assumes a characteristic shape, with a few entries around zero, due to the inefficient region between the SPADs in the pixel center, which is also apparent in Fig. 11. The sum of

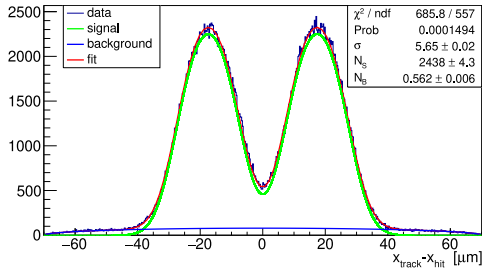


Fig. 15. Distribution of the spatial residual dx , measured at an overvoltage of 2 V and a temperature of -2°C . The fitted sum of the signal $S(dx)$ and background $B(x)$ parameterization, as well as the individual contributions, are also shown.

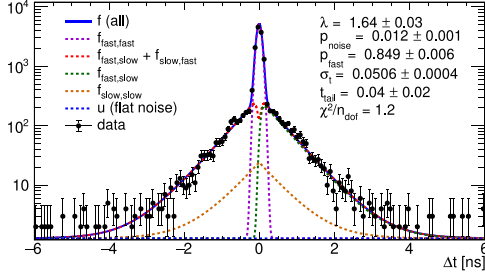


Fig. 16. Distribution of the temporal residual between the two dSiPMs, recorded at an overvoltage of 1.5 V, and temperatures of -7°C and -6°C . The fitted PDF and selected subcomponents, as well as key fit parameters, are shown. The unbinned likelihood fit is constrained to the displayed range, the bin width corresponds to 80 ps.

the signal $S(dx)$ and background $B(x)$ parameterizations, introduced in Section 5.2.1, is fitted to the data and also shown in the figure. The shape of the background distribution is validated by analyzing data from a frame which is chosen to be late with respect to the electron interaction, as done to obtain the cluster size distribution for background events. This allows to determine the standard deviation of the signal contribution, which is found to be $19.2\mu\text{m}$. This result is found to be robust with respect to the fitting procedure. The main sources of uncertainty is a signal contribution to the background $B(x)$ stemming from two pixel clusters. An upper boundary for the latter effect is given by $20.9\mu\text{m}$, the RMS of the full distribution in Fig. 15. Taking a track resolution of the reference system of $3.5\mu\text{m}$ into account, yields a range of $18.9\mu\text{m}$ to $20.6\mu\text{m}$. This is in agreement with the expectation for a sensor with binary readout, without charge sharing, and the given pitch of $69.6\mu\text{m}/\sqrt{12} = 20.1\mu\text{m}$. One should note that the fit parameter $\sigma = 5.65\mu\text{m}$ exceeds the expected track resolution, which might be due to unaccounted material in the setup. On the other hand, also the large amount of deposited charge carriers and diffusion effects might lead to a certain smearing of the distribution at the edge of the SPAD. It is assumed that a combination of these effects leads to the finite amount of entries around zero.

Finally, the trend observed in Fig. 14 suggests a certain dependence of the spatial resolution on the overvoltage. It is found that the spatial resolution deteriorates by about 15 % from the smallest to the largest overvoltage, which is assumed to be due to the fact that the increase in cluster size is dominantly due to noise effects.

6.3. Temporal residuals

The model and fitting procedure introduced in Section 5.2.4 allow quantifying the temporal resolution of the dSiPM. Fig. 16 shows an example for the distribution of the residual between the time stamps provided by the two dSiPMs, including the different contributions from the fit model. The fit results indicate, that the fraction of the fast signal

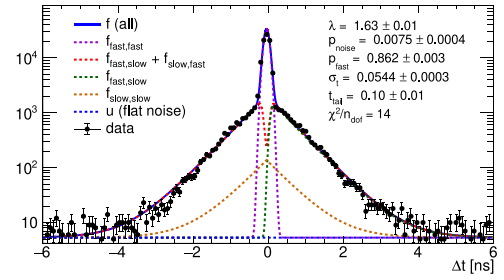


Fig. 17. Distribution of the temporal residual between the two dSiPMs, recorded at an overvoltage of 0.9 V, and temperatures of -7°C and -6°C . The fitted PDF and selected subcomponents, as well as key fit parameters, are shown. The unbinned likelihood fit is constrained to the displayed range, the bin width corresponds to 80 ps.

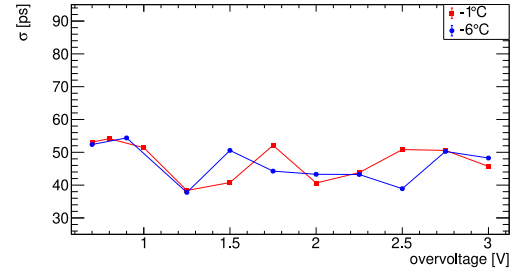


Fig. 18. Fit parameter σ for two different temperatures, and as a function of the overvoltage. This corresponds to the temporal resolution of the dSiPM in about 85 % of the data.

amounts to about 85 % of the total signal. The temporal resolution σ_i of this fast part is 51 ps,⁵ while the exponential tails are characterized by time constant $\lambda = 1/1.64\text{ ns}$. The reduced chi-squared $\chi^2/n_{\text{dof}} = 1.2$ of this fit is the best for all measurements. The fit with the worst value, $\chi^2/n_{\text{dof}} = 14$, is shown in Fig. 17. This might hint towards shortcomings of the empirical fit model, uncovered by the significantly smaller statistical uncertainties.

Fig. 18 shows the width of the central peak for two different temperatures, and as a function of the overvoltage. No trend with respect to these two parameters is observed. However, the data shows fluctuations which are significantly larger than the statistical uncertainties derived from the fit. This could be due to run-to-run variations of the phase between the clock signals. As a consequence, the mean over all data points — $47 \pm 6\text{ ps}$ — is taken as the best estimate of the dSiPMs temporal resolution. The other fit parameters are evaluated in the same fashion. The time constant λ shows a flat behavior within 1.51 ns and 1.64 ns, while the peak fraction p_{peak} is within 83 to 87 %. Only the noise fraction varies as a function of temperature and overvoltage, as could be expected from the previous section. It is found to be within 0.5 to 3.5 %.

The origin of the slow component responsible for the exponential tail was studied in [15,43]. Events with a delay between 1 ns and 5 ns with respect to the trigger time-stamp are selected, and the intercepts of the reconstructed particle tracks with the dSiPM are folded into a single SPAD (similar to Fig. 11). The result is shown in Fig. 19 and indicates that most of the selected events come from a ring than corresponds to the outer region of the SPADs. The reason is assumed to be a low electric-field strength in the edge region of the SPADs, which means that charge carriers generated there require some time to reach the amplification region and to trigger an avalanche.

⁵ The fitted function is defined such that σ_i denotes the resolution of a single dSiPM, assuming that the responses of the 2 dSiPMs are uncorrelated.

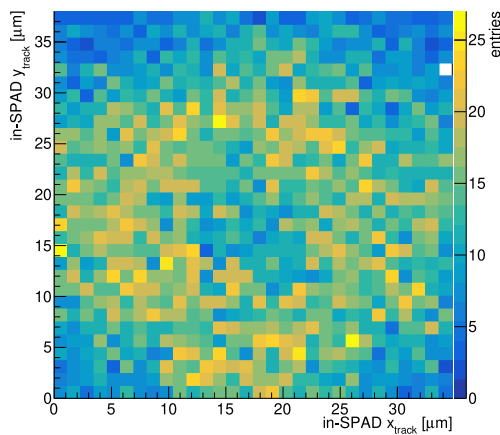


Fig. 19. Number of events with a delay between 1 ns and 5 ns as a function of the intercepts of the reconstructed particle tracks with the dSiPM. Projected into a single SPAD to account for reduced statistics.

Source: Taken from [15].

7. Conclusion and outlook

This work focuses on the test-beam characterization of a dSiPM prototype, designed at DESY in a standard 150 nm CMOS technology, but includes also a discussion of peculiarities of the experimental methods. The results allow an assessment of the MIP detection capability of the prototype. It is shown that the MIP detection efficiency is on the order 31 %, only weakly depending on the overvoltage, and dominated by the fill factor of the prototype. The spatial resolution is found to be around 20 μm , in agreement with expectation. Studies on crosstalk reveal that the large number of charge carriers generated by a MIP comes with an insignificant increase in crosstalk probability. In addition, it is concluded, that allowing to mask individual SPADs would allow a similar reduction of the DCR compared to masking entire pixels, which would reduce the total masked area. Studies on timing performance indicate a fast response, with a temporal resolution of 47 ± 6 ps for about 85 % of the detected events. The remaining 15 % exhibit a slower response, order of nanoseconds, which correlates with particles hitting the edge of the SPADs.

Regarding the prospects of applying the investigated technology in charged particle tracking or track timing, the reported MIP detection efficiency needs to be improved. This could be achieved by increasing the fill factor, moving to larger SPADs, and narrower SPAD isolation, or even charge focusing, as presented in [46]. Moving from the LFoundry 150-nm to the 110-nm CMOS technology, which is optimized for CMOS Image Sensors, would allow profiting from improvements in the SPAD performance, reported to be related to DCR [47], but also photon detection efficiency. A different approach is investigated in [48], where thin (few 100 μm) scintillating crystals are coupled to the same dSiPM prototype, to improve the MIP detection efficiency through the generation of additional photons.

CRedit authorship contribution statement

Finn King: Supervision, Conceptualization, Writing – original draft, Project administration, Investigation, Writing – review & editing, Formal analysis, Software, Methodology, Data curation. **Inge Diehl:** Writing – review & editing, Resources, Investigation, Conceptualization, Validation, Data curation, Visualization, Formal analysis, Methodology. **Ono Feyens:** Visualization, Writing – review & editing, Validation, Investigation, Methodology. **Ingrid-Maria Gregor:** Conceptualization, Funding acquisition, Project administration. **Karsten Hansen:** Resources, Conceptualization, Supervision, Project administration, Methodology, Funding acquisition. **Stephan Lachnit:**

Investigation, Visualization, Software, Formal analysis, Methodology, Validation, Data curation. **Frauke Poblitzki:** Resources. **Daniil Rastorguev:** Validation, Methodology, Resources, Visualization, Software, Investigation, Data curation, Formal analysis. **Simon Spannagel:** Writing – review & editing, Software, Methodology, Supervision, Project administration, Funding acquisition, Conceptualization. **Tomas Vanat:** Conceptualization, Resources, Software, Methodology. **Gianpiero Vignola:** Conceptualization, Writing – review & editing, Validation, Software, Investigation, Data curation, Methodology, Writing – original draft, Supervision, Visualization, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Measurements leading to some of the presented results have been performed at the Test Beam Facility at DESY Hamburg (Germany), a member of the Helmholtz Association (HGF). The authors would like to thank A. Venzmer, E. Wüstenhagen and D. Gorski for test-board and mechanical case design, test setup, as well as chip assembly.

Data availability

Data will be made available on request.

References

- [1] M. Nemallapudi, S. Gundacker, P. Lecoq, et al., Single photon time resolution of state of the art SiPMs, *J. Instrum.* 11 (10) (2016) P10016, <http://dx.doi.org/10.1088/1748-0221/11/10/P10016>.
- [2] F. Acerbi, S. Gundacker, Understanding and simulating SiPMs, *Nucl. Instrum. Methods Phys. Res. A* 926 (2019) 16–35, <http://dx.doi.org/10.1016/j.nima.2018.11.118>, Silicon Photomultipliers: Technology, Characterisation and Applications.
- [3] F. Simon, Silicon photomultipliers in particle and nuclear physics, *Nucl. Instrum. Methods Phys. Res. A* 926 (2019) 85–100, <http://dx.doi.org/10.1016/j.nima.2018.11.042>, Silicon Photomultipliers: Technology, Characterisation and Applications.
- [4] M.G. Bisogni, A. Del Guerra, N. Belcarì, Medical applications of silicon photomultipliers, *Nucl. Instrum. Methods Phys. Res. A* 926 (2019) 118–128, <http://dx.doi.org/10.1016/j.nima.2018.10.175>, Silicon Photomultipliers: Technology, Characterisation and Applications.
- [5] T. Frach, G. Prescher, C. Degenhardt, et al., The digital silicon photomultiplier — Principle of operation and intrinsic detector performance, 2009, pp. 1959–1965, <http://dx.doi.org/10.1109/NSSMIC.2009.5402143>, 2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC).
- [6] G. Torilla, J. Minga, P. Brogi, et al., DCR and crosstalk characterization of a bi-layered 24×72 CMOS SPAD array for charged particle detection, *Nucl. Instrum. Methods Phys. Res. A* 1046 (2023) 167693, <http://dx.doi.org/10.1016/j.nima.2022.167693>.
- [7] L. Ratti, P. Brogi, G. Collazuol, et al., Layered CMOS SPADs for low noise detection of charged particles, *Front. Phys.* 8 (2021) <http://dx.doi.org/10.3389/fphy.2020.607319>.
- [8] P. Fischer, R.K. Zimmermann, B. Maisano, CMOS SPAD sensor chip for the readout of scintillating fibers, *Nucl. Instrum. Methods Phys. Res. A* 1040 (2022) 167033, <http://dx.doi.org/10.1016/j.nima.2022.167033>.
- [9] N. Cartiglia, R. Arcidiacono, M. Costa, et al., 4D tracking: present status and perspectives, *Nucl. Instrum. Methods Phys. Res. A* 1040 (2022) 167228, <http://dx.doi.org/10.1016/j.nima.2022.167228>.
- [10] W. Riegler, P. Windischhofer, Time resolution and efficiency of SPADs and SiPMs for photons and charged particles, *Nucl. Instrum. Methods Phys. Res. A* 1003 (2021) 165265, <http://dx.doi.org/10.1016/j.nima.2021.165265>.
- [11] F. Carnesecchi, G. Vignola, N. Agrawal, et al., Direct detection of charged particles with SiPMs, *J. Instrum.* 17 (06) (2022) P06007, <http://dx.doi.org/10.1088/1748-0221/17/06/P06007>.
- [12] F. Gramuglia, E. Ripicini, C.A. Fenoglio, et al., Sub-10 ps minimum ionizing particle detection with Geiger-mode APDs, *Front. Phys.* 10 (2022) <http://dx.doi.org/10.3389/fphy.2022.849237>.

- [13] E. Vilella, O. Alonso, A. Diéguez, 3D integration of Geiger-mode avalanche photodiodes aimed to very high fill-factor pixels for future linear colliders, *Nucl. Instrum. Methods Phys. Res. A* 731 (2013) 103–108, <http://dx.doi.org/10.1016/j.nima.2013.05.022>, Pixel 2012.
- [14] I. Diehl, K. Hansen, T. Vanat, et al., Monolithic MHz-frame rate digital SiPM-IC with sub-100ps precision and 70µm pixel pitch, *J. Instrum.* 19 (01) (2024) P01020, <http://dx.doi.org/10.1088/1748-0221/19/01/P01020>.
- [15] I. Diehl, F. Feindt, K. Hansen, et al., The DESY digital silicon photomultiplier: Device characteristics and first test-beam results, *Nucl. Instrum. Methods Phys. Res. A* 1064 (2024) 169321, <http://dx.doi.org/10.1016/j.nima.2024.169321>.
- [16] T. Vanat, Caribou – A versatile data acquisition system, PoS TWEPP2019 (2020) 100, <http://dx.doi.org/10.22323/1.370.0100>.
- [17] Caribou developers, The Caribou DAQ System, 2024, URL <https://gitlab.cern.ch/Caribou>. (Accessed June 2024).
- [18] Advanced Micro Devices, Inc., 2024, URL <https://www.amd.com>. (Accessed July 2024).
- [19] P. Ahlburg, S. Arfaoui, J.-H. Arling, et al., EUDAQ — a data acquisition software framework for common beam telescopes, *J. Instrum.* 15 (01) (2020) P01038, <http://dx.doi.org/10.1088/1748-0221/15/01/P01038>.
- [20] EUDAQ developers, EUDAQ code repository, 2024, URL <https://github.com/eudaq/eudaq>. (Accessed March 2024).
- [21] R. Klanner, Characterisation of SiPMs, *Nucl. Instrum. Methods Phys. Res. A* 926 (2019) 36–56, <http://dx.doi.org/10.1016/j.nima.2018.11.083>, Silicon Photomultipliers: Technology, Characterisation and Applications.
- [22] Angelantoni Test Technologies, 2024, URL <https://www.acstestchambers.com>. (Accessed February 2024).
- [23] Keithley, a Tektronix company, 2024, URL <https://www.tek.com/en/products/keithley>. (Accessed February 2024).
- [24] L. Gallego, J. Rosado, F. Blanco, et al., Modeling crosstalk in silicon photomultipliers, *J. Instrum.* 8 (5) (2013) <http://dx.doi.org/10.1088/1748-0221/8/05/P05010>.
- [25] D. McNally, V. Golovin, Review of solid state photomultiplier developments by CPTA and photonique SA, *Nucl. Instrum. Methods Phys. Res. A* 610 (1) (2009) 150–153, <http://dx.doi.org/10.1016/j.nima.2009.05.140>, New Developments In Photodetection NDIP08.
- [26] Z. Liu, S. Gundacker, M. Pizzichemi, et al., In-depth study of single photon time resolution for the Philips digital silicon photomultiplier, *J. Instrum.* 11 (06) (2016) P06006, <http://dx.doi.org/10.1088/1748-0221/11/06/P06006>.
- [27] E. Engelmann, E. Popova, S. Vinogradov, Spatially resolved dark count rate of SiPMs, *Eur. Phys. J. C* 78 (971) (2018) 8, <http://dx.doi.org/10.1140/epjc/s10052-018-6454-0>.
- [28] ZEISS Microscopy, 2024, URL www.zeiss.com. (Accessed April 2024).
- [29] Jenoptik, 2024, URL www.jenoptik.com. (Accessed March 2024).
- [30] R. Diener, J. Dreyling-Eschweiler, H. Ehrlichmann, et al., The DESY II test beam facility, *Nucl. Instrum. Methods Phys. Res. A* 922 (2019) 265–286, <http://dx.doi.org/10.1016/j.nima.2018.11.133>.
- [31] G. Molière, Theorie der Streuung schneller geladener Teilchen II — Mehrfach- und Vielfachstreuung, *Z. Naturforsch. Teil A* 3a (1947) 78–97, <http://dx.doi.org/10.1515/zna-1948-0203>.
- [32] H.A. Bethe, Molière's theory of multiple scattering, *Phys. Rev.* 89 (1953) 1256–1266, <http://dx.doi.org/10.1103/PhysRev.89.1256>.
- [33] H. Jansen, S. Spannagel, J. Behr, et al., Performance of the EUDET-type beam telescopes, *EPJ Techn. Instrum.* 3 (7) (2006) 1–20, <http://dx.doi.org/10.1140/epjti/s40485-016-0033-2>.
- [34] C. Hu-Guo, J. Baudot, G. Bertolone, et al., First reticule size MAPS with digital output and integrated zero suppression for the EUDET-JRA1 beam telescope, *Nucl. Instrum. Methods Phys. Res. A* 623 (1) (2010) 480–482, <http://dx.doi.org/10.1016/j.nima.2010.03.043>, 1st International Conference on Technology and Instrumentation in Particle Physics.
- [35] 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, <http://dx.doi.org/10.1088/1748-0221/14/09/P09019>.
- [36] D. Dannheim, K. Dort, L. Huth, et al., Corryvreckan: a modular 4D track reconstruction and analysis software for test beam data, *J. Instrum.* 16 (03) (2021) P03008, <http://dx.doi.org/10.1088/1748-0221/16/03/p03008>.
- [37] The Corryvreckan Authors, Corryvreckan – The Maelstrom for Your Test Beam Data, 2024, URL <http://cern.ch/corryvreckan>. (Accessed January 2024).
- [38] V. Blobel, C. Kleinwort, F. Meier, Fast alignment of a complex tracking detector using advanced track models, *Comput. Phys. Comm.* 182 (9) (2011) 1760–1763, <http://dx.doi.org/10.1016/j.cpc.2011.03.017>, Computer Physics Communications Special Edition for Conference on Computational Physics Trondheim, Norway, June 23–26, 2010.
- [39] C. Kleinwort, General broken lines as advanced track fitting method, *Nucl. Instrum. Methods Phys. Res. A* 673 (2012) 107–110, <http://dx.doi.org/10.1016/j.nima.2012.01.024>.
- [40] H. Jansen, P. Schütze, Feasibility of track-based multiple scattering tomography, *Appl. Phys. Lett.* 112 (14) (2018) 144101, <http://dx.doi.org/10.1063/1.5005503>.
- [41] G.R. Lynch, O.I. Dahl, Approximations to multiple Coulomb scattering, *Nucl. Instrum. Methods Phys. Res. B* 58 (1) (1991) 6–10, [http://dx.doi.org/10.1016/0168-583X\(91\)95671-Y](http://dx.doi.org/10.1016/0168-583X(91)95671-Y).
- [42] S. Spannagel, H. Jansen, GBL Track Resolution Calculator v2.0, URL <http://dx.doi.org/10.5281/zenodo.48795>.
- [43] S. Lachnit, Time Resolution of a Fully-Integrated Digital Silicon Photo-Multiplier, University of Hamburg, 2024, URL <https://bib-pubdb1.desy.de/record/602203>.
- [44] H. Bichsel, Straggling in thin silicon detectors, *Rev. Modern Phys.* 60 (1988) 663–699, <http://dx.doi.org/10.1103/RevModPhys.60.663>.
- [45] H. Kolanoski, N. Wermes, *Particle Detectors: Fundamentals and Applications*, Oxford University Press, 2020.
- [46] K. Morimoto, J. Iwata, Shinohara, et al., 3.2 megapixel 3D-stacked charge focusing SPAD for low-light imaging and depth sensing, 2021, pp. 20.2.1–20.2.4, <http://dx.doi.org/10.1109/IEDM19574.2021.9720605>, 2021 IEEE International Electron Devices Meeting (IEDM).
- [47] P. Organtini, A. Gola, G. Paternoster, et al., Industrial exploitation of SiPM technology developed for basic research, *Nucl. Instrum. Methods Phys. Res. A* 978 (2020) 164410, <http://dx.doi.org/10.1016/j.nima.2020.164410>.
- [48] I. Diehl, F. Feindt, I.-M. Gregor, et al., 4D-tracking with digital SiPMs, *Nucl. Instrum. Methods Phys. Res. A* 1069 (2024) 169985, <http://dx.doi.org/10.1016/j.nima.2024.169985>.