Technical design of the phase I Mu3e experiment

P. Baesso[a], K. Clark[a], J. Goldstein[a], J. Velthuis[a], A. Bravar[b], K. Briggl[b], F. Cadoux[b], Y. Demets[b], A. Damyanova[b], D. La Marra[b], M. Pohl[b], H. Augustin[c], S. Dittmeier[c], D. Gottschalk[c], J. Hammerich[c,1], A. Herkert[c], S. Hetzel[c], L. Huth[c,2], D. Immig[c], M. Kiehn[c,3], A. Meneses[c], S. Muley[c], A.-K. Perrevoort[c,4], T. Rudzki[c], A. Schöning[c,*], S. Shrestha[c,5], K. Stumpf[c], L. Vigani[c], D. Wiedner[c,6], B. Windelband[c], H. Chen[d], P. Eckert[d], H. Klingenmeyer[d], Y. Munwes[d], H.-C. Schultz-Coulon[d], W. Shen[d], T.C. Zhong[d], F. Ehrler[e], R. Leys[e], I. Perić[e], R. Schimassek[e], A. Weber[e], P. Cooke[f], C. Chavez Barajas[f], A. Groves[f], H. Hayward[f], S. Hughes[f], T. Jones[f], N. Rompotis[f], E. Vilella-Figueras[f], J. Vossebeld[f], A. Wasili[f], G. Hesketh[g], S. Kilani[g], I. Paraskevas[g], N. Berger[h], U. Hartenstein[h], Q.H. Huang[h,7], M. Köppel[h], A. Kozlinskiy[h], M. Müller[h], D. vom Bruch[h,8], F. Wauters[h], K. Arndt[†i], D. Bortoletto[i], M. Jones[i], A. Knight[i], G. Lockwood[i], A. Loreti[i,9], K. Metodiev[i], R. Plackett[i], I. Shipsey[i], D. Fahrni[j], M. Hildebrandt[j], A. Hofer[j], H.-C. Kästli[j], U. Hartmann[j], Z. Hodge[j,10], P.-R. Kettle[j], A. Knecht[j], B. Kotlinski[j], B. Meier[j], F. Meier Aeschbacher[j,**,11], A. Mtchedlishvili[j], A. Papa[j,12], S. Ritt[j,*], G. Rutar[j], A. Stoykov[j], S. Streuli[j], F. Berg[k], S. Corrodi[k,13], L. Gerritzen[k], C. Grab[k], R. Wallny[k], C. Betancourt[l], R. Gredig[l], P. Owen[l], P. Robmann[l], N. Serra[l], O. Steinkamp[l] and U. Straumann[l]

[a] *University of Bristol, H.H. Wills Physics Laboratory, Tyndall Avenue, Bristol BS8 1TL, United Kingdom*
[b] *Département de physique nucléaire et corpusculaire, Université de Genève, 24, quai Ernest-Ansermet, 1211 Genève 4, Switzerland*
[c] *Physikalisches Institut, Universität Heidelberg, Im Neuenheimer Feld 226, 69120 Heidelberg, Germany*
[d] *Kirchoff-Institut für Physik, Universität Heidelberg, Im Neuenheimer Feld 227, 69120 Heidelberg, Germany*
[e] *Institut für Prozessdatenverarbeitung und Elektronik, Karlsruhe Institut für Technologie, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany*
[f] *Department of Physics, University of Liverpool, The Oliver Lodge Laboratory, Liverpool L69 7ZE, United Kingdom*
[g] *Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom*
[h] *Institut für Kernphysik und Exzellenzcluster PRISMA$^+$, Johannes Gutenberg-Universität Mainz, Johann-Joachim-Becher-Weg 45, 55128 Mainz, Germany*
[i] *Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, United Kingdom*
[j] *Laboratory for Particle Physics, Paul Scherrer Institut, Forschungsstrasse 111, 5232 Villigen, Switzerland*
[k] *Institute for Particle Physics and Astrophysics, Eidgenössische Technische Hochschule Zürich, Otto-Stern-Weg 5, 8093 Zürich, Switzerland*
[l] *Physik-Institut, Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland*

ARTICLE INFO

Keywords:
lepton flavour violation
muon decays
monolithic pixel detector
scintillating fibres
scintillating tiles

Abstract

This is revision: `1706706a` on branch `frank`.

The Mu3e experiment aims to find or exclude the lepton flavour violating decay $\mu \rightarrow eee$ at branching fractions above 10^{-16}. A first phase of the experiment using an existing beamline at the Paul Scherrer Institute (PSI) is designed to reach a single event sensitivity of $2 \cdot 10^{-15}$. Here we present the complete technical design of this phase I Mu3e detector. The high rate of up to 10^8 muon decays per second and the low momenta of the decay electrons and positrons pose a unique set of challenges, which we tackle using an ultra thin tracking detector based on high-voltage monolithic active pixel sensors combined with scintillating fibres and tiles for precise timing measurements.

* Spokesperson
** Technical Coordinator
`schoning@physi.uni-heidelberg.de` (A. Schöning); `frank.meier@psi.ch` (F. Meier Aeschbacher); `stefan.ritt@psi.ch` (S. Ritt)
ORCID(s):

[1] Now at University of Liverpool, UK
[2] Now at DESY, Hamburg, Germany
[3] Now at Université de Genève, Switzerland
[4] Now at NIKHEF, Amsterdam, The Netherlands
[5] Now at Arcadia University, Glenside, PA, USA
[6] Now at Technische Universität Dortmund, Germany
[7] Now at Institut Polytechnique de Paris, France
[8] Now at Université de Marseille, France
[9] Also at University of Bristol, UK
[10] Now at University of Washington, Seattle, WA, USA
[11] Also at Physikalisches Institut, Universität Heidelberg
[12] Also at University of Pisa, Department of Physics and INFN, Italy
[13] Now at Argonne National Laboratory, Lemont, IL, USA

1. The Decay $\mu \to eee$ and the Experimental Challenge

1.1. Goals of the Experiment

The goal of the Mu3e experiment is to observe the process $\mu \to eee$ if its branching fraction is larger than 10^{-16}, or otherwise to exclude a branching fraction of $> 10^{-16}$ at the 90 % confidence level. In order to achieve these goals, more than $1 \cdot 10^{17}$ muons have to be stopped in the detector (assuming a total reconstruction efficiency of 20 %) and any background mimicking the signal process suppressed to below the 10^{-16} level. The additional requirement of achieving these goals within a reasonable measurement time of one year of data taking dictates a muon stopping rate of $2 \cdot 10^9$ Hz, along with a high geometrical acceptance and efficiency for the experiment.

The current best source of low-energy muons in the world, the πE5 beam line at PSI, provides muon rates up to $1 \cdot 10^8$ Hz. Higher intensities are possible and currently under study in the *high intensity muon beam* (HiMB) project. However, the new beamlines will not be available before 2025. In order to establish the novel technologies for Mu3e, set first competitive limits and prepare for the very high intensity running, we plan to run a phase I experiment at πE5. The aim of this phase I experiment is a single event sensitivity of $2 \cdot 10^{-15}$ on the branching fraction, which would require $> 2.5 \cdot 10^{15}$ stopped muons[14] or $2.5 \cdot 10^7$ s (290 days) of run time at $1 \cdot 10^8$ Hz stopping rate. The present document describes the technical design of this phase I detector.

For more on the physics motivation and theory predictions, please consult the Mu3e letter of intent [1] and research proposal [2]. This chapter describes the kinematics of the signal and the main background sources, and how these motivate the design of the experiment. Running with $1 \cdot 10^8$ Hz of muon decays also poses challenges for the detectors, the data acquisition and the readout, which will be discusses in later chapters.

1.2. Signal Kinematics

To discriminate the signal from the background, energy and momentum conservation are exploited. The decay $\mu \to eee$ is assumed to be prompt, and the decaying muons are at rest. The vectorial sum of all decay particle momenta $\vec{p}_i$ should therefore vanish:

$$|\vec{p}_{tot}| = \left|\sum \vec{p}_i\right| = 0 \tag{1}$$

and the invariant mass, which is equal to the sum of the energies in the case of vanishing momentum, be equal to the muon mass:

$$m_{inv} = \sum p_i = \sum E_i = m_\mu. \tag{2}$$

The energies of the decay particles range from the electron mass up to half the muon mass, which is about 53 MeV. All decay particles must lie in a plane. Therefore, the decay is described by two independent variables in addition to three global rotation angles describing the orientation in space.

1.3. Modelling of the Signal

The decay dynamics for the $\mu \to eee$ signal are dependent on the unknown lepton flavour violating (LFV) mechanism. We typically assume a phase-space distribution for the signal electrons in our simulations, if not stated otherwise. In order to study effects of different decay dynamics, we utilise the general parametrised Lagrangian proposed by Kuno and Okada [3]:

$$\begin{aligned} L_{\mu\to eee} = & -\frac{4G_F}{\sqrt{2}}\,[m_\mu A_R\, \overline{\mu_R}\sigma^{\mu\nu}e_L F_{\mu\nu} \\ & + m_\mu A_L\, \overline{\mu_L}\sigma^{\mu\nu}e_R F_{\mu\nu} \\ & + g_1\,(\overline{\mu_R}e_L)\,(\overline{e_R}e_L) \\ & + g_2\,(\overline{\mu_L}e_R)\,(\overline{e_L}e_R) \\ & + g_3\,(\overline{\mu_R}\gamma^\mu e_R)\,(\overline{e_R}\gamma_\mu e_R) \\ & + g_4\,(\overline{\mu_L}\gamma^\mu e_L)\,(\overline{e_L}\gamma_\mu e_L) \\ & + g_5\,(\overline{\mu_R}\gamma^\mu e_R)\,(\overline{e_L}\gamma_\mu e_L) \\ & + g_6\,(\overline{\mu_L}\gamma^\mu e_L)\,(\overline{e_R}\gamma_\mu e_R) + H.c.\,] \end{aligned} \tag{3}$$

The form factors $A_{R,L}$ describe tensor type (dipole) couplings, mostly acquiring contributions from the photon penguin diagram, whereas the scalar-type ($g_{1,2}$) and vector-type ($g_3 - g_6$) form factors can be regarded as four fermion contact interactions, to which the tree diagram contributes at leading order. We generate different signal models by varying the relative strengths of the $A_{R,L}$ and $g_1 - g_6$ parameters.

1.4. Signal Acceptance

For a three-body decay with a priori unknown kinematics such as $\mu \to eee$, the acceptance has to be as high as possible in order to test new physics in all regions of phase space. To illustrate the phase space coverage needed, the energy spectrum of the highest energy decay particle (E_{max}) for various LFV coupling amplitudes is shown in Figure 1, and the fraction of events where all decay particles have energies above E_{min} is shown in Figure 2. For these figures, it can be seen that a high acceptance for the signal is only possible if the detector can reconstruct tracks with momenta ranging from half the muon mass down to a few MeV. This must be achieved with large solid angle coverage, limited by the beam entry and exit points preventing instrumentation.

1.5. Backgrounds

The Standard Model branching fraction for the $\mu \to eee$ process is $2.9 \cdot 10^{-55}$ (normal neutrino mass ordering) or $4.6 \cdot 10^{-55}$ (inverted ordering) [5]; the experiment therefore has no physics backgrounds, and the

[14] $N_{\text{required}} = 1/(s \cdot \epsilon)$ for a sensitivity s and a total efficiency $\epsilon \approx 20\,\%$ (phase I)

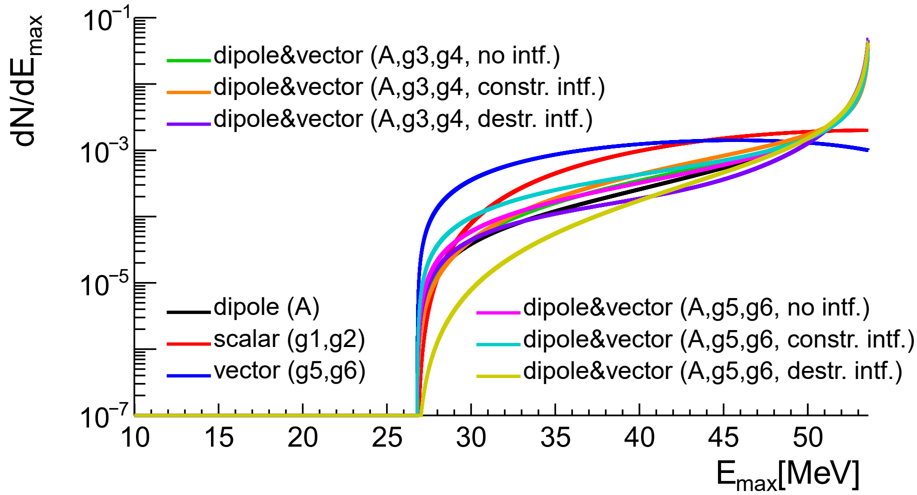


Figure 1: Energy distribution of the highest energy decay particle in the decay $\mu \to eee$ for different effective LFV models. The black line corresponds to pure dipole and the red and blue line to pure four-fermion contact interaction models (no penguin contribution); the other lines correspond to a mixture of dipole and vector interactions. Based on [3].

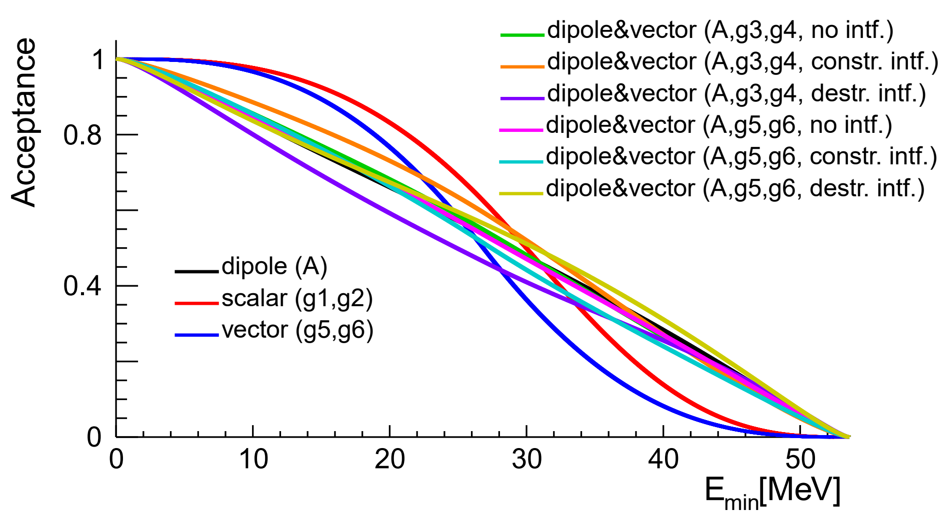


Figure 2: The acceptance, defined as the fraction of $\mu \to eee$ decays in which all decay products have energies above E_{min}, against E_{min} for different effective LFV models. The black line corresponds to pure dipole and the red and blue line to pure four-fermion contact interaction models (no penguin contribution); the other lines correspond to a mixture of dipole and vector interactions. Based on [3].

final sensitivity depends purely on the ability to reduce backgrounds in two categories: overlays of different processes producing three tracks resembling a $\mu \to eee$ decay (*combinatorial background*) and radiative decays with internal conversion (*internal conversion background*) with a small energy fraction carried away by the neutrinos. Combinatorial backgrounds have to be suppressed via vertexing, timing and momentum measurement; momentum measurement is the only handle on internal conversion. In the following sections, these main background sources are discussed.

1.5.1. Internal Conversions

The decay $\mu \to eee\nu\nu$ occurs with a branching fraction of $3.4 \cdot 10^{-5}$ [6]. It can be distinguished from the $\mu \to eee$ process by making use of energy and momentum conservation to infer the presence of the undetected neutrinos: in order to separate the $\mu \to eee$ events from $\mu \to eee\nu\nu$ events, the total momentum in the event is required to be zero and the visible mass

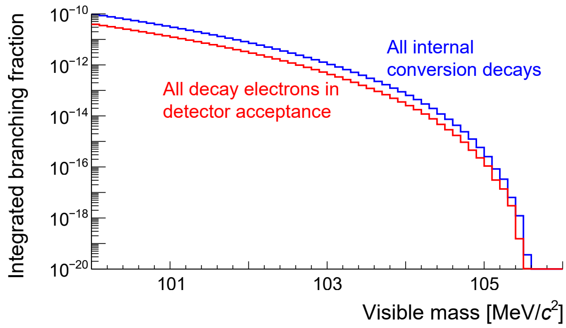


Figure 3: Integrated branching fraction of the decay $\mu \to ee e\nu\nu$ for which the invariant mass of the three decay electrons lies above the x axis value. This is shown for all internal conversion decays (blue line) and those with all three decay particles in the detector acceptance, defined as $E > 10\,\mathrm{MeV}$ and $|\cos\theta| < 0.8$ (red line). The matrix element was taken from [4].

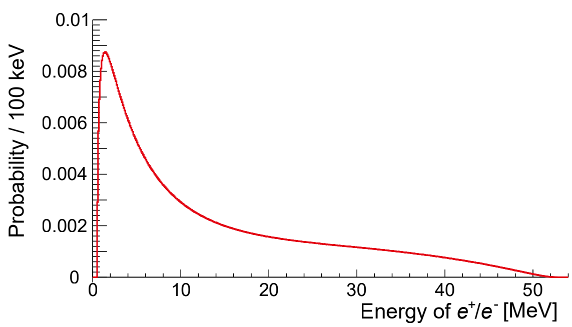


Figure 4: Energy spectrum of all electrons and positron from internal conversion decays.

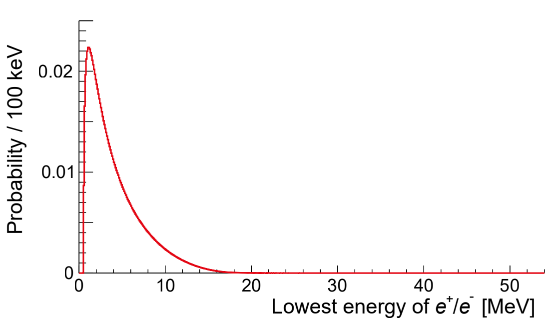


Figure 5: Energy spectrum of the electron or positron with lowest energy from internal conversion decays.

(defined as the invariant mass of the three electrons) equal to the muon rest energy. The branching fraction for $\mu \to eee\nu\nu$ [4] decays above a given visible mass value is shown in Figure 3. Figures 4 and 5 show the energy spectrum of all and the lowest energy electron from $\mu \to eee\nu\nu$ decays calculated with the matrix element from [4]. Recently, NLO calculations of the internal conversion decays have become available [7, 8];

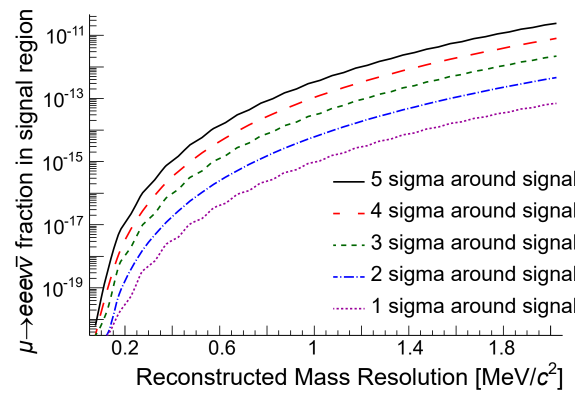


Figure 6: Contamination of the signal region (one sided cut) with internal conversion events as a function of momentum sum resolution.

they predict branching fractions close to the end-point that are about $10\,\%$ smaller than the LO prediction.

Internal conversion is the most serious background for the $\mu \to eee$ search, and the momentum resolution directly determines to what level it can be suppressed and thus the experiment run background free. In order to reach a sensitivity of $2 \cdot 10^{-15}$ with a 2σ cut on the reconstructed muon mass, the average momentum resolution has to be better than $1.0\,\mathrm{MeV}$, see Figure 6, and is not allowed to have sizeable tails towards the high momentum side.

1.5.2. Combinatorial Backgrounds

Michel Decays Using a beam of positive muons, one of the main processes contributing to combinatorial background is that of the Michel decay $\mu^+ \to e^+\nu_\mu\bar{\nu}_e$. This process does not produce a negatively charged particle (electron), so it can only contribute as a background in combination an incorrectly reconstructed track, or with with other processes that “naturally” provide negatively charged particles.

The rate of fake electron tracks being reconstructed from the recurling (incoming) section of a positron track is reduced by a reliable determination of the direction of motion of a particle: achieved by accurate curvature measurements and repeated timing measurements. The main sources of genuine negatively charged particles are Bhabha scattering and radiative decays.

Radiative Muon Decays The process $\mu^+ \to e^+\gamma\nu\bar{\nu}$ (branching fraction $1.4 \cdot 10^{-2}$ for photon energies above $10\,\mathrm{MeV}$ [9]) can deliver an negatively charged electron if the photon converts either in the target region or in the detector. Conversions in the target region generate an event topology similar to the radiative decay with internal conversion $\mu \to eee\nu\nu$ discussed above. Contributions from conversions outside the target region are greatly suppressed both by a vertex constraint and by minimising the material in both the target and detector. However, this process can still contribute to the combinatorial background in combination with an ordinary muon decay.

As for the internal conversion background, a NLO calculation for radiative decay has recently been published [10] and is implemented in our simulation.

Bhabha Scattering Any positron, either from a muon decay on target or in the beam, can undergo Bhabha scattering with electrons in the target material, leading to an electron-positron pair from a common vertex. In combination with a positron from a Michel decay, this can mimic a signal decay. In addition, Bhabha scattering is the main source of electrons for combinatorial background involving two Michel decays. Similarly to the external photon conversion background, the amount of Bhabha scattering is reduced by minimising both the amount, and the average atomic number of the material in the target.

Vertex and Timing Resolution Requirements Separating vertices from different muon decays is a key tool in suppressing combinatorial background. The vertex position resolution is essentially determined by the amount of multiple scattering (and thus material) in the innermost detector layer and the stopping target as well as the average distance between the vertex and the first detector layer.

At high muon rates, good time resolution is essential for reducing combinatorial background, while also facilitating event reconstruction. The combinatorial background has a component scaling linearly with the rate (e^+e^- pair plus a Michel positron) and a component quadratic in the rate (electron plus two Michel positrons). The suppression of these components by timing measurements is also linear and quadratic in the timing resolution. Simulation studies have shown that the linear part is dominating at rates at least up to $2 \cdot 10^9$ muon stops per second. The requirement of reducing the combinatorial background by at least two orders of magnitude puts very tight demands on the resolution of the timing detectors. The timing resolution should be below 500 ps per track to allow for reliable charge identification by time-of-flight and ideally 100 ps or better to identify non-synchronous muon decays.

1.5.3. Other Backgrounds

Pion Decay Certain pion decays, especially $\pi \to eee\nu$ (branching fraction $3.2 \cdot 10^{-9}$ [11]) and $\pi \to \mu\gamma\nu$ (branching fraction $2.0 \cdot 10^{-4}$ [12]) with subsequent photon conversion are indistinguishable from signal events if the momenta of the final state particles fit the muon mass hypothesis. The low pion contamination in the muon beam delivered to the experiment, in addition to the small branching fractions, lead to negligible rates for this background.

Mis-reconstruction Mis-reconstruction of tracks (e.g. from hits created by different particles or noise hits) combined with real tracks from muon decays can fake $\mu \to eee$ decays. Great care is taken in the track reconstruction algorithms to keep a minimal rate of fake tracks, balanced against reconstruction efficiency.

1.5.4. Summary

The sensitivity aims of the Mu3e experiment place strict requirements on the experimental design. Electrons and positrons must be reconstructed down to a few MeV with large solid angle coverage, running at a rate of $1 \cdot 10^8$ Hz of muon decays. The material in the target and deter must be minimised, while achieving excellent momentum, vertex and timing resolution to suppress backgrounds to the necessary level. The rest of this document will discuss how this is achieved.

2. Experimental Concept

Phase I of the Mu3e experiment aims for the background free measurement or exclusion of the branching fraction for the decay $\mu \to eee$ at the level of $2 \cdot 10^{-15}$. As discussed in more detail in chapter 1, these goals require running at high muon rates, excellent momentum resolution to suppress background from internal conversion decay ($\mu \to eee\nu\nu$), and a good vertex and timing resolution to suppress combinatorial background.

The momenta of electrons and positrons from muon decays are measured using a silicon pixel tracker in a solenoidal magnetic field. At the energies of interest, multiple Coulomb scattering in detector material is the dominating factor affecting the momentum resolution. Minimising the material in the detector is thus of the utmost importance.

The detector consists of an ultra-thin silicon pixel tracker, made possible by the High-Voltage Monolithic Active Pixel (HV-MAPS) technology (see chapter 7). Just four radial layers of HV-MAPS sensors around a fixed target in a solenoidal magnetic field allow for pre-

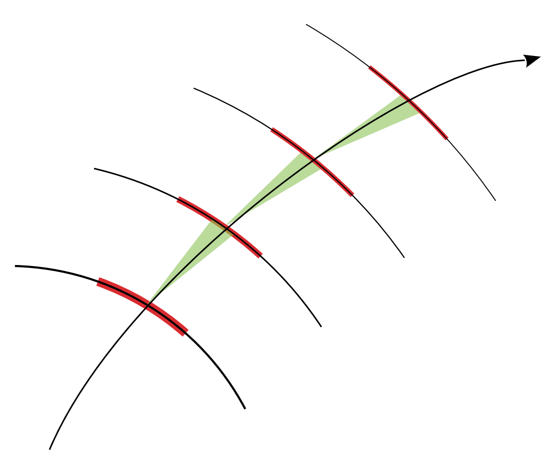


Figure 7: Tracking in the spatial resolution dominated regime

cise momentum and vertex determination. Two timing detector systems guarantee good combinatorial background suppression and high rate capabilities.

2.1. Momentum Measurement with Recurlers

Due to the low momenta of the decay particles, multiple scattering is the dominant effect on the momentum measurement. With a fine-grained pixel detector, we are in a regime where scattering effects dominate over sensor resolution effects, see Figures 7 and 8. Adding additional measurement points does not necessarily improve the precision.

The precision of a momentum measurement depends on the amount of track curvature Ω in the magnetic field B and the multiple scattering angle Θ_{MS}, see Figure 9; to first order:

$$\frac{\sigma_p}{p} \propto \frac{\Theta_{MS}}{\Omega}. \tag{4}$$

So in order to have a high momentum precision, a large lever arm is needed. This can be achieved by moving tracking stations to large radii, which would limit the acceptance for low momentum particles. Instead, we utilise the fact that, in the case of muon decays at rest, all track momenta are below 53 MeV and all tracks will curl back towards the magnet axis if the magnet bore is sufficiently large. After half a turn, effects of multiple scattering on the momentum measurement cancel to first order, see Figure 10. To exploit this feature, the experimental design is optimised specifically for the measurement of recurling tracks, leading to a narrow long tube layout.

Determining the momentum from a particle's trajectory outside the tracker allows us to place thicker timing detectors on the inside both upstream and downstream of the target without significantly affecting the resolution, see Figure 11.

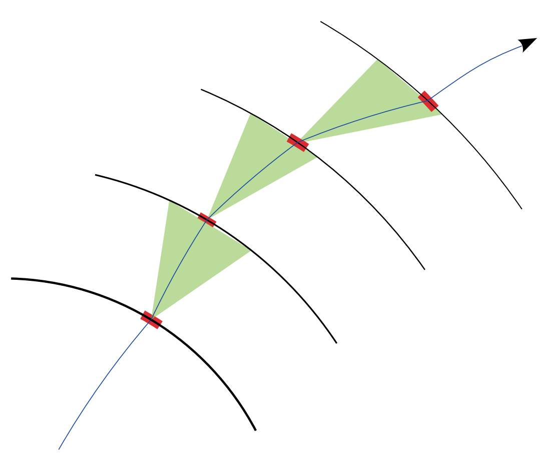


Figure 8: Tracking in the scattering dominated regime

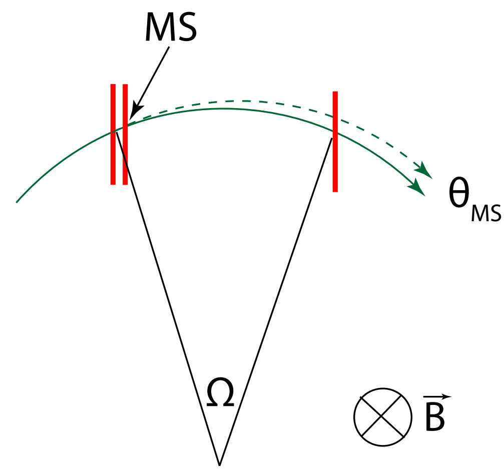


Figure 9: Multiple scattering as seen in the plane transverse to the magnetic field direction. The red lines indicate measurement planes.

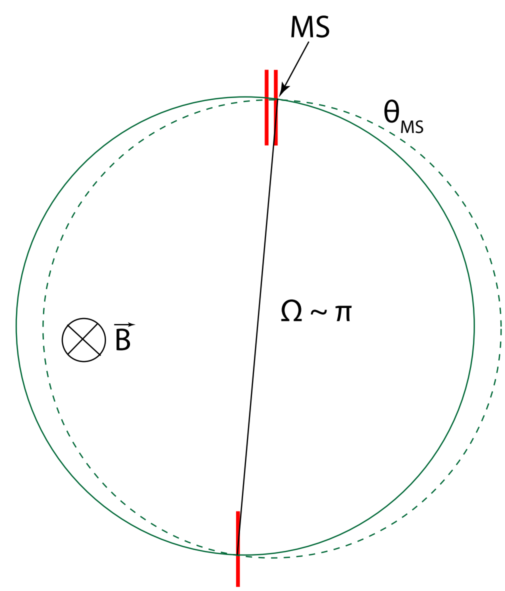


Figure 10: Multiple scattering for a semi-circular trajectory. The red lines indicate measurement planes.

2.2. Coordinate System

The Mu3e coordinate system is centred in the muon stopping target with the z axis pointing in beam direction, the y axis pointing upward and the x axis chosen to obtain a right handed coordinate system. The polar angle measured from the z axis is denoted with ϑ, and measured from the x-y plane denoted with λ. Azimuthal angles are denoted with φ.

2.3. Baseline Design

The proposed Mu3e detector is based on two double-layers of HV-MAPS around a hollow double cone target, see Figures 11 and 12. The outer two pixel sensor layers are extended upstream and downstream to provide precise momentum measurements in an extended region to increase the acceptance for recurling electrons

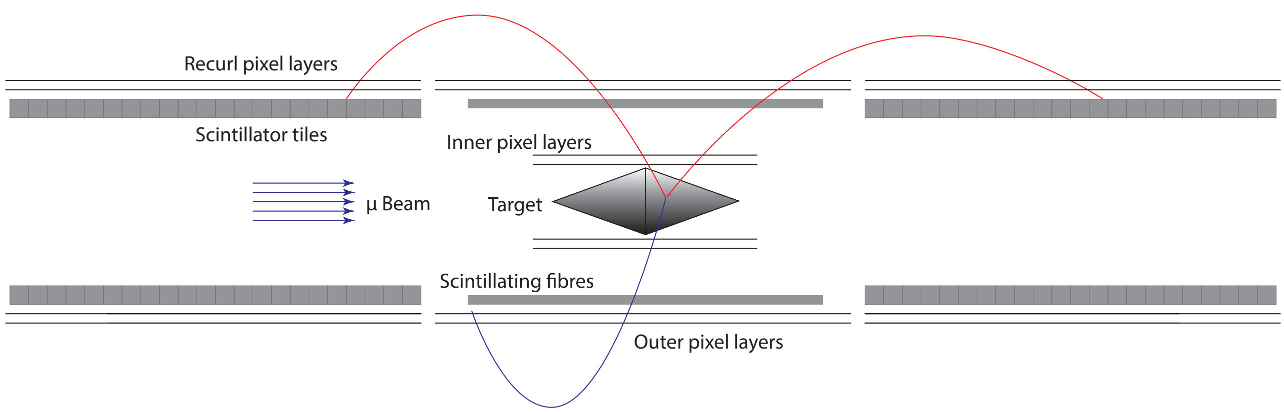


Figure 11: Schematic view of the experiment cut along the beam axis in the phase I configuration.

and positrons. The silicon detector layers (described in detail in chapter 7) are supplemented by two timing systems, a scintillating fibre tracker in the central part (see chapter 10) and scintillating tiles (chapter 11) inside the recurl layers. Precise timing of all tracks is necessary for event building and to suppress the combinatorial background.

2.4. Detector Readout

The Mu3e experiment will run a continuous, triggerless readout, and employs custom ASICs for the pixel and timing detectors which stream out zero-suppressed digital hit data. These hits are collected by FPGAs located on *front-end boards* and then optically forwarded to *switching boards*, which in turn distribute them to a computer farm. This network makes it possible for every node in the farm to have the complete detector information for a given time slice. Decays are reconstructed using graphics processing units, and interesting events are selected for storage. A system overview is shown in Figure 13 and a detailed description can be found in chapter 17.

2.5. Building up the Experiment

One of the advantages of the design concept presented is its modularity. Even with a partial detector, physics runs can be taken. In an early commissioning phase at smaller muon stopping rates, the detector will run with all of the timing detectors but only the central barrel of silicon detectors. The silicon detectors of the recurl stations are essentially copies of the central outer silicon detector; after a successful commissioning of the latter, they can be produced and added to the experiment as they become available. The configuration with two recurl stations (Figures 11 and 12) defines a medium-size setup, well suited for phase I running at the highest possible rate at the πE5 muon beam line at PSI of $\approx 1 \cdot 10^8$ Hz. The sensitivity reach in this phase of the experiment of $\mathcal{O}(10^{-15})$ will be limited by the available muon rate.

2.6. The Phase II Experiment

A new high intensity muon beam line, delivering $> 2 \cdot 10^9$ Hz muons and currently under study at PSI, is crucial for Mu3e phase II. To fully exploit the new beam facility the detector acceptance of phase I will be further enhanced by longer detector stations, see Figure 14. These longer stations will allow the precise measurement of the momentum of all particles in the acceptance of the inner tracking detector. At the same time the longer tile detector stations with their excellent time resolution and small occupancy will help to fight the increased combinatorial backgrounds at very high decay rates. The larger initial muon rate allows for a more restrictive collimation of the beam and thus a smaller (and potentially longer) target region leading to a much improved vertex resolution. The HV-MAPS technology can reach a time resolution of $\mathcal{O}(1\,\text{ns})$ if an adequate time-walk correction is implemented – this would allow to further reduce combinatorial background without adding material and could eventually replace the scintillating fibre detector. Advanced wafer post-processing technologies and chip-to-chip bonding could obviate the need for parts of the flexprints, further reducing the multiple scattering. The combined performance of the enhanced detector setup together with the high stopping rate will allow to search for the $\mu \to eee$ decay with a sensitivity of $\mathrm{B}(\mu \to$

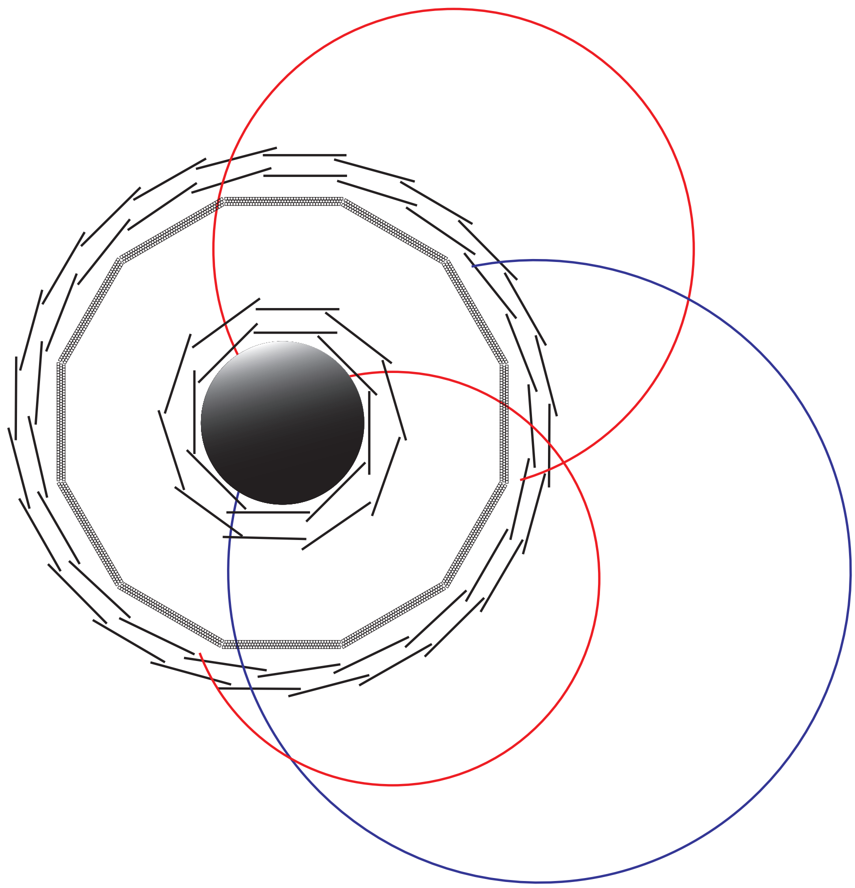


Figure 12: Schematic view of the experiment cut transverse to the beam axis. Note that the fibres are not drawn to scale.

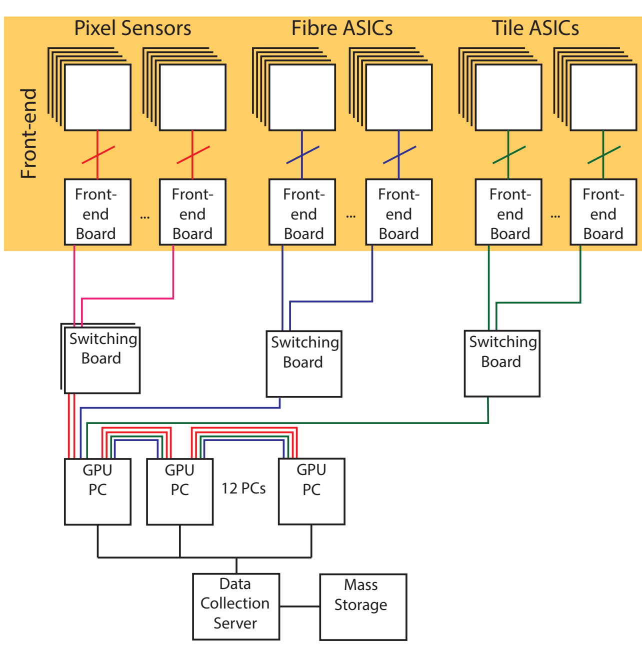


Figure 13: Schematic overview of the Mu3e readout system.

eee)$\leq 10^{-16}$. Whilst we always keep this ultimate goal in mind, the rest of this document is concerned with the phase I detector for existing beamlines.

3. Muon Beam

3.1. Beam Requirements

An experiment such as Mu3e, with a phase I sensitivity goal of $2 \cdot 10^{-15}$ while challenged by combinatorial background, not only requires running at the intensity frontier, but also substantially benefits from a continuous beam structure rather than a pulsed one, allowing a lower instantaneous muon rate. Both of these conditions are satisfied by the high intensity proton accelerator complex (HIPA) at PSI running at 1.4 MW of beam power.

Mu3e requires a muon beam with the highest possible rate of "surface muons", produced from stopped pion decay at the surface of the primary production target [13]. The surface muon yield and hence beam intensity peaks at around 28 MeV/c, close to the kinematic edge of the two-body momentum spectrum of pion decay at rest as can be seen from the measured momentum spectrum in Figure 15.

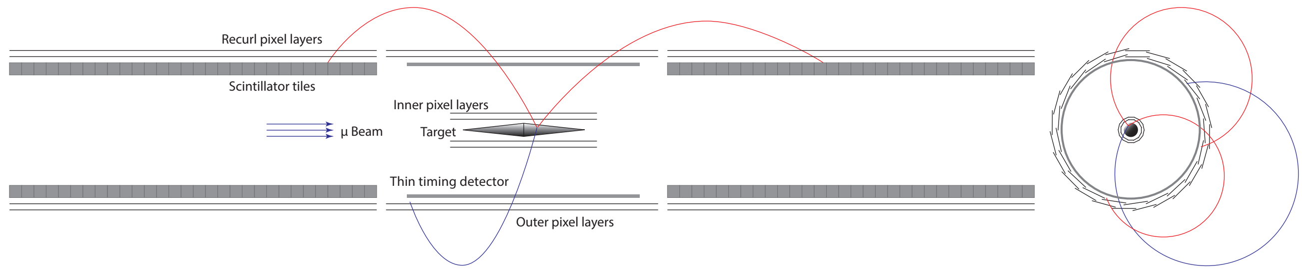


Figure 14: Possible final detector with longer recurl stations, smaller target and more segmented inner layers for high intensity physics runs (phase II).

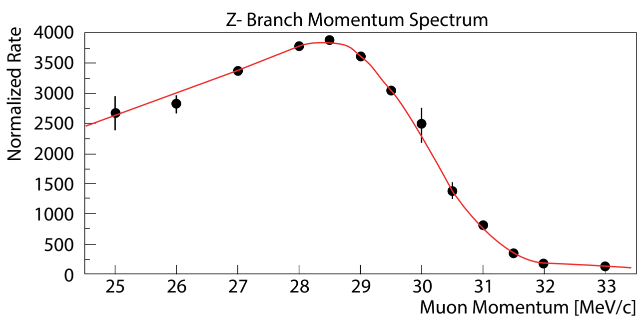


Figure 15: Measured muon momentum spectrum in πE5, with full momentum acceptance. Each point is obtained by optimising the whole beam line for the corresponding central momentum and measuring the full beam-spot intensity. The red line is a fit to the data, based on a theoretical $p^{3.5}$ behaviour, folded with a Gaussian resolution function corresponding to the momentum-byte plus a constant cloud-muon background.

The intensity goal and low energy not only necessitates a beam line capable of guiding these muons to a small, thin stopping target with minimal losses but at the same time minimising beam-related backgrounds. The former requires a small beam emittance and a moderate momentum-byte (full width at half maximum of the momentum acceptance $\Delta p/p$), with an achromatic final focus to balance between beam intensity and stopping density in the target. The minimising of beam-related backgrounds, in the form of Michel e^+ from μ^+-decay, e^+ produced from π^0-decays in the production target, or from decay-in-flight particles produced along the beam line, puts strong restrictions on the amount of material, such as windows and momentum moderators, that can be placed along the beam path, requiring an extension of the vacuum system to just in front of the target.

3.2. The Compact Muon Beam Line (CMBL)

For Mu3e phase I, muon intensities close to 10^8 muons/s will be required, which leaves only one choice of facility in the world, PSI's πE5 channel. This channel will be shared with the upgrade version of the MEG experiment – MEG II [14], whose large detector and infrastructure are permanently located in the rear-part of the πE5 area.

The new CMBL for Mu3e, as presented in the following, not only allows the 3.2 m long Mu3e solenoid to be placed in the front part of the πE5 area – see Figure 16 – but also allows both experiments MEG II and Mu3e to share the front beam transport elements required by both. This solution allows the efficient switching between experiments by only replacing the superconducting beam transport solenoid of MEG II by a dipole magnet (ASL) for Mu3e.

The initial optical design of the CMBL was modelled using the beam optics matrix code programs GRAPHIC TRANSPORT FRAMEWORK [15] and GRAPHIC TURTLE FRAMEWORK [16], while the detailed modelling was undertaken using the newer GEANT4 based simulation software G4BEAMLINE (G4BL) [17]. The 1st-order optical design showing the vertical and horizontal beam envelopes from Target E to the downstream end of the Mu3e detector are shown in Figure 17.

The design includes the elements of the backward (165°) extraction channel πE5 from Target E up to the ASC dipole magnet, the background cleaning-stage including triplet I, the Wien-filter (SEP41), triplet II and the collimator system, used to eliminate the beam-related background. The final injection stage is based on a very compact “split triplet” layout which starts after the 90° dipole ASL41. The “split triplet” consists of the quadrupole doublet QSO41/42 and quadrupole singlet QSM41. In combination with the vertical edge-focusing of the ASK41 65° dipole magnet they serve the same purpose as a total of six quadrupoles that would be needed in a more standard beamline configuration. This allows sufficient space to place the 3.2 m long Mu3e detector in the front area without compromising the optics and physics goals of the experiment.

Based on the GRAPHIC TRANSPORT model, two G4BL models were constructed, one including the full πE5 channel and Target E, simulating the whole pion production process by protons in the primary target, followed by surface muon production and transport to the intermediate collimator. The second shorter version starts from Triplet II, just upstream of the interme-

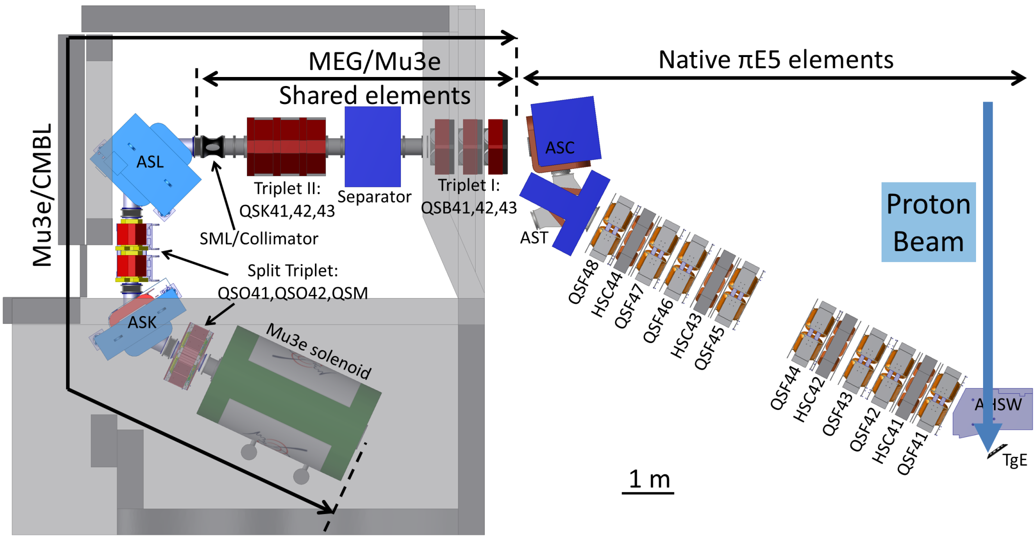


Figure 16: CAD model of the entire πE5 channel & CMBL used as a basis for the G4BL models.

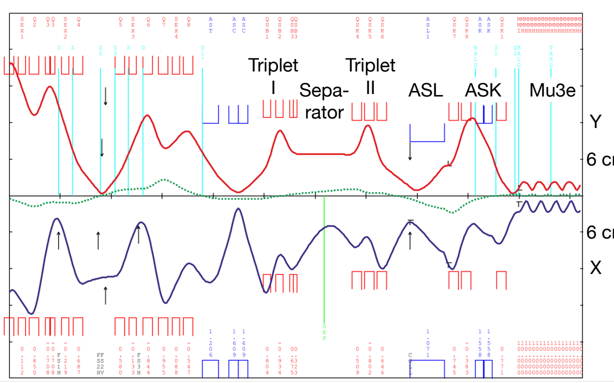


Figure 17: Optical Model of the CMBL from the GRAPHIC TRANSPORT FRAMEWORK program, showing 1st-order vertical and horizontal beam envelopes along the entire beam line from Target E to the end of the Mu3e solenoid with some of the beam elements labelled (note the horizontal scale unit is 2 m, whereas the vertical is 6 cm). The dotted line shows the dispersion trajectory for a 1% higher central momentum.

diate focus at the collimator system, where measured phase space parameters determine the initial beam used for the simulation - see Figure 18. The shorter version predictions were used as a direct comparison to the CMBL commissioning measurements described in the next section.

3.3. CMBL Commissioning Steps

Initial commissioning of the CMBL beam layout was undertaken in two 4-week beam periods in November and December 2014 and May 2015, using mostly existing elements. Figure 19 shows the good agreement between predicted and measured beam sizes at the injection point to the Mu3e solenoid, based on a 1st-order transverse phase space reconstruction. The validated G4BL model was then used, and identified the ASL and ASK dipole apertures as the main limitations for the transmission to the final focus. Consequently, increased pole-gaps and modified vacuum chambers for both dipole magnets allowed for an expected enhanced transmission of 18%, which was proven in the following 2016 measurements [18].

In 2017, the commissioning emphasis was placed on confirmation of increased muon yield using a 60 mm long production Target E instead of the usual 40 mm version. The expectation of only an $\sim$30% increase in muon yield (surface phenomenon) with a full 50% increase in beam positron contamination (bulk phenomenon) for the 165° backward extraction was confirmed. Furthermore, the expected impact on the experiment from an increased beam positron background was also studied and a differential measurement technique developed to distinguish Michel positrons from beam positrons at the final focus [19]. These measurements showed that for the 60 mm Target E a beam-e^+/μ^+-ratio $= 10.1$ was measured, with no Wien-filter in operation, whereas for a 40 mm target the ratio was $\sim$7. However, with the Wien-filter on, an unacceptably high number of beam positrons, seen as a vertically displaced spot, were measured. On investigation it was found that the off-centre, vertically displaced (by the Wien-filter) beam positrons entering triplet II are partially swept-back into the acceptance of the downstream collimator, as demonstrated in Figure 20. The situation was quickly and temporarily solved by placing a lead e^+-stopper between QSK41/42 reducing the contamination by a factor of 15, with a 10% loss of muons. The final solution is the modification of the Wien-filter, which was upgraded in 2019 to have a symmetric electric field with double the present voltage of 200 kV. While not yet experimentally confirmed, this

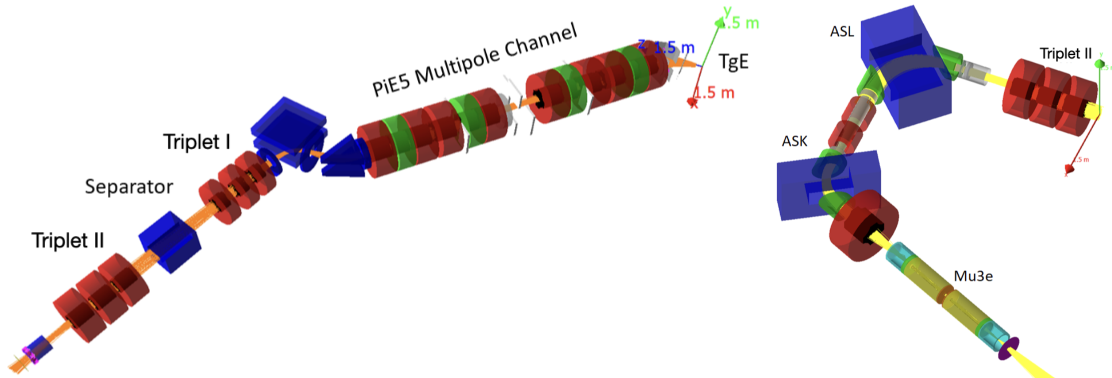


Figure 18: Shows the graphical outputs of the G4BL simulations with some of the beam elements labelled. (Left) – simulation of the full beam line from Target E up to the intermediate collimator system. (Right) – shows the shorter version of the simulation from Triplet II past the intermediate collimator system to the end of the Mu3e Detector solenoid.

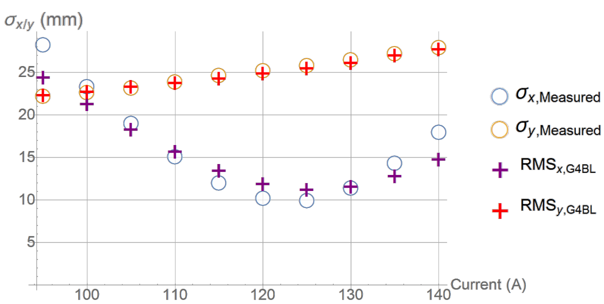


Figure 19: Simulated and measured 2-D beam sizes at the Mu3e solenoid injection point, showing good agreement for a wide range of currents applied to the last quadrupole QSM41.

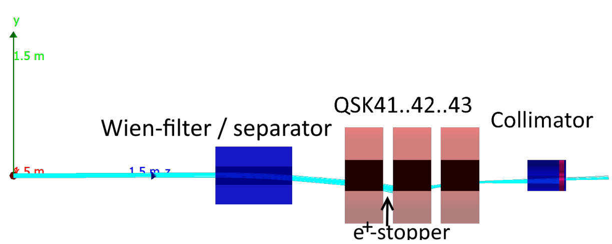


Figure 20: Demonstration of beam positron contamination vertically separated post Wien-filter being swept-back into the acceptance at the collimator by QSK42/43.

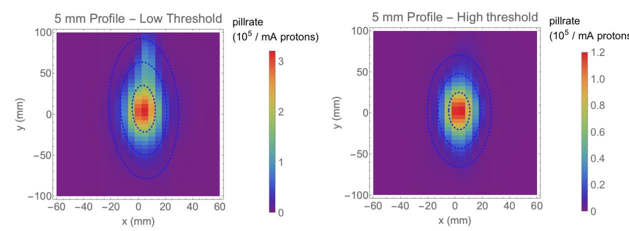


Figure 21: Measured beam spot at the injection point to the Mu3e solenoid triggering on either a low (left: muons + Michels + beam positrons) or high (right: muons only) threshold. A 2D Gaussian fit to the muon data yields $\sigma_x = 8\,\mathrm{mm}$ and $\sigma_y = 23\,\mathrm{mm}$ with a total rate of 1.1×10^8 μ^+/s at a proton current of $2.4\,\mathrm{mA}$ for a $40\,\mathrm{mm}$ long Target E. The vertical beam positron tail in the low threshold profile (top-part) is without the e^+-stopper in triplet II and will be totally removed with the upgraded Wien-filter.

is expected to reduce any beam positron contamination by 3-orders of magnitude.

Finally, using the measured contamination rate, the impact of this on the experiment's sensitivity to combinatorial Michel and beam positron events mimicking a 3-particle signature via Bhabha scattering was investigated [19] and found that only muon decay-in-flight events have a chance of coming close to the reconstructed muon mass region, though occurring at a rate twelve orders of magnitude lower than the most dominant background (Bhabha scattering with overlapping Michel decays).

During the shutdown 2017/18 all magnet power supplies for πE5 were replaced with digitally controlled ones. The better stabilisation of magnet currents contributed to a further increased transmission during the 2018 commissioning run. Optimisation at the injection point of the Mu3e solenoid yielded a final rate of 1.1×10^8 μ^+/s, normalised to the expected future proton beam current of $2.4\,\mathrm{mA}$ for a $40\,\mathrm{mm}$ long Target E, with profile widths of $\sigma_x = 8\,\mathrm{mm}$ and $\sigma_y = 23\,\mathrm{mm}$. The measured high (muons only) and low threshold (muons + Michels) profiles are shown in Figure 21, these $5\,\mathrm{mm}$ raster scan profiles were measured with a 2D automated pill scintillator scanner system with each profile consisting of $\sim$1025 single measurements. The beam intensity is extracted from a 2D Gaussian fit to the profiles.

The 2018 measurements therefore successfully conclude the beam commissioning up to the injection point of the Mu3e solenoid. The final commissioning to the centre of the Mu3e detector will be undertaken when the magnet is placed in the area.

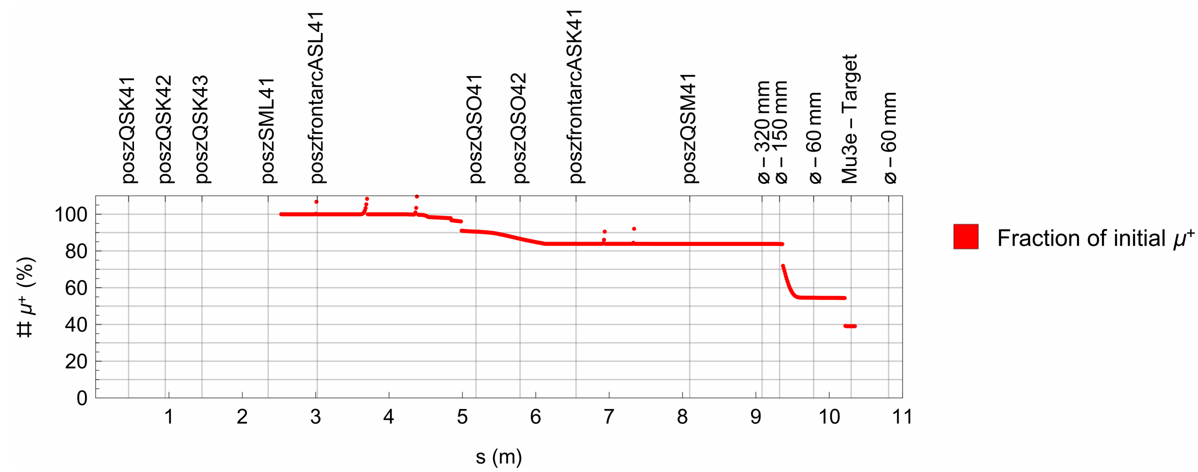


Figure 22: Beam losses along the Mu3e Compact Muon Beam Line (CMBL) starting from the intermediate collimator system to the centre of the Mu3e magnet. In front of the Mu3e target a narrowing of the beam-pipe down to $40\,\mathrm{mm}$ diameter takes place.

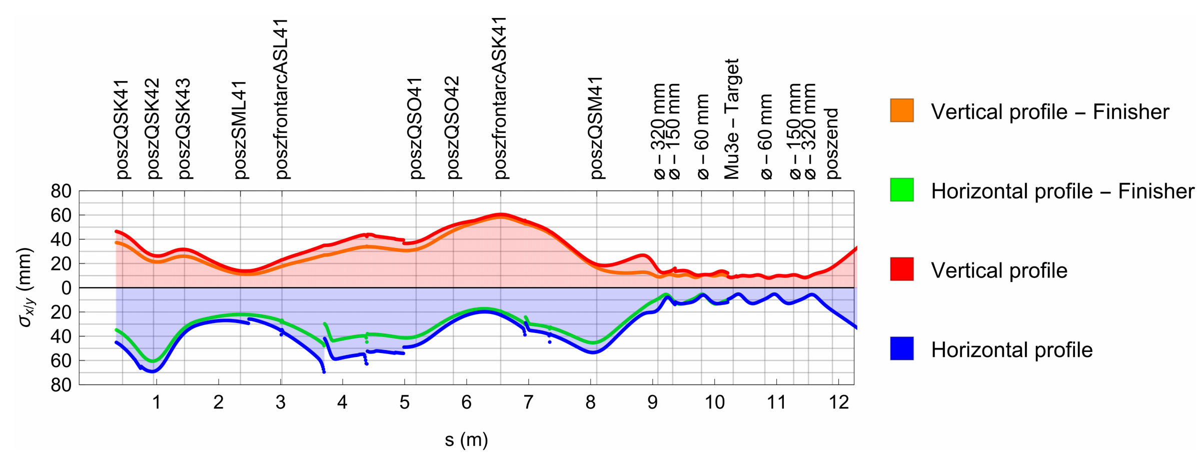


Figure 23: Horizontal and vertical beam envelopes for 'all' particles started in the simulation or only for those that reach the centre of the solenoid ('finishers').

3.4. Expected Muon Rate and Distribution on the Mu3e Stopping Target

As described in the previous section the final optimisation of the CMBL resulted in a surface muon rate at the injection point of the Mu3e magnet of $1.1 \times 10^8\ \mu^+/\mathrm{s}$, normalised to the expected future proton beam current of $2.4\,\mathrm{mA}$ for a $40\,\mathrm{mm}$ long Target E, with profile widths of $\sigma_x = 8\,\mathrm{mm}$ and $\sigma_y = 23\,\mathrm{mm}$.

The coupling to the central detector region inside the solenoid magnet is planned to be with a custom bellows system (see Figure 30) reducing step-wise the aperture to an inner diameter of $60\,\mathrm{mm}$ for the inner vacuum-pipe. This will contain a $600\,\mu\mathrm{m}$ thick Mylar (biaxially-oriented polyethylene terephthalate) moderator located at an intermediate focus point some few hundred millimetres in front of the target and will end with a $35\,\mu\mathrm{m}$ Mylar vacuum window, placed just in front of the Mu3e target, where the aperture narrows down to 40 mm diameter due to the support structure of the inner pixel layers. A double-cone Mylar target of radius $19\,\mathrm{mm}$, length $100\,\mathrm{mm}$ and total thickness of $150\,\mu\mathrm{m}$ (see chapter 6) is located close to the vacuum window at the centre of the solenoid. The warm bore of the solenoid is filled with helium gas at atmospheric pressure to reduce multiple scattering. Furthermore, a $20\,\mathrm{mm}$ thick lead collimator system will be introduced shortly after the moderator to protect the inner pixel layers from hits by the muon beam as well as from particles outside of the target acceptance.

Estimates for the final muon stopping rate on the target are based on the re-measured 1-σ beam emittances at the intermediate collimator system in 2018, corresponding to $\epsilon_x = 950\,\pi \cdot \mathrm{mm} \cdot \mathrm{mrad}$, $\epsilon_y = 490\,\pi \cdot \mathrm{mm} \cdot \mathrm{mrad}$ and the G4BL simulation. The beam losses along the beam line can be seen in Figure 22 and the corresponding beam envelope sizes in Figure 23.

Even though the muon beam intensity at injection into the solenoid achieves the commissioning goal, it is the inner silicon detector diameters and the associated beam-pipe size that determine the stopping tar-

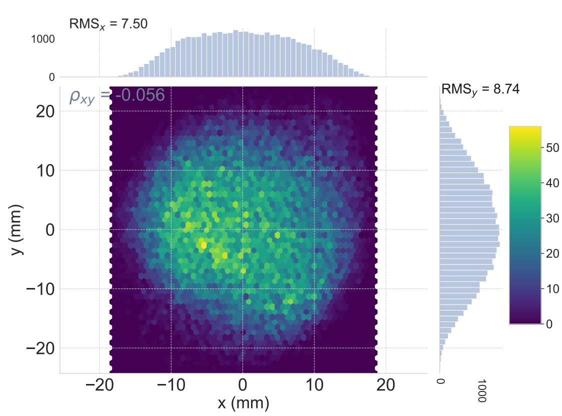


Figure 24: Estimated beam profile at the Mu3e target position.

Magnet parameter	Value
field for experiment	$1.0\,\mathrm{T}$
field range	$0.5 - 2.0\,\mathrm{T}$
warm bore diameter	$1.0\,\mathrm{m}$
warm bore length	$2.7\,\mathrm{m}$
field inhomogeneity $\Delta B/B$ around center	$\leq 10^{-3}$
field stability $\Delta B/B$ (100 days)	$\leq 10^{-4}$
field description $\Delta B/B$	$\leq 2.0 \cdot 10^{-4}$
outer dimensions: length	$\leq 3.2\,\mathrm{m}$
width	$\leq 2.0\,\mathrm{m}$
height	$\leq 3.5\,\mathrm{m}$

Table 1
Requirements for the Mu3e magnet.

get diameter, which has been maximised to a radius of 19 mm. These conditions are a compromise between stopping rate, occupancy and vertex resolution.

The main losses are associated with the transition to the initial diameter of the beam-pipe, and the final narrowing to a 40 mm diameter at its end. The final beam-spot at the target is shown in Figure 24. The beam intensity on the target is expected to be $\sim 5 - 6 \times 10^7\ \mu^+/\mathrm{s}$ at 2.4 mA proton current for the current 40 mm long production Target E. The final muon rate can further be enhanced by the use of the 60 mm production target, or the recently tested 40-mm long slanted target. Both of these targets lead to a further ∼30-40% enhancement, so yielding muon rates on the Mu3e target of about $\sim 7 - 8 \times 10^7\ \mu^+/\mathrm{s}$ at 2.4 mA proton current. Further enhancements are still under study.

4. Magnet

The magnet for the Mu3e experiment has to provide a homogeneous solenoidal magnetic field of $B = 1\,\mathrm{T}$ for the precise momentum determination of the muon decay products. Field inhomogeneities along the beam line are required to stay below 10^{-3} within $\pm 60\,\mathrm{cm}$ around the center. The magnet also serves as beam optical element for guiding the muon beam to the target. To further improve the field homogeneity, and for matching the magnetic field of the last beam elements of the compact muon beam line, compensating coils are included on either side of the magnet.

The basic parameters of the superconducting solenoid magnet are given in Table 1. The outer dimensions also include an iron shield, reducing stray fields to less than 5 mT at a distance of 1 m. This, however, lead to an overall weight of the magnet of 31 tons, 27 of which are due to the iron shielding.

The long term stability of the magnetic field should be $\Delta B/B \leq 10^{-4}$ over each 100 days data-taking period. This is achieved with state of the art power supplies and by permanently monitoring the absolute field with NMR and Hall probes inside the apparatus. The NMR system and hall probes will also be used to map the field. The goal is to measure and describe the field distribution with a precision better than $2.0 \cdot 10^{-4}$.

The tight requirements on the dimensions of the magnet come from the space constraints of the πE5 area as described in the next chapter. In this respect a good compromise had to be found as in particular the total length of the magnet is a critical parameter impacting the specified homogeneity of the field in the central region.

In addition, the magnet also acts as containment for the helium gas used for cooling as described in chapter 12. For this reason, the warm bore is designed with helium-tightness in mind and is sealed off on both ends by removable flanges.

A superconducting magnet design with a closed cooling system was determined to be the most stable and economic solution. The magnet made from niobium-titanium superconductor will operate at nominally 4 K and be cooled by four Gifford McMahon two-stage cryocoolers, each delivering 1.5 W cooling power at their second stage. The cool-down time for the system is about 10 days with liquid nitrogen pre-cooling and the ramp up time to 1 T will be less than 2 hours.

The company Cryogenic Ltd. [15]was tasked to design and produce the Mu3e solenoid magnet. Cryogenic Ltd. has prepared a technical design report for the complete magnet system in 2018. The picture shown in Figure 25 depicts the delivery of the magnet to PSI's experimental hall in July 2020 after initial testing of the magnet at the company showed excellent performance.

Figure 25: Picture of the delivery of the 31-ton Mu3e magnet to PSI's experimental hall where the Mu3e experiment will take place.

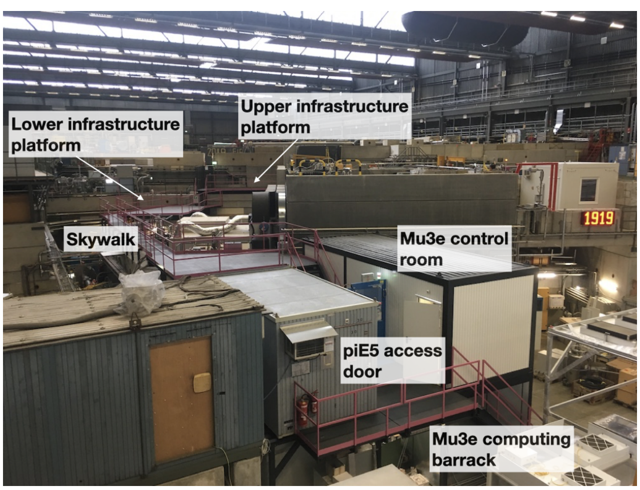


Figure 26: The new rear access to the πE5 Area with the Skywalk and the two new infrastructure platforms for the Mu3e experiment. In the front, the Mu3e control room and computing barrack are located.

5. Area Layout, Infrastructure & Beam Line Connection

Due to the spatial restrictions in the πE5 front area and the substantial infrastructure needs of the experiment, an optimised area layout is necessary. Upgrades were needed to both the electrical installation and cooling-water and, due to safety requirements, an additional access route to the front area had to be added. Figure 26 shows the overview of the new rear access to the πE5 Area via the 'skywalk' with its two new infrastructure platforms and the Mu3e control room and computing farm barrack. The experimental area in the front part of πE5 is located below the two infrastructure platforms and will have a stairway added as a safety requirement, once the large magnet is in place, leading from the lower platform into the experimental area.

The upper infrastructure platform, above the beam entrance wall, is constructed to be removable in order to grant service access the πE5 channel during accelerator shutdown periods, if required. This platform is closest to the magnet and detector and will house the cooling elements such as the compressors for the cryogenic cold-heads as well as the helium and water cooling circuits for the Mu3e detector. The lower, larger platform will not be removable and will carry the magnet power supplies, quench detection system and electronics as well as the power-control circuitry associated with both magnet and detectors.

Also seen in Figure 26 are the two new πE5 barracks located on top of each other. The upper barrack will serve as the Mu3e experiment's control room, while the lower barrack will house the filter farm responsible for the readout of the detector (see chapter 17).

Due to the limited space in the front part of the πE5 area, as can be seen in Figure 27, as well as the fact that the Mu3e magnet is located underneath the roof formed by the πE3 area above, a rail system is required to move the Mu3e magnet from a position where it can be lowered down into the experimental area by crane – shown in Figure 28 – to its final position underneath the roof shown in Figure 27. The crane operation will be a challenging one and extra degrees of rotational freedom included in the rail system are needed to allow for such a movement of the 30-ton magnet to its final

[15] Cryogenic Ltd., Acton Park Industrial Estate The Vale London W3 7QE, UK

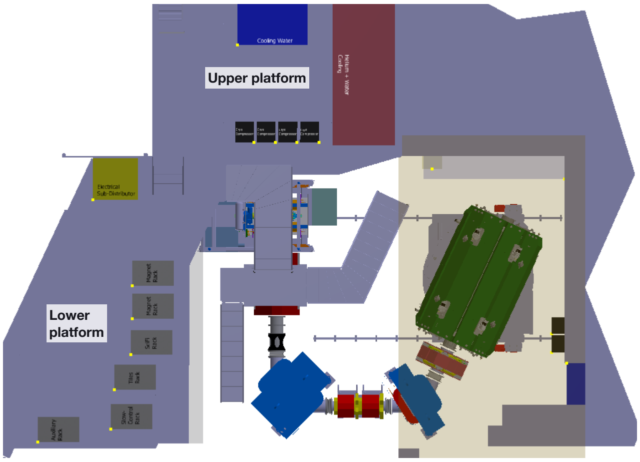


Figure 27: Top view of the πE5 experimental area showing the completed installation. Also visible are the two new infrastructure platforms located on the shielding blocks above the area and the stairs leading down to the experiment. The transparent beige-area marks the roof underneath which the Mu3e magnet has to be installed.

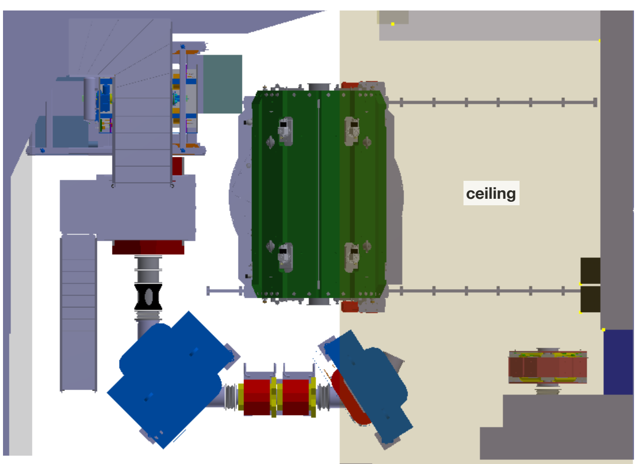


Figure 28: Position of the magnet when lowered into the experimental area onto its rail system. The rail system allows to move the 30-ton magnet underneath the ceiling and turn it in line with the rest of the beam elements.

position under the roof. In addition, a small crane is needed to move the last quadrupole QSM41 away from its position along the beamline in order to allow the free movement of the magnet.

Figure 29 shows the magnet in its maintenance position. This position allows the Mu3e solenoid to be rotated in such a way that the full detector-cage can be extracted onto its transport support structure for repairs, maintenance or transportation. A detailed description on how the detector can be extracted onto the support structure can be found in chapter 13.

Finally, the detailed coupling mechanism of the beam line to the solenoid magnet is described. The components are shown in Figure 30. The standard ISO-320-K beam line vacuum tube, with its upstream bellows connection protrudes into the magnet allowing a maximum acceptance of the converging beam envelopes before entering a custom 150-mm diameter intermediate bellows connection to the final 60-mm diameter beam tube of the Mu3e detector. The mounting sequence is as follows: In a first step the detector system is mounted with its cage inside the magnet bore and fixed in position. Subsequently, the internal beam line elements are mounted onto the inside end of the He-tight flange of the magnet bore, which is then bolted onto the cryostat. In order to achieve a vacuum tight connection between the custom bellows assembly and the Mu3e detector cage beam-flange, the final screws are tightened from the inside of the ISO-320-K vacuum tube, so pressing on the O-ring seal. As a last step, internal tensioning supports for the bellows are mounted and securely fixed in place to prevent the bellows from collapsing when evacuated.

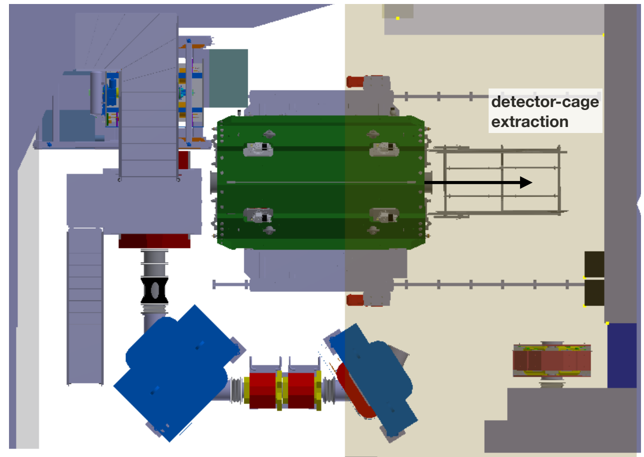


Figure 29: Maintenance position of the magnet on the rail system used to extract the detector-cage onto its transport unit.

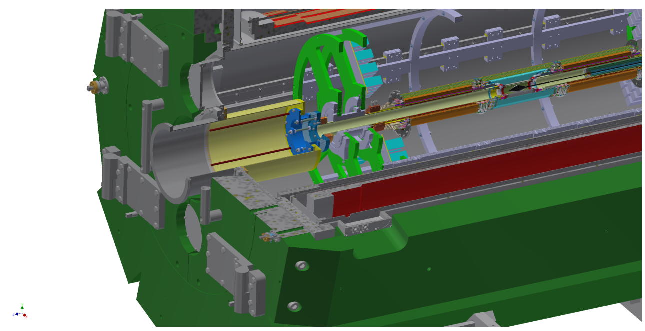


Figure 30: Connection of the Mu3e detector to the muon beam line is achieved through a custom bellows assembly inside the solenoid.

6. Stopping Target

The main challenge for the design of the stopping target is to optimise the stopping power, while also minimising the total amount of material in order to reduce both backgrounds and the impact on the track measurement. Therefore the stopping target should contain just enough material in the beam direction to stop most of the muons, which is facilitated by a moderator in the final part of the beam line, but should be as thin as possible to minimise the material in the flight direction of decay electrons entering the detector acceptance. Usage of a low-Z material is advantageous as photon conversion and large-angle Coulomb scattering are suppressed. In addition, the decay vertices should be spread out as wide as possible in order to reduce accidental coincidences of track vertices and to produce a more or less even occupancy in the innermost detector layer.

6.1. Baseline Design

These requirements can be met by a hollow double cone target à la SINDRUM [20, 21]. In our baseline design (see Figure 31), the target is made from 70 µm of Mylar in the front part and 80 µm Mylar in the back part, with a total length of 100 mm and a radius of 19 mm. This leads to an incline of 20.8° of the target surface with regards to the beam direction. The projected thickness is thus 197 µm for the front and 225 µm for the back part, giving a total of 422 µm of Mylar corresponding to 0.15% of a radiation length. The mass of the Mylar in the target is 0.671 g. The total area of the target is 6386 mm^2.

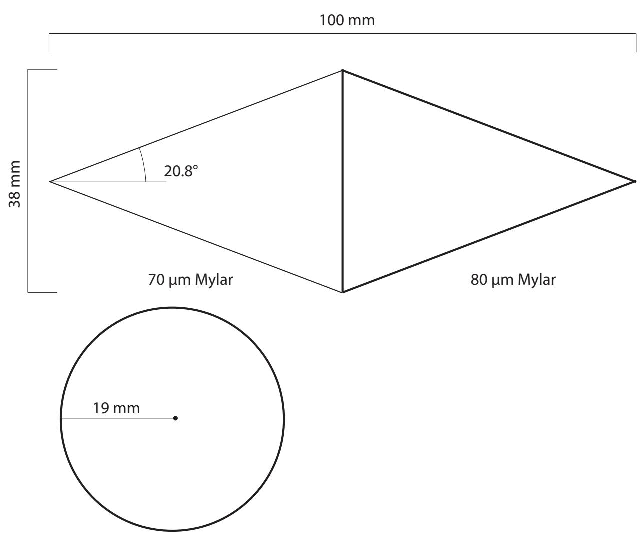


Figure 31: Dimensions of the baseline design target. Note that the material thickness is not to scale.

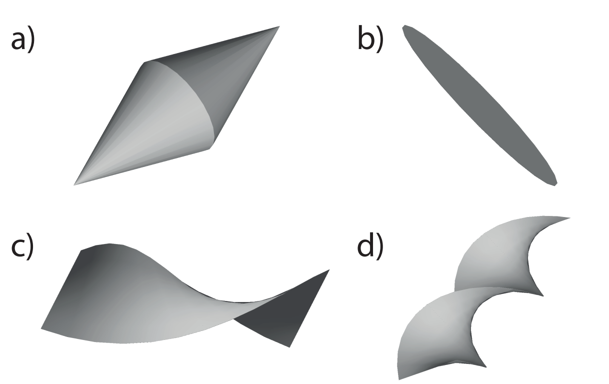


Figure 32: Target shapes studied. a) Is the default hollow double cone, b) a simple plane, c) a single-turn garland and d) a double-turn garland. For the chiral shapes c) and d), both senses of rotation were tried.

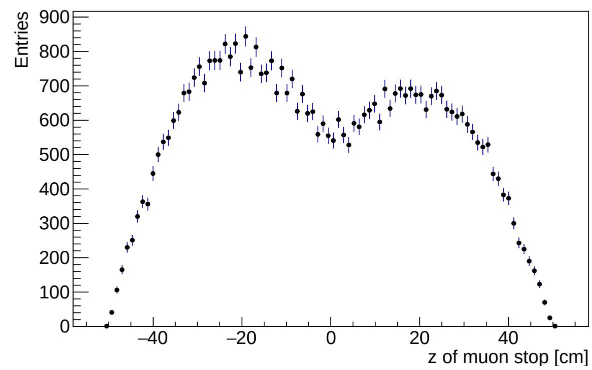


Figure 33: Simulated stopping distribution along the beam (z) axis for the baseline target.

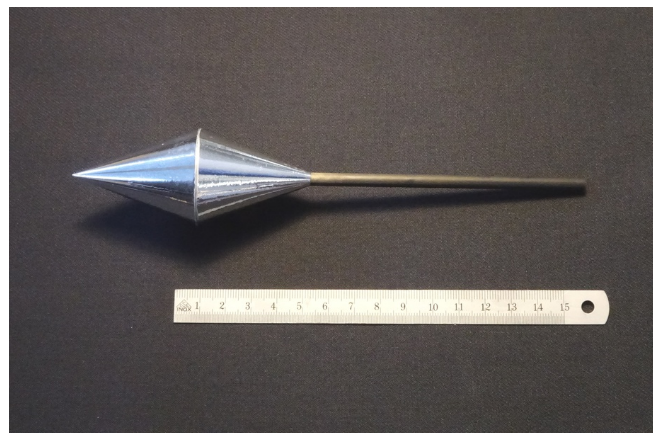


Figure 34: Hollow double-cone muon stopping target made of aluminized Mylar foil.

We have studied the stopping power and material budget for a variety of target shapes (see Figure 32) and found that for the given beam parameters and geometrical constraints, the double cone offers the highest stopping fraction with the least material. The simulation was performed with Mylar as the target material, a previous study using aluminium however gave very similar results. The stopping distribution along the beam axis for the baseline target is shown in Figure 33.

6.2. Production

At PSI, a manufacturing procedure was developed and a complete target was produced, see Figure 34. Each single hollow cone of the double cone structure

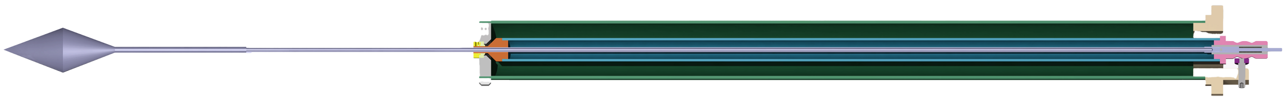


Figure 35: Cross section of target support and alignment mechanism. Muons hit the target from the left. The stopping target is mounted on a thin carbon tube which is steered and fixed in the support structure. The rear end of the support structure consists of an alignment mechanism to adjust the position of the target.

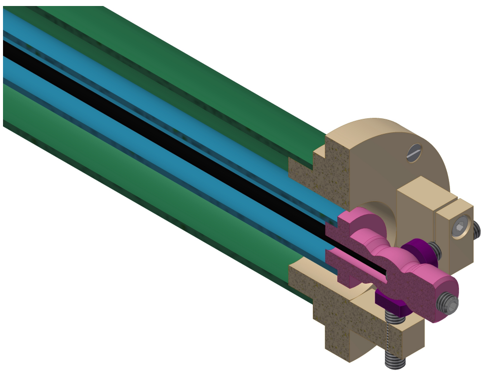


Figure 36: Cross section of alignment mechanism. The setup is spring-loaded towards the two screws and allows an adjustment of the target position in 3 coordinates. The direction of the spring is in the bisecting line with respect to the two screws and is in this view therefore hidden by the holder.

is manufactured separately and is a sandwich structure consisting of 2 or 3 rolled up thin Mylar foils glued together with epoxy glue. The thickness of the individual Mylar foils and the combination of several foils are chosen to match best with the desired final thickness. Finally, the two individual cones are glued together to build up the hollow double cone structure.

The inner and the outer foil in each sandwiched stack is aluminium coated and the orientation of the aluminium layers is such that the inner and outer surface of the cones features an aluminium layer. The conductive surfaces, in combination with the mounting on a conductive carbon tube avoid a possible charging up of the target due to the high stopping rate of positive muons.

6.3. Support

The double cone structure will be glued on a carbon tube which will be fixed in a dedicated support structure with an alignment mechanism. Figure 35 shows a cross section of the complete target system consisting of stopping target, carbon tube and support, while Figure 36 shows an enlarged view of the rear end of the support structure consisting of the alignment mechanism. The target support structure will be placed on the downstream side of the experiment in order not to disturb the incident muon beam.

The carbon tube (silver / black rod in Figures 35 and 36) has an inner diameter of 5 mm and will be glued on the tip of the downstream cone of the stopping target. Along the first 10 cm downstream of the target the original wall thickness of 0.5 mm of the carbon tube is reduced to $\sim$0.125 mm by means of centerless-grinding in order to reduce the material budget in the central region of the detector.

To avoid possible vibrations of the target due to a long lever arm the carbon tube is not only rigidly fixed at the rear end of the support structure (pink part in Figure 36), but also guided in a joint at the front end of the structure close to the target itself.

The alignment mechanism (see Figure 36) allows an adjustment of the target position in all 3 coordinates. To ensure sufficient clearance between the target and the innermost layer of the silicon detectors, the range of movement for the target is limited to $\pm$ 2 mm in x- and y-directions, and $\pm$ 4 mm in z-direction. This is achieved with a limited range for the adjustment screw at the rear end of the support structure, in conjunction with the transformation ratio due to the different lengths of carbon tube and support structure.

The central tube (turquoise part in Figure 36) of the support structure hosting the carbon tube and connected to the holder at the end (pink part in Figure 36) is spring-loaded towards the adjustment screws to allow for a hysteresis-free adjustment of the target.

7. Pixel Tracker

The Mu3e pixel tracker provides precision hit information for the track reconstruction of the electrons produced in muon decays. Achieving the best possible vertex and momentum resolution measurements for these electrons is of key importance to the success of the experiment. Due to the dominance of multiple scattering, a rigorous minimisation of the material in the active region of the tracking detector is critical. For this reason, the tracker relies on High-Voltage Monolithic Active Pixel Sensors (HV-MAPS) [22], thinned to 50 µm and mounted on a low mass service flex circuit. The detector is operated inside a dry helium atmosphere and cooled by helium gas flow to further reduce multiple scattering.

7.1. Overview of the Pixel Tracker

The Mu3e pixel tracker consists of three parts, the central pixel tracker and two recurl stations, see Figure 37. Pixel layers in the central tracker provide the main hits used for the reconstruction of tracks and of the decay vertex associated with multiple tracks. The hits detected in the recurl stations allow us to reconstruct tracks with higher purity and improved momentum resolution.

Throughout the pixel tracker all MuPix sensors have the same dimensions, with an active area of $20 \times 20\,\mathrm{mm}^2$. A small non-active area of the sensor chip houses peripheral digital and analogue circuitry, enlarging the chip in one dimension to about 23 mm. The chips are mounted on High Density Interconnect (HDI) circuits, which incorporate both signal and power lines as aluminium traces on thin polyimide substrates. The HDIs provide power and bias voltage, and transmit control signals and data. The latter over, up to, 3 differential lines per chip at a bandwidth of 1.25 Gbit/s per line.

The MuPix chips are bonded to the HDI using *Single-point Tape Automated Bonding* (SpTAB) without the need for additional bonding material [23]. Pixel modules are constructed from self-supporting sensor-HDI-polyimide ladders. These host between 6 and 18 sensors, and represent a total radiation length of approximately $X/X_0 = 0.115\%$ per layer.

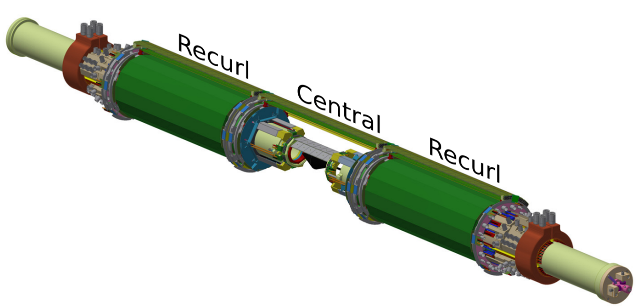


Figure 37: The Mu3e pixel tracker with the central pixel tracker in the middle and the two recurl stations down- and upstream. Some modules in the central pixel tracker have been removed for visibility.

7.1.1. Pixel Tracker Layout

The central pixel tracker has four layers of MuPix sensors: two inner layers (layer 1 and 2) at small radii and two outer layers (layers 3 and 4) at larger radii. The inner and the outer layers are both arranged as double layers, pairs that work together to provide a track trajectory. The layout of the central pixel tracker is shown in Figure 38 and the corresponding geometrical design parameters are listed in Table 2. The recurl stations have only two pixel layers (layers 3 and 4) which are identical in design to the outer layers in the central tracker.

Each tracking layer is composed of mechanically robust modules which integrate 4 or 5 of the more fragile sub-modules (ladders). Ladders represent the smallest mechanical unit in the tracker.

The inner tracking layers, 1 and 2, are of equal length, 12 cm, hosting 6 chips per ladder. These provide the vertexing in Mu3e. The inner layers have full overlap in z with the muon stopping target, which has a length of 10 cm. The outer and recurl pixel tracker modules are significantly longer and provide a larger acceptance for downstream and upstream going particles. The outer and recurl layers are critical for selection of high-quality tracks and for the momentum resolution in Mu3e. The outer layers instrument a region with a length of 34 cm (layer 3) and 36 cm (layer 4), corresponding to 17 and 18 MuPix chips, respectively.

The MuPix ladders are mounted with a small overlap, in the radial direction, of the active area with the adjacent ladder, see Figure 39. The lateral overlap is 0.5 mm, which ensures high acceptance for low momentum tracks and also helps with the alignment of the pixel tracker. There is a small physical clearance between overlapping sensors of $\approx 200\,\mu\mathrm{m}$.

7.1.2. Signal path

The signal connection between the front-end FPGA board, located on the service support wheels (SSW, section 13.3), and the MuPix chips is purely electric and differential with impedance-controlled lines. A schematic path of a differential signal is shown in Figure 40. The FPGA board is plugged into a back-plane where basic routing is performed. The distance to the detector (about 1 m) is bridged with micro-twisted pair cables, each consisting of two copper wires with $127\,\mu\mathrm{m}$ diameter, insulated with $25\,\mu\mathrm{m}$ polyimide and coated together with a polyamide enamel. The differential impedance of this transmission line is $Z_{\mathrm{diff}} \approx 90\,\Omega$. 50 such pairs are combined to a flexible bundle with a diameter of less than 2 mm. At both ends, the wires are soldered onto small PCBs, plugged into zero-insert-force (ZIF) connectors. On the detector end, the signals are routed on flexible PCBs to the HDI (see section 7.2.5). The connections between the components use industry-standard parts (back-plane connectors, gold-ball/gold-spring array interposers) and SpTAB bonding, as shown in the figure.

7.2. Pixel Tracker Modules

The pixel tracker modules of all layers have a very similar design. They consist of either four or five instrumented ladders mounted to a polyetherimide (PEI) end-piece at the upstream and downstream ends. The ladders host between 6 and 18 MuPix chips glued and electrically connected to a single HDI circuit. For the inner two layers, self-supporting half-shells define a module, with each half shells comprising four (layer 1) or five (layer 2) short ladders with six MuPix sensors.

For the outer two layers, a single module is an arc-segment, corresponding to either 1/6th (layer 3) or 1/7th (layer 4) of a full cylinder. Outer layer modules comprise four ladders with either 17 (layer 3) or 18 (layer 4) MuPix sensors.

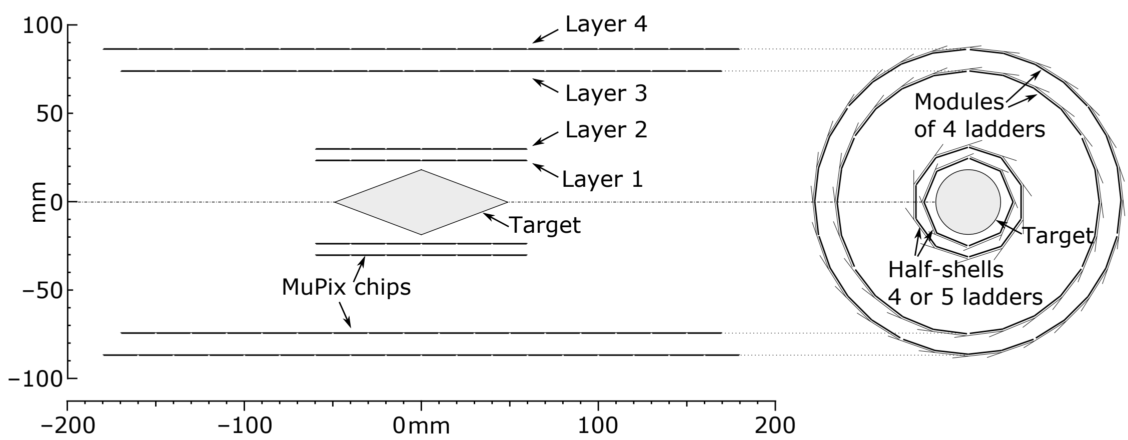


Figure 38: Geometry of the central pixel tracker including the target.

layer	1	2	3	4
number of modules	2	2	6	7
number of ladders	8	10	24	28
number of MuPix sensors per ladder	6	6	17	18
instrumented length [mm]	124.7	124.7	351.9	372.6
minimum radius [mm]	23.3	29.8	73.9	86.3

Table 2
Pixel tracker geometry parameters of the central barrel. The radius is defined as the nearest distance of MuPix sensor w/o polyimide support to the symmetry axis (beam line).

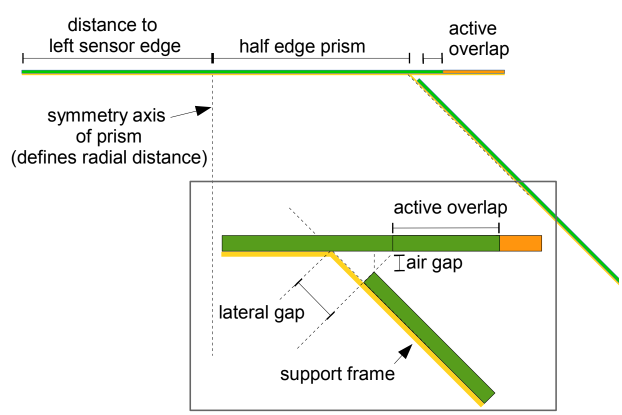


Figure 39: Design of the MuPix ladder overlap region of pixel layer 1. The green region indicates the active part of the sensor, the orange part the inactive periphery and the yellow part the polyimide support structure. The lower edge shows a zoom into the overlap region. Note that the sensor thicknesses are not to scale.

The longer outer layer ladders require additional reinforcement to achieve a mechanical stability comparable to the shorter ladders of the inner layers. For this purpose, two polyimide strips folded into a v-shape (yellow structure in Figure 43) are glued to each ladder on the inner side. The obtained v-channels also serve as high flux helium cooling channels.

7.2.1. Inner Layer Modules

Modules for layers 1 and 2 are constructed by mounting for or five ladders to upstream and downstream half-shell end-pieces. The half shells are strengthened with a 25 µm polyimide foil, glued to the MuPix ladders on the inward facing side. The foil also restricts helium from flowing through the gaps between ladders. After mounting and gluing, the half module represents a mechanically robust structure. To construct the full layers 1 and 2, half-shell modules are mounted on two end-rings, see Figure 41.

The electrical connection to the outside is made through multilayer copper-polyimide interposer flexprints which are SpTAB-bonded to the HDI just outside the active region at the position of the end-pieces. The interposer flex is connected to repeater PCBs via the interposer, which provides a 7×12 micro grid array of gold-spring contacts. The repeater PCB distributes all the signal and power lines.

7.2.2. Outer Layer Modules

The thirteen outer tracking modules in the central detector, see Figure 42, have a modular structure. Each module comprises four MuPix ladders which are glued to upstream and downstream module endpieces. As with the inner layers, the HDI circuit is SpTAB bonded to a multilayer copper-polyimide interposer flex circuit that connects to the 7×12 micro grid interposer plate. Connections from four ladders are combined on

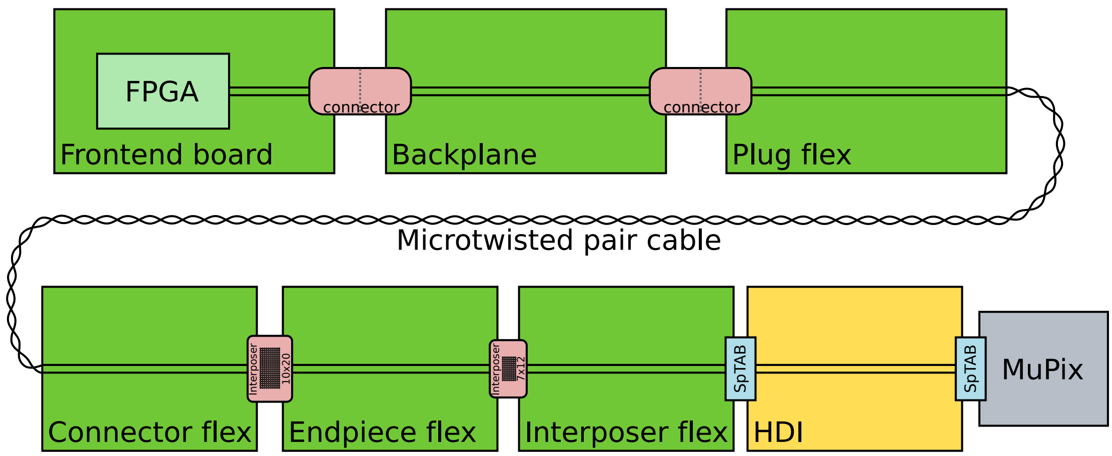


Figure 40: Signal path between MuPix chip and FPGA for a differential readout line. The parts on the top are located on the SSW, the ones on the bottom in the tracker barrels.

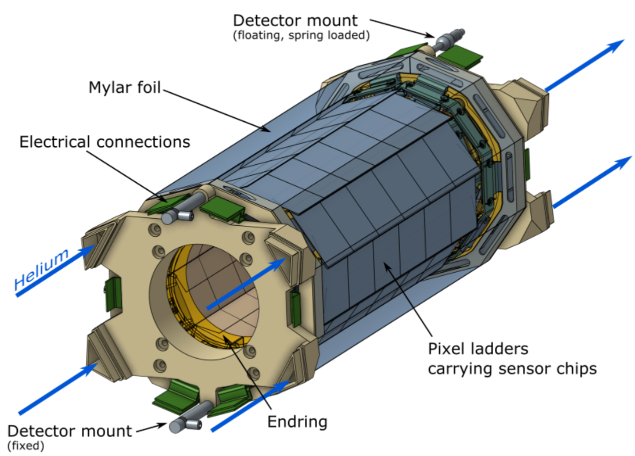


Figure 41: Schematic view of the vertex pixel barrel. Each of the 8+10 ladders carry 6 sensor chips. End-rings (split in halves) provide the mechanical support to the ladders.

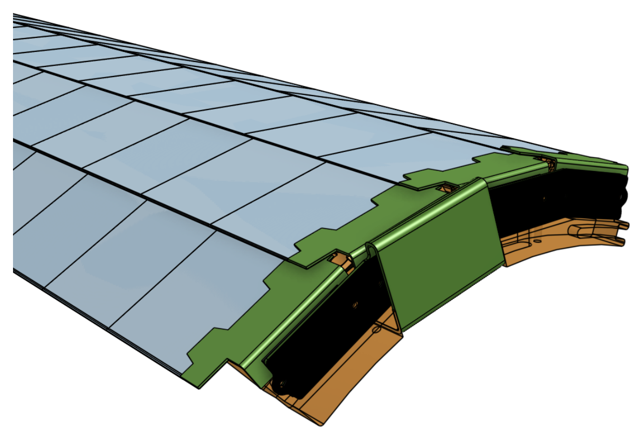


Figure 42: Schematic representation of a layer 4 module, integrating four long ladders with 18 MuPix sensors each. The picture shows one end, including the holding end-piece which also provides the electrical connections. An exploded view can be found in Figure 49.

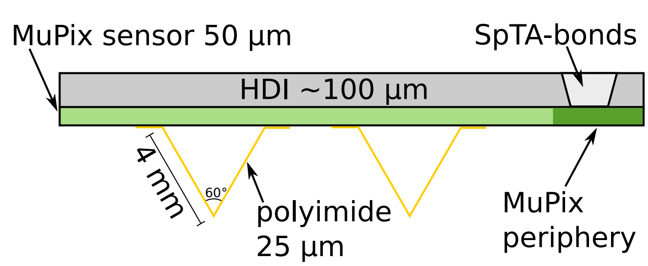


Figure 43: Cross section of an outer layer ladder. From top to bottom: HDI, MuPix sensor, polyimide support structure. Not to scale.

the endpiece flex circuit. The final connection from a module to the outside world is made through a further 10×20 interposer plate, combining gold-spring and ball-grid array contacts, to which the module connects when mounted on its endrings. Layer 3 and 4 modules are assembled into full cylinders by mounting to a PEI endring. The design foresees a swing-in mechanism for installation, where modules are located by a dowel pin on each end-ring and fixed by two screws on the endpieces at either end. Modules for the recurl stations are identical to the outer layer modules in the central region.

7.2.3. Pixel Ladder Design

The MuPix ladders integrate and support the pixel sensors. They have a compound structure optimised for a minimal material budget. The material composition for the inner and outer MuPix ladders is listed in Table 3 and amounts to a radiation length of approximately $X/X_0 = 0.115\%$ per layer.

The mechanical stability of the outer MuPix ladders is mainly determined by the two v-fold channels on the inner side which also serve as cooling channels.

	thickness [µm]	Layer 1-2 X/X_0	thickness [µm]	Layer 3-4 X/X_0
MuPix Si	45	$0.48 \cdot 10^{-3}$	45	$0.48 \cdot 10^{-3}$
MuPix Al	5	$0.06 \cdot 10^{-3}$	5	$0.06 \cdot 10^{-3}$
HDI polyimide & glue	45	$0.18 \cdot 10^{-3}$	45	$0.18 \cdot 10^{-3}$
HDI Al	28	$0.31 \cdot 10^{-3}$	28	$0.31 \cdot 10^{-3}$
polyimide support	25	$0.09 \cdot 10^{-3}$	≈ 30	$0.10 \cdot 10^{-3}$
adhesives	10	$0.03 \cdot 10^{-3}$	10	$0.03 \cdot 10^{-3}$
total	158	$1.15 \cdot 10^{-3}$	163	$1.16 \cdot 10^{-3}$

Table 3
Material budget of a MuPix ladder. The thicknesses and radiation length are given as an average over the 23 mm width of the ladder.

The inner layers do not have v-folds and are supported by the polyimide support structure, see Figure 39.

Every ladder is electrically divided into two halves and MuPix sensors are read out from both ends of the ladder, i.e. three sensors per half ladder for the inner layers and eight or nine sensors per half for the outer layers. The components of the MuPix ladders and modules are described in the following in more detail.

7.2.4. Sensors

The MuPix sensors are monolithic pixel sensors in HV-CMOS [22] technology. A full discussion of the functionality of the MuPix sensors as well as detailed performance results can be found in chapter 8. For the purpose of this section we discuss geometric properties and aspects relevant to the physical connectivity between the sensors and the outside world. Each sensor has a sensitive area of $20.48 \times 20.00\,\mathrm{mm}^2$ equipped with pixels of size $80 \times 80\,\mu\mathrm{m}^2$, corresponding to 256×250 pixels. The overall dimensions of each chip are $20.66 \times 23\,\mathrm{mm}^2$, where the additional non-sensitive area hosts a comparator and memory cells for each pixel, as well as voltage regulation and digital logic circuits. All MuPix sensors will be thinned to a thickness of 50 µm. The MuPix sensors can send data over up to three serial links, each running at a data rate of 1.250 Gbit/s. The sensors require an operating voltage of 1.8 V, a sensor bias voltage of up to −100 V, ground potential, and differential signal traces for the readout, clock, slow control and monitoring. Bond pads of size $200 \times 100\,\mu\mathrm{m}^2$ provide all electrical connections. All pads are arranged in one row in the inactive peripheral area of the chip.

7.2.5. High Density Interconnects

The High Density Interconnects (HDI) provide all electrical connections for the MuPix sensors which are directly glued onto the HDI after which electrical connections are made using SpTAB-bonding. In order to achieve the target material budget of one per mille of a radiation length per layer, the HDIs have to be very thin and must not contain any high Z materials. The HDIs are produced by LTU Ltd. (Ukraine) [24], who offer thin aluminium/polyimide technology as well as preparing the HDI for SpTAB-bonding. With the latter, aluminium traces are directly connected through vias either to chip pads or to other aluminium layers, see Figure 44 for an image of such bonds. This technique avoids the use of fragile wires and also saves material. Tests with prototypes circuits have shown good results [25].

The performance of all electrical lines on the HDI is critical to the successful performance of the MuPix ladders. The traces for power and ground have to be large enough to provide the required power of up to 30 W on the longest MuPix ladders. On the other hand, all traces should be as small as possible to fit them in the two aluminium layers available within the material budget. All fast signals (serial links, clocks, resets, etc.) are implemented using the LVDS standard to minimise

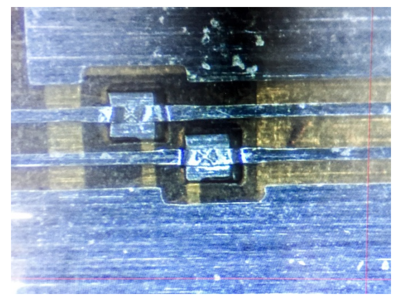


Figure 44: Photograph of two SpTAB bonds on a test flexprint produced by LTU Ltd [24].

Figure 45: Stack chosen for the LTU produced 2-layer HDI circuits. PI=polyimide, Al=aluminium.

Material	Thickness [µm]	X/X_0
upper Al layer	14	$1.57 \cdot 10^{-4}$
insulator (PI)	35	$1.22 \cdot 10^{-4}$
glue	10	$0.25 \cdot 10^{-4}$
lower Al layer	14	$1.57 \cdot 10^{-4}$
lower PI shield	10	$0.35 \cdot 10^{-4}$
total	83	$4.96 \cdot 10^{-4}$

Table 4
Material composition of the HDI for a $Z_{\text{diff}} = 100\,\Omega$ prototype.

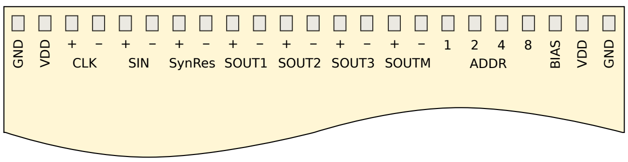


Figure 46: Conceptual MuPix pad layout on the HDI. Depending on location, either SOUT1 to SOUT3 or SOUTM is connected to accommodate for different data rate needs (vertex or recurl layers, respectively). Power and ground have multiple pads to reduce effects of voltage drop.

cross-talk. The differential impedances of all fast differential transmission lines are designed to match the specification. Differential bus terminations are foreseen on the last chip in a row, i.e. at the centre of the HDI. All fast differential transmission lines are laid out underneath wide ground and VDD potential traces which serve as shielding and define the impedance. With a minimum possible width of the aluminium traces of 63 µm (LTU), the distance between signal and shielding layer is 45 µm with polyimide as the insulator. The thickness of the insulator and aluminium layers and the outer shielding define the total thickness and thus the material budget of the HDI. The main parameters of the HDIs are listed in Tables 4 and 5 and a schematic stack is shown in Figure 45.

Space constraints on the HDI have motivated the use of a minimal set of traces for power, control and readout. Differential buses are used for slow control, clock and reset. Global power and ground lines are foreseen. Voltage gradients between sensors due to path length dependent ohmic losses are minimised by design. The remaining small voltage differences are planned to be equalised using voltage regulators implemented in the MuPix chip. A fall-back design solution is to place these voltage regulators on the end-ring and to route the power and ground lines point-to-point to every single – or a small group – of sensors.

Every pixel sensor is electrically connected by only 21 pads, see Figure 46. Four address bits, selected by SpTAB bonding pads to ground or the supply voltage bus are used to set the chip address for the uplink communication bus. All bond pads have a relatively large size of at least $150 \times 150\,\mu\text{m}^2$ to fulfil the specification requirement for SpTAB-bonding and to ensure a high production yield.

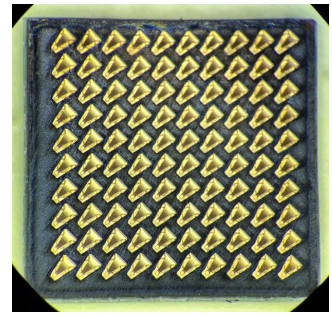


Figure 47: Interposer ZA8 from Samtec, version with 10×10 connections. The pins have a pitch of 0.8 mm

The MuPix ladders will be electrically connected to the end-rings by interposers. The chosen interposer from Samtec (ZA8H)[26] is a type of flat connector allowing for high speed electrical signal transmission up to 30 GHz, see Figure 47, with a high density of connections and a total thickness (compressed) of 0.3 mm.

7.2.6. *Module End-Pieces*

The end-pieces of the inner layers consist of half-arcs made from PEI[16] to which the polyimide support structure and the MuPix ladders are glued. The end-pieces of layer 1 (layer 2) have a four (five)-fold segmentation, see Figure 48.

The end-pieces of the outer layers have a fourfold segmentation and include an internal open volume to distribute helium gas for the cooling-system, see Figure 49. To realise the internal volume for helium distribution the endpieces are manufactured out of a main part with a thin lid, glued on, to seal the open volume after machining. The endpieces also accommodate the

[16] We use this insulator material to mitigate the risks of eddy currents in case of magnet quenches.

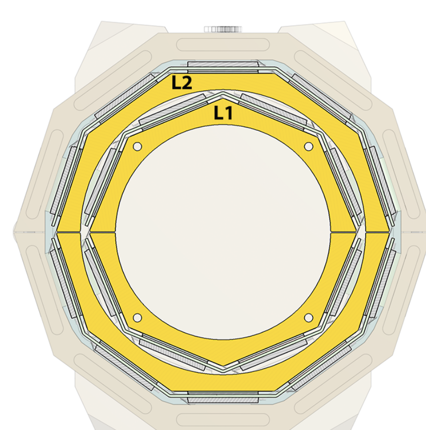


Figure 48: Holding end-pieces and end-rings of the inner layers with the octagonal (layer 1) and decagonal (layer 2) geometry, shown in yellow.

	layer 1&2	layer 3	layer 4
HDI length [mm]	140	360	380
instrumented area [mm^2]	120×20	340×20	360×20
number of MuPix chips	6	17	18
the following numbers refer to ladder halves			
number of traces:			
bias (HV) (BIAS)	1	1	1
$1.8\,\mathrm{V}$ (VDD)	1	1	1
ground (GND)	1	1	1
number of differential pairs:			
data (SOUT)	3×3	9×1	9×1
clock bus (CLK)	1	1	1
reset bus (SynRes)	1	1	1
slow control bus (SIN)	1	1	1

Table 5
Specification of the HDIs for inner and outer layers. All lines have a shared bus topology except for the fast data lines (SOUT), which are point-to-point. Note that the numbers of electrical lines refer to each half of the HDI.

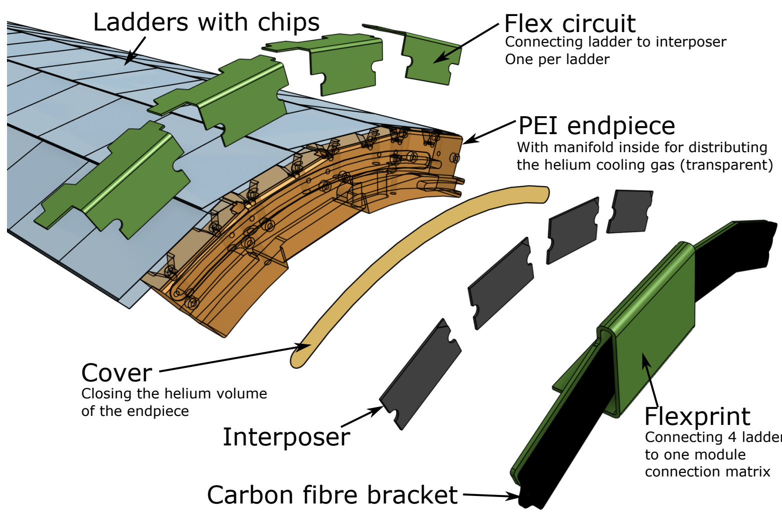


Figure 49: Exploded view of the outer layer module assembly. The end piece region provides a manifold for distributing the helium cooling gas and holds the flexible circuits to connect ladders to a single electrical connection matrix on the bottom.

interposer connectors for power, control and data transmission.

7.2.7. MuPix *Ladder Integration and Chip Bonding*

During fabrication of the MuPix ladders, chips are placed accurately on the HDI by use of fiducial marks on the chips and cut-outs on the HDI. The chips are then glued using an epoxy (Araldite 2011) and the positions are checked again. The flex circuit for connecting to the interposer is placed and glued in a similar manner. After curing, all connections between the HDI and each chip, and between the HDI and the interposer are SpTAB-bonded (any vias on the HDI are bonded beforehand by the manufacturer). Once all the connections are in place, the unit is electrically fully functional. This allows for the comprehensive quality testing of a MuPix ladder before they are assembled into modules.

7.3. Pixel Tracker Global Mechanics

Pixel tracker inner and outer layer modules are integrated into the full cylindrical tracking layers by mounting the modules to the inner or outer layer pixel endrings. The latter in turn are connected to the up- and downstream beam-pipes. Like the module endpieces these are manufactured out of PEI. For the inner layers the endrings have gas inlets and outlets to provide the helium flow between layers 1 and 2.

A drawing of an outer layer end-ring equipped with a layer 3 module is shown in Figure 50. The outer endring supports six modules of layer 3 and seven modules of layer 4. The end-rings have dowel pins for every module to guide the module when it is rotated into

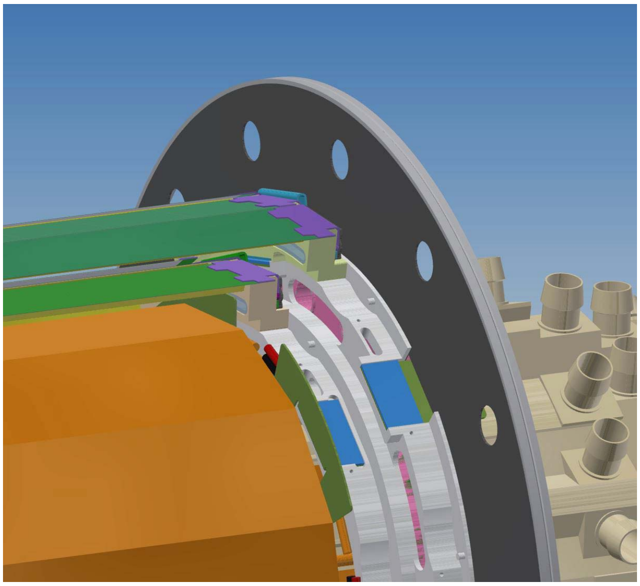


Figure 50: End-ring situation shown with modules inserted in layers 3 and 4.

position, ensuring no accidental contact is made with already mounted modules during installation. The final mechanical connection is done with screws. This also secures the contact through the end-ring interposer which provides the further electrical connection from the module, via the front-end flexprints, to the front-end boards, from where the steering and control of the MuPix sensors is handled and where the signal processing is done.

The outer layer end-rings provide conduits for the helium gas flow between layers 3 and 4 and to the module end-pieces for the gas flow in to the v-fold channels. At the upstream end, the inner and outer end-rings are rigidly connected to the beampipe. At the downstream end, the end-rings are supported by bearings and connected via a small spring tension such that the downstream end-rings can move along the beam direction to accommodate thermal expansion of the ladders.

7.4. Pixel Tracker construction and Quality Control

The production workflow of the pixel detector parts consists of manufacturing steps and quality control points, shown in Figure 51. The manufacturing steps make use of custom-made tooling for careful picking and accurate placing of parts. To protect parts from damage and contamination, manufacturing will take place in controlled environments, e.g. cleanrooms of suitable levels, and standard ESD protection procedures will be in place. Polyimide expands when exposed to humidity. All manufacturing steps crucial to defining tolerances will be carried out in environments with strict temperature and humidity control and material will be stored therein for proper equilibration prior to use. Raw parts are either obtained from suppliers (e.g. MuPix, HDI, interposer, etc.) or made in-house using custom tooling (e.g. polyimide folds) or CNC machines (e.g. endpieces).

Quality control takes place before and after every manufacturing step. Tests include (but are not limited to): visual inspection, dimension control, electrical testing, and gas leak testing. All components and their test results are tracked and documented in a production database. Raw parts will be acceptance tested upon receipt. In case of the MuPix chips, electrical testing will take place on the wafer, and on single die after dicing, using appropriate probe cards. Thanks to the modular design of the process, full electrical testing of all intermediate products is possible and foreseen. This includes the possibility to check sensor response using lasers or lab-grade radioactive sources.

7.4.1. Inner Pixel Layers: Ladder and Module Production

The full inner pixel production and assembly takes place at Heidelberg. The small nature of this detector part (18 ladders with 6 chips per ladder) makes a manual procedure a cost-effective choice.

Chips are positioned relative to each other and to the interposer flexes on a custom jig. The interposer flexes define the position of a ladder on the endrings. The positioning is done by moving the chips with a sliding block and fixating each chip at the desired position by vacuum (Figure 52). While the position of the first chip is defined by a stop edge, following chips are placed using a micrometer screw and by monitoring the chip-to-chip gap with a microscope[17].

Epoxy (Araldite 2011) is applied to the chips and the interposer flexes manually in small dots. The HDI is aligned to the chips and flexes by fiducial marks on both parts under the microscope. Weights ensure flatness and a uniform distribution of glue. A finished prototype ladder on the jig is shown in Figure 53. Prototype construction has demonstrated a placement precision of $\sigma < 5\,\mu\mathrm{m}$ and an average glue thickness of $(5 \pm 4)\,\mu\mathrm{m}$.

After curing, connections between the chips and the HDI and between the interposer flexes and the HDI are made using SpTAB bonding. From this point on, the ladder is electrically fully functional. Each ladder undergoes a basic functionality test including powering, configuration and the readout of each MuPix chip.

Ladders that pass all QA checks are mounted into half-shells on custom assembly tools (Figure 54). These tools, for layer 1 and layer 2, accommodate the module endpieces and are designed such that each facet can be brought into the horizontal position for ladder placement. The ladders are glued consecutively to the polyimide flap of the previously positioned ladder. Again, weights ensure flatness and a uniform glue distribution. At the same time, the ladders are attached to the PEI endpieces by clamping them to a stack comprising the

[17] Dino-Lite AM4515T8-EDGE, resolution of $1.5\,\mu\mathrm{m}$

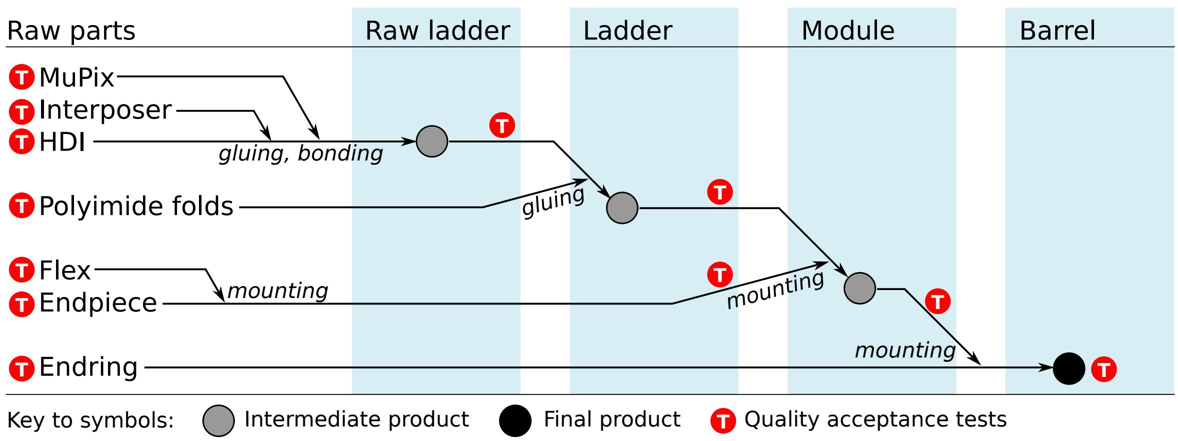


Figure 51: Module manufacturing workflow and quality points. Only main steps are shown.

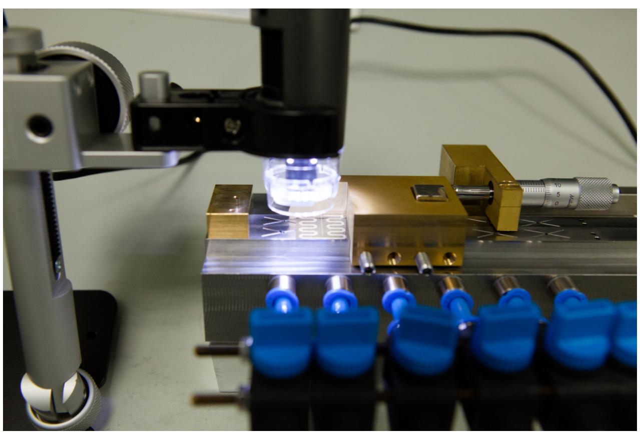


Figure 52: Assembly tool for the inner tracking ladders. Brass sliding block in the middle is guiding a prototype chip into position. Brass stop edge to the left. Micrometer screw to the right. Microscope to monitor position at the top.

end of the ladder, the interposer and the endpiece flex held by a carbon fibre bracket.

7.4.2. Outer Pixel Layers: Ladder and Module Production

Ladder assembly for layers 3 and 4 of the MuPix tracker takes place at the Oxford Physics Microstructure Detector (OPMD) Laboratory. To make a ladder, 18 (17) chips are positioned on a vacuum jig using a 4-axis gantry positioning system, integrated with vision and electro-valve controls, and custom built tooling (see Figure 55). A positioning accuracy within 10 microns (see Figure 56) is achieved. After this interposer flex circuit is added, located by the jig, glue is deposited by a commercial machine vision guided liquid dispensing robot and the HDI is glued to the chips using a counter-jig. Connections between the sensor chips and the HDI circuits are made using SpTAB bonding. The completed assembly is reinforced with two V shaped, folded polyimide support structures glued to each ladder. The liquid dispensing robot is used to accurately apply the required epoxy to achieve 5 micron thick glue layers to adhere sensor chips to the flexprints and polyimide V-folds to the ladders.

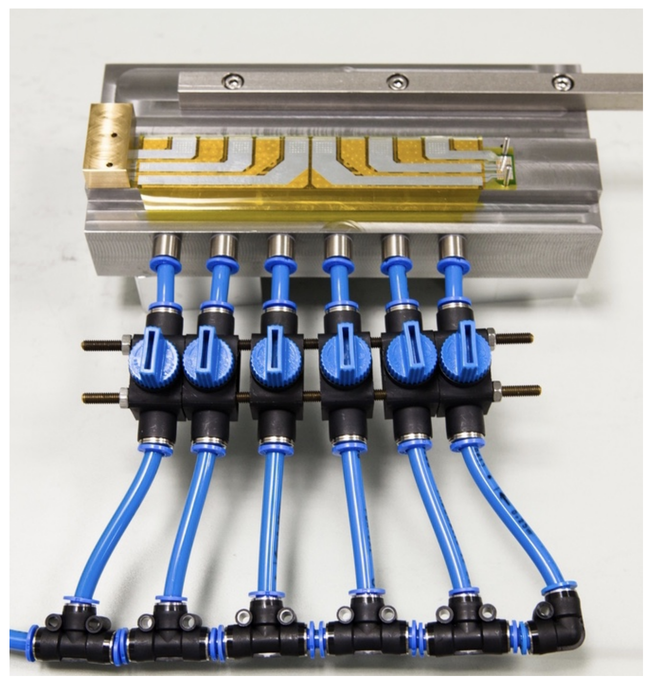


Figure 53: Prototype ladder for the inner pixel layers after gluing on mounting tool.

The polyimide V-folds are repeatably aligned and joined to the ladders using a custom jig with linear rails and micrometer adjusters. Semi-automated non-contact metrology of components and completed ladders is performed with an optical probe on a coordinate measuring machine.

After testing, MuPix ladders are shipped to Liverpool, for the assembly into modules. Upstream and downstream module endpieces are mounted to a custom jig that defines the overall length of a module. Ladders

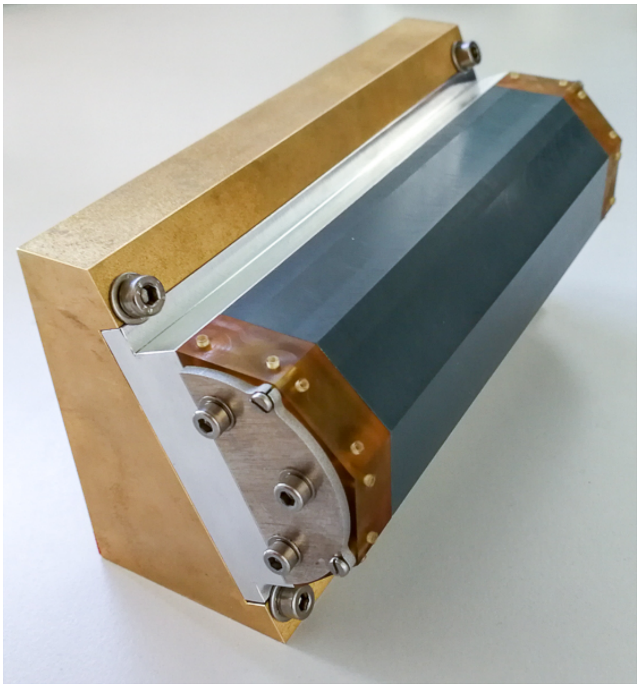


Figure 54: Assembly tool for layer 2 modules. Tilting of the full tool and sliding of the grey block allows to bring every facet into the horizontal position.

are positioned and glued onto the endpiece. Glue is applied manually to the surface of the endpiece and inside the v-shaped cut-out in the endpiece, as well as to the underside of the end of each ladder and the outside of the polyimide v-channels. Weights are used to ensure flatness and a uniform distribution of glue. After four ladders are assembled into a module, the v-channels on the inward facing side of the module are sealed with additional adhesive and electrical connections are made by fixing a stack of the ends of the four ladders, four interposers and the endpiece flex with a single carbon fibre bracket. Modules are checked for gas flow and leaks and for electrical conformity.

7.5. Prototyping and System Tests

A programme of manufacturing thermo-mechanical prototype modules for both the inner and outer layers of the MuPix tracker has been used to develop and commission the assembly tooling and processes. At the same time the built modules are intended to provide a testbed, called the *thermo-mechanical mockup* (TMM), to develop and demonstrate the helium cooling concept for the MuPix tracker. Modules for the TMM provide a close match to the final detector in terms of their mass and materials used and provide the means to dissipate heat loads, matching those in the real detector into the structure. Circuitry to monitor temperatures is also incorporated. Modules are built out two types of ladders:

Silicon heater ladders closely match the material stack of the final detector. Silicon heater chips (Figure 57) have been manufactured at the Max-Planck Halbleiterlabor in Munich using sputtered aluminium on silicon without a passivation layer. A meander with $R = 3.24\,\Omega$ allows heat to be generated in the chip in the range of 1 to 1.6 W with similar voltages as for MuPix chips. An additional meander with $R \approx 1000\,\Omega$ is used as a resistance temperature detector (RTD) to measure the temperature in situ. The chips are thinned to 50 µm thickness. Ladders are fabricated using adapted versions of the HDI with the same stack as foreseen for the detector (Figure 58). Connections are made with SpTAB bonding. The manufacturing steps needed and tooling used are the same as for detector fabrication, providing the ideal test bed to develop, commission and qualify the tooling and processes for the final detector production.

Figure 55: Robotic gantry (upper figure) for placement of 17 or 18 MuPix chips on the vacuum jig (lower figure).

Tape heater ladders are simpler objects, based around an aluminium-polyimide laminate (Figure 59) resistive heating circuit that has the same shape as the HDI plus interposer flex assembly in the final detector. Laser cutting and etching are used to manufacture the tape heater flexes in sizes corresponding to inner ladders ($R \approx 0.5\,\Omega$) and outer ladders ($R \approx 3.7\,\Omega$). To create a more realistic mechanical model and material budget, 50 µm thick stainless steel dummy chips can be attached if needed for specific test purposes. This more

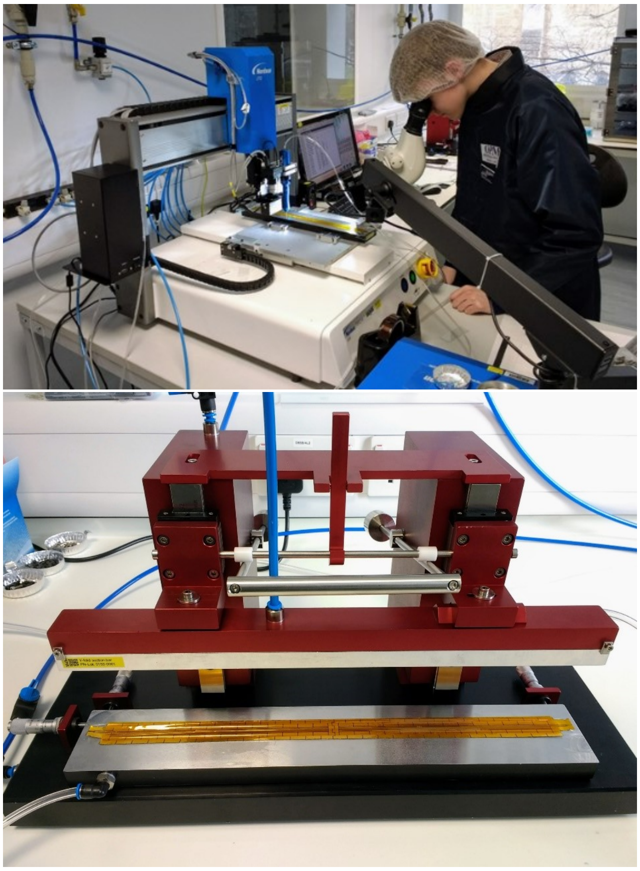


Figure 56: Glue dispensing robot (upper figure) and custom tooling for the gluing of v-channel reinforcements (lower figure).

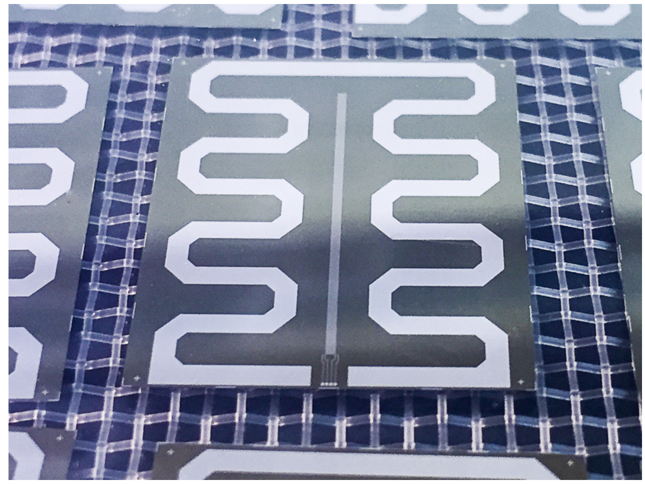


Figure 57: Silicon heater chip. The large meander for heating the chip and a narrow meander used as an RTD can both be seen. Contact pads are arranged on the bottom edge corresponding to the final chip connection locations.

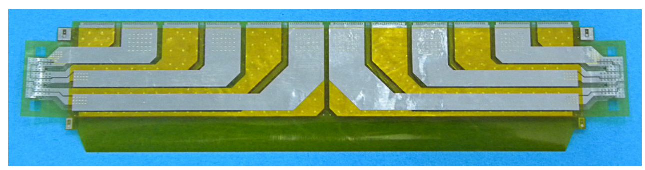


Figure 58: HDI for silicon heaters, layer 1 and 2. Six silicon heaters can be mounted on the back side.

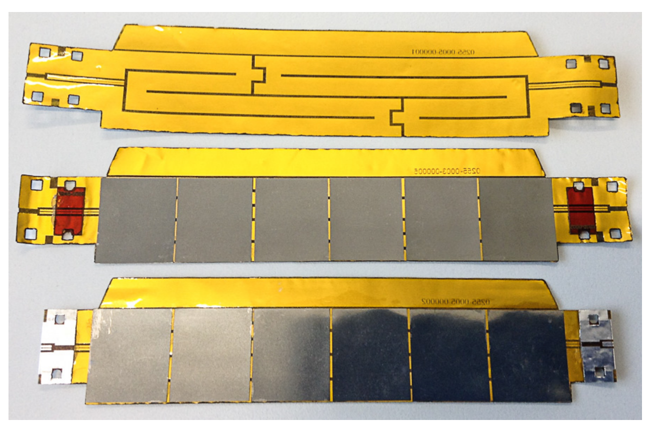


Figure 59: Tape heaters for layer 1 and 2. Top to bottom: bare heater with meander, stiffener attached to match final dimensions, dummy chips glued on. Large contact pad pair on both ends used for powering. Chip size $20 \times 23\,\mathrm{mm}^2$.

cost-effective option is used for simpler manufacturing tests and to instrument most of the full TMM.

The TMM is assembled by assembling the silicon heater and tape heater ladders into modules and barrels using the same mechanical components as are use in the final detector. With the heating capabilities and all the cooling facilities in place, realistic measurements of the cooling and mechanical stability will be possible. Tests stands for intermediate and final assemblies have been developed in preparation for the final testing of detector assemblies. All manufacturing steps are taking place at the locations foreseen for detector fabrication.

An example of ladders manufactured for the TMM is shown in Figure 60.

7.6. Pixel Tracker Cooling

The full pixel detector will dissipate about $4.55\,\mathrm{kW}$ of heat[18] in a *conservative scenario* assuming $400\,\mathrm{mW/cm}^2$. The latest chip versions have shown a heat dissipation below $250\,\mathrm{mW/cm}^2$. This heat load is used for our most *realistic scenario*. Table 6 shows expected heat load in each layer of the tracker under these two scenarios. The cooling system must keep the maximum temperature, anywhere in the pixel detector, safely below $70\,°\mathrm{C}$, given by the glass-transition temperature of the adhesives used for construction.

We use gaseous helium at ambient conditions[19] as coolant. The helium is distributed in separate circuits, serving different parts of the detector separately. The

[18] Throughout this section, heat from chips and losses in conductors inside the HDI are taken into account, summing up to the heat density used in the scenarios.

[19] This means temperatures above $0\,°\mathrm{C}$ and the absolute pressure around 1 bar.

Detector Part	Area [cm^2]	$250\,\mathrm{mW/cm^2}$ [W]	$400\,\mathrm{mW/cm^2}$ [W]
layer 1	192	48	77
layer 2	240	60	96
layer 3	1632	408	652
layer 4	2016	504	807
Recurl Station ($2\times$)	3648	912	1459
total	11376	2844	4550

Table 6
Heat dissipation of the pixel detector for a power consumption of $250\,\mathrm{mW/cm^2}$ (realistic scenario) and $400\,\mathrm{mW/cm^2}$ (conservative scenario).

No.	Description	#	Inlet			Outlet	
			$\dot{m}$ g/s	Δp mbar	v m/s	$\dot{m}$ g/s	Δp mbar
1	Gap flow vertex detector	1	2.0	+40	10	2.0	-40
2	Gap flow b/w SciFi and L3	1	6.9	+25	10	0	0
3	Gap flow b/w SciTile and L3	2	5.7	+28	10	0	0
4	Gap flow b/w L3 and L4	3	7.6	+25	10	0	0
5	Flow in V-folds L3	3	1.3	+90	20	1.3	-90
6	Flow in V-folds L4	3	1.5	+80	20	1.5	-80
7	Global flow, $D \approx 300\,\mathrm{mm}$	1	4	+0.04	var.	45	-0.04
	Total	14	56			56	

Table 7
List of helium circuits inside the experiment. Pressures are given relative to ambient in the experiment and were obtained from CFD simulations. Circuits with outlet flows and pressures of 0 vent into the main volume, collected in the global flow outlet. The total flow corresponds to about $20\,\mathrm{m^3/min}$ under standard conditions. Column # gives number of identical circuit copies in the detector.

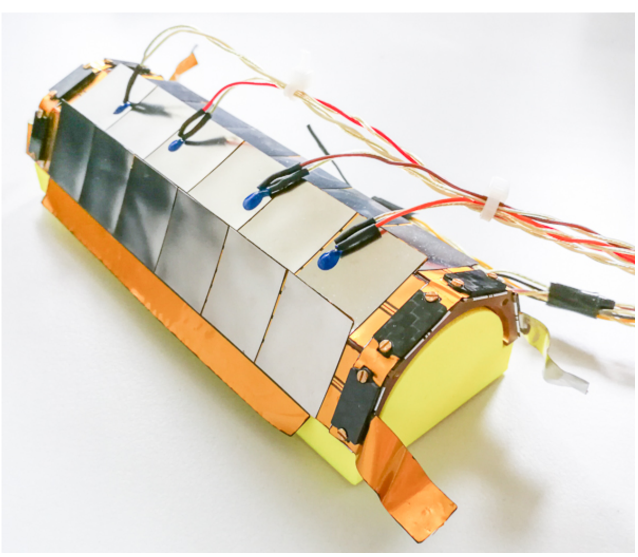
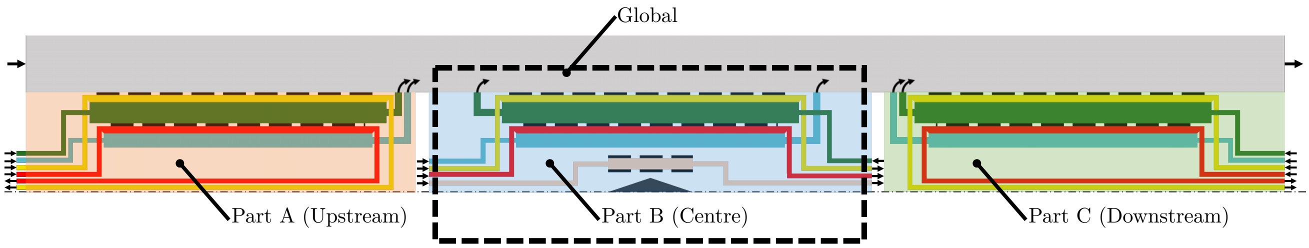


Figure 60: Layer 1 half shell made with tape heaters and stainless steel dummy chips, placed on a handling block (yellow). RTDs attached with conductive glue for temperature profiling in cooling tests.

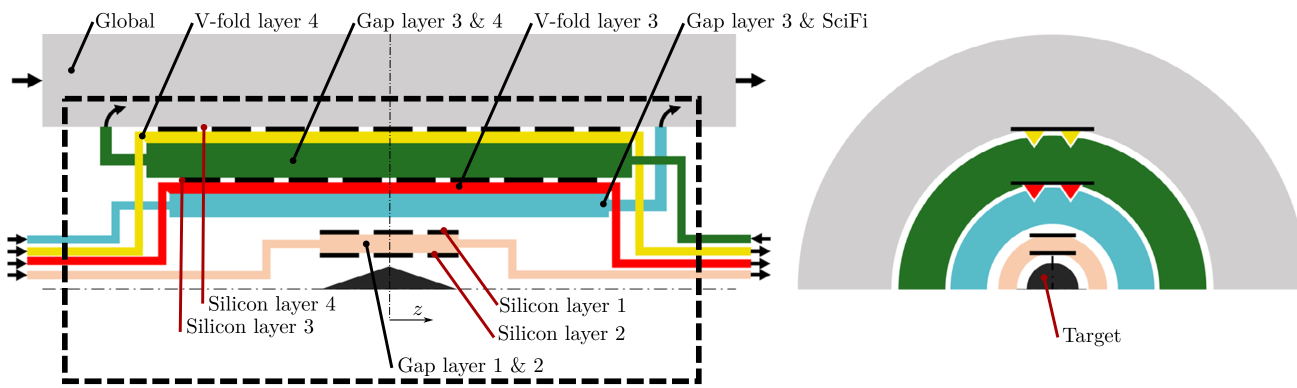
concept is shown in Figure 61 and the different helium circuits are listed in Table 7. The flow in the global circuit increases along z because of other circuits directly venting into the global flow. The global flow is constrained by a thin mylar foil (thickness $5\,\mu\mathrm{m}$) surrounding the full pixel detector in a conical shape that keeps the helium velocity near constant along z (see Figure 41.

The system described is the result of a process of optimisations through simulation studies using computational fluid dynamics (CFD)[20] and verifications in the laboratory [27, 28, 29, 30, 31, 32]. The models used in both simulation and laboratory measurements progressed in detail and the final mock-up models described in section 7.5 match the final detector to a great extent in shape, materials and heat-density. The heat-load density distribution used in the simulations take care of the uneven heat density on the pixel chip. Half of the power dissipation on the chip is expected to be located on the periphery, the remaining half within the pixel matrix, equivalent to $200\,\mathrm{mW/cm^2}$ and $1730\,\mathrm{mW/cm^2}$ respectively, for an averaged $400\,\mathrm{mW/cm^2}$ in the conservative scenario. Simulation results for the pixel tracker are shown in Figure 62, confirming a safe ΔT even in the conservative scenario. In simulations with the realistic scenario the obtained ΔT values have been found to scale down linearly with the reduced

[20] Autodesk® and ANSYS CFX® CFD software were used.



(a) Schematic of the helium flows in the full pixel detector



(b) Central detector cooling detail

Figure 61: Sketch of the helium cooling system for the pixel detector. (a) shows all volumes with its flow directions in a cut view. The system is cylindrically symmetric around the long dashed-dotted line. Some volumes vent into the global flow inside the experiment, indicated by bent arrows. Every circuit is individually controlled for flow and pressure inside the detector volume. (b) shows a cut in the transverse direction as well. The triangles (in red and yellow) indicate the V-fold channels, which exist in pairs for every ladder.

power dissipation and are therefore not shown.

The experimental cooling tests were performed inside a cylindrical closed volume with a diameter of 22 cm and a length of approximately 1 m. Helium was initially provided by compressed gas bottles, limiting measurements to a few minutes. This was overcome by using a miniature turbo compressor (described in chapter 12) allowing for helium recirculation and hence continuous operation. The agreement between simulation and mock-up measurements are good, see Figure 63 for an example comparison for the vertex detector [33].

Vibrations induced by the helium flows must not damage the structures or have a substantial impact on the hit resolution. Such vibrations were studied using a setup based on a Michelson interferometer pointing to reflective surfaces on a realistic mock-up. For velocities up to 20 m/s, average amplitudes of 2 µm were observed, with peaks of 10 µm [34, 30, 35]. This is well below the single hit resolution of the pixel sensors. An excitation spectrum using a speaker showed resonances between 50 Hz to 1000 Hz with no major peaks. No damage to the test structures has been observed during these studies.

Table 8
This is a placeholder table for technical reasons and will go away in the final version. Please disregard.

8. MuPix Pixel Sensor

This section is currently missing and will be handed in later. It will describe the MuPix sensor in great detail, including its functionalities, properties and results from performance measurements.

9. MuTRiG

A common Application Specific Integrated Circuit (ASIC) has been developed for both the fibre and tile detectors in Mu3e, capable of operating with the rather different conditions of the two systems.

9.1. Introduction

MuTRiG (Muon Timing Resolver including Gigabit-link) is a 32 channel, mixed-signal Silicon photo-multiplier (SiPM) readout ASIC designed and fabricated in UMC 180 nm CMOS technology. It has been developed to read out the fibre and tile detectors in Mu3e, and is designed to achieve the required timing resolution for

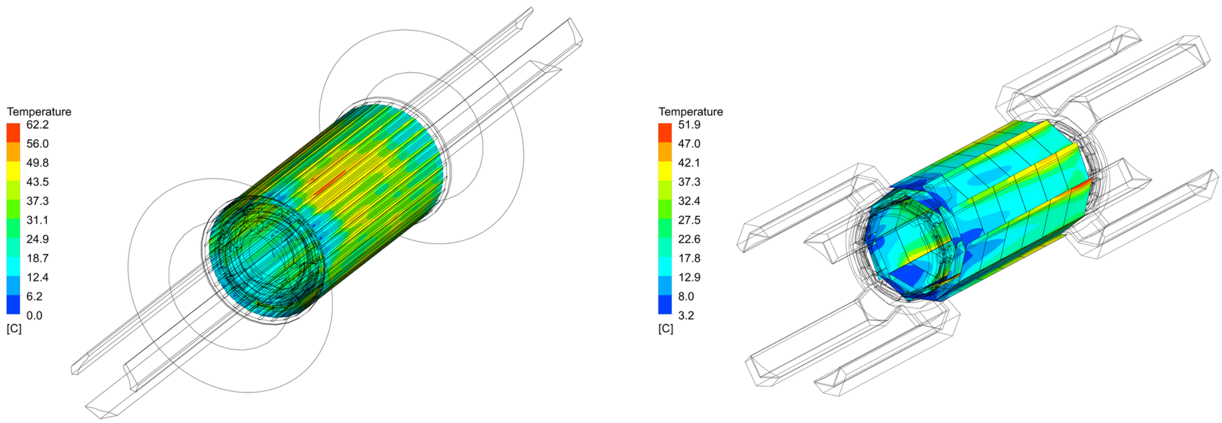
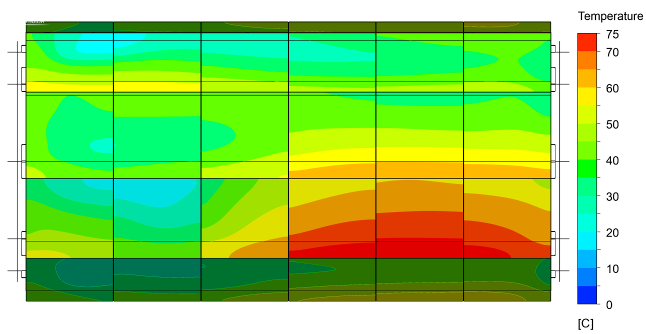
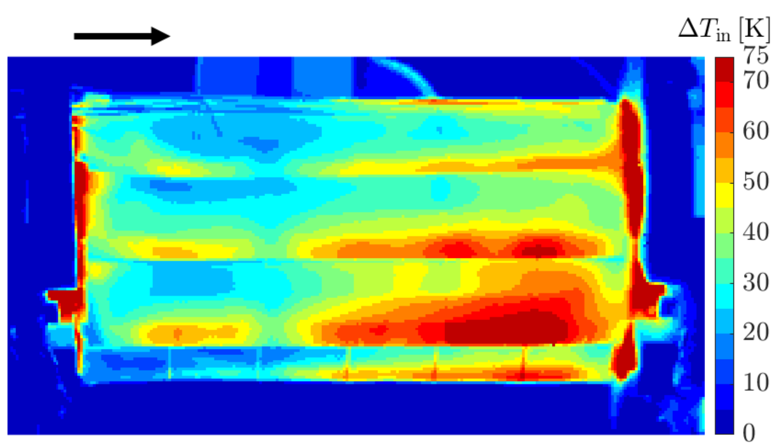


Figure 62: Simulated ΔT distribution of the silicon in the tracking detector for a power dissipation of $400\,\mathrm{mW/cm^2}$ (non-uniform distribution, more heat power at periphery edge). Inlet gas temperature is $T = 0\,°\mathrm{C}$ Left: full barrel. Right: vertex barrel inside the full barrel.



(a) Simulated temperature on the outer layer of the mock-up.



(b) Measured temperature on the outer layer of the mock-up using an infrared camera.

Figure 63: Temperature obtained by measurement and CFD-simulation. Angle of view of simulation has been carefully matched to the camera view. Cold helium enters from the left. Hot zones on the right of (b) are cable connections from the setup not present in the final detector.

Figure 64: This is a placeholder figure for technical reasons and will go away in the final version. Please disregard.

both systems while keeping up with the high event rate in the scintillating fibre detector.

9.2. ASIC Description

MuTRiG is an evolutional development from the STiCv3.1 chip developed at the Kirchhoff Institute in

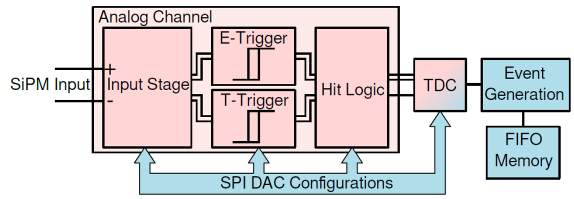


Figure 65: Diagram of a MuTRiG channel. After taking signal from SiPM by the input stage, separate are provided to the T-Trigger and E-Trigger branches for time and energy discrimination respectively. The discrimination signals are encoded in the hit logic module to generate the combined hit signal, and then converted to digital time stamps after the TDC module. The signal is then buffered in the on-chip memories before being transferred out of the chip. The analogue front-end, TDC and digital modules are configured using a Serial Peripheral Interface (SPI) interface

Heidelberg for medical applications of SiPMs (EndoTOFPET-US [36]).

The analogue processing building blocks of MuTRiG inherit from the STiCv3.1 chip, whose satisfactory performance has been validated in several testing conditions. However, the STiCv3.1 chip is only capable of transferring ∼50 kHz per channel through the 160 Mbit/s data link, which is too slow for the Mu3e timing detectors, especially for the fibre detector which is required to handle 1 MHz/channel event rate to achieve 100% data acquisition efficiency. The MuTRiG chip extends the excellent timing performance of the STiCv3.1 chip with a newly developed fast digital readout for high rate applications.

Figure 65 shows the channel diagram of the MuTRiG chip and Figure 66 shows the sketch of the chip functionality. (More details can be found in [37].)

The good timing resolution of MuTRiG derives from its differential analogue front-end and the 50 ps binning time-to-digital converter (TDC), which were inherited from the STiCv3.1 chip. The working principle of a TDC is shown in Figure 67. At the arrival of a hit signal over threshold, the TDC module samples the state of a *coarse-counter*, which is incremented at 625 MHz by a reference clock. A *fine counter* with 50 ps bins is then used to make a more precise measurement of the hit time within the 1.6 ns coarse counter bin. The coarse and fine counter values are then recorded as the time stamp of the hit signal. The time the signal drops back below threshold is similarly recorded. The *Global TimeBase Unit* provides common coarse and fine counter values to all the channels for time stamping, as shown in Figure 68. The TDC requires ∼30 ns to reset after a hit.

In order to fulfil the high rate data readout, a double data rate serialiser and a customised low-voltage differential signalling (LVDS) transmitter were developed to establish a gigabit data link with the data acquisition

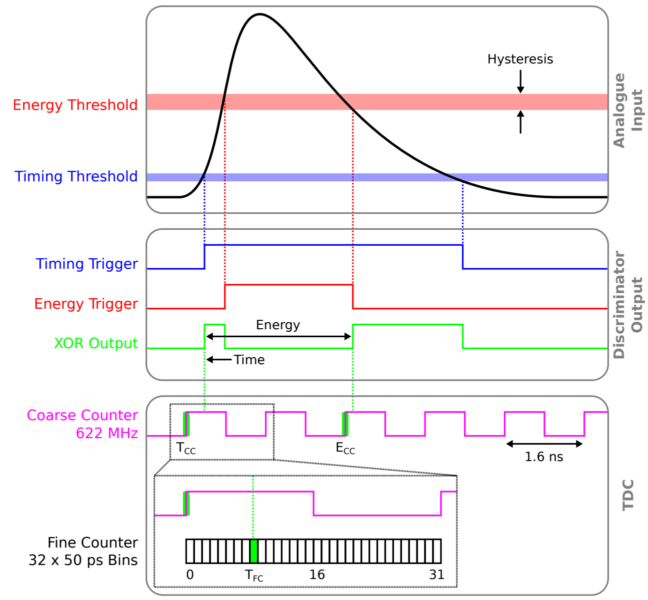


Figure 66: Sketch of the functionality of the MuTRiG chip. The time and energy information of the analogue input signal is obtained via two discriminator units. The discriminator output is processed by a TDC with a 625 MHz coarse counter and a fine counter with a bin size of 50 ps.

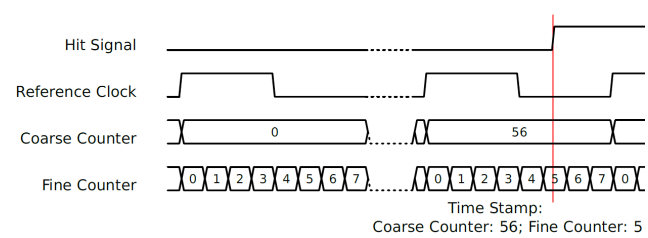


Figure 67: Working principle of the TDC, showing the fine and coarse counters, reference clock and example arrive of a hit signal.

system (DAQ) for data transmission. The event data from all the channels are buffered and sent out in frames via the 1.25 Gbps LVDS serial data link. In order to increase the event rate capability of the MuTRiG chip, the output event structure can be switched from the standard 48 bits, containing both the time stamps a hit signal passes above and back below threshold, to a short event structure of 27 bits, containing only the first of these times and a 1 bit energy flag of the hit.

A few more new functionalities were implemented in the digital logic circuit of the MuTRiG chip for convenient and reliable operation of the chip. Table 9 shows a summary of the major differences in event and data handling capabilities of the STiCv3.1 and MuTRiG chips.

9.3. Characterisation Measurement

9.3.1. Rate Limitation Measurement

The event rate limit of the chip is measured by injecting test pulses to multiple channels and measuring

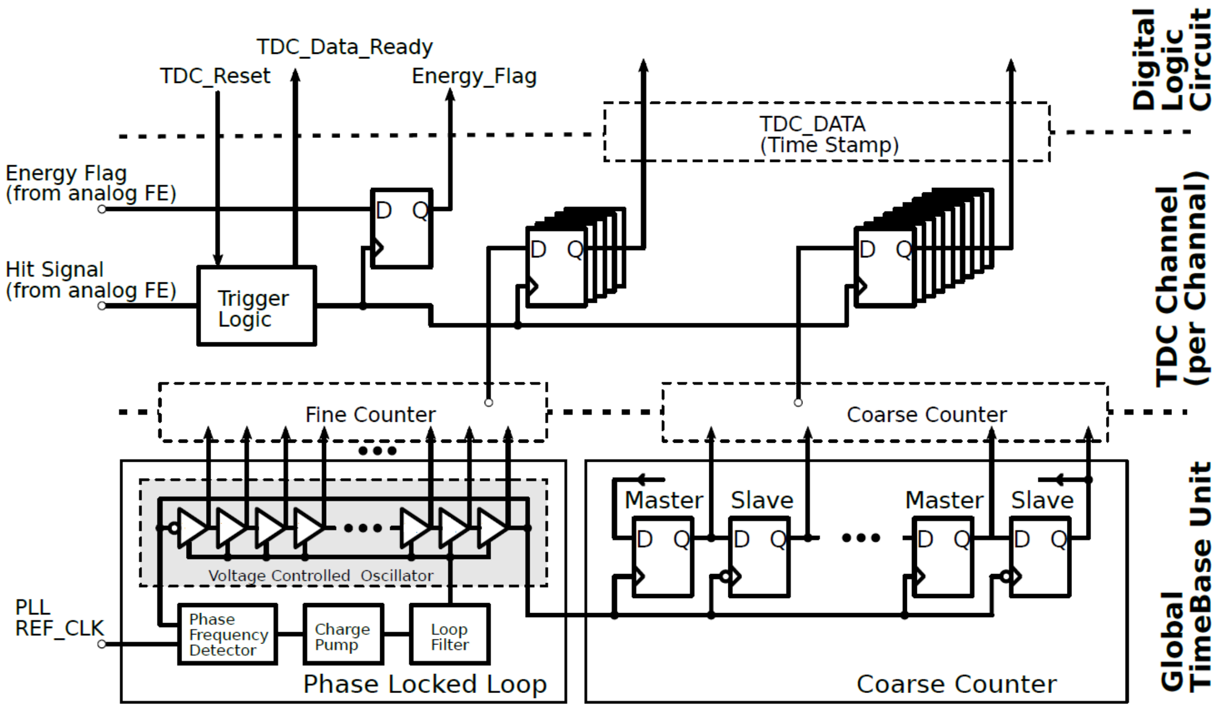


Figure 68: Schematic of the MuTRiG TDC.

	STiCv3.1	MuTRiG
number of channels	64	32
LVDS speed [Mbit/s]	160	1250
8b/10b encoding	yes	yes
event size [bit]		
standard event	48	47
short event	-	26
event rate / chip [MHz]		
standard event	∼2.6	∼20
short event	-	∼38
event rate / channel [kHz]		
standard event	∼40	∼650
short event	-	∼1200
power per channel [mW]	35	35
size [mm x mm]	5x5	5x5
number of PLLs	2	1

Table 9
Comparison of STiCv3.1 and MuTRiG.

the output event rate for a serial data link bit rate of 1.25 Gbps. Results are shown in Figure 69. For the standard event structure configuration of 48 bits, the output event rate is limited to 20.24 MHz (on average 632 kHz/channel) by the bit rate of the serial data link. The maximum event rate for the 27 bits short event configuration is 25 MHz (781 kHz/channel), 1/5th of the system clock frequency(125 MHz).

9.3.2. Jitter Measurement

The jitter found in just the front-end, and in a full channel (front-end, TDC and the digital part of the chip), have been measured. The front-end jitter was measured by charge injection over a 33 pF capacitor. The time difference between the marker signal from the arbitrary waveform generator and the MuTRiG timing trigger signal was then measured using a high bandwidth oscilloscope. The front-end jitter in five different cases as shown in Figure 70. The jitter on a full channel was measured with input charges of 1 pC and an optimised time threshold.

9.3.3. Test-beam Result

In order to verify the functionality and the timing performance of the MuTRiG chip under realistic experimental conditions, the MuTRiG chip was tested with the Mu3e Tile detector prototype in an electron test beam campaign at DESY (Feb. 2018). The setup, shown in Figure 71, was the same as a tile detector submodule: 16 scintillator tiles arranged in a 4 by 4 matrix and read out by SiPM photon detectors. Example time-over-threshold spectra and coincidence time resolutions are given in Figures 72 and 73. Excellent channel-to-channel timing resolutions of <50 ps were been obtained over a large chip configuration parameter range, confirming the performance and functionality of the chip.

10. The Fibre Detector

To suppress all forms of combinatorial background, a very thin detector with good spatial and very good timing resolution, very high efficiency, and high rate capability is required in the central region of the Mu3e

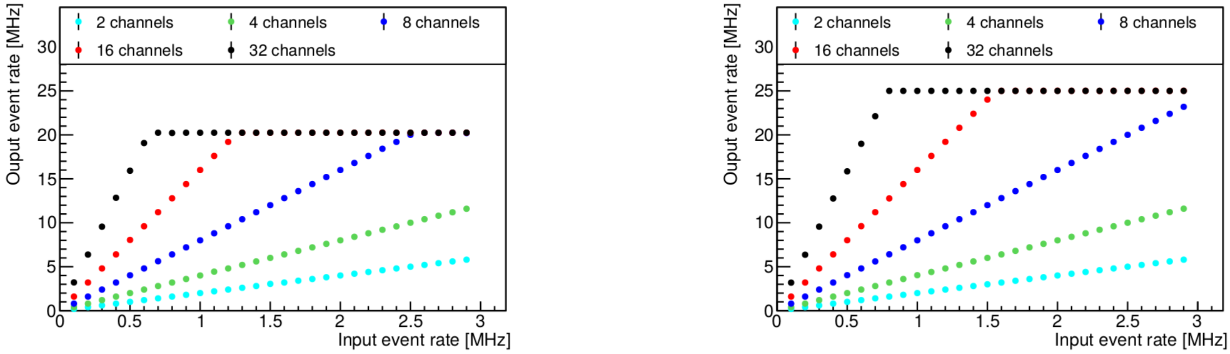


Figure 69: Event rate measurements for the standard output event structure (left) and short output event structure (right).

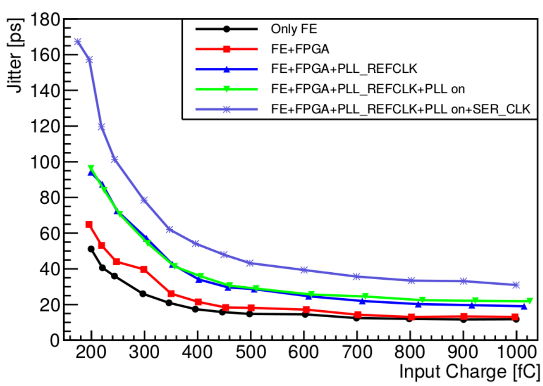


Figure 70: MuTRiG front-end jitter measurement by injecting charge over a 33 pF capacitor.

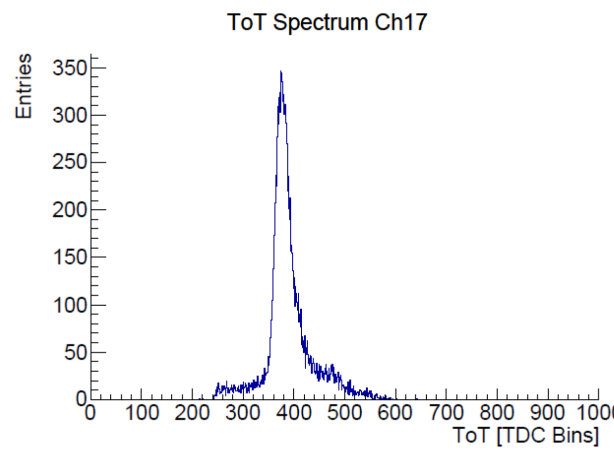


Figure 72: The time-over-threshold (ToT) of minimum-ionising electrons recorded on channel 17.

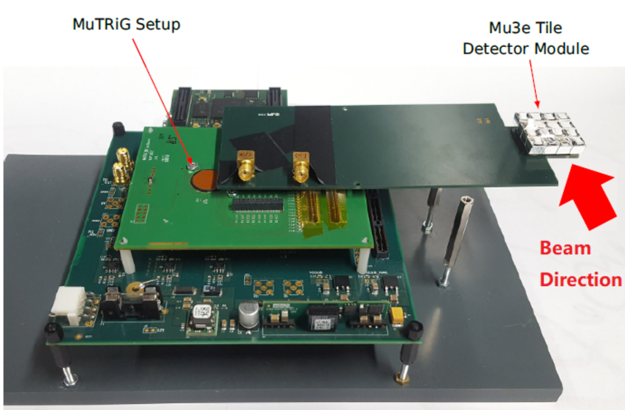


Figure 71: The MuTRiG and Mu3e Tile Detector test beam setup.

apparatus. With this in mind, a thin Scintillating Fibre (SciFi) detector with a time resolution of a few 100 ps, an efficiency in excess of 95 %, a spatial resolution around 100 µm, and a thickness of $X/X_0 < 0.2\,\%$ has been developed. Figure 74 shows the SciFi detector inside the Mu3e experiment. In particular, the space constraints in the central part of the Mu3e experiment impose a very compact design on this sub-detector. In addition to timing, the SciFi detector helps resolve the direction of rotation (i.e., the charge) of the recurling tracks in the central region of the Mu3e detector by time of flight measurements.

The SciFi detector is roughly cylindrical in shape, with a radius of 61 mm and a length of about 300 mm (280 mm in the Mu3e acceptance region). It is composed of 12 SciFi ribbons, each 300 mm long and 32.5 mm wide[21]. The width of the ribbons matches the size of the photo-sensor (see below). The detector is located just inside the outer silicon pixel double-layer.

To a SciFi ribbon consists of three layers of scintillating fibres that are staggered in order to assure continuous coverage and high detection efficiency. Figure 75 shows a full size SciFi ribbon prototype. 250 µm diameter round multiclad fibres from Kuraray, type SCSF-78MJ, were selected. Both ends of the SciFi

[21] This particular value is set by the size of the photo-sensor: the radius of a circle inscribed inside a regular dodecagon with side 32.5 mm, i.e., the size of the photo-sensor, is indeed 61 mm.

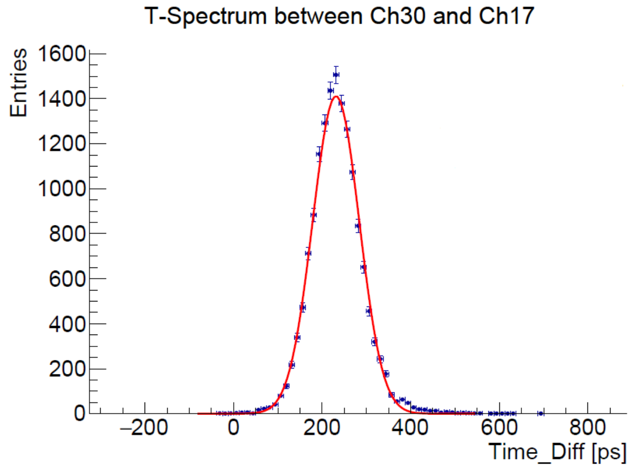


Figure 73: Coincidence timing spectrum between channel 30 and 17.

Figure 75: Full size SciFi ribbon prototype with preliminary holding structure. The SciFi ribbon is formed by staggering three layers of round scintillating fibres.

ribbons are coupled to silicon photomultiplier (SiPM) arrays. After careful evaluation the 128-channel Hamamatsu S13552-HRQ SiPM array, that has also been used in the LHCb experiment, was selected. The SiPM arrays are read out with a dedicated mixed-mode ASIC, the MuTRiG (chapter 9).

By far the largest source of background to the $\mu \rightarrow eee$ search comes from the accidental combination of positron tracks from muon decays, in which two muons decay very closely in space, such that the decay vertices cannot be resolved, with at least one decay positron undergoing Bhabha scattering and ejecting an electron from the target, thus mimicking the topology of a single three-prong decay. Such backgrounds can be efficiently suppressed by timing. Figure 76 shows the background suppression power of the SciFi detector as a function of the detector time resolution. Exploiting the fibre detector alone, with an estimated time resolution of 250 ps and a 90 % overall efficiency, leads to a suppression of the accidental background of $\mathcal{O}(3 \cdot 10^{-2})$. Combining the fibre and tile (see chapter 11) timing detectors the background is further suppressed to $\mathcal{O}(1.4 \cdot 10^{-2})$. For this study, we simulated a Bhabha electron/positron pair plus a Michel positron emerging from the same vertex and distributed in a 50 ns time window, assuming a beam intensity of 10^8 stopping μ^+ per second. The three outgoing tracks are required to pass the selection criteria described in chapter 22.

Figure 77 shows the time difference (time of flight) between two consecutive SciFi detector crossings of recurling track candidates. The correlation between the time difference and the reconstructed trajectory length allows one to determine the sense of rotation of the track (and thus the charge) and/or to reject mis-reconstructed tracks with confused recurling track segments.

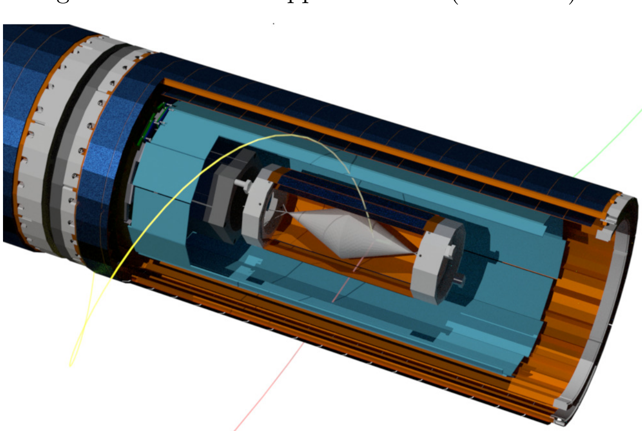


Figure 74: Open view of the central part of the Mu3e detector. The SciFi ribbons are depicted in light blue.

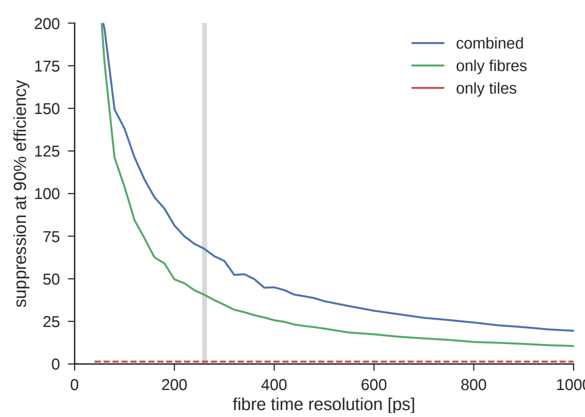


Figure 76: Suppression of Bhabha e^+/e^- pairs plus Michel e^+ accidental background as a function of fibre detector time resolution if only the fibre detector (green) is used or both timing detectors (blue) are used. A time resolution of 60 ps for the tile detector and a working point with a 90 % overall signal efficiency are assumed. The vertical line (in grey) corresponds to a 250 ps time resolution for the fibre detector.

10.1. Scintillating Fibre Ribbons

Three considerations determine the SciFi detector location. Firstly, no material should be placed outside of the fourth silicon pixel layer, where the main momentum measurement is performed. Secondly, it has to be in close proximity to a pixel layer, as the track finding algorithm accounts for multiple Coulomb scattering only in the tracking layers. And thirdly, with larger the radius the SciFi detector occupancy is reduced along with the resulting detector pile-up. The best performance is obtained with the SciFi detector positioned just inside the third silicon pixel layer.

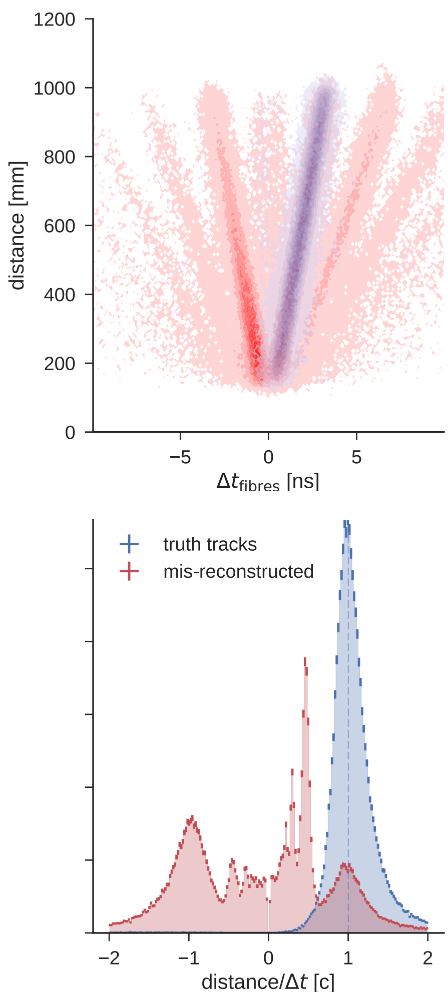


Figure 77: (top) Correlation between the time difference between two consecutive crossings of the fibre detector and the length of the trajectory of a recurling track. The different branches correspond to the combination of different track segments. The correctly reconstructed tracks with the correct charge assignment are shown in blue, while tracks with wrong charge assignment and/or mis-reconstructed tracks are shown in red. (bottom) Speed v = track length/$\Delta t \times c$ of recurling tracks. The different branches in the top plot correspond to the peaks in the bottom spectrum. Track candidates with $\Delta t < 0$ ($v < 0$) have a wrong charge assignment.

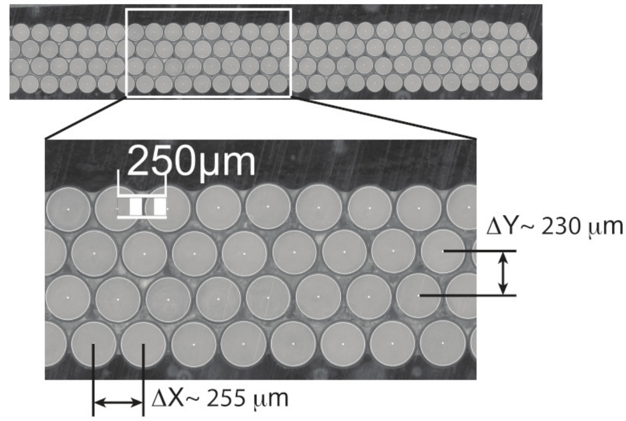


Figure 78: Front view of a SciFi ribbon prototype. A good uniformity can be achieved by this ribbon construction technique. Note that the photograph shows a four-layer SciFi ribbon, while in Mu3e three-layer ribbons are used.

characteristic	value
cross-section	round
emission peak [nm]	450
decay time [ns]	2.8
attenuation length [m]	>4.0
light yield [ph/MeV]	n/a (*high*)
trapping efficiency [%]	5.4
cladding thickness [%]	3 / 3
core	Polystyrene (PS)
inner cladding	Acrylic (PMMA)
outer cladding	Fluor-acrylic (FP)
refractive index	1.59/1.49/1.42
density [g/cm^3]	1.05/1.19/1.43

Table 10
Properties of the 250 µm diameter round multi-clad Kuraray SCSF-78MJ scintillating fibres as quoted by the manufacturer.

Each SciFi ribbon is formed by staggering three layers of 250 µm diameter round fibres (there are 128 fibres in a layer) with a length of 300 mm. After careful evaluation Polytec EP 601-Black epoxy was selected (this is a two component, low viscosity, black-coloured adhesive) for the assembly of the final SciFi detector. Figure 78 shows the cross-section of a fibre ribbon prototype. As can be observed, the fibres in a layer are separated by $\sim$ 255 µm centre to centre with a very good uniformity and the separation between the layers is $\sim$ 230 µm, which gives an overall thickness of approximately 700 µm for a three-layer ribbon.

10.1.1. Scintillating Fibres

The constraints on the material budget, the occupancy, and position resolution require the use of the thinnest available scintillating fibres. In extensive measurement campaigns, a detailed comparison was undertaken of different types of 250 µm diameter round scintillating fibres produced by Kuraray (SCSF-78, SCSF-

81 and NOL-11) and Saint-Gobain (BCF-12), as well as square cross-section fibres by Saint-Gobain (BCF-12). Scintillating fibre ribbon prototypes coupled to SiPM arrays have been tested in test beams at the CERN PS (T9 beamline) and PSI (πM1 beamline) and with ^{90}Sr sources. The detailed results of these studies are reported in [38, 39, 40, 41]. Based on their performance with respect to light yield and time resolution, round double-clad SCSF-78MJ fibres from Kuraray were chosen. Table 10 summarizes the characteristics of these fibres. Novel NOL fibres, based on Nanostructured Organosilicon Luminophores, give the best performance, but will only become commercially available in the years to come and will be considered for future SciFi detector upgrades.

10.1.2. Number of SciFi Layers

A critical point of optimization is the number of staggered fibre layers. More layers lead to an improved timing resolution and a higher detection efficiency but reduces the momentum resolution of the pixel tracker due to multiple Coulomb scattering. Since the particles cross the SciFi ribbons at an angle, more layers lead also to a larger cluster size (i.e., the number of channels in the SiPM array excited by the scintillating light) and therefore to a larger occupancy.

Using the physical characteristics of the SciFi ribbons extensive simulation studies were performed on the impact of this sub-detector on the momentum resolution, efficiency and track reconstruction (details on the complete detector simulation, reconstruction algorithm and event selection can be found in chapters 18, 19 and 22).

The amount of multiple Coulomb scattering generated by the fibre detector is shown in Figure 79. Note that a ribbon of three layers of 250 µm round fibres corresponds to $X/X_0 \approx 0.2\,\%$. Multiple Coulomb scattering affects the momentum resolution (Figure 80)

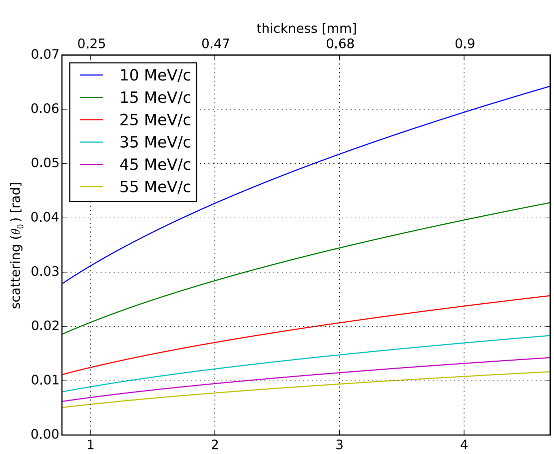


Figure 79: Multiple Scattering θ_0 depending on electron/positron momentum and fibre ribbon thickness.

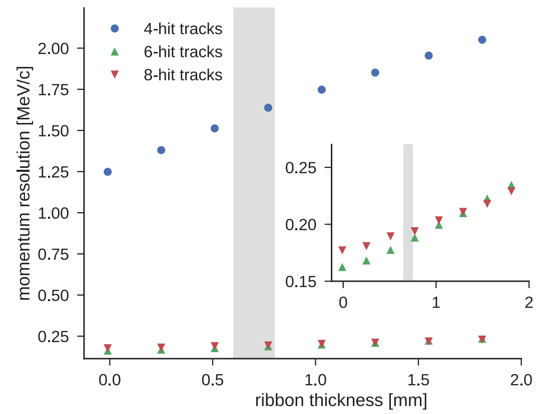


Figure 80: Momentum resolution for short (outgoing only) and long (outgoing and recurling) tracks as a function of fibre ribbon thickness using simulated Michel decays. The highlighted region corresponds to a three-layer SciFi ribbon thickness of ~ 0.7 mm. The momentum resolution of long (6- and 8-hit) tracks is improved over short (4-hit) tracks due to recurling (more measured points).

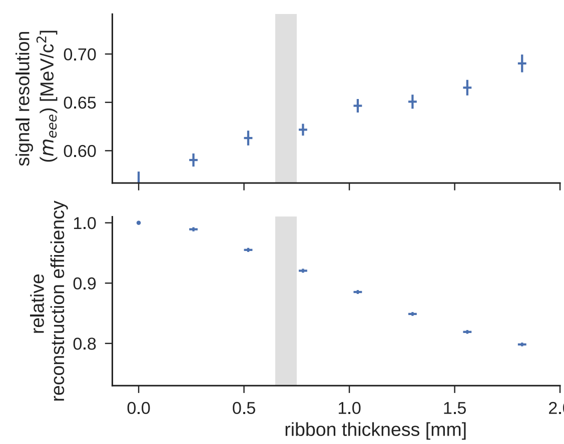


Figure 81: (top) Signal resolution in terms of the invariant mass of the three tracks of a candidate m_{eee} decay and (bottom) loss in reconstruction efficiency as a function of the fibre ribbon thickness. The highlighted region corresponds to a three-layer SciFi ribbon thickness of ~ 0.7 mm.

and thus the $\mu \to eee$ signal invariant mass resolution and reduces the overall reconstruction efficiency (Figure 81). More fibre layers also lead to a larger occupancy in the SciFi detector.

As a compromise between these constraints, ribbons consisting of three staggered layers of 250 µm diameter round fibres are chosen. With a thinner detector it would be challenging to fulfill the efficiency requirements and the time resolution would not be sufficient to effectively reject accidental backgrounds, reliably determine the sense of rotation of tracks and reject misreconstructed track candidates.

10.2. Silicon Photomultiplier Arrays

The light produced in the scintillating fibres is detected in SiPM arrays at both fibre ends. Acquiring the signals on both sides increases the time resolution (two time measurements instead of one), helps to distinguish between noise and signal and increases the detection efficiency of the whole system (because of the noise rejection). Moreover, by taking the mean time of the two time measurements, the timing measurements is made independent of the hit position (assuming that light propagates at the same speed to both fibre ends) and thus no position correction is necessary.

The Mu3e fibre detector is read out with Hamamatsu S13552-HRQ SiPM arrays, with a high quenching resistance. The segmentation of the sensor is obtained by arranging the individual SiPM pixels into independent readout columns (channels). Each channel consists of 104 pixels, each measuring 57.5 µm × 62.5 µm, arranged in a 4×26 grid. The sensitive area of one channel is therefore 230 µm × 1625 µm. The pixels are separated by trenches of the fifth generation Hamamatsu low-crosstalk development (LCT5). A 20 µm gap separates the array's columns, resulting in a 250 µm pitch. Each sensor comprises 64 such channels, which share a common cathode. Two sensors, separated by a gap of 220 µm, form the 128 channel device shown in Figure 82. The overall current consumption of one array is expected to be below 1 mA even for heavily irradiated sensors. The sensors are delivered wire-bonded on a PCB with solder pads on the backside. The sensors are covered with a 105 µm thick protective layer of epoxy resin. Table 11 summaries the most important features of the sensor.

characteristic	value
breakdown voltage	52.5 V
variation per sensor	±250 mV
variation between sensors	±500 mV
temperature coefficient	53.7 mV/K
gain	$3.8 \cdot 10^6$
direct crosstalk	3 %
delayed crosstalk	2.5 %
after-pulse	0 %
peak PDE	48 %
max PDE wavelength	450 nm
mean quench resistance R_Q	490 kΩ at 25 °C
recovery time τ_{recovery}	(68.9 ± 2.1) ns
short component τ_{short}	< 1 ns
long component τ_{long}	(50.1 ± 4.1) ns

Table 11
SiPM array ((model S13552-HRQ) characteristics at $\Delta V = V_{\mathrm{op}} - V_{\mathrm{breakdown}} = 3.5$ V and $T = 25$ °C from [42].

This sensor was developed for the LHCb experiment and matches the requirements of the Mu3e fibre detector. The photon detection efficiency (PDE[22]) of up to 50 %, single photon detection capabilities and very fast intrinsic time response (single photon jitter of approximately 200 ps) are the key features for the use in the Mu3e fibre detector. The SiPM arrays are read out with a dedicated mixed-mode ASIC, the MuTRiG (see chapter 9). The high gain ($> 10^6$) allows for the use of the MuTRiG without any pre-amplification. Typical dark-rates are around 100 kHz at room temperature per SiPM array channel for unirradiated sensors. In contrast to LHCb, where the SiPM arrays are operated around −40 °C, the Mu3e sensors are being operated at a temperature close to 0 °C, but in a less intense radiation field. The moderate cooling of the detector is required to further reduce the dark count rate and mitigate the radiation damage effects.

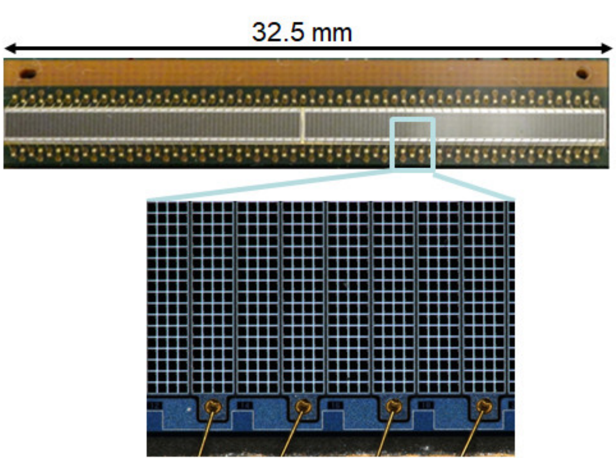


Figure 82: Picture of a Hamamatsu S13552-HRQ SiPM column array including a close view showing the pixel structure of the sensor.

Figure 83 shows the I-V curves for one SiPM array for each channel of the sensor. All breakdown voltages are comprised within ±0.25 V of the central value of 52.5 V. The best performance is obtained for an operational voltage (V_{op}) 3.5 V above the breakdown voltage ($V_{\mathrm{breakdown}}$), but the sensor can also be operated at higher voltages for an increased gain. Since all channels share a common cathode, the sensor is usually operated at a common voltage for all channels. The performance of the photo-detector can be further improved by adjusting V_{op} individually for each channel. The MuTRiG readout ASIC allows for the fine tuning of the bias voltage around a common value for each individual channel of the sensor.

The fibre ribbons are coupled directly to the surface of the SiPMs on both sides. Figure 84 shows a possible mapping of the SciFi ribbon on the SiPM array. As can be seen, no one-to-one matching is possible between the fibres and the SiPM columns because of the

[22] With contributions from quantum efficiency and geometrical fill factors.

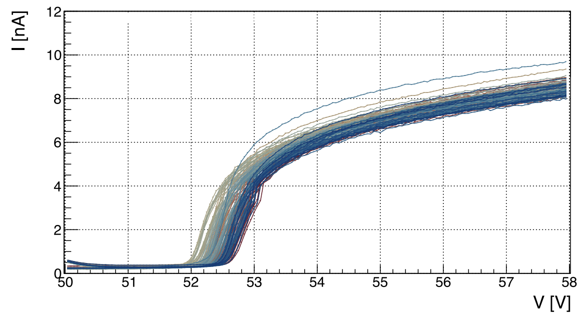


Figure 83: I-V curves for each channel of the SiPM array. All breakdown voltages are comprised within $\pm 0.25\,\mathrm{V}$ of the central value of $52.5\,\mathrm{V}$.

staggering of the fibres. To ease detector assembly and maintainability, the coupling is realised by only mechanical pressure without the use of optical interfaces.

10.3. SciFi Readout Electronics

The Mu3e scintillating fibre detector requires the digitization of the crossing time information at a single photon level. That leads to very high rates per SiPM channel coming from the particles crossing the SciFi ribbons ($\sim 200\,\mathrm{kHz}$ signal rate) and the dark noise ($\sim 1\,\mathrm{MHz}$ for irradiated sensors). The latter is reduced by clustering during the real-time processing of the data.

For the readout of the 3072 SiPM channels we use the mixed-mode MuTRiG ASIC with 50 ps TDC time binning (see chapter 9 for a detailed description of the ASIC). Each ASIC comprises 32 fully differential input analogue channels, therefore four MuTRiG ASICs are required for the readout of one SiPM array. Although the ASIC has a fully differential input, single ended signals are used, because the SiPM array channels share a common cathode. When operated with the SiPM arrays, the signal is compared to two thresholds: a low one for timing (a time stamp is generated) and a high one for hit selection (single flag).

The analog signals from each SiPM array (128 channels) are digitised by one SciFi module board (SMB) hosting four MuTRiG ASICs. Figure 85 shows the first version of the SMB. The space limitations in the Mu3e setup require a very compact design of the board. The ASICs are wire bonded directly to the board. The final version of the board will measure $26\,\mathrm{mm} \times 45\,\mathrm{mm}$ and is currently under development. The electrical connection between the SiPM sensors and the readout electronics is realised through flex-print circuits. A 128-channel SiPM array is soldered to a support PCB with an embedded flex-print, which continues to a second PCB hosting the MuTRiG ASICs. In addition to the MuTRiG ASICs, the SMB hosts the clock and reset distribution circuits, components for the control of the MuTRiG, LDOs for power distribution, and temperature probes. In total 24 such SMBs are needed, one per SiPM array. Finally all SMBs are connected to front-end FPGA boards (see section 17.2) via micro twisted-pair cables.

10.3.1. Power requirements

The power requirements of the MuTRiG ASICs are given in chapter 9. The powering of one SciFi front-end board requires 2 V at 2.5 A, 3.5 V at 0.1 A, and a bias line (around 55 to 57 V) for the SiPM array. Each SciFi front-end board generates around 5 W of thermal output, which has to be cooled.

10.4. SciFi Detector Performance

Figure 86 shows the light yield in a *cluster* excited by a minimum-ionizing particle crossing a three-layer SCSF-78MJ fibre ribbon prepared with clear epoxy. A cluster is defined as the sum of all consecutive SiPM channels with an amplitude larger than a specific threshold (in this case 0.5 photo-electrons) and a cluster mul-

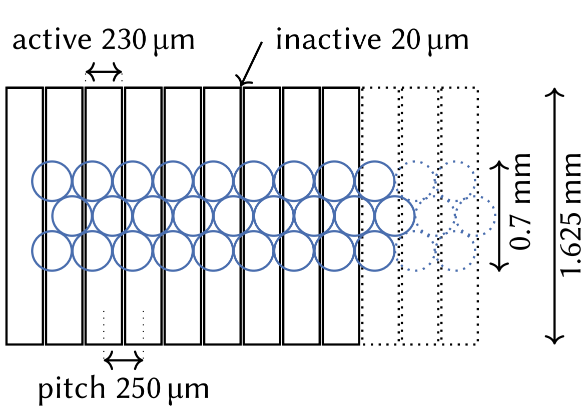


Figure 84: Mapping of the SciFi ribbon on the SiPM array. No one to one matching is possible between the fibres and the SiPM columns.

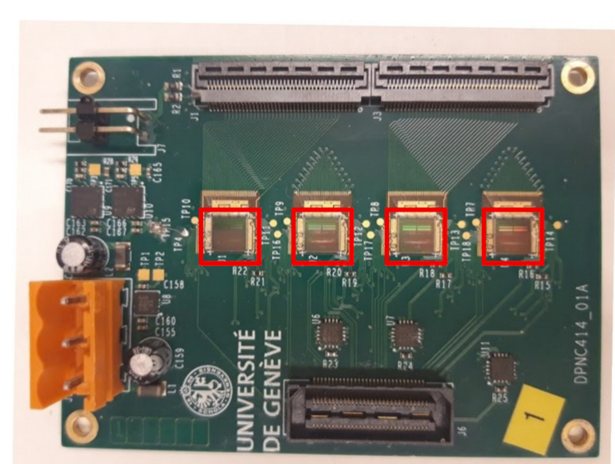


Figure 85: First version of the SciFi front-end board hosting 4 MuTRiG ASICs (outlined in the red boxes) wired bonded directly on the board.

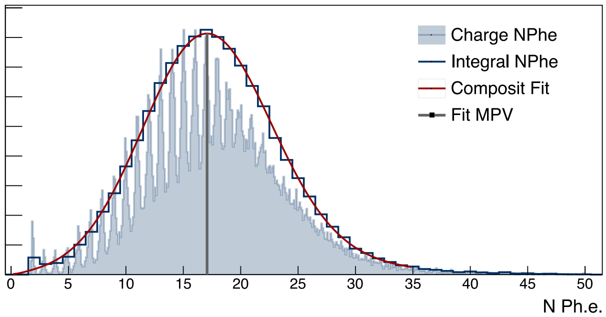


Figure 86: Light yield of a cluster (see text) for a m.i.p. crossing a three-layer SCSF-78MJ fibre ribbon prepared with clear epoxy. The integral NPhe is obtained by integrating the charge in a region of ± 0.5 ph.e. around each peak (integer). A convolution of a Gaussian and of a Landau is used to fit the data and the MPV of the spectrum is marked with the vertical line.

tiplicity of at least two adjacent SiPM channels above the same threshold. The number of photo-electrons (ph.e.) is defined by the charge sum of all channels in a cluster at one side of the SciFi ribbon matched to a crossing track. The light yield is measured with respect to the centre of the fibre ribbon (i.e., 150 mm from the edge). A convolution of a Gaussian and of a Landau distribution is used to fit the data. The fit provides also the most probable value (MPV) for the number of detected ph.e., which is of about 17 for this configuration. This ph.e. spectrum, however, is not accessible in the experiment since the MuTRiG provides only the timing information and no charge information. Test-beam data were recorded using a fast pre-amplifier and readout digitizing electronics based on the DRS4 ASIC. The recorded waveforms were then processed using timing algorithms close to the MuTRiG functioning (i.e., 0.5 ph.e. low threshold leading edge discriminator).

The cluster size distribution for the same SciFi ribbon is shown in Figure 87. Typical cluster sizes are around 3.5 for a threshold of 0.5 ph.e., for a particle crossing the ribbon at 0° (i.e., perpendicularly to the ribbon). The cluster size can be reduced by increasing the detection threshold to e.g., 1.5 ph.e. or higher. The figure shows also cluster sizes for particles crossing the ribbon at an angle of 30°, which is close to the mean crossing angle in Mu3e of 25°[23]. A larger crossing angle increases the average cluster size.

The detection efficiency of the SciFi detector depends on the applied thresholds, minimal cluster multiplicity and the requirement of time matched clusters at both SciFi ribbon ends. For the selected working point, which requires a threshold of 0.5 ph.e., with a minimal cluster multiplicity of two and a 5 σ timing cut on the matched clusters, where σ is the intrinsic time resolution of the SciFi detector, the detection efficiency is around 95 %. Without the timing cut, the detection

[23] Due to spiralling tracks in the magnetic field

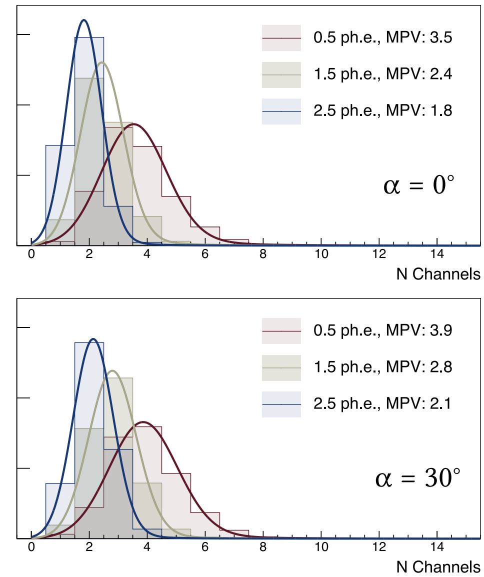


Figure 87: Cluster size for a particle crossing the ribbon at two different angles and different thresholds. Electrons from a radioactive ^{90}Sr source are used for this measurement. An angle of $\alpha = 0°$ describes a perpendicular crossing.

efficiency increases close to 100 %. It should be noted that the cluster matching and the timing cut can only be applied in the offline analysis of the SciFi data and can be tuned to optimize the detection efficiency.

Finally, an example of the timing performance of the SciFi detector is shown in Figure 88. This measurement has been performed using the MuTRiG evaluation board, shown in Figure 71. The measurement has been performed using a four-layer SciFi ribbon with a ^{90}Sr source requiring a minimal cluster multiplicity of two neighbouring channels with an amplitude of at least 0.5 ph.e. Similar results have also been obtained with the analogue electronics (DRS4-based DAQ) men-

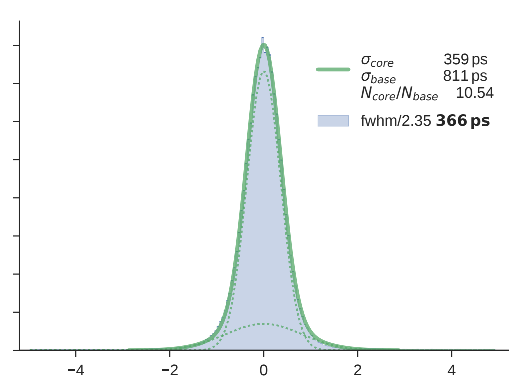


Figure 88: Time resolution of a 4 layer SCSF-78MJ SciFi ribbon extracted from clusters with at least 2 active columns. No channel by channel time offset correction has been applied.

tioned above and particle beams [40]. The spread of the time difference distribution from the two ribbon sides $\sigma(t_{\text{left}} - t_{\text{right}})$ corresponds to twice the intrinsic detector resolution (mean time). For example, the FWHM/2.35 of the distribution obtained in this measurement is 366 ps implying a resolution on the mean time around 200 ps. For a three-layer ribbon as used in Mu3e, the time resolution is slightly worse, at around 250 ps.

10.5. SciFi Detector Mechanics

Figure 89 shows the overall structure of the SciFi detector. The detector is composed of 12 SciFi ribbons, 300 mm long and 32.5 mm wide. The ribbons are staggered longitudinally by about 10 mm (Figure 91) in order to minimise dead spaces between the ribbons and to provide sufficient space for the spring loading of the ribbons. To avoid sagging and to compensate for the thermal expansion the ribbons are spring loaded on one side of the structure (6 ribbons on one side and the other 6 on the other side).

A detailed study to determine the effects of the thermal expansion and sagging has been performed. A thermal expansion coefficient for the 300 mm long SciFi ribbon of $(65 \pm 16) \cdot 10^{-6}$/K has been measured. Therefore, for a 50 °C thermal excursion, an elongation of the ribbons of around 1 mm is expected. This elongation effect can be compensated by spring-loading the ribbons as mentioned above. Figure 90 shows the sag of a three-layer 300 mm long and 32.5 mm wide SciFi ribbon as a function of the temperature for different values of the applied tension. Figure 90 also shows that a tension of 8 N is required to prevent sagging over the whole temperature interval and to guarantee the correct positioning of the detector.

To ease the sub-detector installation, the SciFi ribbons are assembled in modules. Each module consists of two SciFi ribbons, as shown in Figure 91.

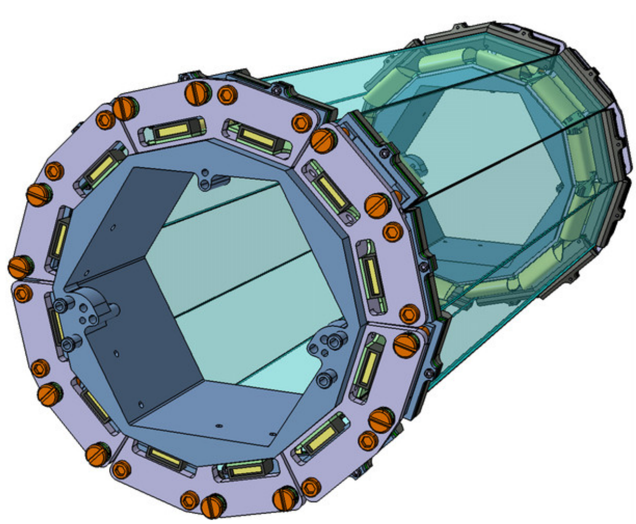


Figure 89: Overall structure of the scintillating fibre detector.

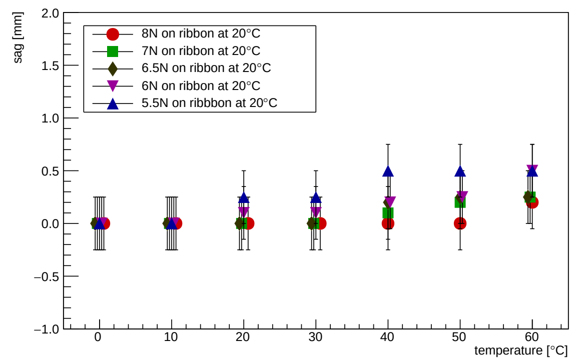


Figure 90: Sag of the SciFi ribbon as a function of the temperature for different values of the applied spring tension.

The SciFi ribbons are coupled to the SiPM arrays by simple mechanical pressure (no grease or other optical interface). Each SiPM sensor is connected to a front-end digitizing board via a flex-print circuit. Figure 92 shows an expanded view of the assembly structure: the SciFi ribbons are attached to the SiPM arrays, which in turn are supported by stiffeners fixed to L-shaped supports, where the assembly is also spring loaded. The same L-shaped supports are also used to mount the SciFi front-end boards.

The L-shaped supports are fixed to a hollow dodecagonal prism as shown in Figure 93, 45 mm tall with an outer diameter of 100 mm, which also provides the necessary cooling for the front-end electronics. Two such cooling structures are attached to the beam pipe on each side of the Mu3e detector and connected to the pipes of the Mu3e cooling system. This cooling structure is created by 3D printing in aluminium with inner piping for the circulation of the coolant. Each MuTRiG ASIC generates about 1 W of thermal output, therefore around 50 W has to be dissipated on each side of the SciFi detector. Since the SiPM arrays are in thermal contact with the L-shaped supports, they are cooled by the same cooling structure. The goal is to cool the SiPM arrays down to 0 °C.

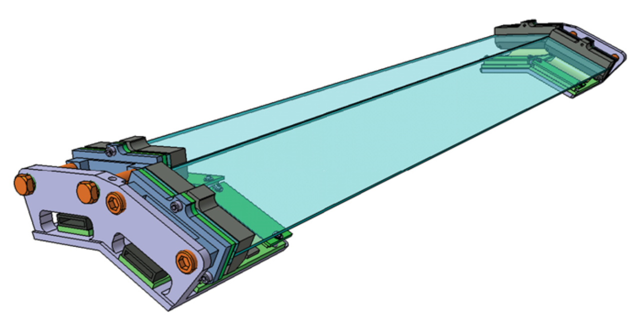


Figure 91: A fibre module consists of two SciFi ribbons with the associated support structure. The ribbons are staggered longitudinally to minimise dead spaces between the ribbons and are spring loaded alternately on opposite sides of the structure.

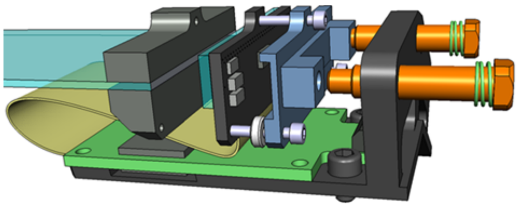


Figure 92: Expanded view of the SciFi support structure, showing all the elements of the detector: SciFi ribbon, SiPM sensor, SciFi front end board and the L-shaped support structure.

11. Tile Detector

The tile detector aims at providing the most precise timing information of the particle tracks possible. As it is located at the very end of recurling particle trajectories, there are no constraints on the amount of detector material; the placement inside the recurl pixel detectors however implies very tight spatial constraints. The detector consists of plastic scintillator segmented into small tiles. Each tile is read out with a silicon photomultiplier (SiPM) directly attached to the scintillator. The main goal of the tile detector is to achieve a time resolution of better than 100 ps and a detection efficiency close to 100% in order to efficiently identify coincident signals of electron triplets and suppress accidental background.

11.1. Detector Design

The tile detector is subdivided into two identical stations – one in each recurl station. Each tile detector segment has the shape of a hollow cylinder enclosing the beam-pipe. The length of a segment is 34.2 cm along the beam direction (z direction) including the endrings, while the outer radius is 6.4 cm, which is limited by the surrounding pixel sensor layers. The detector in each recurl station is segmented into 52 tiles in z direction and 56 tiles along the azimuthal angle (ϕ direction). This is the highest feasible channel density, considering the space requirements for the readout electronics. The high granularity is essential in order to achieve a low occupancy as well as a high time resolution.

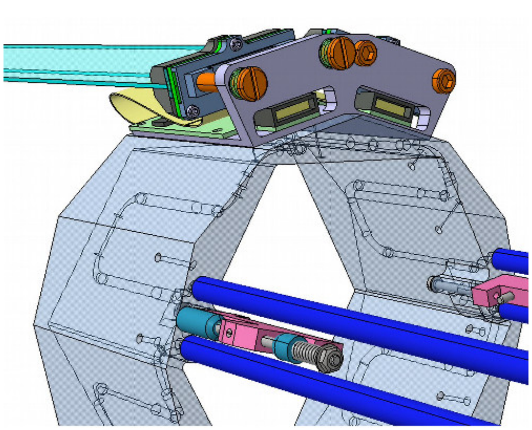


Figure 93: Support structure of the SciFi detectors, which serves also as cold mass for the dissipation of the heat generated by the SciFi front-end electronics and the cooling of the SiPM arrays.

The technical design of the tile detector is based on a modular concept, i.e. the detector is composed of small independent detector units. The base unit of the tile detector, referred to as *sub-module*, is shown in Figure 96a. It consists of 32 channels arranged in two 4×4 arrays. The tiles are made out of Eljen technology EJ-228 plastic scintillator and have a size of $6.3 \times 6.2 \times 5.0\,\mathrm{mm}^3$, see Figure 94. The edges of the two outer rows of an array are bevelled by 25.7°, which allows for seven base units to be arranged approximately in a circle.

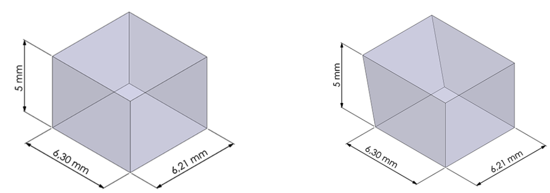


Figure 94: Tile scintillator geometry: (left) central tile, (right) edge tile.

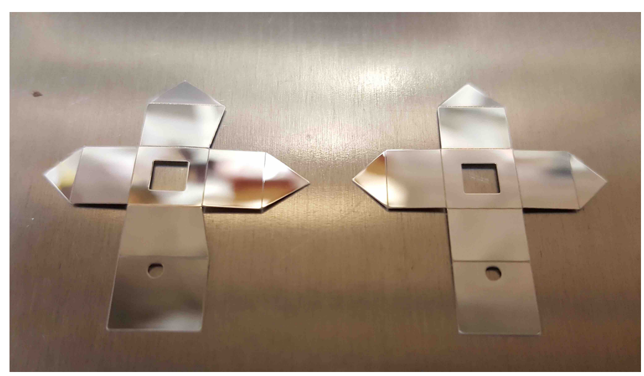
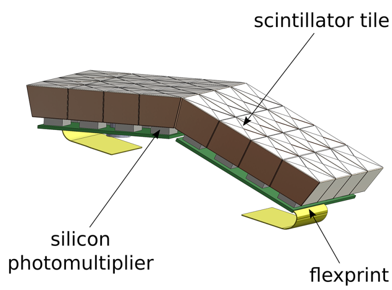


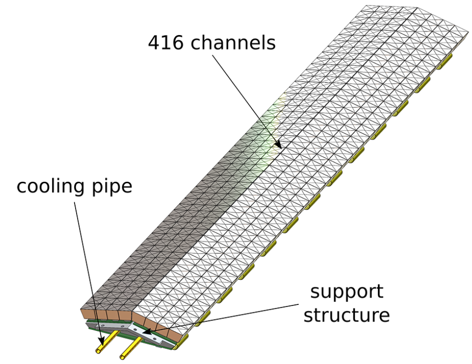
Figure 95: Individual ESR reflective foils for two types of scintillator tiles:(left) edge tile, (right) central tile.

The individual tiles are wrapped with Enhanced Specular Reflector (ESR) foil. In order to increase the light yield and optically isolate the channel, the foil is designed to cover the entire tile except for an opening window of the size of the SiPM surface, as can be seen in Figure 95. Every tile is read out by a $3 \times 3\,\mathrm{mm}^2$ SiPM with 3600 pixels, which is glued to the bottom $6.3 \times 6.2\,\mathrm{mm}^2$ side of the tile. The SiPMs are soldered to a printed circuit board (PCB), which is connected via a flexible PCB (flexprint) to one of the ASICs on the readout board, the Tile Module Board (TMB).

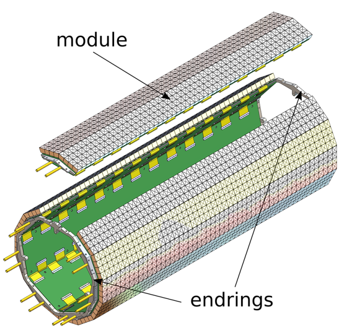
A tile *module* is comprised of 13 sub-modules, and contains a total of 416 channels. A CAD rendering of such a module is shown in Figure 96b. The sub-modules



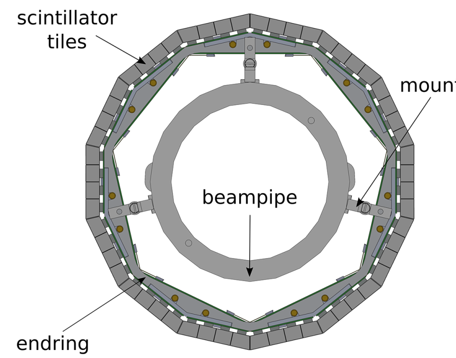
(a) The tile detector base unit consisting of 32 scintillator + SiPM channels. The sensors are mounted on a flex-rigid PCB. CAD rendering.



(b) Module (416 channels) of the tile detector consisting of 13 base units, which are mounted on a support structure. A copper pipe for cooling liquid is placed inside the support structure to cool the readout chips and SiPMs. CAD rendering.



(c) Full tile detector (CAD rendering, exploded view).



(d) Full tile detector (front view). The detector modules are mounted on two endrings connected to the beampipe. CAD rendering.

Figure 96: CAD rendered views of the tile detector.

are mounted on a water-cooled aluminium support structure and are read out by 13 MuTRiG ASICs assembled on one TMB, which collects the analog signals of the SiPMs and forwards the digitised signals to the front-end FPGAs. The subsequent data transmission is discussed in chapter 17. The heat of the readout chips is dissipated via liquid cooling through a copper tube, with an outer diameter of 2.5 mm and an inner diameter of 2.0 mm, which is placed in a U-shaped groove on the bottom side of the support structure.

Figure 96c shows an exploded view of a full tile detector recurl station, which consists of seven modules. The modules are assembled on two endrings, which in turn are mounted on the beam pipe, as shown in Figure 96d.

Based on previous studies [43], the best timing resolution is achieved with the plastic scintillator BC418 (equivalent to EJ-228), which has both a high light yield and a fast response time, and therefore is chosen as the baseline material for the tile detector. This scintillator has a nominal light output of about 10 200 photons per MeV, a rise time of 0.5 ns and a decay time of 1.4 ns. The emission spectrum of the scintillator peaks at a wavelength of 391 nm, which roughly matches the maximum spectral sensitivity of the SiPM. This allows the direct read-out of the scintillation light

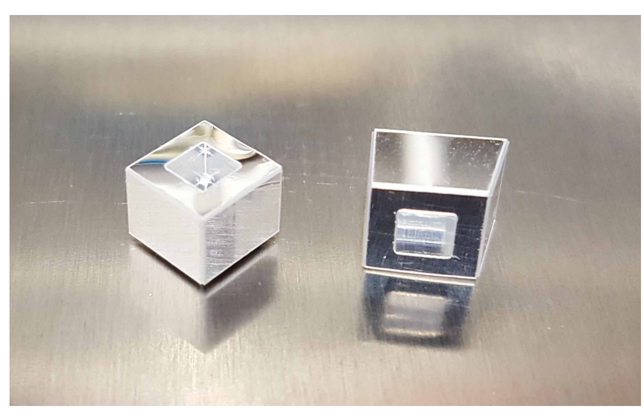


Figure 97: Individual tiles wrapped with ESR reflective foil.

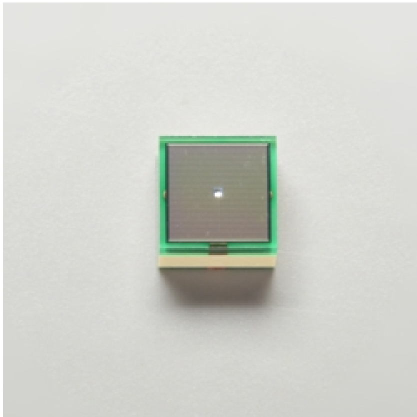


Figure 98: Hamamatsu MPPC S13360-3050VE.

without the need of an additional wavelength shifter.

Different SiPM types have been compared in simulation studies in order to find the best suited device for the tile detector. Based on the simulation studies, a $3 \times 3\,\mathrm{mm}^2$ SiPM with $50\,\mu\mathrm{m}$ pixel size is chosen as the baseline photo-sensor. A respective SiPM from Hamamatsu (MPPC S13360-3050VE, see Figure 98) has been successfully tested in the tile detector technical prototype (see section 11.5).

11.2. SiPM Radiation Hardness

Ionising radiation can have a large impact on the SiPM characteristics and performance. The most prominent effect caused by irradiation is a strong increase in the SiPM dark-rate. Furthermore, there are several studies (e.g. [44, 45]), which have observed a slight decrease in the detection efficiency after exposure of the SiPM to radiation. A possible explanation for this effect is the progressively larger amount of pixels in a permanent off-state [45]. Both an increasing dark-rate and a reduced signal amplitude directly influence the time resolution of the sensor. The exact amount of signal degradation caused by radiation depends on the particle energy and type, as well as the specific SiPM device.

During the data taking period of phase I of the Mu3e experiment, the SiPMs will be exposed to a total radiation dose of about $10^{10}\,\mathrm{e}^+/\mathrm{mm}^2$. So far, no conclusive experimental data of the SiPM signal degradation is available for the given irradiation dose, particle type and energy. First studies of the radiation damage in SiPMs using a ^{90}Sr source indicate that the degradation of time resolution during the Mu3e phase I is of the order of a few percent. However, more detailed studies are required in order to precisely predict the SiPM performance over phase I runtime and to obtain a comprehensive picture of the radiation effects. These studies will be performed in parallel for the tile and fibre SiPMs.

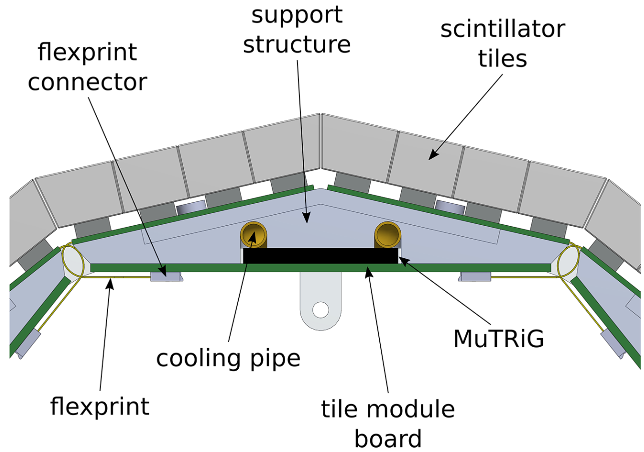


Figure 99: Tile detector with readout electronics. The tile module is divided into a PCB hosting the MuTRiG chips and 2×13 PCBs hosting the SiPMs. The SiPM boards are connected to the MuTRiG board via flex cables. The tile readout board is placed on the cooling structure, connecting the tile sub-modules to the front-end FPGA readout board.

11.3. Tile Readout

The tile detector will use the same MuTRiG ASIC as the fibre detector, see chapter 9 for details. The output signals of 32 tile SiPMs are connected via a flexible printed circuit board to a MuTRiG chip. The arrangement of the SiPMs and the readout electronics around the cooling structure is shown in Figure 99. The MuTRiG will be operated in two-threshold mode (see Figure 66), allowing for time-walk correction. The data are then forwarded via the TMB, mounted on the detector module, to the FPGA front-end boards similar to those planned for the pixel detector, see chapter 17.

11.4. Assembly Tools and Productions steps

As a first step, the SiPMs are sorted by breakdown voltage and preselected in groups of 32 SiPMs with a spread of the breakdown voltage smaller than 100 mV. This will allow the operation of each sub-module with the same operating High Voltage (HV).

The tiles are manufactured in the Kirchhoff-Institute for Physics in Heidelberg. The scintillator material is mounted on a vacuum plate, where the full matrix is milled from the top, only leaving a 0.5 mm base. The plate is flipped by 180 degrees on to an ice-vice, which freezes the matrix to mill off the base, as sketched in Figure 100. Using this method, a relatively fast production rate with very high accuracy of several micrometers is achieved.

After cutting the tiles to the required shape, the tiles' length and width are measured using a digital micrometer before wrapping them with the reflective foil. In order to wrap the tiles, a semi-automatic tool was designed that allows for an easy wrapping of such small

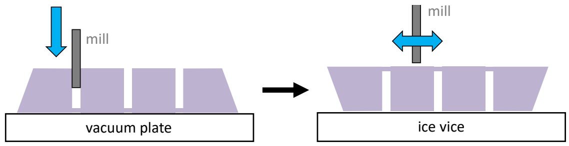


Figure 100: Tile milling procedure: (left) milling the matrix shape on a vacuum plate, (right) flip the matrix, freeze on the ice-vice, and mill from top.

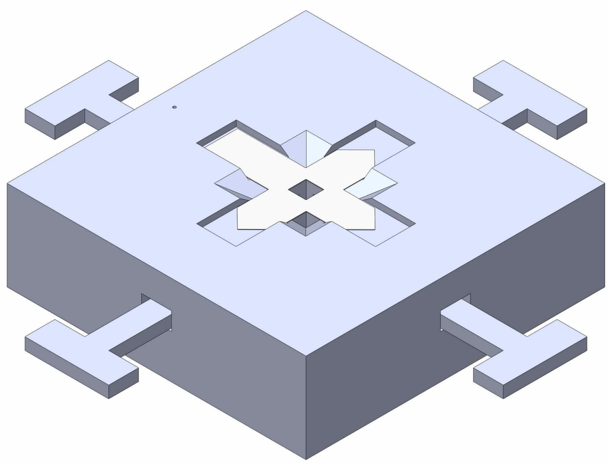


Figure 101: Sketch of the tile wrapping tool with a foil on top.

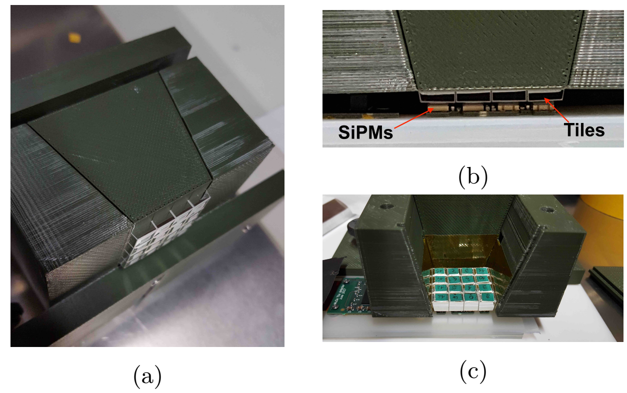


Figure 102: Gluing tool design to glue a 16 channel matrix on one side of tile matrix board: (a) the tools with 16 tiles pressed before gluing, (b) tiles pressed on SiPMs during the gluing stage, (c) glued tile matrix after curing.

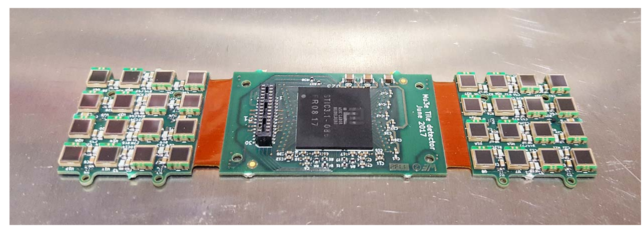


Figure 103: Tile matrix board assembled with SiPMs and a BGA packaged STiC 3.1 ASIC.

tile sizes. A sketch of the wrapping tool is presented in Figure 101. The foil is placed into a dedicated groove on top of the tool; then the tile is placed onto the foil. By pushing the tile down into a customised funnel, the foil side walls are folded around the tile. Using the side rods of the tool, the wrapping is folded like an envelope and a small sticker is placed on top to close it. The resulting wrapped tiles are shown in Figure 97.

In the following step, the tiles must be glued to the SiPMs. This is done on matrix level in order to avoid tolerance issues. A gluing tool was designed with the emphasis of allowing a small degree of freedom with respect to the height of the individual tiles in order to compensate different SiPM heights due to soldering paste and tolerances of the SiPM manufacturing. The scintillator tiles are manually arranged inside the tool and are pressed from the back side and the top such that half of the tiles' height is outside of the tool as shown in Figure 102a. The matrix board is mounted on a pedestal and the glue is dispensed onto the SiPMs. At this stage, the tool is pressed onto the SiPMs as shown in Figure 102b, where the x-y position is set using alignment pins. After a curing time of 24 hours, the outer wall of the gluing tool is taken out (see Figure 102c) and the gluing tool can be removed.

11.5. Technical Prototype

A technical prototype of the tile detector has been developed and tested. The goal of this prototype was to evaluate the detector performance and cooling concept, develop production tools and finalize assembly procedures. This detector has a similar design to the one described in section 11.1, with a few modifications in the sub-module layout that were done in a later stage based on the experience from this technical prototype. For this prototype, the endrings, the cooling support structure and the tile matrix readout board were produced. At the time of production, the MuTRiG ASIC was not available. Therefore, a BGA packaged STiC 3.1 was used, which has the same front-end as the MuTRiG ASIC. In addition, a first version of the TMB, which allows the readout of a full module, was produced. In Figure 103, a tile matrix board assembled with SiPMs and a BGA-packaged STiC 3.1 ASIC is shown. In this design, eight digital temperature sensors were placed between the SiPMs and used for monitoring.

The scintillator material was manually cut in the workshop of the Kirchhoff-Institute for Physics in Heidelberg, with a tolerance of 180 µm for two different tile geometries, as presented in Figure 94. The tiles were individually wrapped with ESR reflective foil that was designed in a way to maximize the light yield while at the same time minimizing optical cross-talk between the channels. The foils were cut to the desired shape

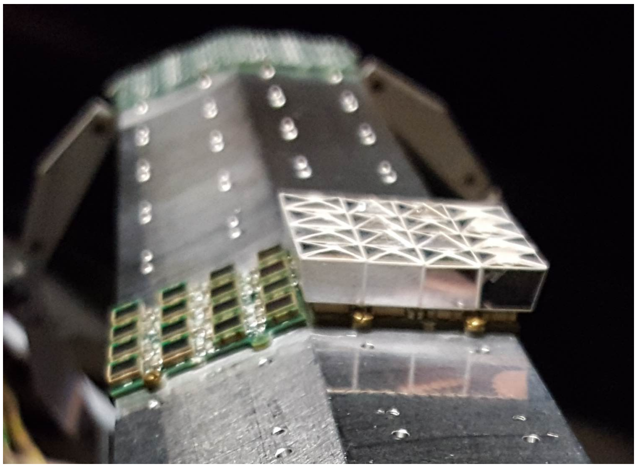


Figure 104: First half-assembled sub-module mounted on the cooling structure.

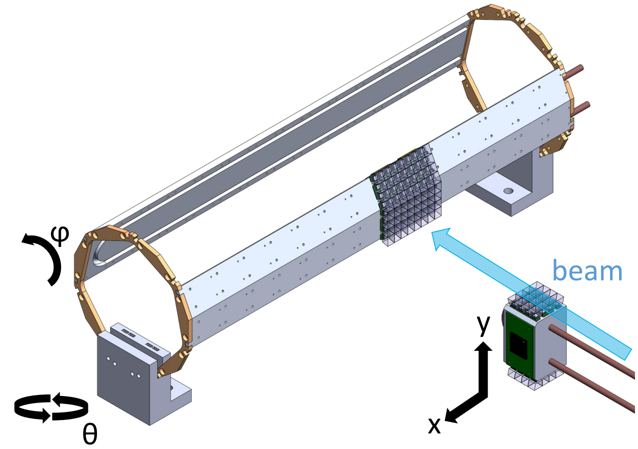


Figure 105: Schematic drawing of the test-beam setup at DESY, which includes three sub-modules.

using a laser cutter. For this prototype, an additional hole on the top side was added in order to monitor the gluing quality as shown in Figure 95.

In total, three sub-modules consisting of 96 channels, were assembled and tested in test-beam conditions. In Figure 104, the first half of a sub-module assembled on the cooling structure is presented.

11.5.1. Prototype Performance

In order to evaluate the detector performance, the timing and detection efficiency were measured with an electron beam at the DESY test-beam facility. A schematic view of the test setup is shown in Figure 105. For the measurements, one sub-module array of 4×4 scintillator tiles was positioned in parallel to the beam and served as a reference, such that the incident particles traversed four tiles in a row for each electron event. The other two sub-modules were assembled on the cooling structures and used as devices under test (DUTs). The devices under test could be rotated in ϕ and θ with respect to the beam and were read out using the prototype TMB board. The reference detector and the prototype TMB board were connected with 50 cm cables to a test FPGA board, which merged the data from the three ASICs. During the test-beam, both the reference matrix and all the channels of DUT_0 were calibrated, while for DUT_1 only a single row was optimized.

Figure 106 shows a typical time-over-threshold (ToT) spectrum. Several distinct features are visible: The most prominent feature is the peak at a ToT of about 610 CC bins (coarse counter bins), in the following referred to as Landau peak. This peak originates from electrons which fully traverse the tile. The second peak at a ToT of 210 CC bins originates from cross-talk between neighbouring scintillator tiles. This can be shown by selecting hits where at least one direct neighbour in the rows above or below the selected tile has a large signal with an energy deposition in the Landau peak. The corresponding events are shown by the red curve in

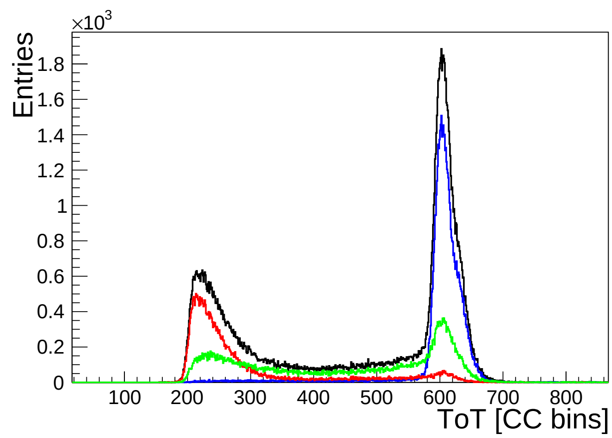


Figure 106: Energy deposition in a scintillator tile (black). The spectrum is composed of the Landau peak (blue), a plateau arising from edge effects (particles grazing the tile, green) and a peak from optical cross-talk (red).

Figure 106. The green line refers to a plateau arising from edge effects, where particles pass only partially through the tile. The large plateau and gap between the two peaks indicate the excellent light collection and low optical cross-talk between tiles, which shows the benefit of the individual tile wrapping.

11.5.2. Detection Efficiency

The detection efficiency is determined using electrons which traverse all four tiles of a certain row. Such events are selected by requiring at least three hits in the row, two of which must be in the first and last tile. The detection efficiency ε is then given by the probability to detect a hit in the remaining channel of the row with an energy deposition above the cross-talk level.

Due to the large light yield, which guarantees the signal to be well above the detection threshold, the efficiency is expected to be $\varepsilon \approx 100\,\%$. The resulting efficiency can be seen in Figure 107. An efficiency between $\varepsilon = 93.8\,\%$ and $\varepsilon = 98.7\,\%$ is achieved. In a small fraction of the events, a hit prior to the expected event was observed, which screens the expected hit thus causing an inefficiency in the channel. Correcting for this screening effect leads to an efficiency above 99 %. The remaining inefficiency can presumably be attributed to edge effects and misalignment of the tiles and inefficiency of data acquisition. For a better efficiency estimation, the measurement will be repeated using a tracker.

11.5.3. Time Resolution

The detector was optimized for timing measurements by fine-tuning the SiPM bias voltage and the timing thresholds. In order to evaluate the time resolution, a channel-to-channel time delay calibration was performed. These time delays are arising from different path lengths of the signal lines on the PCB and can vary up to 600 ps. When measuring the timing using threshold discrimination, an additional time delay caused by time walk effects needs to be corrected. For this correction, a tight ToT cut is applied on the reference channels in order to minimize time walk effects from the reference side.

In order to estimate the time resolution of a single channel, coincidence time distributions between at least three channels are used. The channel time resolution can then be extracted by:

$$\begin{aligned} \sigma_{i,3}^2 &= \sigma_i^2 + \sigma_3^2, \quad i = 1,2 \\ \sigma_{1,2}^2 &= \sigma_1^2 + \sigma_2^2 \\ \sigma = \sigma_3 &= \frac{1}{\sqrt{2}}\sqrt{\sigma_{3,1}^2 + \sigma_{3,2}^2 - \sigma_{1,2}^2} \end{aligned} \tag{5}$$

where $\sigma_{1,2}^2$, $\sigma_{1,3}^2$ and $\sigma_{2,3}^2$ are the three widths extracted from the coincidence time distribution between different pairs of channels.

The internal channel resolution, calculated with Equation 5 using channels in the same sub-module, is presented in Figure 108 for the DUTs. For these results, runs with tracks parallel to the DUTs are used in order to have at least three channel hits for the same electron event within a sub-module matrix by requiring events in the Landau distribution. A similar average resolution was measured both for the reference sub-module and for the two DUTs, where the average time resolution measured is 46.8 ± 7.6 ps. However, when repeating the same calculation using channels from different sub-modules, an additional jitter between the sub-modules is observed. The extra jitter between the reference sub-module and the DUTs of 45.5 ± 3.2 ps leads to a worse time resolution as shown in Figure 108 (blue). The main contribution to this arises from non-optimal design of the test board used for the read out of all sub-modules.

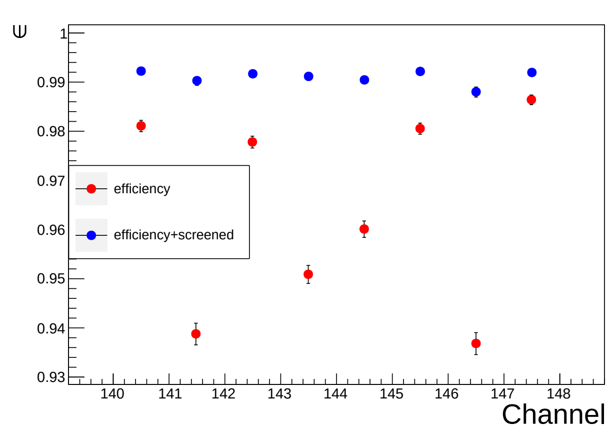


Figure 107: Efficiency calculated for the reference sub-module before and after the correction of the screening effect.

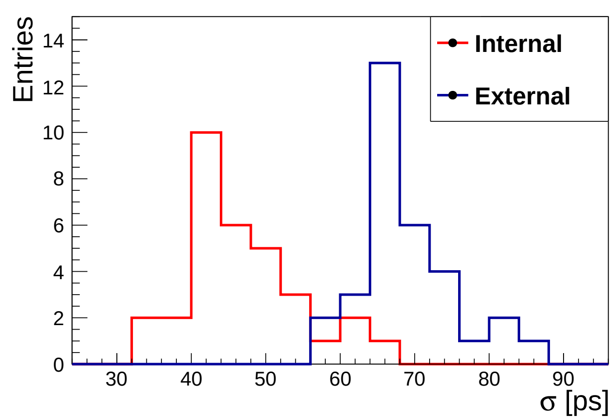


Figure 108: DUT channel resolution: (red) internal, (blue) external.

The expected event multiplicity during phase I of the experiment is presented in Figure 109a. While the average cluster size is ≈2, also cluster sizes higher than 9 can be observed. In order to evaluate the time resolution as a function of cluster size, a run with 0° incident angle was used, where the electron can pass through four channels in the reference matrix and up to four channels in DUT_0. The time resolution is evaluated using an even-odd analysis. For a given electron track, all hits are grouped into 'odd' or 'even' based on their channel position and the time difference is defined by:

$$\Delta t_{even-odd}(N_{hits}) = \frac{1}{N_{hits}} \left\{ \sum_{i=1}^{N_{even}} t_{2i} - \sum_{i=1}^{N_{odd}} t_{2i-1} \right\} \tag{6}$$

where N_{hits} is the sum of all hits. In order to avoid the additional jitter between the reference and the DUT, the sums in Equation 6 can be arranged such that the subtraction is only done within a sub-module, which leads to a requirement for an even total number of hits within each sub-module. In Figure 109b, the result for the even-odd analysis is shown. For this result only a single tower of the reference matrix (meaning four channels) and a single row of DUT_0 were used. The average resolution of these channels was measured to be 45 ± 4 ps, see Figure 108. The resolution as a function of clusters is extracted from Figure 109b by fitting it with the following function:

$$\sigma_t(N_{hits}) = \sigma_t^{single}/\sqrt{N_{hit}} \oplus \sigma_t^{const} \tag{7}$$

where $\sigma_t(N_{hits})$ is the time resolution for events with N_{hits}, σ_t^{single} is the time resolution of a single channel, and σ_t^{const} is an additional jitter that can be caused by misalignment between the channels. From the fit, a single channel resolution of ≈ 45 ps is measured, which is in agreement with the value extracted from the single channel measurements. In addition, a small misalignment is also observed. Furthermore, it can be seen that a time resolution better than 20 ps can be reached for events with high multiplicities.

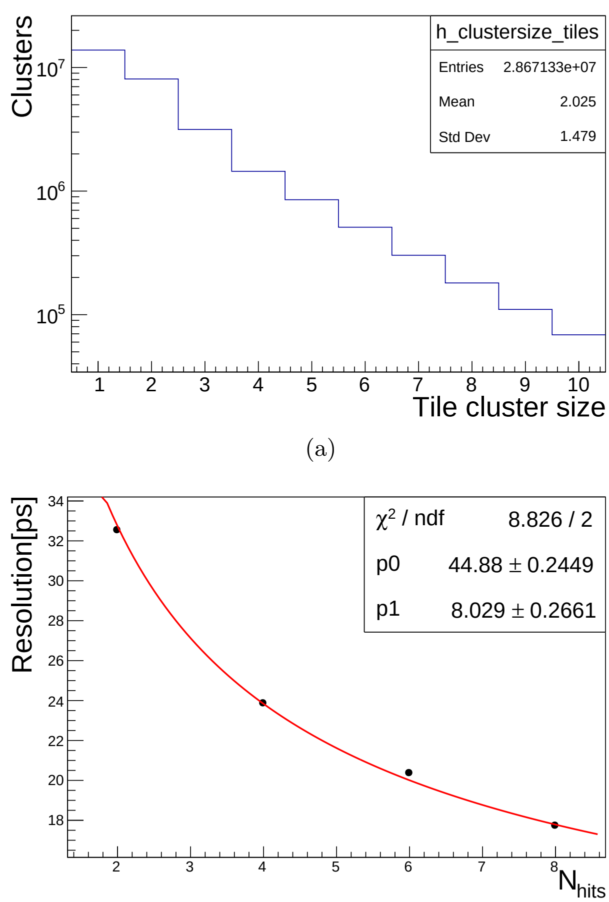


Figure 109: Cluster size impact on time resolution: (a) Simulated phase I cluster size per track. (b) Measured time resolution as a function of number of hits using even-odd analysis.

11.6. Cooling Simulation of the Tile Detector

To study the feasibility of the cooling system, thermal simulations were performed using the CAD implementation of the technical prototype, while in parallel, several measurements of the prototype in the laboratory environment were undertaken. After calibrating the simulation settings to the laboratory conditions, it was shown that the measurements can be reproduced in the simulation [46]. The simulation was therefore modified to investigate the cooling performance of a full module operating at the MuTRiG working power consumption of 1.2 W. Furthermore, the temperature of the water was adjusted to 1 °C to be closer to the operating conditions foreseen for the tile detector within the experiment, while the environment temperature was increased to 50 °C in order to subject the system to a stress test. The temperature of the SiPMs and the MuTRiG ASICs was investigated under these conditions. In Figure 110, the temperature of the PCBs on which the SiPMs are assembled is examined. While the temperature on the single PCBs is uniform down to a few tenths of a degree, the temperature range across the full length of the module spans about 2 °C. Considering the SiPM temperature coefficient $\Delta T_{Vop} = 54\,\text{mV}/°\text{C}$, these differences can be compensated by adjusting the high voltage of the individual SiPMs. Overall, the temperature is clearly reduced compared to the environment temperature, demonstrating the influence of the cooling system. Furthermore, the maximum temperature of the ASICs can be extracted from the simulation as ≈ 42 °C. This is still well within the safe margin of operation.

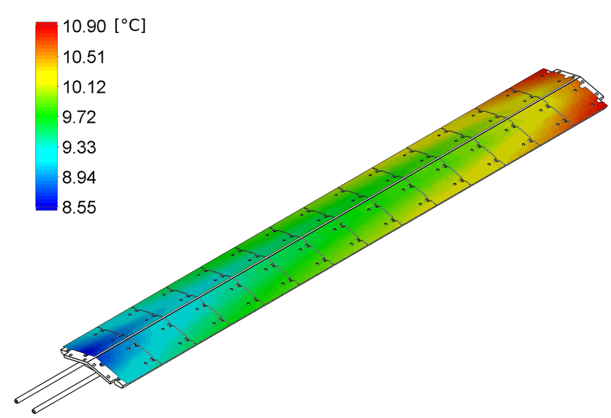


Figure 110: Simulated temperature of the SiPM PCBs. The temperature of the cooling water was set to 1 °C at a flow speed of 1 m/s, while the environment temperature was set to 50 °C.

12. Cooling infrastructure

The detectors, their electronics, the power converters and the data acquisition systems are located inside the densely spaced Mu3e magnet. The heat they produce is transferred to the outside by forced convection cooling.

System	Est. power W
Crate (front end FPGA boards)	2700
DC-DC converters	1500
Copper rods	200
Fibre detector (MuTRiG, SiPM)	120
Tile detector (MuTRiG, SiPM)	420
Total	4940

Table 12
List of systems requiring water cooling inside the experiment, with a conservative estimate of the heat dissipation. All circuits will be run independently.

Except for the pixel sensor chips, we are using water cooling everywhere. For the pixels, a novel gaseous helium cooling has been developed.

12.1. Water cooling

Water cooling is used to cool all the front-end electronics which are located outside the active volume of the detector, i.e. the front-end ASICs of the timing systems, the front-end FPGA-boards, the DC-DC converters, voltage regulators, etc. The anticipated heat load per source is listed in Table 12 and totals to about 5 kW. To protect the detector from ice buildup, the water inlet temperature is required to be above 2 °C, although the helium atmosphere provides a dry environment with a dew point below −40 °C. Pipe systems inside the experiment distribute the water to the heat sinks, see Figure 111. The FPGA boards are cooled via a manifold embedded into the circularly shaped crates. The DC-DC converters for the pixel powering are directly connected to a cooling loop. Heat dissipation for the low-voltage power distribution between the DC-DC converters and the front-end electronics (MuTRiG and MuPix ASICs) is a potential issue for the copper rods around the beam-pipe. Due to this issue active cooling of them is provided through a dedicated cooling ring thermally coupled to the rods. The timing detectors have their own cooling loops to dissipate the heat from the front-end ASIC and to keep the SiPM at a controlled low temperature. Further details on detector cooling of the timing systems can be found in chapters 10 and 11, and on cooling of the FPGA boards inside the crate in chapter 17.

Chilled water will be used from the PSI main supply via heat exchangers. Additional chillers are in place for circuits requiring lower set temperatures. The timing detectors will receive their independent chilled water loops for enhanced control of their temperatures.

12.2. Helium cooling

All MuPix chips of the pixel tracker are cooled by gaseous helium of $T_{\mathrm{He,in}} \gtrsim 0\,°\mathrm{C}$ at approximately ambient pressure. Assuming a maximum power consumption of the pixel sensors of $400\,\mathrm{mW/cm^2}$ the helium gas system is designed for a total heat transfer of 5.2 kW, which increases the averaged gas temperature by about 18 °C.[24] For this, the helium cooling system has to provide a flow of about $20\,\mathrm{m^3/min}$ (equal to 56 g/s of helium) under controlled conditions split between several cooling circuits (see section 7.6).

A process flow diagram for the helium plant is shown in Figure 112. Helium is pumped using miniature turbo compressors run at turbine speeds of up to 240 krpm. These units provide compression ratios up to ≈ 1.2 at mass flows in the range up to 25 g/s, depending on supplier and model. The energy consumption of the compressors for the full system is estimated to be around 6 kW in total. The helium circuits are designed with minimised pressure drops for a most economic system layout. The combination of a compressor and a valve for every circuit allows the control of the mass flow and the pressure differential applied individually. Compact, custom made Venturi tubes will be used to monitor the mass flows of every circuit. Leaks lead to losses and will contaminate the helium with air. In addition, outgassing organic residues from electronic components and adhesives need to be removed. Hence a cold trap is included in a by-pass configuration to keep the helium pure enough. An expansion volume will be present to compensate for the compression and expansion of the gas volume during ramp-up and ramp-down of the gas flows. A low pressure drop shell-and-tube heat exchanger is used to remove the heat from the helium.

13. Mechanical Integration

The detector is maintained at its nominal position inside the magnet by a removable frame called the *detector cage*. The cage also carries infrastructure such as crates for the power converters and the front-end FPGA boards, and provides support for all cabling and piping.

[24] The pixel detector consists of 2844 chips (108 in the vertex detector, 3 × 912 in the outer layers), giving about $1.14\,\mathrm{m^2}$ of active instrumented surface ($20 \times 20\,\mathrm{mm^2}$ active area per chip, neglecting the chip periphery) or about $1.3\,\mathrm{m^2}$ including chip peripheries. The conservative (optimistic) scenario leads to about 5.2 kW (3.3 kW) of dissipated heat. The specific heat capacity of gaseous helium is 5.2 kJ/(kg K).

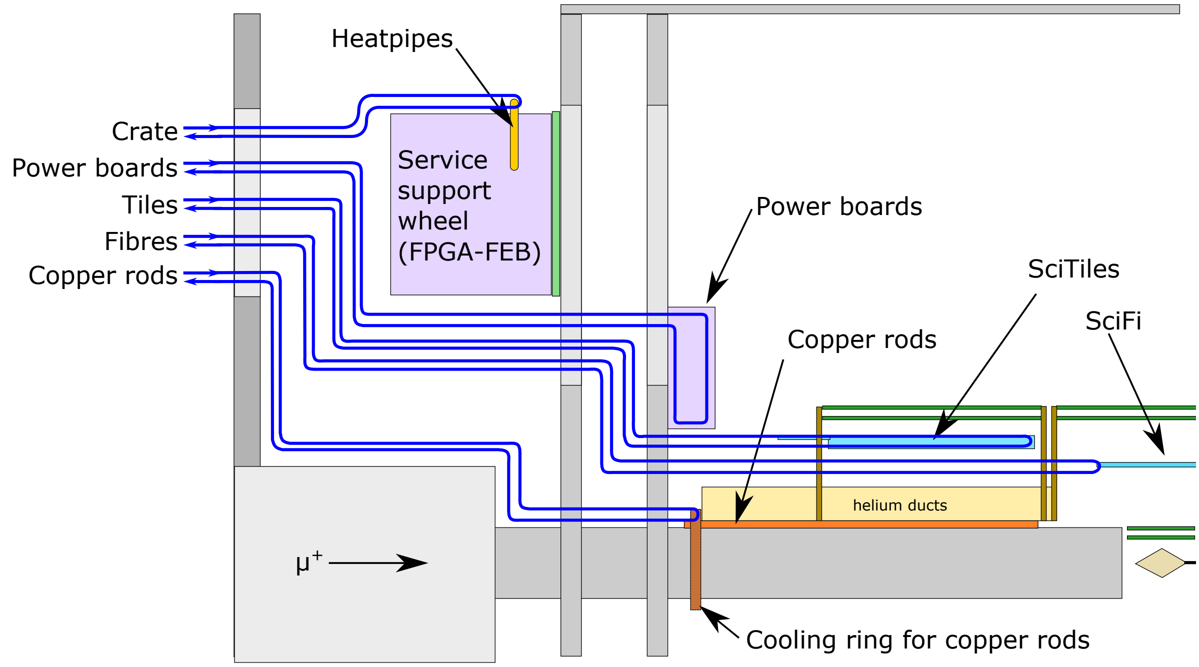


Figure 111: Schematic view of the water cooling topology for one quadrant of the experiment inside the magnet.

13.1. Detector Cage and Rail System

The detector cage has the shape of a hollow cylinder with its axis horizontal, as shown in Figure 113. At each end, a ring frame made of pairs of glass-fibre reinforced polymer wheels has a clamp at its centre for the beam-pipe. Aluminium struts connect the two ring frames and form the cylinder. Gliders on the wider struts (at the 3- and 9 o'clock positions) guide the cage on the rail system inside the magnet.

To compensate for possible thermal expansion in the x (horizontal, perpendicular to the beam-pipe) direction, the gliders on the left rail are floating whilst on the other rail they are kept at a defined position. In the y (vertical) direction the position is defined by the top surface of the rail. The z position is kept fixed by screws.

The clamps in the centre of the rings at either end hold the two beam-pipes in position and take all the weight of the detector. Mechanisms to fine-adjust the beam-pipe pointing angles are built into the clamps. Finite element simulations were performed to test the sturdiness of the design. Load tests have been carried out on a full-scale mock-up, confirming the simulation results of a deflection of 0.3 mm under a typical detector load of 10 kg at the beam-pipe tips. The connection of the beam-pipes to the beam line is described in chapter 5.

13.2. Mechanical support of detector stations

The detector components are mounted on the beam-pipes, see Figure 114. As shown in the previous chapters, both pixel and timing detectors follow a barrel concept. They are mounted on pairs of end rings, supported on the beam-pipes. Whilst the recurl stations have their support on one beam-pipe, the central barrel has one mechanical support on the upstream beam-pipe and the other on the downstream beam-pipe. To compensate for any tilt of the end rings and movements due to thermal expansion, the detector mounts are spring-loaded at one end.

Detectors can be mounted and dismounted in sequence from inner to outer without the need to retract the beam-pipes. For example to mount the central barrels, the vertex half-shells of layers 1 and 2 will be installed first, followed by the fibre ribbons. Finally, the pixel modules for layers 3 and 4 will be mounted. For this sequence, the cage can be placed on a special extraction cart on wheels. It has the same rail system as that inside the magnet. For better access, the cart has rollers allowing the rotation of the cage around its own z-axis in a safe manner.

The beam-pipes also provide support for other services. The copper bars to supply power to the central detectors (section 14.4) are glued onto the beam-pipe with a custom procedure to ensure proper electrical insulation, and manifolds for the helium cooling system are attached to the ends of the beam-pipes.

13.3. Supply systems and cable routing

Service support wheels (SSW) are situated outside either end of the detector cage. They are loosely coupled to the cage in the z direction and have their own gliders to decouple mechanical forces from the cage.

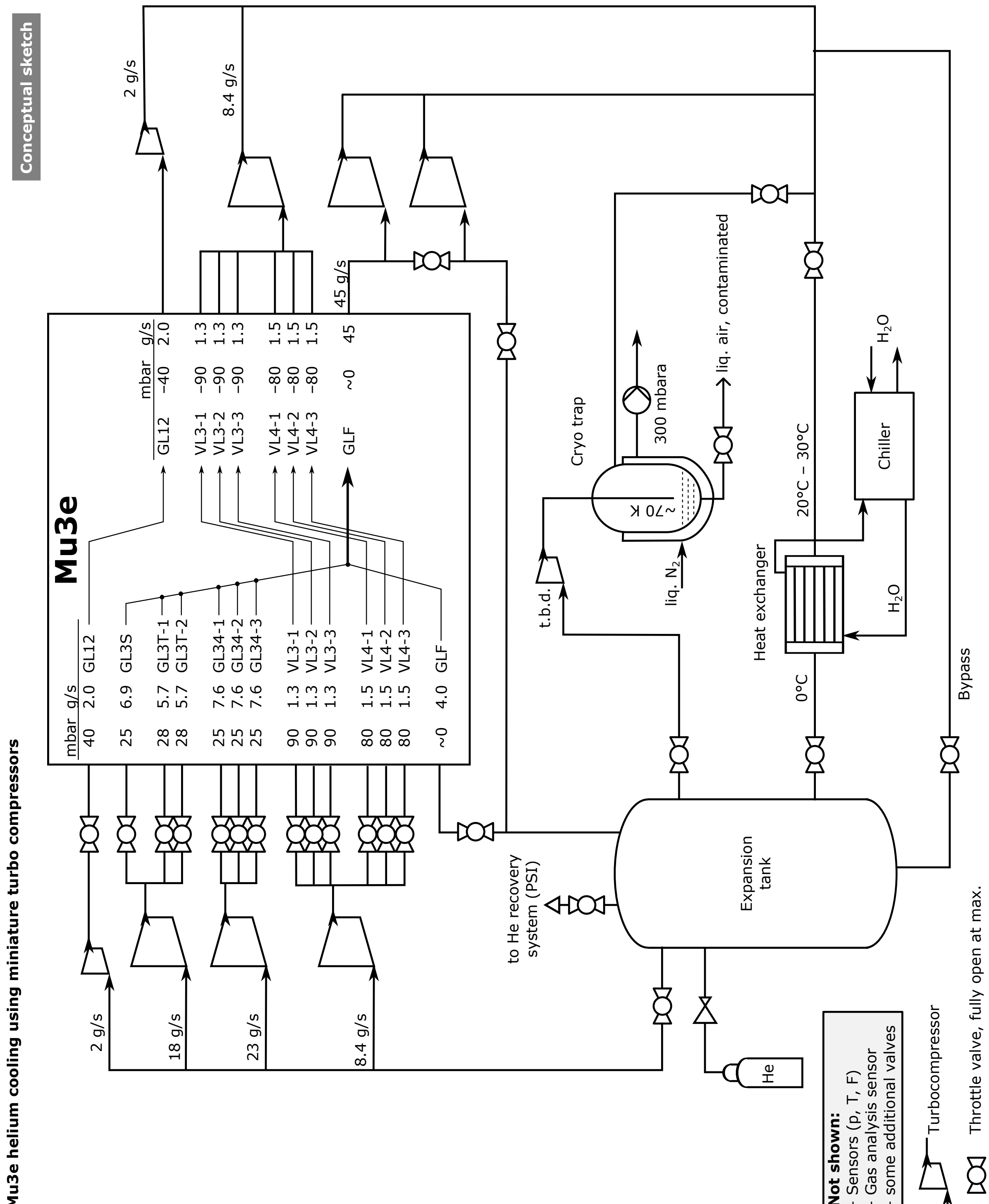


Figure 112: Conceptual process flow diagram of the Mu3e helium cooling infrastructure. Miniature turbo compressors in the circuit may be implemented using multiple units operated in parallel or in series, depending on needs.

The SSWs hold crates for the front end boards, patch panels for the power connections and routing for the cooling pipes (water and helium). The DC-DC converter boards (low voltage supply) and the bias voltage generators are mounted on the inner side of the glass-fibre wheels. All services have connections at the outward facing planes of the SSWs. Figure 116 shows a conceptual view.

Services have to be routed from the inside to the outside of the experiment through flanges sealing the internal dry helium atmosphere from the ambient environment. Four identical flange plates are mounted on

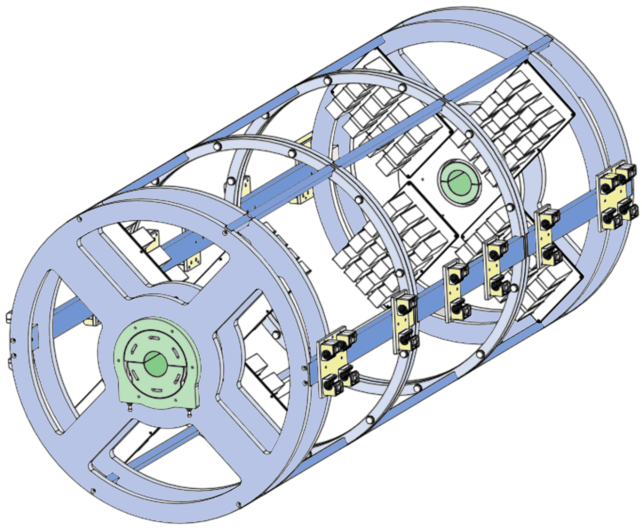


Figure 113: Detector cage structure consisting of two ring frames (light blue) connected by struts (dark blue). The clamps holding the beam-pipes are inside the ring pairs at either end (shaded green). The gliders (yellow) allow the cage structure to be moved into the magnet on the rail system.

four turrets at the end plates of the magnet, two at either end. Ports for all media are present and provide suitable connectors. For the power connections, sealed heavy-duty double-sided 56 pin connector assemblies are used[25]. Tubes for the helium and water coolants are welded into the flange and will use industry standard fluid connectors. The fibre bundles are sealed with epoxy into brackets that are sealed with an O-ring to the flange. A drawing is shown in Figure 115.

13.4. Access to the Mu3e detector

Extracting the experiment from the warm bore of the magnet requires an orchestrated procedure, which essentially looks as follows:

1. Detach the beam line, secure cables and hoses. Temporarily remove beam line parts as needed to make space.
2. Move the magnet into the extraction position.
3. Open the magnet doors. Remove access plates from the helium sealing plate.
4. Disconnect all cables and hoses though access holes.
5. Safely remove the sealing plates, secure cables and hoses while doing this.
6. Place extraction cart in front of experiment. Engage rail coupling. Carefully remove experiment, guided by the rails.

For detector insertion, the procedure is reversed. The extraction cart is the same as described in the previous section. Guide pins and clamps help to safely couple the cart to the rail system in front of the magnet.

For servicing the detector, a protective tent will be available that can be used either inside the area for quick work, or outside the area in a secure space. External crane attachment points are provided for transferring the experiment to outside of the beam area.

[25] Supplier: Souriau-Sunbank

14. Power Distribution and Cabling

With a power consumption from the pixel tracker, the SiPM readout electronics, front-end board, and step-down converters (see Table 12) of up to 10 kW, the Mu3e detector needs a robust but also compact power-distribution system. The conceptual design for such a system is shown in Figure 117. Power supplies located on the lower infrastructure platform deliver 20 V DC, a voltage high enough to allow for a compact and flexible set of power cables, which are brought into the experiment through a high-density power connector. From there, the power is distributed to either the front-end board crates with embedded buck converters, or to the power boards which step down the voltage for the MuPix chips, and the tile and fibre readout boards. In addition, separate power is provided to the slow control systems which need to run when the main detector power is switched off.

14.1. Power Partitions and Grounding

The Mu3e experiment is divided into 112 detector partitions, which also act as independently controlled power partitions (see Table 13). The DC power supplies for these partitions will be the TDK-Lambda GENESYS low-voltage power supply, which are known to be reliable, for example they are being used in the MEG experiment. Each supply can provide up to 90 A / 2700 W, which is distributed to several power partitions via a power relay bank. A massive common return line per supply minimizes the voltage drop. Each power supply output is floating, and the return line is referenced to the common ground inside the experimental cage. Slow control systems such as the alignment system, environment monitoring, the controller boards regulating the detector power, and all safety critical systems are powered separately. This enables the powering of all diagnostic tools of the experiment prior to the turn on of the high-power detector electronics.

This powering scheme means that care has to be taken to not introduce ground loops when connecting the various detector partitions to e.g. a slow control bus or a high-voltage input. To avoid this all data connections to the outside go via optical fibres, the readout is therefore fully electrically decoupled.

14.2. DC-DC conversion

Switching power converters will be used to step down the 20 V to the voltages needed by the detector and

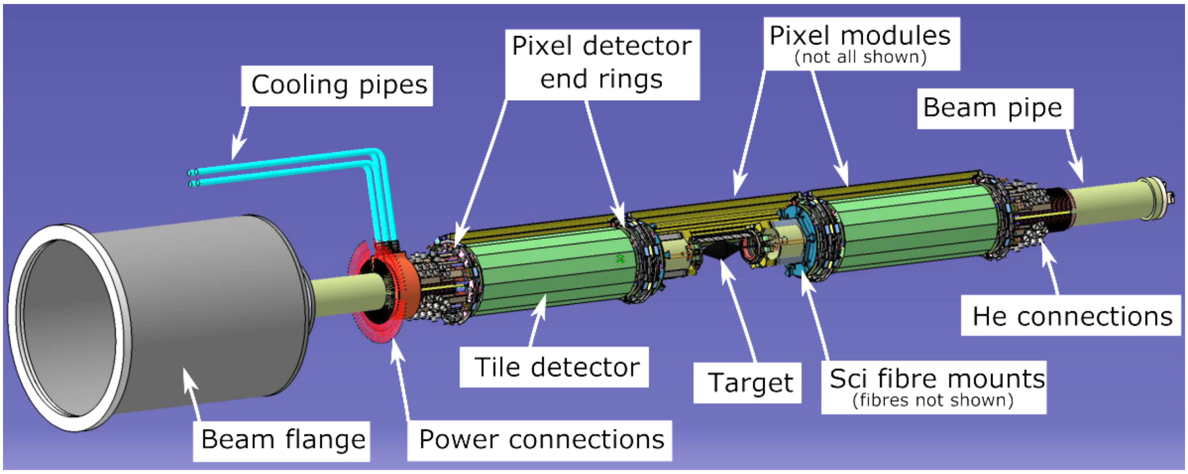


Figure 114: The Mu3e experiment mounted on the beam-pipes. Not shown are the detector cage and supplies. Some parts have been partially removed for visibility. LV: low voltage power.

Partition type (ASIC)	#partitions	#ASICS/ partition	Maximum power per partition [W] Excluding	Including DC-DC	Total Power including DC-DC losses [W]
Pixel(MuPix)					
layer 1	4	12	19.2	25.6	102
layer 2	4	15	24.0	32	128
layer 3	3×12	32, 36	51.2, 57.6	68.3, 76.8	2660
layer 4	3×14	36	57.6	76.8	3230
Fibre(MuTRiG)	12	8	9.6	12.8	153
Tile(MuTRiG)	14	13	15.6	20.8	291
Front-end board	8	14 boards	266	350	2800
Total					**9370**

Table 13
Power partitions for the Mu3e detector ASICs and electronics inside the magnet bore. The high-power elements on the front-end board are the Arria V FPGA, clock chip, and the transceivers. A respective maximum power consumption of 1.2 W and 1.6 W for the MuTRiG and MuPix chips is assumed. The upper limits on the power figures are driven by the cooling system, and depend on power losses in the entire power distribution system. For the total power budget, a 75% efficiency of the DC-DC converters is assumed.

electronics (Table 14) Typical efficiencies are of the order of to 70% to 90%, depending on the current and the voltage step. Compact high-power converters typically used for FPGA boards such as the LTM4601 (Analogue Devices) have a ferrite core inductor, with is incompatible with the high magnetic field environment of the experiment. Mu3e has selected the following solution: a commercial synchronous buck converter combined with a custom air coil, where the coil properties and the switching frequency are optimized for the required output voltage and current. As they are mounted outside the active area of the detector, these converters don't have to be radiation hard.

14.2.1. Front-end board converters

The front-end boards with an Arria V FPGA, and the LVDS and optical tranceivers (section 17.2) require several DC voltages at typical currents of 1-3 A. Three switching DC-DC converters will generate 1.1, 1.8, and 3.3 VDC with Peak-Peak ripple below 10 mV (see Table 14). Passive filters and active filtering with devices such as the LT3086 (Analogue Devices) further reduce the voltage ripple, and allow intermediate voltages to be generated. The switching converters on the front-end board are based on a compact TPS548A20 synchronous buck converter with integrated switches (Texas Instruments), combined with single layer cylindrical air coils. Figure 118 shows a stand alone 2x4 cm 1.8 V prototype, which has demonstrated good performance at operating conditions. The converter embedded on the front-end board will have a similar footprint, with an additional copper shielding box covering the coil to reduce EMI and improve mechanical stability [47].

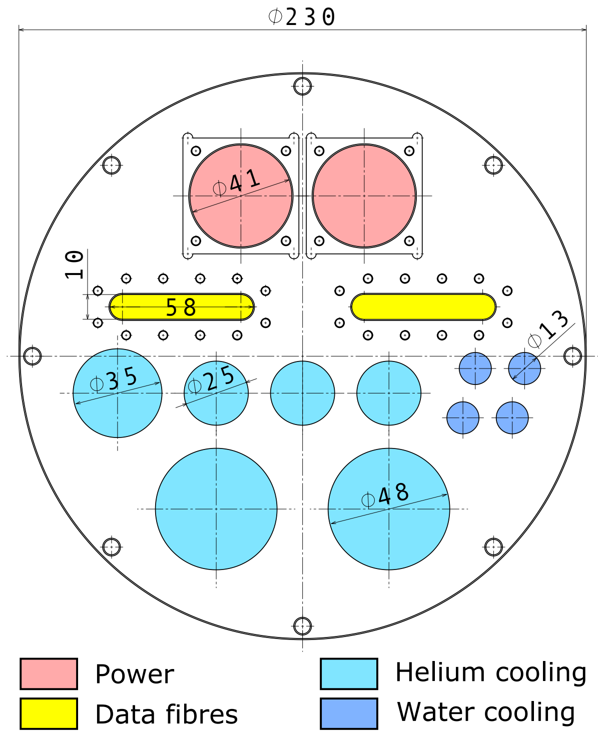


Figure 115: Drawing of the cabling flange. Additional ports will be added for auxiliary use. Dimensions in mm.

14.2.2. Power boards

With currents potentially up to 30 A and very few options for additional filtering further down the line, the requirements for the active detector DC-DC converters are more challenging. The TPS53219 buck controller and CSD86350Q5D power MOSFET switch from Texas Instruments were identified as meeting these requirements. A first prototype was developed (see Figure 119). This board has space for various input and output filter configurations, and has dimensions close to the final form factor. The configuration shown, with a toroidal coil in combination with a secondary LC filter, has the best noise figure with a Peak-Peak ripple of approximately 10 mV. The board was stress tested in a magnetic field, and successfully used to power a MuPix 8 pixel detector during a DESY testbeam campaign.

The final power board has a secondary output filter, and several additional features such as current monitoring, and interface connector for the back plane, and an embedded temperature interlock connected to a temperature diode on the MuPix sensor [48].

In the experiment, 16 power boards are mounted in a crate on the SSW (see Figures 117 and 116), with a MSCB slave (chapter 16) as controller. This controller adjusts the output voltage, switching frequency, and monitors several parameters. It also interfaces the DC-DC converters with an external interlock system.

14.3. Bias voltage

Bias voltages between 50 V and 120 V are required for the SiPMs used in the fibre tracker and the tile detector as well as for the MuPix chip. As only moderate currents of few µA per channel are needed, these voltages can be generated with a Cockroft-Walton chain. Converters supplying positive voltages have been developed and optimized in the context of the MEGII experiment. For Mu3e this design is carried over to a new board which will be mounted inside the magnet volume.

The pixel tracker requires a negative bias voltage of up to -100 V for each chip. For economic reasons, a set of four power groups is provided with a common voltage with dedicated current measurements and the possibility to turn off each power group individually. Since voltage generators which run at the high magnetic field are not available commercially, a custom board based on the Cockroft-Walton voltage multiplier design has been created. Figure 121 shows the simplified block schematic of this device. A micro-controller connected to the MSCB slow control system operates the DAC, ADC and switches of the voltage generator. It is capable of generating a bias voltage from $0 \ldots -150\,\mathrm{V}$ out of a single power supply of 5 V. First tests with a prototype indicated that an absolute voltage accuracy of $\pm 1\,\mathrm{mV}$ at a current of 2 mA can be achieved with a residual ripple below 10 mV. Each channel contains a shunt resistor and an ADC, which can measure the individual current. High voltage CMOS switches operated by the micro-controller can switch off individual channels in case the corresponding pixel chips would have a problem.

Figure 120 shows the top and bottom sides of a prototype of the high voltage board. It has a size of $30 \times 60\,\mathrm{mm}^2$. The Cockroft-Walton chain can be identified on the bottom side of the board. No magnetic components have been used in the design, making it possible to operate the board in magnetic fields of up to 2 T.

14.4. Cabling

The basic concept of the cabling inside the detector is shown in Figure 116. From the power boards, the connections to the detector components are carried out with minimal possible length using solid copper cable of $2.5\,\mathrm{mm}^2$ gauge. Because all connections have to be done outside the detector acceptance, only the space around the beam pipes is left. Copper rods with a cross-section of $5 \times 2.5\,\mathrm{mm}^2$ are used to bridge the connection between the detector endring mount and the outer end of the beam pipe. These rods are individually insulated using a polyimide foil wrap, and held in place by epoxy. The rods are in a densely packed environment, hence the dissipated power will be actively removed using a copper cooling ring (see section 12.1). The cables are connected using screw-mounted copper

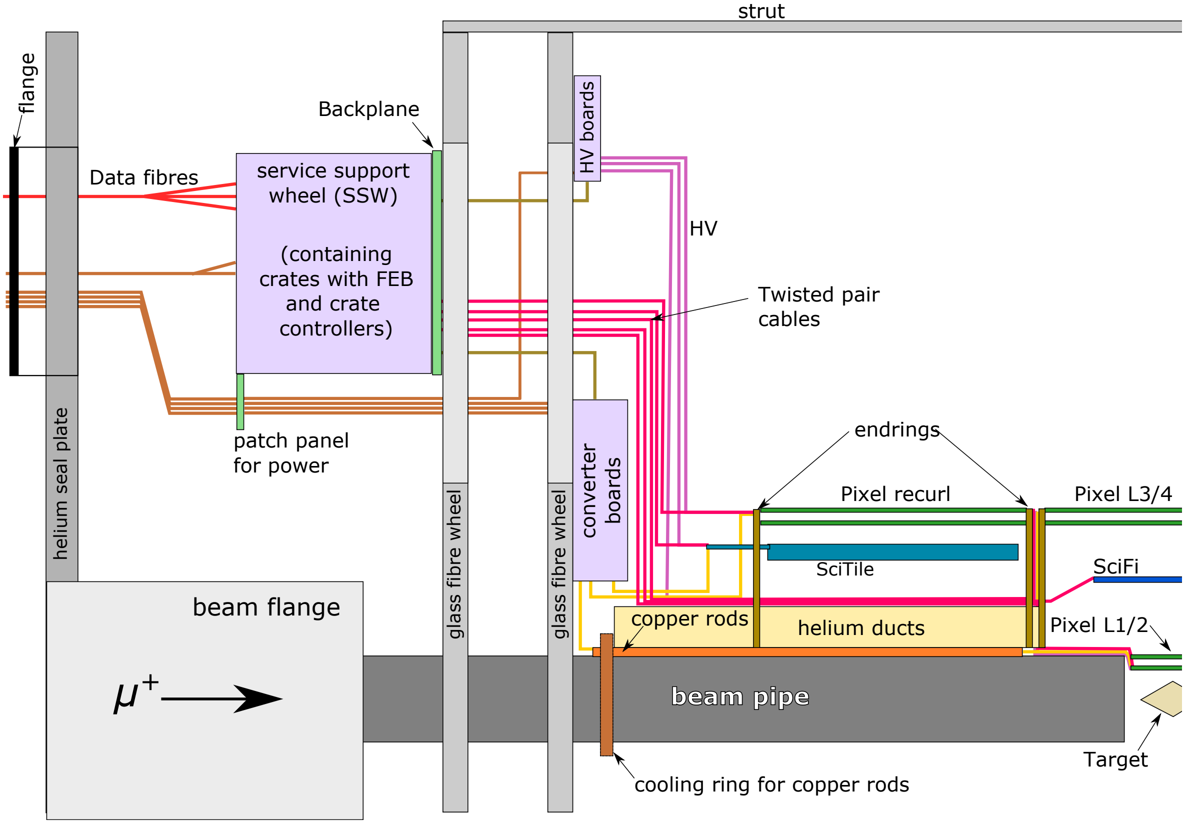


Figure 116: Conceptual view of supply system positioning and cable routing (rz-view, not to scale). All supplies can be disconnected for the extraction of the experiment. The feed-throughs on the helium seal plate are gas-tight.

Component	Voltage [V]	Typical current [A]	Min. inductance air coil [µH]	coil design
Front-end board	1.1	2	2	cylindrical
Front-end board	1.8	1.7	6	cylindrical
Front-end board	3.3	2.2	4	cylindrical
MuPix partition (layer 1,2)	ca. 2.3	10	0.5	toroid
MuPix partition (layer 3,4)	ca. 2.3	21	0.4	toroid
Fibre partition	ca. 2.0	7	0.7	toroid
Tile partition	ca. 2.0	9	0.7	toroid
Tile partition	3.3	3	3.3	toroid

Table 14
Specifications for different buck converter channels stepping down the voltage from 20 V with an efficiency >70%. The quoted MuPix voltage takes into account an anticipated voltage drop of 200 to 300 mV between the converter and the chip.

clamps.

Data cables between detectors and the front-end boards are micro-twisted pair wires: AWG 36 wires with a Polyimide isolation and an impedance of 90 Ω from *Heermann GmbH*. Each bundle of up to 50 pairs has a typical outer diameter of 2 mm. The bundles are arranged around the water cooling pipes: see Figure 122 for a sketch of the arrangement. The data cables are attached to the detector elements using soldered connections on flexible printed circuit boards, which connect to the PCB or HDI via interposers. The attachment to the frontend boards takes place on the patch panel of the SSW using zero-force connectors.

15. Clock Distribution

The precise timing measurement as well as the operation of many Gbit/s links in the experiment requires a very stable clock distribution. In order to ensure synchronisation between the timestamps of all sub-detectors, a global synchronous reset is also required.

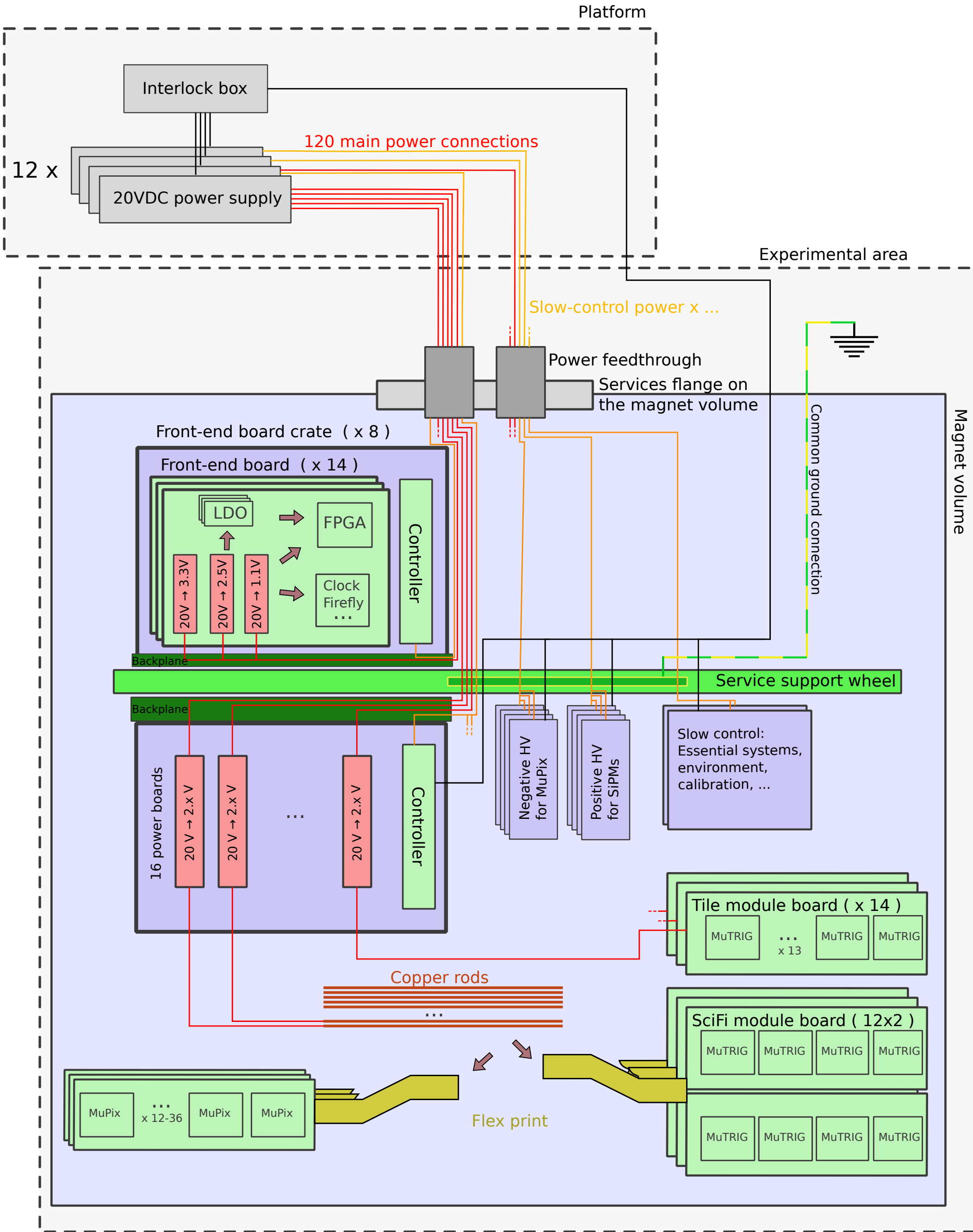


Figure 117: Schematic view of the power distribution inside the detector. Floating 20 VDC supply lines provide up to 12 A of current each (red lines). Custom DC-DC converters on the front-end board and close to the active detector step this 20 VDC down to the required voltages. Separate power is provided to the slow control systems (orange lines), which need to run independently from the main detector power. Note that these services are distributed over the upstream and downstream Service Support Wheel (SSW). This SSW also acts as a common ground plane, with a single ground connection to the outside.

The frequency of the clock distribution system is chosen to be 125 MHz; other frequencies can be derived locally by phase-locked loops. To meet the timing resolution requirements for all of the detector subsystems, the phase stability of the clock distribution has to be better than 10 ps over the complete system. The jitter requirements of the global reset ($\sim$ half a clock period) are more relaxed.

The overall block diagram of the clock and reset system can be seen in Figure 123. The 125 MHz clock is generated by a commercial low-jitter clock oscillator on a dedicated board. This board is controlled by an FPGA which also controls the reset signal. The board provides both clock and reset signals to opti-

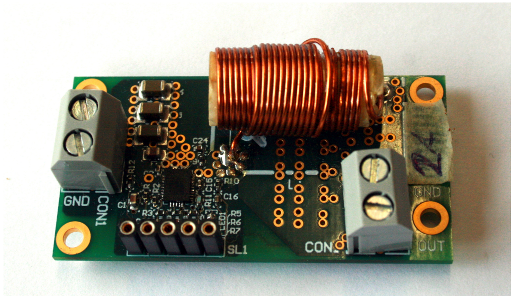


Figure 118: The second prototype for the buck converters for the frontend board. Good performance with efficiencies >75% in a 0.7 T magnetic field was demonstrated.

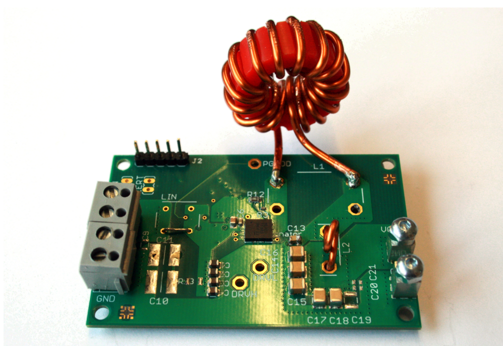


Figure 119: A 4.5 x 7 cm prototype for the power board, with a 0.5 µH toroidal inductor and a secondary LC filter at the output. With an output ripple of 10-20 mV, this has been used to succesfully power MuPix 8 sensors.

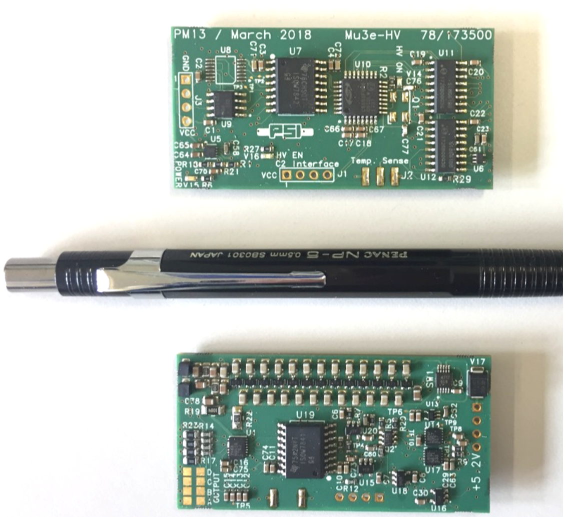


Figure 120: Prototype of the high-voltage generator board with top side (upper picture) and bottom side (lower picture).

cal transmitters, which are then passed to a number of bespoke boards. These actively split the signal and supply the clock and reset lines on optical fibres to local clock distribution boards inside the warm bore of the magnet, as well as to the DAQ switching boards (see section 17.4). Inside the magnet, optical receivers forward the signals to a jitter cleaner and fan-out chip on the front-end board, which then drives LVDS signals to the FPGA and front-end ASICs. For the filter farm PCs, each FPGA board inside a PC receives the global clock via a Clock Transmission Board (CTB), a small custom board which converts the optical clock signal to an electronic one. Reset and state changes are communicated to the farm via ethernet.

The reset signal is synchronised to the clock and uses 8 bit datagrams in 8 bit/10 bit encoding [49, 50], to induce not only resets but also changes of operation mode and the synchronising of the jitter cleaners. In idle mode, a comma word is sent, allowing for word alignment. Resets of different subsystems and changes between idle and running modes are triggered by sending one of the 256 possible data words.

The front-end boards then have the task of distributing the clock and reset (here the reset is an on/off signal) to all ASICS (MuPix and MuTRiG). A dedicated jitter cleaner and fan-out component is used, which will also be used to generate the 625 MHz clock needed by the MuTRiG Phase-Locked Loop (PLL).

Slower clocks, required e.g. for slow control and configuration signals, will be generated by clock dividers and/or PLLs in the FPGAs.

15.1. FMC Distribution Board

At the heart of the clock and reset system is the FPGA Mezzanine Card (FMC) distribution board, shown in Figure 124. The distribution board connects to the FPGA development board via an FMC connector [51] which allows access to the 10 Multi-Gigabit Transceiver (MGT) lines the Xilinx Kintex-7 FPGA offers. The low-jitter IC which is used to generate the 125 MHz clock is the Silicon Labs Si5345 which can generate up to 10 any-frequency clock outputs with an ultra-low jitter of 90 fs RMS. The SI5345 also offers in-circuit non-volatile programming which ensures a regular power up with a known frequency. The distribution board uses 8 MGTs for the reset lines and all 10 clock outputs from the clock generator, two of which are used to generate the MGT lines and the remaining 8 are used for the clock lines. The clock and reset lines are routed to the inputs of two optical Firefly transceivers. Such a configuration allows the individual control of the 8 pairs of the clock and reset lines. In the experiment, these can be translated as 8 individually controllable partitions, though only 4 are needed for the phase I Mu3e DAQ. In addition, the distribution board provides an I^2C interface which is used to communicate with the active splitting boards, which among many other functions, can provide the ability to disable/enable individual clock and reset lines.

15.2. Active Splitting

The optical splitting of the clock and reset lines is modular in design. It allows the active splitting to be versatile and can be potentially used as a generic active

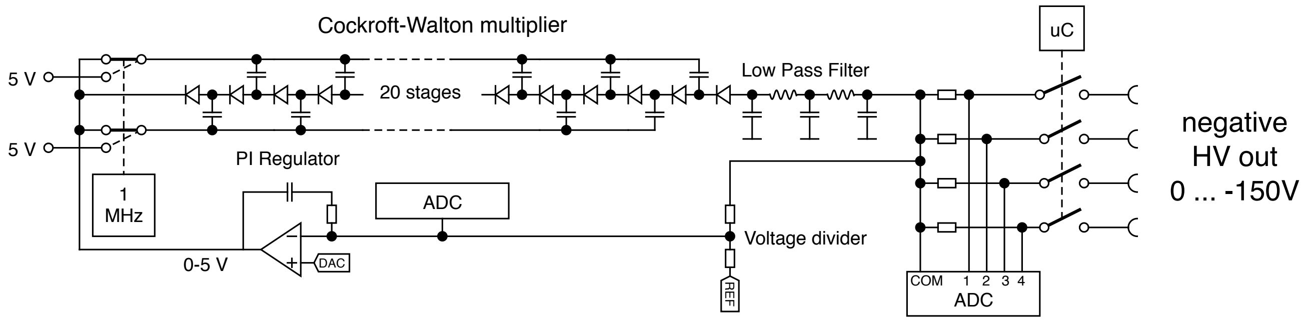


Figure 121: Block schematic of the pixel high voltage generation.

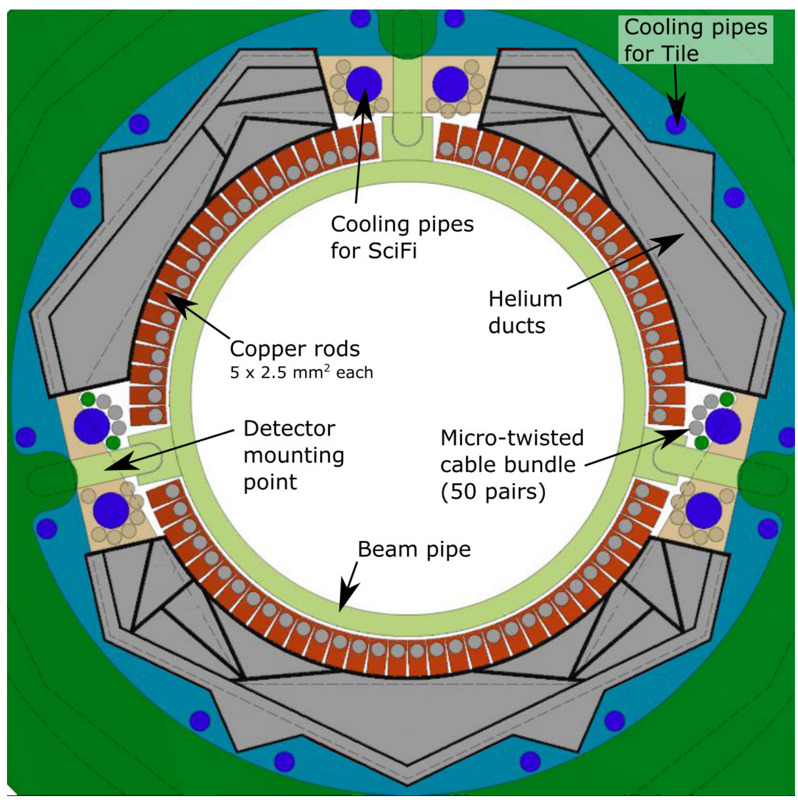


Figure 122: Cross section of a recurl station. The micro-twisted pair cables come in bundles and are shown as circles around the cooling pipes for the fibre detector (the colour code shows detector assignment: green for fibres, grey for vertex layers, brown for pixel outer layers. The helium ducts have separated channels for different destinations. Their cross sections are optimised for minimal pressure drop.

optical splitting solution. The system unit comprises one motherboard and 8 daughter boards. The motherboard takes 8 optical inputs and electrically routes each of the 8 signals to one of the eight daughter boards. The daughter board, which connects to the motherboard via a high-speed mezzanine connector, creates 36 copies of the input signal for a total of 288 optical copies per system unit. One system unit is sufficient for all clock reset lines required by the Phase 1 Mu3e DAQ. A 3D representation of the system can be seen in Figure 125.

15.2.1. Board designs

The daughter board, shown in Figure 126, utilises the OnSemi NB7L1008M fan-out chip; a 1:8 6.5 Gbit/s differential fan-out buffer with a random clock jitter < 0.8 ps RMS. Overall, 5 fan-out chips are used per board, creating a total of 40 replicated signals. However, only 36 are used as the three on-board Firefly transceivers have a total of 36 optical transmitters. The 36 optical lines are carried by three 12-fibre OM3 MTP cables. The board also has a low-noise DC-DC converter with power monitoring and is connected to an I^2C bus which allows this power monitor to be read, in addition to

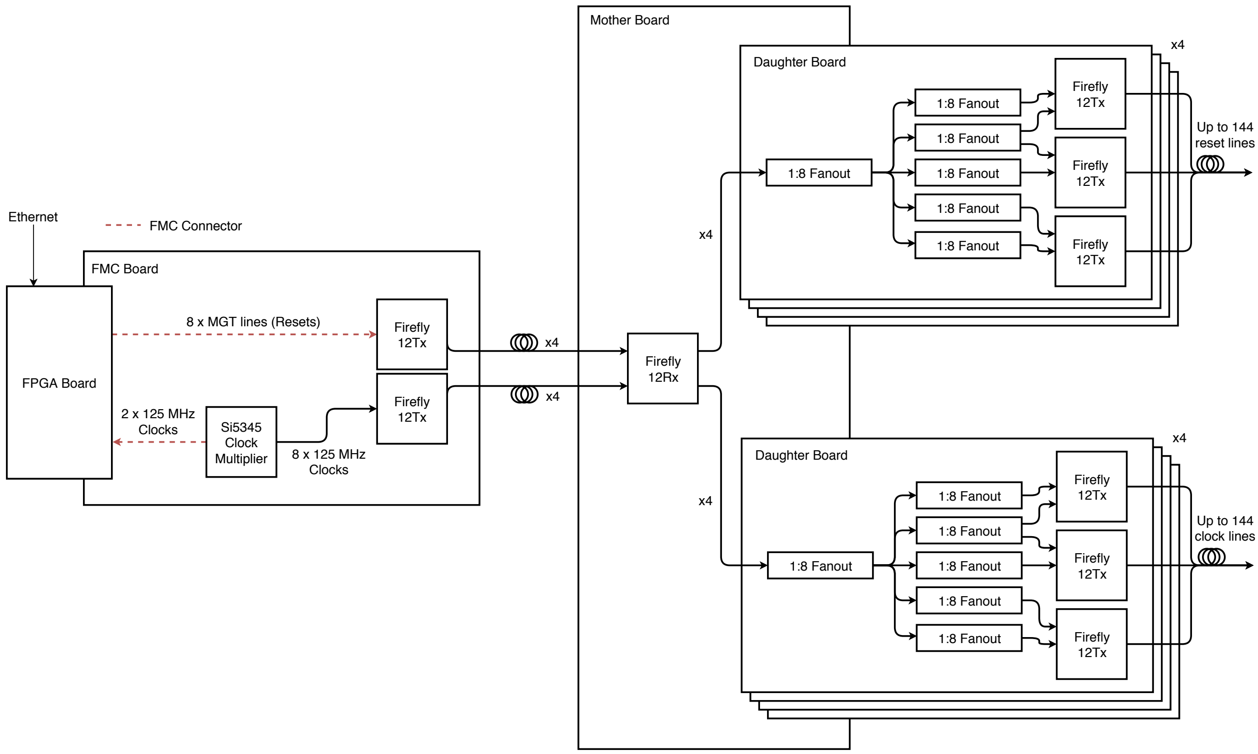


Figure 123: A block diagram illustrating the connections and relationships between the various boards of the clock and reset distribution system. The block diagram shows how the network-controlled FPGA can be paired with a bespoke FMC board to create 8 reset and 8 clock optical lines (only 4 of each are used in the phase I DAQ). These are then split with bespoke fan-out boards, effectively creating a 1:36 active optical splitter.

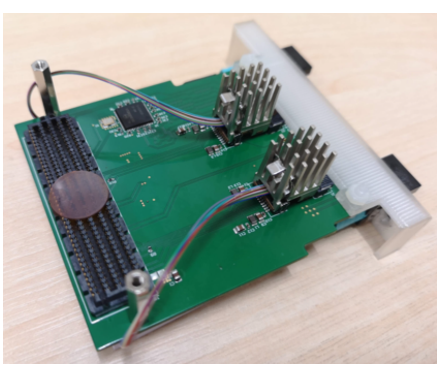


Figure 124: The Clock and Reset FMC distribution board has two Firefly optical transceivers. The transceiver on the top connects directly to the Silicon Labs Si5345 clock generator IC. The bottom optical transceiver connects to the MGT lines coming from the FPGA via the FMC connector. Both Firefly transceivers have 12 optical transmit lines but only 8 of each are used.

controlling the Firefly transceivers (e.g. disabling and enabling individual optical channels).

The daughter board receives its high-speed signals, communication bus and power from the motherboard, as has been described and seen in Figure 125, where the full system integrated into a 19 inch rack-mountable box is also shown.

15.3. Clock and Reset Operation

The clock and reset firmware is developed on the Digilent Genesys 2 board [52], a Xilinx Kintex-7 evaluation board. The custom firmware implements the IPBUS protocol [53], allowing the end-user to modify registers through a network connection and thereby control the clock and reset system. The firmware provides two IPBUS-based interfaces: to the FPGA MGT 8b/10b transceivers; and to the I^2C control of the Firefly transceivers, clock generator, and the power and cooling systems. The IPBUS protocol is also implemented in the MIDAS DAQ system, which provides a control page for the clock and reset distribution system, see Figure 127.

15.4. Performance

The full clock and reset system have undergone extensive testing. All optical outputs are fully operational, as are the configuration and monitoring of the system. The Firefly transceivers have good thermal performance and stability with cooling provided by fans in the rack-mountable box. The relative phase of clocks from different daughter boards transmitted via separate optical fibre assemblies to two different receivers has been measured to have a jitter of less than 5 ps, including a sizeable contribution from the measurement set-up, as shown in Figure 128.

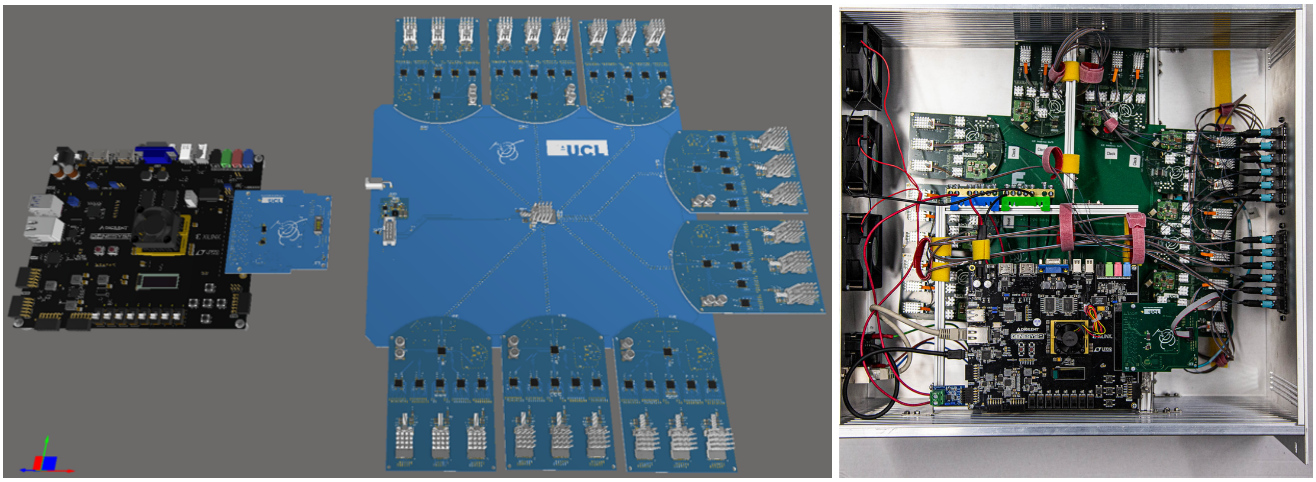


Figure 125: Left: A 3D representation of the full clock and reset distribution system. On the left side is the Genesys 2 board with the clock and reset FMC distribution board. On the right side is the active splitting motherboard with has an optical receiver (centre of the motherboard) that accepts the clock or reset lines from the FMC board via an optical fibre. The motherboard electrically routes the 8 signals to the fan-out daughter boards where each board generates 36 optical copies of the routed signal. Right: The full clock and reset system in a 19-inch rack-mountable box.

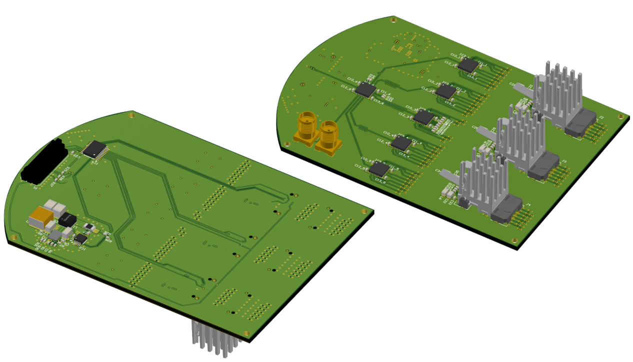


Figure 126: A 3D representation of the clock and reset daughter board. Left: Bottom view of the board showing the mezzanine connector and DC-DC circuitry. Right: Top view of the board showing the fan-out ICs and the three Firefly optical transceivers.

Figure 127: MIDAS page for control and monitoring of the clock and reset distribution system.

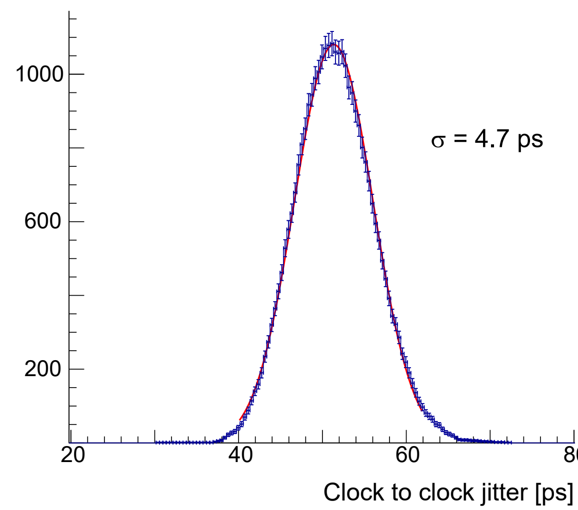


Figure 128: Rising edge time difference between two different clocks distributed via different daughter boards and different optical cable assemblies (leading to the $50\,\mathrm{ps}$ offset) as determined via a fast oscilloscope. The fit is a simple Gaussian.

16. Slow Control

The slow control system deals with all "slow" data such as high voltages for the SiPMs and silicon sensors, ambient temperatures and pressures, the beam line magnet settings and parameters of the cooling system. The configuration of the MuPix and MuTRiG ASICs is handled separately, as described in section 16.2.

For the slow control parameters it is important to have all data and control functionality in a homogeneous single system. This makes the maintenance of the system much simpler, since only a limited number of different hardware standards have to be taken care of. The integration of all data enables us to define control loops between otherwise completely different subsystems. Examples are regulating or switching off the detector power in the case of overheating of the pixel sensors or irregularities appearing in the helium cooling system, or adjusting the high-voltage on the basis of detector data such as energy spectra or hit rates.

The integration of all systems will be done through the MIDAS DAQ system, and as much as possible combined with the associated MIDAS Slow Control Bus (MSCB) system [54], which is discussed in section 16.1. In addition to the MSCB system, the MIDAS DAQ system receives and sends slow control data to the various layers of FPGAs and GPUs through the main fast data links (chapter 17). The slow control system also contains interfaces to the PSI beamline elements via the EPICS system [55]. This allows monitoring and control of the beamline from the main DAQ system, which has proven very versatile in other experiments using this scheme.

The full state of the system is kept in the MIDAS Online Data Base (ODB), and all slow control data is stored in the history system of the MIDAS system, so that the long term stability of the experiment can be effectively verified. The slow control data is also fed into the main event data stream, to be available for offline analysis.

All data fed into the MIDAS system is accessible by the MIDAS distributed alarm system. This system allows upper or lower limits to be set on all slow control data in a flexible way through the MIDAS web interface. In the event of an alarm, shift crews can be notified through spoken alarm messages and contacted via mobile phones. Scripts can be triggered which put the whole experiment in a safe state in order to avoid damage from excessive temperatures or other dangerous conditions. In addition to this MIDAS-based alarm system, an interlock system that is fully independent from the DAQ handles the most critical parameters of the apparatus (see section 16.3).

16.1. Midas Slow Control Bus

The MSCB system uses a serial differential bus for communication, with two data lines (positive and negative polarity) and a common ground. Over long distances, such as between crates, the physical standard for this bus is RS-485, running at a relatively low speed of 115.2 kbit/s in half-duplex mode. The slow speed makes this bus highly immune against improper termination or electrical interference, while the short commands of the MSCB protocol still allow the readout of many hundreds of nodes per second. This optimised protocol allows the monitoring of many thousands of channels with repetition periods in the 100 ms range, which is more than sufficient.

The MSCB bus uses a single-master, multiple-slave architecture, where all slave nodes on the bus only have to reply to requests sent by the master node, thus making the bus arbitration very simple. Many devices already exist for this system, such as the SCS-3000 units, as shown in Figure 129. Since the system was developed at PSI, it can be quickly adapted to new hardware.

The MSCB nodes inside the experiment are either dedicated 8-bit microcontrollers or soft-core microcontrollers instantiated on the FPGAs, connected to the RS-485 bus via insulated transceivers to avoid ground loops and noise. These microcontrollers perform local control loops, such as high-voltage stabilisation, power conversion or environmental control, and send measured values to the central DAQ system for monitoring. Custom high-voltage boards mounted inside the magnet have an embedded microcontroller acting as an MSCB node, thus no high-voltage cables have to be fed into the experimental volume (section 14.3). The DC-DC power converters (section 14.2.2) controller board also act as an MSCB node.

A dedicated slow control segment is connected to the environment sensors inside the magnet, monitoring parameters such as the temperature and pressure of the helium flows, the humidity inside the cage, the magnetic field at various positions and the temperature of various detector components. The monitoring data processed by the detector ASICs and the FPGAs will primarily be read out through the main data stream, with the MSCB-based readout as a backup system.

Microcontroller-based bus adapters are used to bridge between each RS-485 segment and optical fibres, which allows us to route all segments from inside the magnet to the outside world via optical fibres, where they are connected to the experiment Ethernet network. In this way, all MSCB nodes can be accessed from any computer connected to the experiment's Ethernet network.

16.1.1. Frontend board control

All front-end boards are connected to the MSCB bus via 3.3 V RS-485 optically isolated transceivers. Since the MSCB protocol is very simple, using only a few bytes for addressing, data and redundancy, its implementation requires less than 700 lines of C code. This makes it possible to run the MSCB core inside a NIOS II soft-processor on every FPGA used in the experiment.

Test implementations have shown that this needs only a few percent of the available FPGA resources, which can be easily accommodated. Having a dedicated slow control link to all FPGAs in the experiment is a powerful tool for debugging and configuration, since this allows the management of the FPGAs even if the optical data links are down.

16.2. ASIC Configuration

The configuration of the pixel detectors is a special case as it requires many millions of parameters, e.g. the tune-DAC values for each pixel. Since this amount of

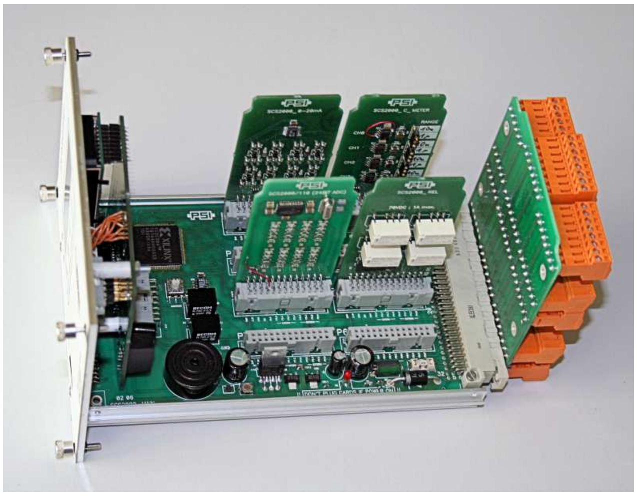


Figure 129: SCS-3000 unit as part of the MSCB slow control system. This unit has 64 input/output channels, which can be configured via plug-in boards as digital or analogue channels. Many plug-in boards exist already such as PT100 temperature sensor readout cards, analogue high resolution inputs (24 bit resolution), valve control outputs and many more.

data is considerably larger than the total for all other systems (∼120 MB for the full phase I detector), an extension of the slow control system is implemented. A dedicated program manages, visualises and exchanges the pixel detector configuration parameters between an optimised database and the pixel hardware. In this way the time required to configure the pixel detectors can be minimised, while this program is still connected to the main DAQ system. It can be synchronised with run starts and stops, and can inject pixel monitoring data periodically into the event data stream for offline analysis. The regular slow control data stream contains a pointer to the relevant state of the pixel configuration database.

The configuration of the individual pixel sensors is written via a dedicated differential configuration bus with up to 9 sensors (one electrical group in the outer layers) connected in parallel. This corresponds to approximately 20 Mbit of configuration data, which in turn dictates the need for configuration speeds above 10 MHz in order to guarantee fast run starts. Sensors on different ladders (different configuration buses) can and have to be programmed in parallel. Slow control data output from the sensors is sent using the fast LVDS data link. As the on-chip memory of the front-end FPGA is too small to hold the complete configuration data for all connected sensors, it has to be delivered just in time from the switching boards (see section 17.4). Bandwidth is not an issue (a 6 Gbit/s optical downlink is available), but the data stream has to be synchronised such that no buffering on the front-end is required, which necessitates a careful interplay between the software driving the data into the switching boards, the switching board firmware and the front-

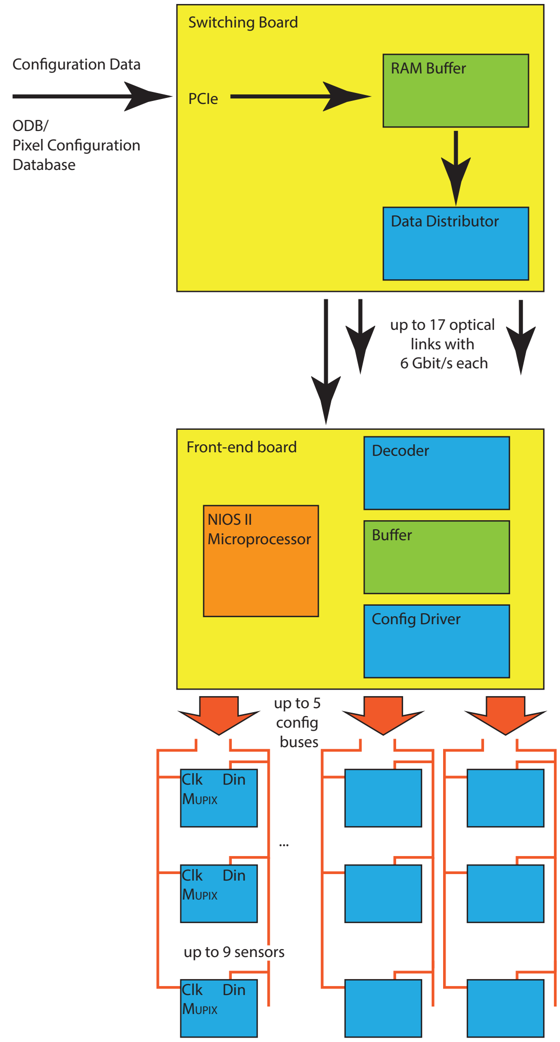


Figure 130: Data flow for the front-end ASIC configuration for the pixel detectors.

end firmware. An overview of the data flow for the pixel configuration is shown in Figure 130.

For the timing detector MuTRiG ASICs, the same configuration path is used - the configuration data is however much more compact than in the pixel case and can be stored in the MIDAS ODB.

16.3. Interlock System

As 10 kW is dissipated in a small volume, continuously carried away by water and helium cooling systems, an additional interlock system fully independent from the MIDAS DAQ controls the critical parameters of the experiment. This system returns the experiment to a safe state in case of an emergency or critical failure, and prevents unsafe transitions between operating modes requested by the user. For example, the detector power can only be turned on when the helium and water cooling systems are fully functional. The central controller of the interlock system is a commercial pro-

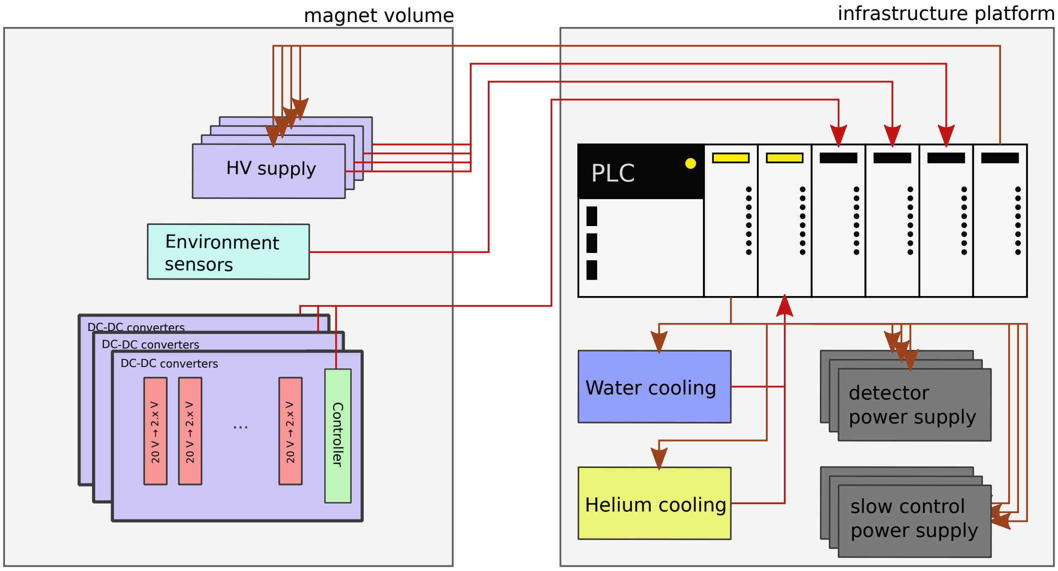


Figure 131: Layout of the Mu3e interlock system, directly interconnecting the safety critical sub-systems of the experiment. The PLC is a Siemens SIMATIC S7-series controller.

grammable Logic Controller (PLC) with fail-safe IO. Figure 131 shows a conceptual wiring diagram of this system. It is located on one of the infrastructure platforms, and connected to the subsystems via closed-loop electrical circuits, which are decoupled from the detector and other slow control power circuits. This system is designed as an additional safeguard; during normal operation all transitions are instigated by the MIDAS DAQ system.

17. Data Acquisition

The Mu3e data acquisition (DAQ) system works without a hardware trigger on a push basis, i.e. the detector elements continuously send zero-suppressed hit information. The DAQ consists of three layers, namely front-end FPGAs, switching boards and the filter farm. The topology of interconnects is such that every farm PC receives the complete detector information for a certain time slice. See Figure 132 for an overview of the readout scheme.

Hits in all subsystems are timestamped and the front-ends ensure that time-ordered information is forwarded to the rest of the readout system. At the input to the farm PCs, data from several timestamps is merged to form overlapping reconstruction frames, as shown in Figure 133. In this scheme, the latency of individual detector elements is not critical, as long as the latency differences do not exceed the buffering capacity at each step.

17.1. Bandwidth Requirements

The bandwidth requirements of the data acquisition are largely determined by the expected detector occupancy, as all the Mu3e subdetectors produce zero-suppressed output.

Occupancies have been estimated with the full simulation for a rate of muons stopping on target of $1 \cdot 10^8$ Hz, and pessimistically estimating the beam-related background by assuming another $0.9 \cdot 10^8$ Hz of muons stopping along the last metre of beam line.

17.1.1. Front-end bandwidth requirements

The pixel sensors contain electronics for hit detection, as well as time and address encoding. The hits are then serialised and sent to the front-end FPGA board via a 1250 Mbit/s low voltage differential signalling (LVDS) link.

The sensors at the centre of the innermost layer have the highest occupancy, about 1.3 MHz/cm^2 or 5.2 MHz per sensor. The protocol implemented in the MuPix 8 prototype and all subsequent chips allows a maximum of 74% of the available time slots for sending hit information. With 8 bit/10 bit encoding, this leads to a maximum hit bandwidth of 740 Mbit/s, equivalent to $23 \cdot 10^6$ 32 bit hits per link per second. This gives a safety factor of four even for the busiest sensors, which will use three parallel links. The total bandwidth requirements for the phase I pixel detector up to the front-end boards are shown in Table 15.

The average occupancy determines the bandwidth requirements, but fluctuations are also modelled in the simulation in order to optimise the system design. In particular, online buffer sizes must be large enough to allow the latency required to absorb the highest expected peaks in hit rate.

The MuTRiG ASIC foreseen for both timing detectors will also output zero-suppressed hit data with timestamps over a 1250 Mbit/s LVDS link. The average hit rate per channel of the fibre detector is estimated from the simulation as 620 kHz, with a hit size of

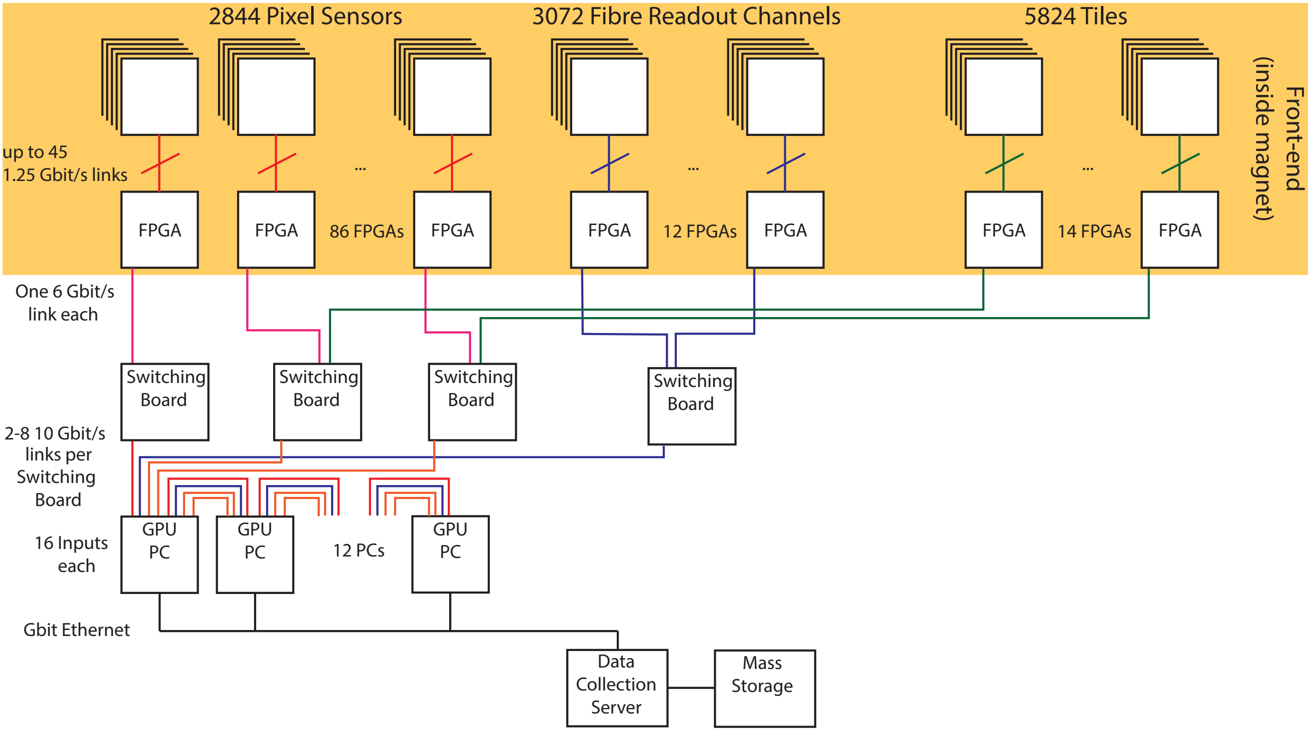


Figure 132: Overall Mu3e readout scheme.

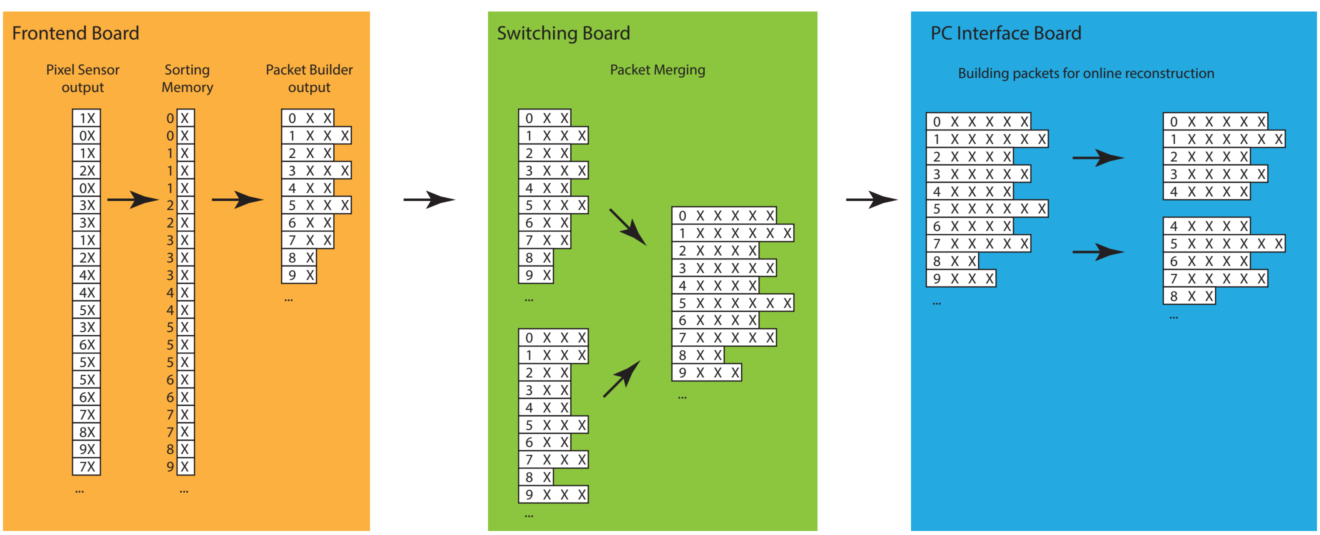


Figure 133: Schematic flow of pixel time information through the Mu3e readout system. Numbers stand for hit timestamps, X stands for the remaining hit information (address, charge).

28 bits [41]. With 32 channels per ASIC, this uses about 700 Mbit/s of the link bandwidth, limiting the acceptable dark count rate to roughly 300 kHz per channel.

The tile detector, operating at a relatively high threshold and an expected total hit rate of roughly 180 MHz, will contribute very little to the overall bandwidth requirements and is very far from saturating single channel limits.

17.1.2. Optical link bandwidth requirements

The hits are collected on a front-end FPGA and transmitted off the detector using optical links. The corresponding bandwidth requirements are listed in Table 16. For the fibre detector, clustering is assumed to take place on the front-end FPGA, although unclustered data could be sent out using twice the number of front-end boards with two optical links each.

Four switching boards will collect the data from the front-ends, one for the central pixel detector, one each for the up- and downstream recurl stations (pixels and

	Sensor Chips	Max Hits /Chip/s 10^6	Average Hits /Layer/s 10^6	Chip→FPGA link capacity Mbit/s needed/available	Chip→FPGA total in Layer Gbit/s needed/available	Front-end FPGAs
Layer 1	48	5.2	194	281/3750	10.5/180	8
Layer 2	60	5.2	195	281/3750	10.5/225	
Layer 3	408	1.2	266	65/1250	14.4/510	12
Layer 4	504	1.2	248	65/1250	13.4/630	14
Recurl IU	408	0.15	41	8.1/1250	2.2/510	12
Recurl OU	504	0.14	44	7.6/1250	2.4/630	14
Recurl ID	408	0.11	28	5.9/1250	1.5/510	12
Recurl OD	504	0.10	29	5.4/1250	1.6/630	14
Total	2844		1045		56.4/3825	86

Table 15
Pixel front-end readout requirements (10^8 muon stops/s). The recurl station layers are labelled by inner/outer (I/O) and up- and downstream (U/D). The rates include protocol overhead and 8 bit/10 bit encoding, and assume $32\,\text{bit}$ hit size.

Subdetector	Maximum rate/FPGA MHz	Hit size Bits	Bandwidth needed Gbit/s	FPGAs
Pixels	58	48	4.6	86
Fibres	28	48	2.3	12
Tiles	15	48	1.2	14

Table 16
FPGA bandwidth requirements. For the fibre detector, clustering in the front-end FPGA is performed. For the bandwidth, $75\,\%$ protocol efficiency and 8 bit/10 bit encoding are assumed. The pixel hit size assumes, conservatively, that the full hit and address information including time is transmitted for each hit. This can be reduced by time-grouping hits and encoding parts of the address in the link.

	Rate MHz	Bandwidth Gbit/s
Central Pixels	905	58
Upstream Recurl	85 + 106	12
Downstream Recurl	58 + 73	8.4
Fibres	337	21.5
Total	1564	100

Table 17
Switching board bandwidth requirements. $48\,\text{bit}$ hit size and $75\,\%$ protocol efficiency are assumed.

tiles) and one for the fibre detector. The corresponding bandwidths passing through these boards are listed in Table 17.

17.2. Front-end FPGA Boards

The front-end boards have to collect the data sent from either the MuPix or the MuTRiG chip, sort and package it, and then forward it to the switching boards on a fast optical link. In the case of the fibre MuTRiG data, preliminary clustering will be applied in order to reduce the data rate taken up by dark counts. In addition, the boards have to provide the

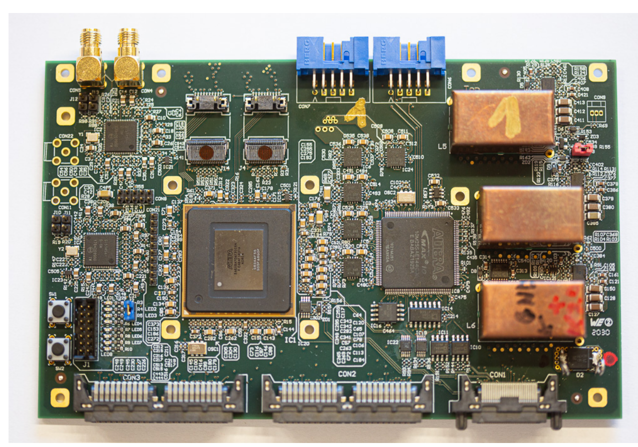


Figure 134: Prototype front-end board based on an Intel Arria V FPGA. The FPGA in the centre left is surrounded by connectors to the crate backplane (leading to the detector ASICs) at the bottom, a Intel MAX10 CPLD for configuration and monitoring in the centre right, clocking circuitry at the the left, two connectors for Firefly optical transceivers on the top left, blue JTAG connectors for programming on top and the DC/DC converter circuitry on the right. The copper boxes contain and shield the air coils.

sensors with control signals and monitor the environment. The space constraints inside the magnet necessitate small, highly integrated boards incorporating FPGAs and optical modules with a small footprint and limited power consumption. A working prototype can be seen in Figure 134.

The boards feature an Intel Arria V FPGA (5AGXBA7D for data processing as well as a flash-based Intel MAX10 FPGA (10M25SAE144C8G) for configuration and monitoring. For the optical data transmission we use Firefly transceivers by Samtec (ECUO-B04-14), each of which provides four transmitting and four receiving links at up to 14 Gbit/s in a very small footprint (20.3 mm×11.25 mm and a power consumption of roughly 1 W. A single link per board is sufficient for the bandwidth requirements of phase I; we nevertheless foresee the option to install two Fireflys and thus obtain 8 outgoing links. The incoming links are used for the clock and reset distribution (see chapter 15) as well as the slow control and pixel configuration. Clocks received by the Fireflys are conditioned by two Si5345 jitter attenuator/ clock multiplier chips and forwarded to the FPGAs as well as the detector ASICs. The front-end firmware receives detector data, performs time-sorting and multiplexes the data from all connected ASICs to an optical link. Synchronisation and run transitions are controlled by the reset link, as described in section 17.6.2.

The boards are connected to a backplane, which forwards the detector signals and provides control and monitoring signals via a separately powered crate controller. The boards are cooled by a custom-made aluminium cooling plate connected to the water-cooled frame of the crate via a heat pipe.

17.2.1. Slow-control and configuration integration

For slow control and pixel configuration data, two paths are foreseen. Firstly, surplus bandwidth on the optical links to and from the switching board can be used, which is especially useful for large volume data such as pixel tune values. Secondly, a separate differential line for use of the MSCB protocol (see chapter 16) is foreseen for monitoring the status of optical links, switching power, etc.. The interface to MSCB and the slow control-related tasks on the FPGA will be implemented in a NIOS II soft processor core [56] on the FPGA [57].

The FPGA firmware can be updated by writing a Serial Peripheral Interface (SPI) flash memory from either the optical slow control link or the MSCB connection via the MAX10 FPGA. On power-up or a reconfiguration command, the MAX10 then reprograms the Arria V FPGA.

17.3. Read-out Links

Electrical links are used between the detector ASICs and the front-end FPGAs, all other data links are optical. The data links are complemented by a (smaller) number of slow control links in the opposite direction [58].

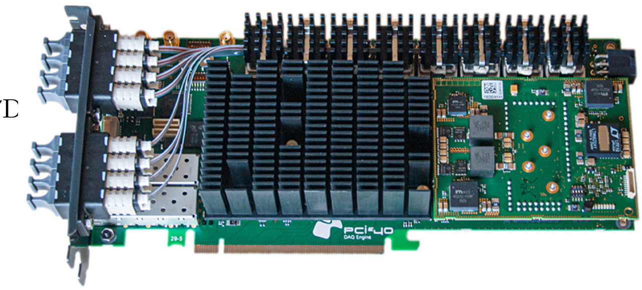


Figure 135: *PCIe40* board as developed for the LHCb and ALICE upgrades [61, 62] and employed as switching board in Mu3e. The large Arria 10 FPGA with two PCIe Gen3 8 lane interfaces is complemented by 48 fast optical receivers and 48 fast optical transmitters.

The data from the MuPix and MuTRiG chips will be transmitted to the front-end FPGAs via LVDS links at 1250 Mbit/s. The link is physically implemented as a matched differential pair of aluminium traces on the sensor HDI, followed by a micro twisted-pair cable connected to the detector side of the backplane, see section 7.1.2.

There are two types of optical high speed data links. The first one goes from the front-end FPGAs to the switching boards, the second from the switching boards to the FPGA PCIe boards in the event filter farm PCs. The optical links from the front-end FPGAs to the switching boards have a bandwidth of 6 Gbit/s, which fits well with the FPGA specifications. Each FPGA has nine fast transceiver blocks, which connect to the Firefly optical assemblies. The laser has a wavelength of 850 nm and the optical fibre is of 50/125 multi-mode OM3 type, since this is a standard both in industry and in particle physics detector readout.

The links from the switching boards to the filter farm are implemented as 10 Gbit/s high speed links. The PCIe FPGA board is fitted with four quad small form-factor pluggable (QSFP+) optical modules.

All the links have been tested using the development hardware and were found to have bit error rates low enough for stable and consistent running of the experiment [59, 60]; typically, no errors were found in a few days of running, leading to upper limits on the bit error rate of $1 \cdot 10^{-14}$ down to $1 \cdot 10^{-16}$.

17.4. Switching Boards

The main task of the switching boards is to act as switches between the front-end FPGAs on the detector and the online reconstruction farm, thus allowing each farm PC to see data from the complete detector. The board design and choice of FPGAs is dominated by the number of fast links required.

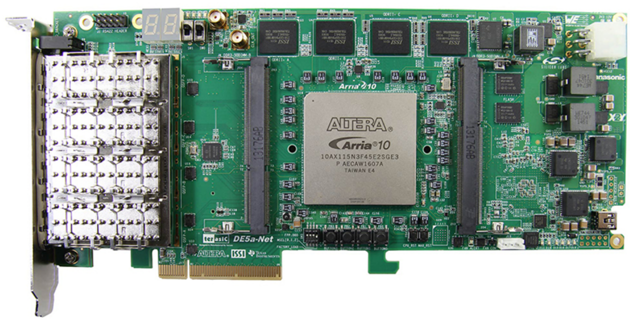


Figure 136: The DE5a NET board built by Terasic Inc. and used as a receiving board in the filter farm PCs. The Arria 10 FPGA with a PCIe Gen3 8 lane interface is complemented by a 16 duplex fast optical links and several GB of DDR4 memory.

We use four *PCIe40* boards (see Figure 135) developed for the LHCb and ALICE upgrades at the LHC [61, 62, 63]. These boards provide up to 48 full duplex optical links at up to 10 Gbit/s, plus two eight lane PCIe 3.0 interfaces bundled to a sixteen lane interface by a switch. The FPGA is an Altera Arria 10. The PCs hosting the boards are used to store and transmit the extensive pixel configuration and tuning data as well as the timing detector ASIC configuration via PCI express, and link the boards to the experiment control and monitoring system via standard Ethernet.

The firmware for the switching board has to receive several data streams in parallel, merge them synchronously and then forward them to the event filter. Additional firmware is needed in the case of the fibre tracker, where hits from both ends of the fibre have to be matched in order to suppress dark counts.

17.5. Event Filter Interface

The filter farm PCs are equipped with FPGA boards in PCIe slots and optical receivers. The boards are commercial DE5a NET boards (see Figure 136) built by Terasic Inc. They are equipped with four QSFP+ quad optical transmitters/receivers, two laptop-compatible DDR3/DDR4 memory interfaces and a large Altera Arria 10 FPGA with an 8 lane PCIe 3.0 interface. This FPGA performs the event building and buffering, and also allows simple clustering, sorting and selection algorithms to be run. The event data is then transferred via Direct Memory Access (DMA) over the PCIe bus[26] to the main memory of the filter farm PC and subsequently copied to the memory of a GPU, where the fitting and vertex selection algorithms are run. The GPU then posts IDs of selected events to the main memory of the PC, which triggers a transfer of the respective data from the FPGA buffer memory via the PC main memory and Ethernet to the central DAQ computer running the MIDAS software. At that computer, the data streams from the farm PCs are combined into a single data stream, combined with various slow control data, compressed and stored. The maximum data rate over an eight-lane PCIe 3.0 bus is 7.88 Gbyte/s of which we are able to use 4.8 Gbyte/s for user data, amply sufficient for phase I.

The GPU boards will be obtained commercially as late as possible in order to profit from the ongoing rapid development and sinking prices. Current high-end GPUs already have enough floating point capability for high rate running. Newer boards are, however, expected to offer higher memory bandwidth and better caching. For example, between the GTX 680 and the GTX 980 GeForce GPUs, both the compute power and the copy speed increased by 30 % [64, 65]. The GTX 1080Ti cards we obtained in 2017 were sufficient to run the full Phase I selection load on 12 nodes [66].

The farm PCs are hosted in individual rack-mounted tower casings, ensuring enough space for the FPGA board, the high-end GPU and a custom clock receiver board [67] whilst allowing for air cooling. Each tower consumes around 0.7 kW, so active cooling of the racks and the counting house is necessary.

17.6. Run Control, Data Collection and Storage

17.6.1. The MIDAS System

The filter farm outputs selected events at a data rate of the order of 50-100 MBytes/s in total. This data rate is low enough to be collected by a single PC connected to the filter farm by common Gbit Ethernet and written to local disks. Then the data is then transferred to the central PSI computing centre, where it is stored and analyzed. For the central DAQ, the well-established MIDAS (Maximum Integrated Data Acquisition System) [68] software package is used. This software is currently used in several major experiments such as the T2K ND280 detector in Japan [69], ALPHA at CERN and the MEG experiment at PSI [70]. It can easily handle the required data rate, and contains all necessary tools such as event building, a slow control system including a history database and an alarm system. A web interface allows the control and monitoring of the experiment through the Internet. The farm PCs use MIDAS library calls to ship the data to the central DAQ PC. The framework also offers facilities to send configuration parameters from a central database (the "Online DataBase" or ODB) to all connected farm PCs and to coordinate common starts and stops of acquisition (run control).

For the purpose of monitoring and data quality control of the experiment, the MIDAS system offers the capability to tap into the data stream to connect analysis and graphical display programs.

[26] Note that PCIe is actually not a bus protocol, but offers switched point-to-point connections. The *bus* designation is due to the software-side backwards compatibility to the original PCI bus interface.

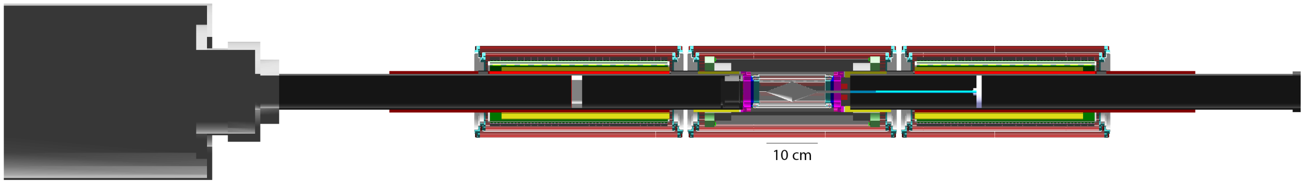


Figure 137: Side view of the simulated phase I detector cut along the beam axis.

17.6.2. Run start/stop synchronisation

In traditional DAQ systems, starting and stopping is controlled by enabling and disabling trigger signals. In a streaming system such as in Mu3e this is not an option, and great care has to be taken to synchronise data across the complete detector at run start and ensure that the frame numbers in all subsystems are in agreement.

To this end, a global reset signal is distributed together with the global clock. At the front-end, the reset signal is forwarded to the pixel sensors and there sets the timestamp counters to zero as long as it is on (note that the pixel sensors cannot be inactivated, so even during a reset they will still collect, process and send hits, however all with timestamp zero). At the start of a run, the reset signal is released synchronously for all sensors, which then start counting timestamps. The front-end firmware will ignore all hits with timestamp zero at the beginning of the run and start sending packets into the switching network as soon as non-zero timestamps arrive. All subsequent stages in the network then synchronise on the first packet and from then on stay in sync using consistent packet numbering. A similar synchronization mechanism is implemented for the MuTRiG ASICs.

At the end of the run, the global reset goes high and the front-end continues forwarding packets until timestamp zero is detected.

18. Simulation

This chapter describes the Geant4 [71, 72] based simulation used to study and optimise the detector design, to develop the reconstruction code and to estimate signal efficiency and background rates.

The Mu3e software stack consists of the simulation described here, which includes generators for many different muon decays, the track reconstruction described in the following chapter, a vertex fit program and a range of analysis codes. Besides Geant4, root [73], which is used for storage, histogramming and related analysis tasks, is the other mayor external dependency. Core code is written in C++, python is used for some of the analysis and plotting code.

18.1. Detector Geometry

The simulated detector geometry closely follows the planned detector geometry described in earlier chapters. The simulated volume extends for three metres in all directions from the target centre. The magnet metal and the surrounding volume is only used in the cosmic ray simulation. Figures 137 and 138 show the simulated detector geometry.

18.1.1. Beam Delivery

In the detector simulation, the beam starts 1 m in front of the target inside the beam pipe. Beam particles are generated with a profile and momentum spectrum taken from the beam simulation at the same point. The beam passes a 600 µm Mylar moderator followed by thick lead collimator removing particles undergoing large angle scattering in the moderator. It then exits the beam vacuum through a 35 µm vacuum window, the holding structure of which serves as the final collimator.

18.1.2. Target

The target is simulated as a hollow Mylar double cone supported from the downstream side by a thin carbon fibre tube, see also chapter 6.

18.1.3. Pixel Detector

The simulated geometry of the pixel detector includes the sensor, the flexprint (with an average trace density assumed and represented as thinner metal layers) and the polyimide support structure. The plastic end-pieces and support wheels are also simulated in detail, including flex prints, interposers and screws.

18.1.4. Scintillating Fibres

The fibre ribbon simulation implements fibre shape, cladding thickness, staggering as well as optional fibre coatings and glue. The fibres are matched to SiPM arrays at both ends. Parameters of the geometry can be easily changed to study different options.

The baseline setting consists of 12 ribbons in 3 layers. Each layer consists of 128 round 250 µm thick fibres read out by SiPM arrays with 250 µm column width.

18.1.5. Tile Detector

In the simulation, the tile Detector consists of the scintillating tiles, SiPMs, two layers of PCBs hosting the SiPMs and the readout chips, respectively, as well as the support structure.

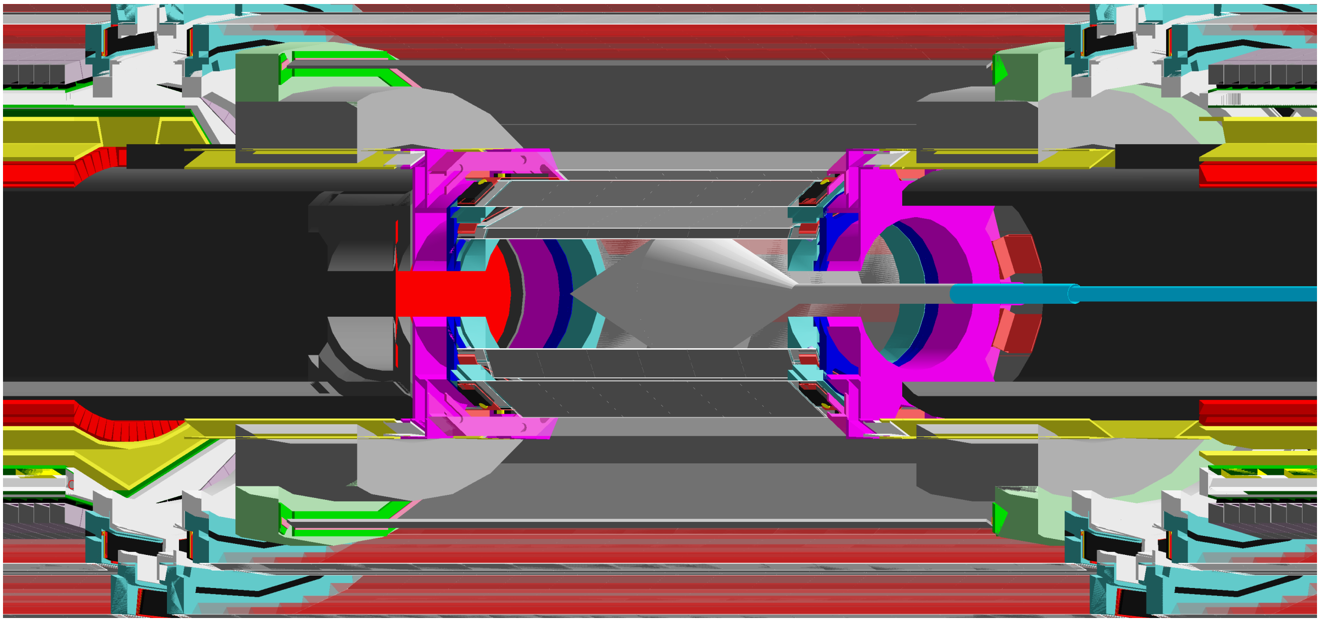


Figure 138: Perspective view of the central part of the simulated phase I detector cut open at $x = -19.0\,\mathrm{mm}$.

18.2. Magnetic Field

The magnetic field in the simulation is taken from an azimuthally symmetric field map with 10 mm step size calculated by the magnet manufacturer. Linear interpolation is used between the field map grid points. It will be replaced by a measured map as soon as this becomes available.

18.3. Physics Processes

18.3.1. Multiple Coulomb Scattering

Multiple Coulomb scattering is the main limiting factor for the resolution of the experiment; an accurate simulation is thus crucial. Per default, we use the Urban model [74] as implemented and recently improved [75] in Geant4. Alternative models are also implemented, including one derived from the results of a dedicated study of multiple Coulomb scattering in thin silicon at the DESY electron test beam [76]. This however still needs to be validated at the low momenta expected in Mu3e.

18.3.2. Muon Decays

Geant4 implements the Michel decay including polarization of both the muon and the positron based on Scheck and Fischer [77, 78]. The radiative decay of the muon in Geant4 was implemented by the TWIST collaboration [79] based on Fronsdal et al. [80]. This code has been adapted for the simulation of the Mu3e experiment using the differential branching fraction provided by Kuno and Okada [3].

A unified description of radiative corrections for photons below a soft cut-off and photons that are tracked within Geant4 above a cut-off based on calculations by Fael, Mercolli and Passera [10] has been implemented.

The radiative decay with internal conversion is simulated using the matrix element of Signer et al. [7], with the option of using the NLO version [7, 8] when the accuracy is required.

Signal The signal kinematics are highly model-dependent, see chapter 1. If not stated otherwise, we have used three particle phase space distributions in the simulation, following the practice of SINDRUM and earlier experiments. We have also implemented the general matrix element by Kuno et al. [3] in order to study the kinematics of different decay dynamics.

Special Decays In order to study accidental background whilst factoring out timing and vertex suppression, the simulation code allows for more than one muon to decay at a single vertex. This is beneficial for studying the overlap of an internal conversion and a Michel decay.

Cosmic Muons As the detector alignment will rely in part on the high momentum tracks of cosmic ray muons, we have implemented a cosmic muon generator based on the spectrum and angle parametrisation of Biallass and Hebekker [81].

18.4. Time Structure and Truth Information

The Mu3e experiment operates with a quasi continuous beam, which has necessitated adaptations of the Geant4 package in order to take into account particles crossing boundaries of reconstruction frames. For every interaction with active detector material, both the particle of origin and the sequence of interactions is saved, we thus have the full simulation truth available at every level of reconstruction and analysis.

18.5. Detector Response

18.5.1. Pixel Detector Response

The response of the pixel detector can be simulated by either setting a threshold on the charge deposited or by defining a single hit efficiency, which is then applied by randomly discarding hits. Noise is simulated by randomly creating extra hits at an adjustable rate.

The simulation does include effects of charge sharing between pixels. δ-electrons are simulated if they have a range above 50 µm, the associated (large) clusters should thus be correctly simulated. The response simulation is constantly adapted to the measured properties of the pixel sensors.

Pixel Readout Simulation The readout of the pixel detector is not strictly in order of timestamp (see chapter 8) and very large clusters of hits can lead to overflows in the sorting algorithm on the front-end FPGA (see section 17.2). These effects are simulated by treating each column as a queue, into which hits are pushed at creation. A fixed number of hits is then removed from these column queues for every time slot. Hits are time-sorted in a separate programme and those that are too far out-of-time are dropped; alternatively we can run a bit-accurate simulation of the front-end board firmware. We do currently not simulate the dead-time caused by hits stored in the pixel cell.

18.5.2. Fibre Detector Response

In a first step the response of the scintillating fibres to an incident particle is simulated. Since simulating single photon propagation inside fibres is not feasible in the main simulation, the response of the scintillating fibres is parametrised. The number of arriving photons at both fibre ends can either be parametrised in deposited energy (E_{dep}) and hit position or simply generated according to measured efficiencies.

In a second step the SiPM response to the arriving photons is simulated and the distribution of photons into the different SiPM cells is modelled. The main parameter for this process is the photon distribution at the fibre ends and propagation in the epoxy layers before the SiPM active layer as well as optical cross-talk. The SiPM response depends on an adjustable photon detection efficiency (PDE) and is mixed with a constantly present dark rate and its own pixel to pixel cross-talk. The time distribution of the detected events bases on the measured time resolution of fibre ribbons and photon time of flight in the fibres. In a last step pile-up events are merged.

18.5.3. Tile Detector Response

The tile detector will record the timestamps of the scintillation signals, as well as the energy deposition in the tiles, which is proportional to the number of scintillation photons. In the simulation, the scintillation process and photon propagation is not simulated, in order to maintain a reasonable computation time. Instead, the response characteristics of the tile, including the readout electronics, is parametrised, using the true timestamp and energy deposition of a hit as an input. The response is described by the following parameters: time resolution, energy resolution, jitter of the readout electronics, channel dead-time and energy threshold.

In order for a signal in the tiles to be detected by the readout electronics, a minimum energy deposition is required. This corresponds to the energy threshold of the MuTRiG chip, which is assumed to be roughly $E_{thresh} = 0.1\,\mathrm{MeV}$. Due to the linearised ToT method implemented in the MuTRiG chip, the digitised energy information is approximately proportional to the energy deposition in the tile. The energy deposition of consecutive hits (pile-up events) which occur within the dead-time of the channel is assigned to the original hit. This reflects the behaviour of the MuTRiG chip. The channel dead-time is determined by two parameters: the intrinsic dead-time of the MuTRiG TDC and the dead-time related to the ToT of the analogue input signal. The time resolution is parametrised by the intrinsic jitter of the MuTRiG chip and the energy dependent resolution of the tile.

19. Reconstruction

The reconstruction algorithm has to efficiently identify the tracks of particles from muon decays, while dealing effectively with the combinatorial background to keep the rate of incorrectly reconstructed tracks to an acceptable level. The main challenges are the high event rate and resulting occupancy, and the curvature of trajectories of low momentum particles in the 1 T magnetic field. Particle trajectories can make several turns in the detector, and hit combinations can span distances of more than half a meter with hits on opposite sides of the detector. This is of particular importance for the determination of the direction of travel and therefore the charge of the particle, as fully reconstructing tracks is critical in correctly applying the information from the timing systems.

As the detector readout is triggerless, all muon decays have to be fully reconstructed in the filter farm, setting high demands on the speed of the online track reconstruction algorithm.

Multiple Coulomb scattering (MS) in the detector layers is the dominant uncertainty. The track finding and initial fitting is thus built around a fast three-dimensional MS fit, which is based on fitting the multiple scattering angles at the middle hit position in a hit triplet combination (see [82] for a detailed description). In the most basic implementation of the fit, spatial uncertainties of the hit positions are ignored. This is a good approximation in the case of Mu3e, as the pixel resolution uncertainty ($80/\sqrt{12} \approx 25\,\mu\mathrm{m}$) is much smaller than that from multiple scattering (typically several hundred µm).

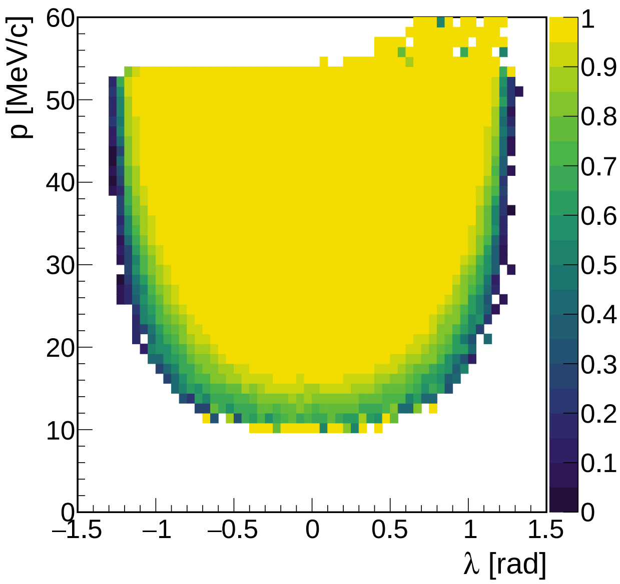


Figure 139: Ratio of reconstructed short tracks to generated particles producing a hit in each of the four detector layers as a function of momentum p and angle λ. The entries at high momentum in the forward direction are from decays in flight.

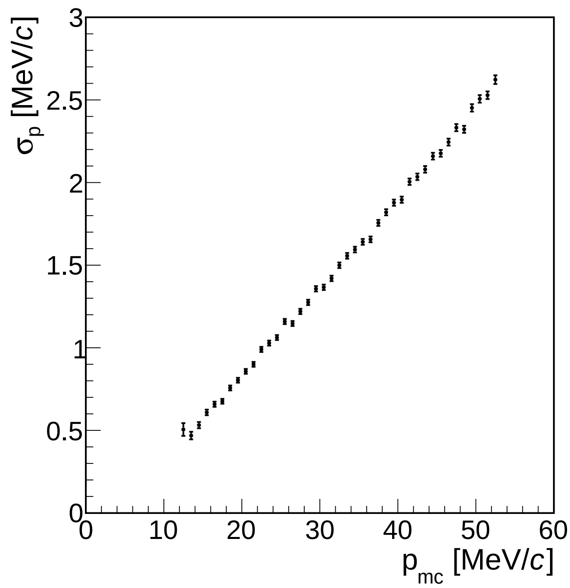


Figure 140: Momentum resolution σ_p as a function of the generated momentum p_{mc} of short tracks.

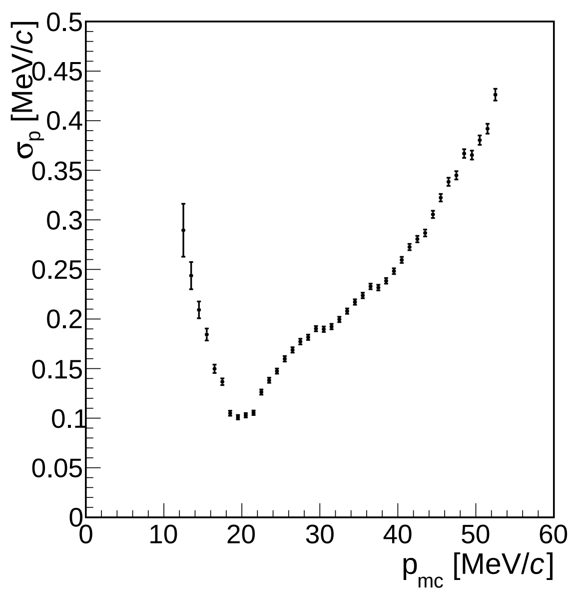


Figure 141: Momentum resolution σ_p as a function of the generated momentum p_{mc} of 6-hit long tracks. The momentum resolution has a minimum for tracks that traverse exactly half a circle outside the outermost pixel layer.

In order to achieve the best possible resolution, a general broken line (GBL) fit [83, 84] can be used. This technique determines hit positions and scattering angles simultaneously and also incorporates energy loss in the detector material, but requires knowledge of the assignments of hits to tracks from a preceding linking step as well as an approximate track trajectory. Therefore, it can only be used as a final step. Currently a GBL fit is used for detector alignment (see chapter 21), and it will be used in offline analysis.

The track finding and fitting studies presented in the following are all based on a fast MS fit that also implements energy-loss corrections and takes into account hit position uncertainties. Events are generated with the full Geant4 simulation (see chapter 18). The beam intensity is set such that $1 \cdot 10^8$ muons decay in the target region per second, corresponding to an optimistic estimate for the rate achievable at πE5. In theses studies the track reconstruction is performed on 50 ns non-overlapping frames. Studies of the tracking acceptance and efficiency are performed on a sample without simulated signal decays.

19.1. Track finding

In the first step, triplets of hits in the first three layers consistent with tracks originating from the target are identified. These triplets are fit with the fast MS fit and if the fit χ^2 is sufficiently good, they are extrapolated to the fourth layer, where the presence of an additional hit compatible with the triplet is required. Again a fast MS fit is performed and a χ^2 cut applied.

The resulting short tracks, with four hits each, are the input for the vertex fit in the online reconstruction; see Figure 139 for the reconstruction efficiency and Figure 140 for the momentum resolution of 4-hit tracks.

For the full offline reconstruction, the short tracks are extended to long tracks with 6 hits (forward and backward going tracks) or 8 hits (tracks close to per-

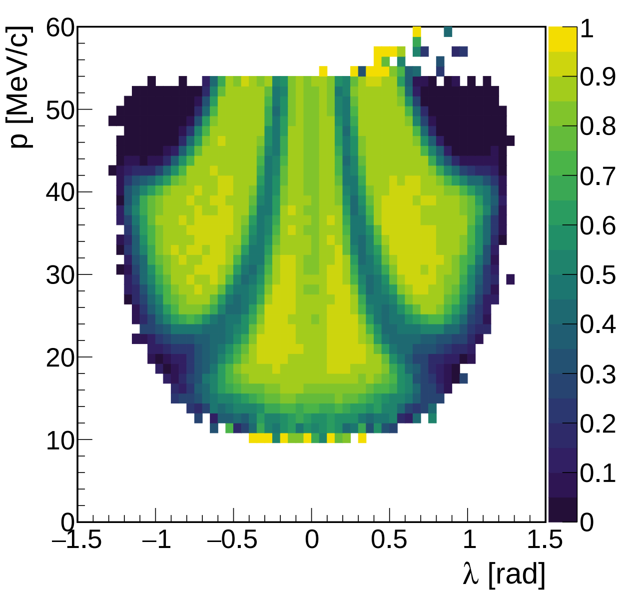


Figure 142: Number of reconstructed long tracks relative to the number of short tracks as a function of momentum p and inclination angle λ. The entries at high momentum in the forward direction are from decays in flight.

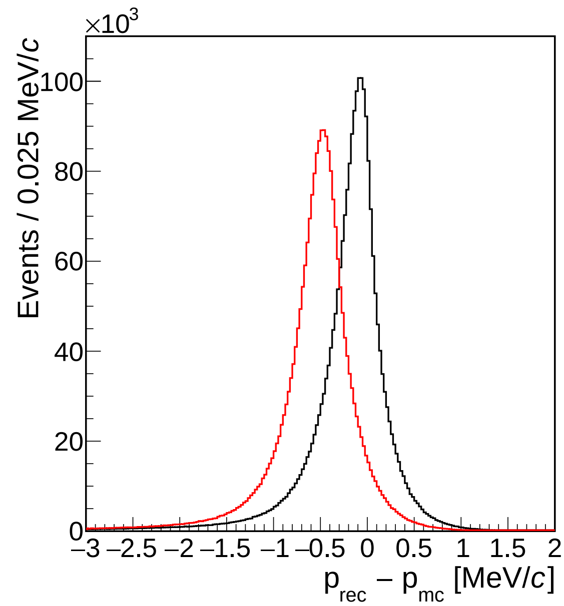


Figure 143: Difference between reconstructed and generated momentum after the first detector layer, for long tracks with (black) and without (red) energy loss correction.

pendicular to the beam, passing the vertex layers repeatedly) incorporating the recurling parts of the track. These long tracks have a much larger lever arm for momentum measurement and thus provide a much enhanced momentum precision, as shown in Figure 141. With the phase I detector setup, there is however a limited acceptance to see the recurling part of the track for low polar angles and also in the gaps between the central part and the recurl stations, as shown in Figure 142.

Additional algorithms are designed to remove incorrectly reconstructed tracks. One algorithm is performed after the reconstruction of short tracks. A graph is constructed where nodes represent tracks and edges correspond to intersections (common hits) between those tracks. A subset of nodes is selected that maximises the number of unconnected nodes (i.e. the maximum number of tracks that do not share hits). A second algorithm is run after the reconstruction of long tracks. Chains of long tracks are constructed where each pair of long tracks shares a short segment. These chains are required to have no intersections, with a maximum length and minimum total χ^2. In very dense regions with many recurling tracks, machine learning techniques can be used to correctly identify the sequence of track segments [85].

19.2. Energy loss

For long tracks, the momentum resolution becomes comparable with the total energy loss suffered by particles traversing the detector, resulting in an observable shift in the momentum. The energy loss correction is implemented by adjusting the curvature of each helix according to the sum of the most probable energy losses in each layer passed by the particle up to that point in the helix. Figure 143 shows the momentum resolution for long tracks before and after the implemented energy loss correction.

19.3. Timing Detectors

Reconstructed tracks are extrapolated to the fibre and tile detectors and the closest hits (within a maximum distance) are assigned to the tracks. The timing from fibre hits allows the determination of the direction of rotation (and thus the charge) of recurling particles. Figure 144 shows the time versus distance determined from two linked clusters of fibre hits for all 8-hit tracks, and demonstrates the potential for charge identification by timing.

The tile hits have the best timing resolution, providing an important constraint on the accidental combinatorial background by allowing timing information to be compared for different tracks assigned to a single vertex. The efficiency of the assignment of tile hits to tracks that pass through the tile detector is 98%.

20. Online Event Selection

The full data rate of the detector needs to be greatly reduced before permanent storage – only physically relevant event candidates can be kept. Requiring three tracks coincident in time is not sufficient to reduce the data rate by three to four orders of magnitude. Con-

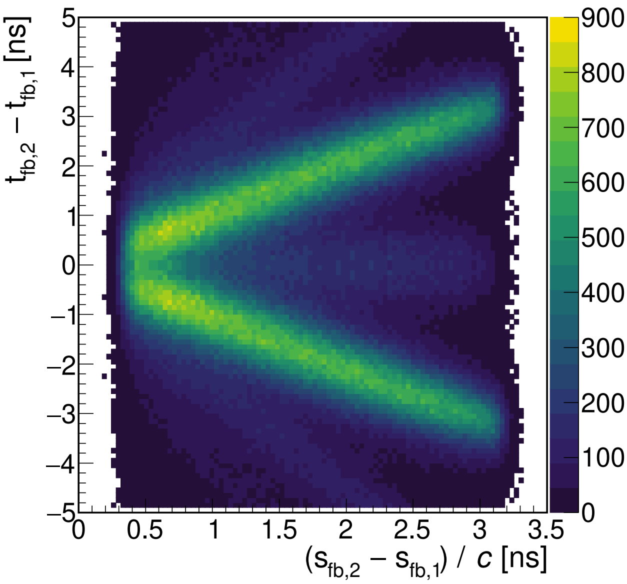


Figure 144: Time difference between fibre clusters assigned to 8-hit long tracks as a function of distance along the trajectory. The upper branch corresponds to the correct charge assignment and direction of rotation, and the lower branch to the wrong charge assignment.

sequently no hardware trigger is employed and instead the online filter farm reconstructs all tracks and applies a selection algorithm in software. The selection requires three tracks coincident in time, and consistent with originating at a common vertex and with the expected kinematic properties of signal events. The computing power required for this is provided by *Graphics Processing Units* (GPUs), where we profit from the very high rate of technological advance driven by the gaming and deep learning markets.

A simple version of the fast linear fit based on multiple scattering (see chapter 19) is implemented on the GPUs for quick track fitting. In addition, events with at least two positive and one negative electron tracks are checked for a common vertex and signal-like kinematics. This selection is applied on a frame by frame basis on individual farm PCs and the selected frames are merged into the global data stream, see Figure 145 for an overview. The technical implementation of the event filter is described in section 17.5.

For the online reconstruction, only hits from the central station of the pixel detector are considered, since matching recurling tracks and time information from the tiles and fibres is computationally too expensive and also not necessary for a first selection.

Combinations of hits from the first three detector layers are matched to form triplets. Before the actual fitting procedure, a number of simple geometrical selection cuts are applied at the FPGA stage in order to reduce the number of combinations by a factor of about 50.

The fitting of triplets is non-recursive and linear, and hence can be done in parallel for all hit combinations. It is therefore an ideal candidate for parallelisation on GPUs. With their many computing kernels but small memories, they perform well at tasks where many similar computations are performed on the same memory content. For a muon rate of 10^8 Hz and 50 ns time frames, we expect $\mathcal{O}(10)$ hits per layer leading to $\mathcal{O}(10^3)$ combinations. With code optimised for these conditions, we have achieved $1 \cdot 10^9$ fits/s on a NVIDIA GTX 980 GPU (released 2014), which is sufficient to handle this level of combinations.

For each triplet passing the χ^2 and radius cuts, the track is extrapolated to the fourth detector layer. If at least one hit exists within a certain transverse radius and distance in z, the hit closest to the extrapolated position is used to form a second triplet from hits in layers two, three and four and give an improved value for the curvature of the track from an updated fit. Finally the charge of the particle is derived from the track curvature and all combinations of two positive tracks and one negative track are examined with respect to a common vertex.

The vertex position is calculated from the mean of two-circle intersections of the tracks in the transverse plane (perpendicular to the magnetic field), weighted by the uncertainty from multiple scattering in the first layer and hit resolution. A χ^2-like variable is defined using the distances of closest approach of each track to the mean intersection position and their uncertainties, both in the transverse and the $r - z$ plane. Vertices are selected based on their proximity to the target and the χ^2 value. In addition, cuts on the total kinetic energy and combined momentum of the three tracks at the points of closest approach are applied. After all cuts, the frame rate is reduced by a factor ≈200, which is further reduced by the full online reconstruction as described in chapter 19.

In addition to signal candidate events, cosmic ray muon candidates and random frames will be saved for calibration, alignment and studies of the selection efficiency. The parameters of all reconstructed tracks are histogrammed for monitoring as well as for searches e.g. for two-body muon decays [86].

The triplet fit, the propagation to the fourth layer and the vertex fit as well as the monitoring have been implemented, optimized for performance and tested on GTX 1080Ti cards (of 2016/17 vintage). It was shown that 12 of these cards are sufficient for the phase I load [66].

21. Detector Alignment

In order for the reconstruction algorithms to work optimally, the position and orientation of all active detector elements and the stopping target have to be

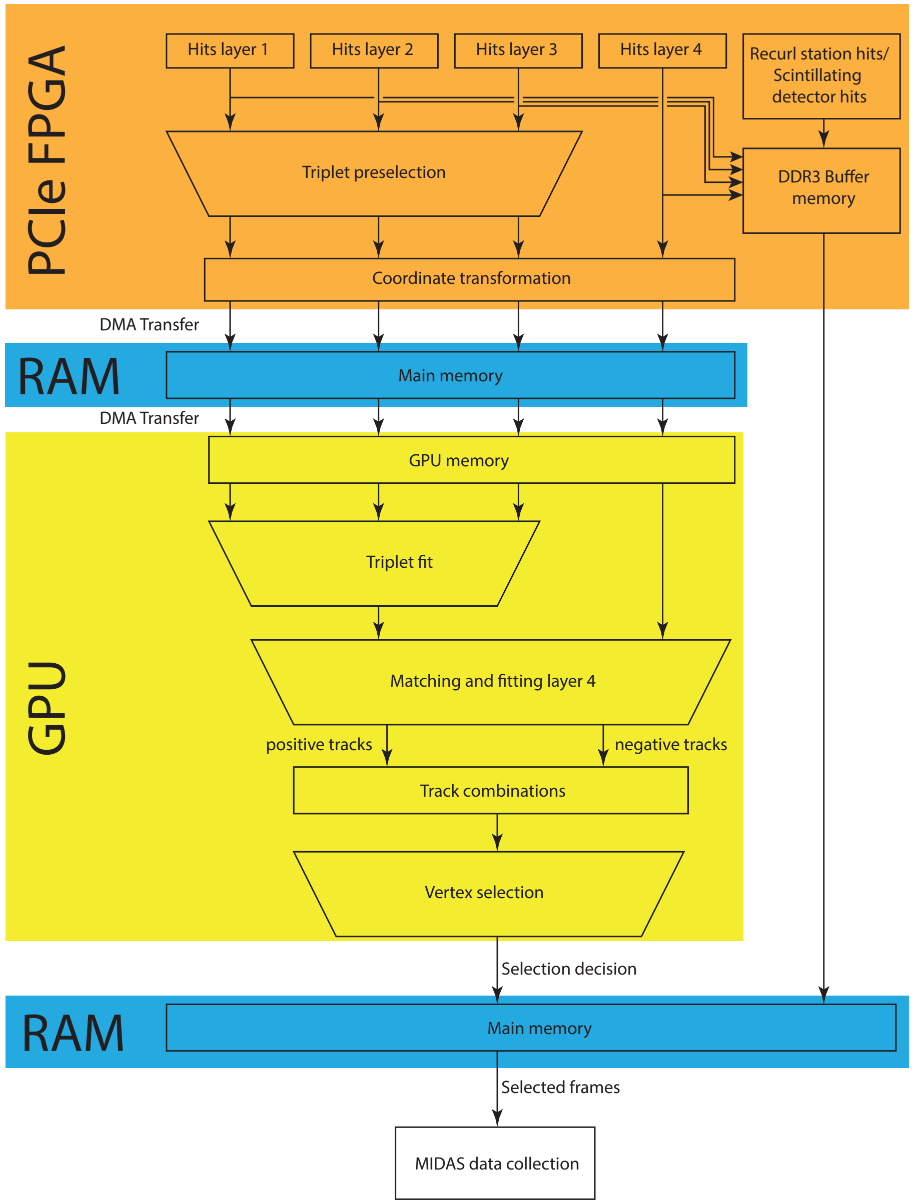


Figure 145: Flow diagram of the online reconstruction software and firmware.

known to good precision. The position of the pixels inside a sensor is given by the tolerances of the manufacturing process, which are much better than the minimal feature size of 180 nm; compared to all other sources of misalignment, this is completely negligible. The task of detector alignment is thus to determine the position, orientation and potentially deformation of all active detector parts (HV-MAPS chips, fibres, tiles).

The first step in ensuring a well-aligned detector is the careful assembly of modules and layers using precision tools, followed by a detailed survey. After detector installation, movements of larger detector parts (e.g. the recurl stations with regards to the central detector) can be followed by a system of alignment markers observed by digital cameras inside the magnet. The ultimate alignment precision is however only reached with track-based alignment methods, starting with cosmic ray tracks and refining using beam data.

21.1. Effects of Misalignment

We have studied the effects of a misaligned pixel detector on the reconstruction efficiency and tracking resolution using the full detector simulation [87]. For technical reasons the sensors are all in their nominal position for the Geant4 simulation, and the reconstruction is then performed with different sensor positions.

The hierarchical mechanical structure of the pixel detector with stations, modules, ladders and sensors is

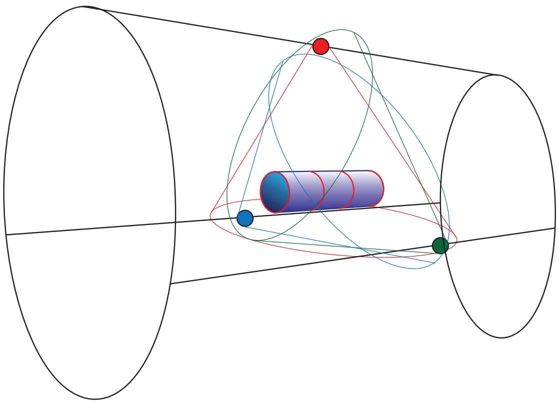


Figure 146: Schematic view of a possible alignment system using three cameras. The detector support cage is shown in black, the blue tube represents the detector stations with the end-rings shown in red. The three cameras and their fields of view are shown in red, green and blue.

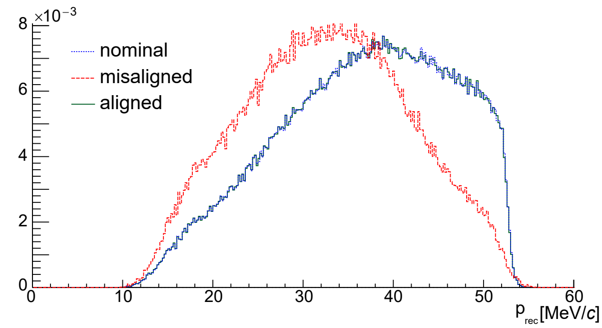


Figure 147: Reconstructed momentum of Michel positrons (using only long tracks) for the nominal detector versus the (estimated) detector after assembly and for the aligned detector.

expected to be reflected in the misplacements of all detector parts after assembly. To reproduce this, the sizes of various misalignment modes (i.e. rotations and shifts of all structures, and deformations of individual sensors) are estimated and then applied in a randomised way. The result is an average absolute offset of about 450 µm of single sensor corners[27] with respect to their nominal position for the estimated misalignment after detector construction. This leads to a worsening of the reconstruction efficiency; there is however an efficiency plateau if the overall error on the corners of the sensors is less than 100 µm. Far more relaxed criteria apply to global movements of detector stations (e.g. recurl stations, vertex layers) with regards to each other.

The constraints on the alignment accuracy for achieving optimal momentum resolution are much tighter than for the efficiency – positions and rotations should be known well enough to achieve an error smaller than 50 µm for the sensor edge positions.

21.2. Position Monitoring System

The positions of the detector stations relative to each other are monitored by a series of cameras mounted to the detector cage. They are complemented by light sources (the detector is usually operated in the dark) and alignment marks on the end-rings of the detector stations. Cameras with a 85° field of view are sufficient to view all end rings in the phase I detectors when mounted to the detector cage. A system of three cameras (e.g. top and ± 60° from the bottom) also allows tracking the relative movements of the cameras, as they can see each other. Sub-millimetre resolution requires fairly high resolution cameras (2K or 4K) or the use of separate cameras with zoom lenses focused on the station-station transitions. A possible three-camera system is shown in Figure 146.

[27] Studying the sensor corners has the advantage of covering shifts as well as rotations of sensors with respect to their nominal position.

21.3. Track-Based Alignment

The fine alignment of the silicon sensors (as well as the fibres and tiles) will be performed using track-based methods initially developed in the H1 experiment [88] and subsequently successfully applied to a variety of large and very large tracking systems, e.g. CMS at the LHC [83, 89, 90, 91, 92, 93, 94, 95].

The alignment of the complete detector is a large minimisation problem, where, for a very large sample of tracks, the residuals from the measured hits to the fitted tracks have to be minimised under variation of both all track and all alignment parameters (suitably parametrised detector positions). If a rough detector alignment is known, corrections will be small and the minimisation problem can be linearised.

To this end, tracks reconstructed with the standard reconstruction algorithms described in chapter 19 are re-fitted using the general broken lines (GBL) algorithm [83, 84]. The GBL software can calculate and output the complete covariance matrix between track and alignment parameters. As the track parameters are not correlated between tracks and only relate to the alignment of the small subset of sensors which are hit by the track, the resulting matrix is sparse. There are efficient algorithms for the inversion of such gigantic sparse matrices, one of which is implemented in the *Millepede II* programme [96], which we are using.

Whilst the sensor alignment is locally well constrained via the overlap of the sensors in the azimuthal direction and the closeness of the double layers, overall deformations such as shifts of the top part with regards to the bottom part, an overall torsion or the position of the recurl stations are only weakly constrained by using tracks from muon decays. These so-called *weak modes* need additional input from tracks which correlate distant parts of the detector. These tracks are provided by cosmic ray muons. As the cosmic rate is tiny com-

Parameter	Nominal	Misaligned	Aligned
Efficiency (short) [%]	100.00	59.09 ± 0.08	100.01 ± 0.03
Efficiency (long) [%]	100.00	46.72 ± 0.12	100.05 ± 0.14
Momentum resolution (short)	2.628 ± 0.003	4.271 ± 0.006	2.635 ± 0.003
Momentum resolution (long)	1.341 ± 0.002	1.645 ± 0.003	1.337 ± 0.002

Table 18
Tracking performance (using Michel positrons) for nominal, misaligned and aligned configurations of the pixel detector. The efficiencies are relative to the nominal configuration. The misaligned version corresponds to an estimate of the expected sensor misplacements after assembly. Momentum resolutions show the RMS of the distributions.

Parameter		Nominal	Misaligned	Aligned
Efficiency (short) [%]		100.0	5.9	99.7
Efficiency (long) [%]		100.0	2.2	100.1
$x_{rec} - x_{true}$ [mm]	Mean	-0.002 ± 0.011	0.029 ± 0.068	-0.021 ± 0.011
	RMS	0.553 ± 0.008	0.724 ± 0.048	0.550 ± 0.008
$y_{rec} - y_{true}$ [mm]	Mean	-0.010 ± 0.012	-0.188 ± 0.050	0.048 ± 0.012
	RMS	0.555 ± 0.008	0.687 ± 0.035	0.552 ± 0.008
$z_{rec} - z_{true}$ [mm]	Mean	0.003 ± 0.009	0.105 ± 0.067	-0.005 ± 0.009
	RMS	0.356 ± 0.006	0.813 ± 0.048	0.355 ± 0.006

Table 19
Signal reconstruction efficiency and vertex resolution for nominal, misaligned and aligned configurations of the pixel detector. The efficiencies are relative to the nominal configuration. The misaligned version corresponds to an estimate of the expected sensor misplacements after assembly.

pared to the beam muon rate, it is imperative to have a special trigger stream to collect enough cosmics for alignment.

Our strategy is to perform a preliminary alignment of the detector using cosmic muons, which will have to fulfil the efficiency requirements. Michel tracks can then be used until the required resolution is reached. The requirements seem well within reach: in simulation, an average error on the sensor edge position of about 110 µm has already been achieved[28]. In addition the effects of the residual misalignments do not significantly deteriorate the performance of the tracking detector (see Table 18 and Table 19). The general reconstruction efficiencies and momentum resolutions for short and long tracks, as well as the signal reconstruction efficiencies for short and long tracks, in the re-aligned detector are almost identical to the values for the nominal detector.

In Figure 147 the distributions of the measured momenta of positrons originating from a Michel decay for the misaligned and aligned pixel detector are compared to the result for the nominal detector. The misalignment applied in Figure 147 corresponds to expectations about detector misplacements right after assembly. Where a misaligned detector geometry causes a clear distortion of the momentum distribution, we are able to recreate the nominal spectrum almost perfectly by applying the track-based alignment.

We have also implemented sensor deformations and temperature dependent sensor expansion in the alignment software. The fibre and tile detectors will also be aligned using track-based methods, using the pixel detector as a reference. The Millipede II algorithm can then also be used for a precise time alignment of all detector parts.

21.4. Target Alignment

The position of the target needs to be known with very high accuracy to allow placing requirements on the distance between the vertex and target. As the target is passive, the residual-based method described in the previous sections is not applicable. The overwhelming majority of decay positrons originate on the target surface, however, and will thus have a point of closest approach (POCA) to the beam axis inside the target.

This can be used to determine the target position by plotting the distribution of the POCAs in the transverse plane in slices of z for many tracks, which will give an accurate determination of the position of the outer target edge, as shown in Figure 148 and Figure 149. The target thickness has to be determined during manufacture or using photon conversions.

[28] These results are based on an estimate of the misalignment right after detector assembly.

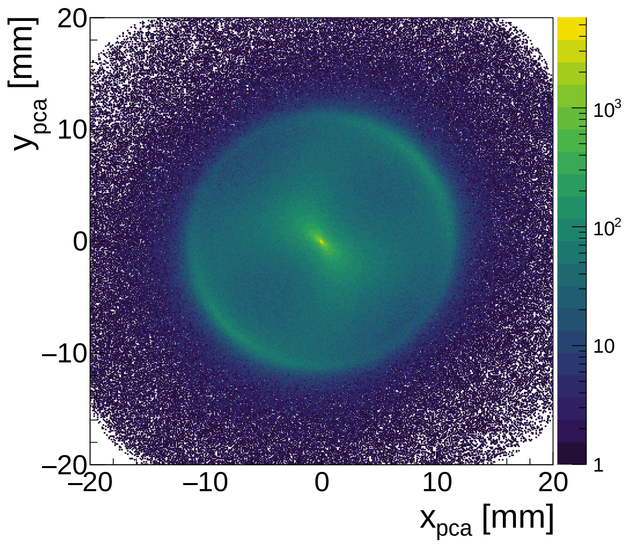


Figure 148: Position in x and y of the points of closest approach to the beam line for a $1\,\text{mm}$ slice in z at $-20\,\text{mm}$ for $3.84 \cdot 10^8$ stopped muons. The target is clearly visible at its nominal position. The accumulation of entries towards the origin is a feature of the reconstruction method.

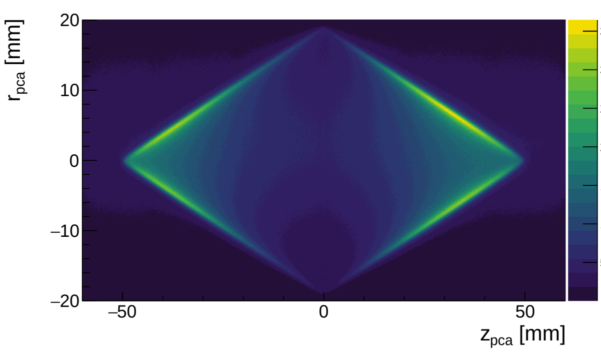


Figure 149: Position in r and z of the points of closest approach to the beam line for $3.84 \cdot 10^8$ stopped muons; the target is clearly visible. Negative radius is defined to be when the beam line is inside the track circle, positive is outside.

22. Perfomance Simulation

We study the performance of the detector described in the preceding parts by running the Geant4 simulation and the reconstruction programme. Even under optimistic assumptions, only a handful of signal decays are expected in the data. Nonetheless, we use relatively large signal samples to study the detector performance and deduce a very preliminary and rough event selection.

For the various expected backgrounds, in principle the simulation of several times the expected number of decays in data is required. This is impractical both in terms of processing time and available storage space. We thus try to identify important sources of background from either general considerations (internal conversion) or from simulating a few seconds of run time (accidentals).

From these starting points we generate special background samples. In the case of accidental background samples we can make use of the approximation that timing suppression is independent of vertexing and event kinematic suppression.

We use a simple cut based analysis in order to show that background free running with the conditions at πE5 is possible. An analysis with optimised cuts or based on likelihoods can likely deliver a higher signal efficiency and thus final sensitivity per running time.

22.1. Signal Performance

We study the nominal performance of the detector setup using about 8.5 million signal decays. The decay electrons are generated with a phase space distribution. Efficiencies are determined relative to all muons decaying inside a cylinder with the outer dimensions of the stopping target.

In the first step, all three tracks from the signal decay have to be reconstructed to at least short (4-hit) tracks; for the efficiency and resolution of the track reconstruction, see chapter 19.

22.1.1. Vertex Fit

The three tracks from signal decays should intercept at a common point on the surface of the target. We look at all combinations of a track with negative charge and two tracks of positive charge. In order not to fit recurling tracks with themselves, the track tangent vector at the point of closest approach is determined. If the cosine of the opening angle between two tracks is more

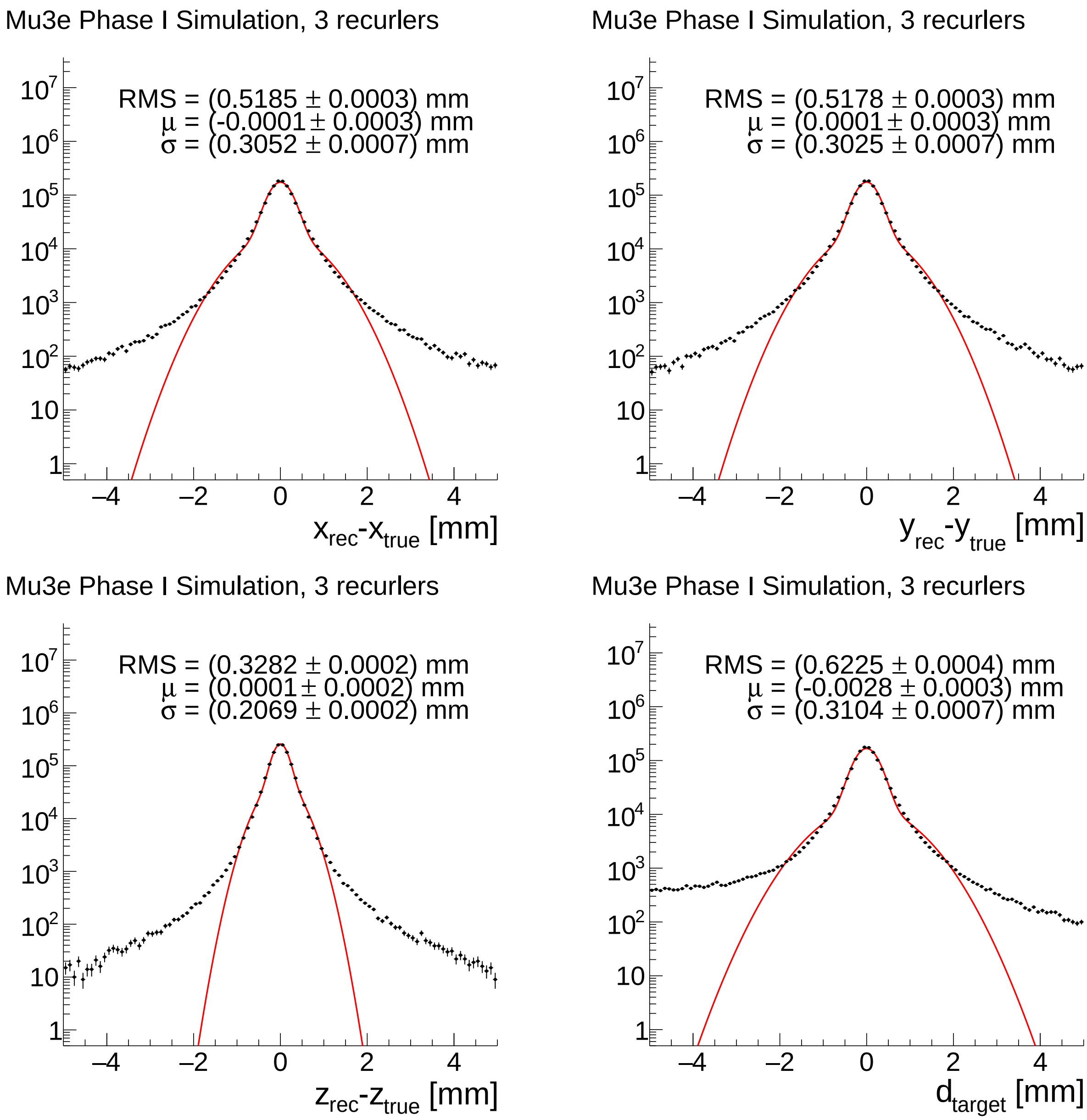


Figure 150: Vertex resolution for signal decays. Three tracks with recurlers are selected. The fits are the sum of two Gaussian distributions and the quoted σ is the area-weighted mean. Top left in x, top right in y, bottom left in z and bottom right in the distance to the target; negative target distances denote a reconstructed vertex position inside the target.

than 0.99 and the momentum difference is less than 1 MeV/c, the combination is not further considered.

Starting from the track positions and directions in the first detector plane, we perform a vertex fit by forcing three tracks to intersect in a common point in space, taking multiple scattering in the first detector layer as the only degree of freedom [97]. The χ^2 of the fit and the distance of the vertex to the target surface are two handles for suppressing accidental background; the performance of the vertex reconstruction is illustrated in Figure 150.

22.1.2. Mass and Momentum Reconstruction

For all candidates with a vertex fit $\chi^2 < 30$ the tracks are extrapolated to the vertex and four-vectors are constructed with an electron mass assumption. From the three four-vectors, the mass of the decaying particle (should correspond to the muon mass) and the momentum of the center-of-mass system (CMS) in the detector frame (should be zero for decays at rest) are determined.

The resolution for the muon momentum is depicted in Figure 151. The magnitude of the reconstructed mo-

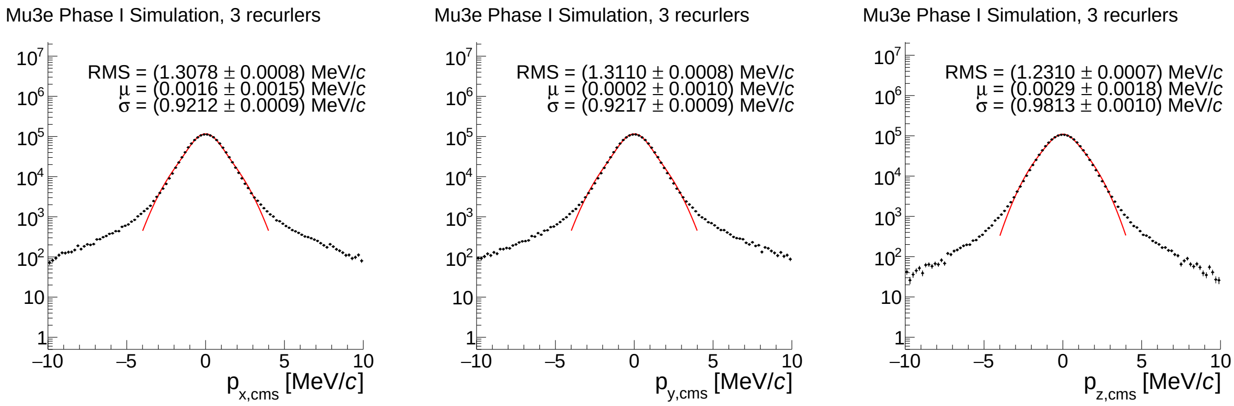


Figure 151: Reconstructed decay muon momentum in x, y and z direction (which corresponds to the resolution for p_x, p_y and p_z for muons decaying at rest). Only long tracks enter the analysis.

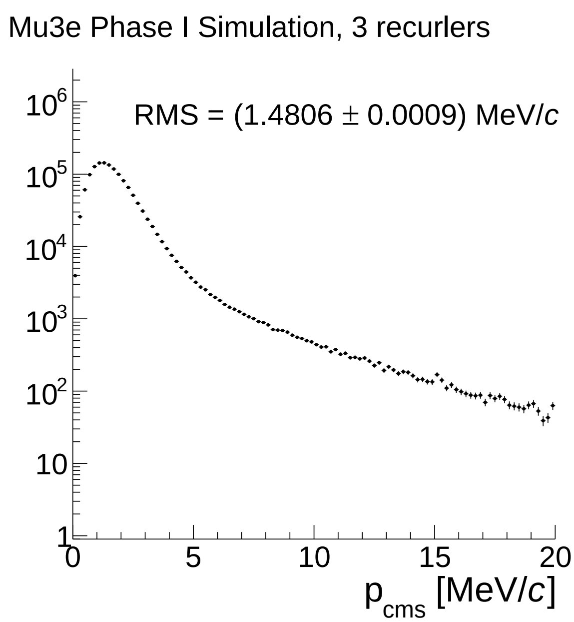


Figure 152: Center of mass system momentum reconstructed for signal events with three recurlers required.

mentum is shown in Figure 152.

We define the decay plane from the three momenta $\vec{p}_i$,

$$\vec{a} = \frac{(\vec{p}_1 - \vec{p}_2) \times (\vec{p}_3 - \vec{p}_2)}{|(\vec{p}_1 - \vec{p}_2) \times (\vec{p}_3 - \vec{p}_2)|}. \tag{8}$$

$\vec{a}$ is a vector perpendicular to the decay plane (if the tracks are from a muon decaying to the signal channel at rest).

The SINDRUM experiment based their selection on the projection of the CMS momentum onto this vector, called acoplanar momentum

$$\vec{p}_{acoplanar} = \vec{p}_{CMS} \cdot \vec{a}. \tag{9}$$

and the coplanar momentum

$$\vec{p}_{coplanar} = \vec{p}_{CMS} \times \vec{a}. \tag{10}$$

To first order, the resolution for the acoplanar momentum is only dependent on the measurement (and thus resolution) of the track angle, whereas the coplanar momentum is dominated by the absolute momentum resolution.

The corresponding distributions are shown in Figure 153; for the Mu3e setup (similar to SINDRUM) the resolution in the acoplanar momentum is superior to the coplanar momentum resolution. No detailed optimization of the momentum selection has been performed for Mu3e so far, so for the distributions shown in this report, we used the requirement of $p_{CMS} < 8\,\mathrm{MeV/c}$.

Finally, we show the resolution for the reconstructed mass in Figure 154. As the distributions show, the core of the mass resolution fulfils the criteria set out in chapter 1 and especially if requiring recurling tracks. Sizeable Landau-like tails only appear on the low mass side.

22.1.3. Signal Efficiency

For every reconstruction step, there is a possibility of signal loss; the largest loss is due to the geometrical acceptance of the detector. For phase-space signal decays in the target, approximately 38.1% have all three electrons traverse the four layers of the central detector in the active region. If recurling tracks are required, the acceptance is further reduced. There are also inefficiencies in the reconstruction and vertex fit, especially due to the χ^2 cuts, which mostly get rid of tracks with large angle scattering, which preclude a reliable and precise reconstruction. The overall efficiency after applying all mentioned cuts as well as a veto on events where the tracks have inconsistent timing is shown in Figure 155 in dependence of the required number of recurling tracks.

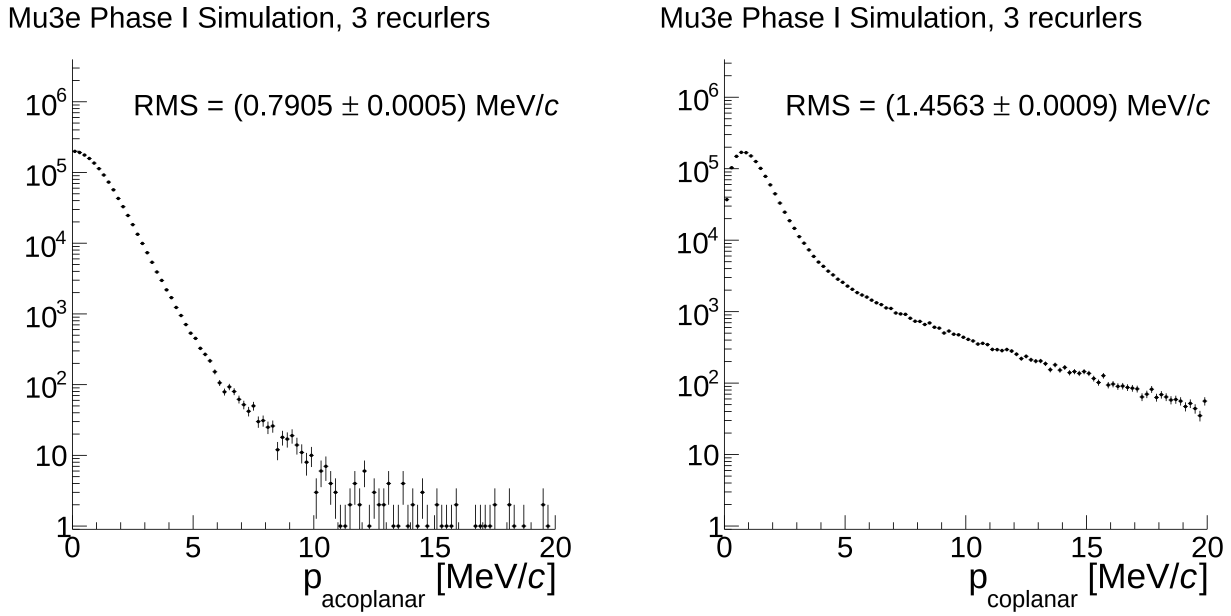


Figure 153: Acoplanar momentum (left) and coplanar momentum (right) reconstructed for signal events with three recurlers required.

Step	Step efficiency	Total efficiency
Muon stops	100%	100%
Geometrical acceptance, short tracks	38.1%	38.1%
Geometrical acceptance, long tracks	68.0%	25.9%
Short track reconstruction	89.5%	34.1%
Long track reconstruction[1]	67.2%	17.4%
Vertex fit	99.4%	17.3%
Vertex fit $\chi^2 < 30$	97.6%	16.9%
CMS momentum $< 8\,\mathrm{MeV/c}$	97.6%	16.5%
Timing	90.0%	14.9%

Table 20
Efficiency of the various reconstruction and analysis steps.
[1]: Note that the efficiency of this step is quoted relative to the acceptance for long tracks.

With the used selection criteria, the overall efficiency is 14.9 % when three recurling tracks are required. The efficiency losses are listed in Table 20. Further gains are expected from a through optimisation of the cuts; on the other hand, imperfections of the real detector will likely lead to some additional losses. The selection can optimised for efficiency or mass resolution e.g. by requiring recurling tracks only above a minimum momentum, see Figure 156.

22.2. Backgrounds

22.2.1. Internal Conversion Background

We simulate the internal conversion background as described in section 18.3.2 using the matrix element provided by Signer et al. [7]. The total branching fraction for this decay is $3.4 \cdot 10^{-5}$ [6], so a complete simulation is challenging. We are however mostly interested in the region of phase space were the neutrinos carry little momentum; the branching fraction for this high visible mass region (we used a lower cutoff of $90\,\mathrm{MeV/c^2}$ for the studies presented here), is strongly suppressed and we can generously oversample in the simulation. In addition, we use weighted events in order to better populate the high mass tail. Migrations from lower masses than $90\,\mathrm{MeV/c^2}$ into the signal region are very strongly suppressed if three recurling tracks are required, see Figure 157.

22.2.2. Accidental Background

Accidental background arises from the combination of two Michel positrons with an electron. It is thus important to understand and limit electron production in the target region. This is of particular importance for processes such as Bhabha-scattering, where the electron and positron tracks intersect in space and time and only the separation to the second positron remains as a suppression criterion.

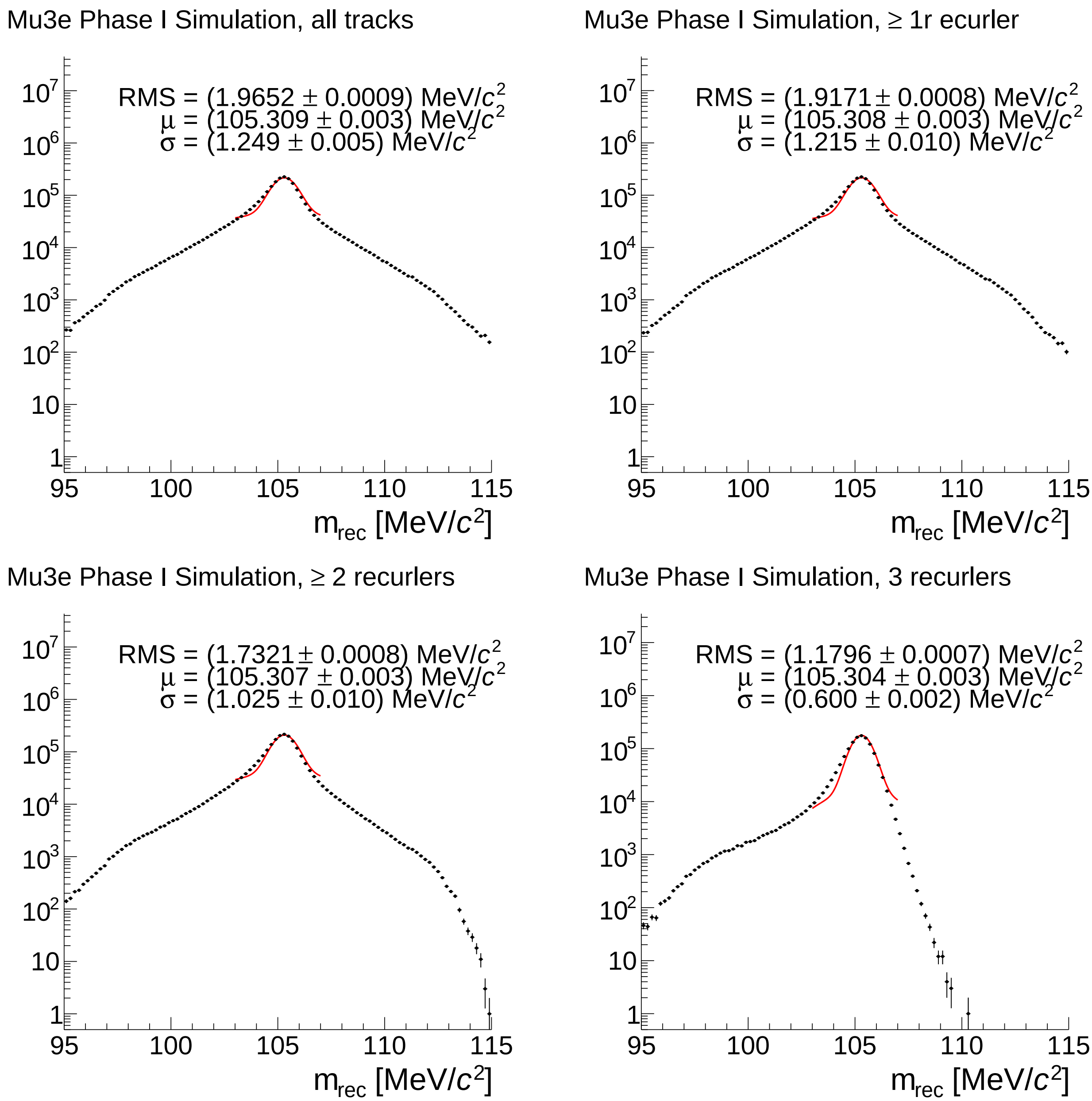


Figure 154: Reconstructed muon mass for all tracks (top left), at least one recurler (top right), at least two recurlers (bottom left) and three recurlers (bottom right). The fits are the sum of two Gaussian distributions and the quoted σ is the area-weighted mean; the main purpose of the fit is to guide the eye and highlight the non-symmetric resolution distribution.

Electron Production in the Target The default target is part of the Geant4 detector simulation as described in chapter 18. The material of the target is a place where electrons from Bhabha and Compton scattering as well as from photon conversion can be produced and contribute to accidental background. Bhabha scattering needs special attention, as very often both the electron and the positron partaking in the scattering process end up in the detector acceptance; the corresponding vertices are shown in Figure 158. As shown in Figure 159, almost all the corresponding primary positrons come from muon decays in the target and can thus not be further reduced or shielded.

The total number of electrons produced per Michel decay is shown in Table 21. As can be seen, Bhabha scattering is the most important background process. The reason that there are significantly lower number of electron reconstructed is because of the the momentum spectrum falling fast, see Figure 160. The means that many of the electrons end up at or below the low edge of the detector and reconstruction acceptance.

Timing Suppression Time information from hits in the fibre and tile detectors provides an important handle for the suppression of accidental backgrounds. If

Electron source	Produced in inner detector	Produced in target region	Reconstructed inner detector short tracks	Reconstructed target region short tracks	Reconstructed inner detector long tracks	Reconstructed target region long tracks
Bhabha scattering	$5.5 \cdot 10^{-4}$	$1.1 \cdot 10^{-4}$	$2.7 \cdot 10^{-4}$	$5.7 \cdot 10^{-5}$	$2.3 \cdot 10^{-4}$	$4.4 \cdot 10^{-5}$
both visible	$4.3 \cdot 10^{-4}$	$7.7 \cdot 10^{-5}$	$1.5 \cdot 10^{-4}$	$2.6 \cdot 10^{-5}$	$1.1 \cdot 10^{-4}$	$1.7 \cdot 10^{-5}$
Photon conversion	$2.3 \cdot 10^{-5}$	$2.1 \cdot 10^{-6}$	$1.1 \cdot 10^{-5}$	$1.0 \cdot 10^{-6}$	$9.2 \cdot 10^{-6}$	$8.0 \cdot 10^{-7}$
both visible	$5.7 \cdot 10^{-6}$	$4.6 \cdot 10^{-7}$	$1.5 \cdot 10^{-6}$	$1.3 \cdot 10^{-7}$	$1.2 \cdot 10^{-6}$	$9.3 \cdot 10^{-8}$
Compton scattering	$3.6 \cdot 10^{-5}$	$4.3 \cdot 10^{-6}$	$1.7 \cdot 10^{-5}$	$2.2 \cdot 10^{-6}$	$1.4 \cdot 10^{-5}$	$1.7 \cdot 10^{-6}$
Internal conversion	$3.1 \cdot 10^{-5}$	$2.9 \cdot 10^{-5}$	$1.7 \cdot 10^{-5}$	$1.6 \cdot 10^{-5}$	$1.3 \cdot 10^{-5}$	$1.3 \cdot 10^{-5}$
two visible	$1.1 \cdot 10^{-6}$	$1.0 \cdot 10^{-6}$	$3.6 \cdot 10^{-7}$	$3.3 \cdot 10^{-7}$	$2.3 \cdot 10^{-7}$	$2.2 \cdot 10^{-7}$
Total	$6.4 \cdot 10^{-4}$	$1.5 \cdot 10^{-4}$	$3.2 \cdot 10^{-4}$	$7.6 \cdot 10^{-5}$	$2.6 \cdot 10^{-4}$	$5.9 \cdot 10^{-5}$

Table 21
Electrons with transverse momentum larger than $10\,\mathrm{MeV}$ created per Michel decay in the target region. The inner detector region is a cylinder including the vacuum window and the first pixel layer, the target region is a cylinder just containing the target.

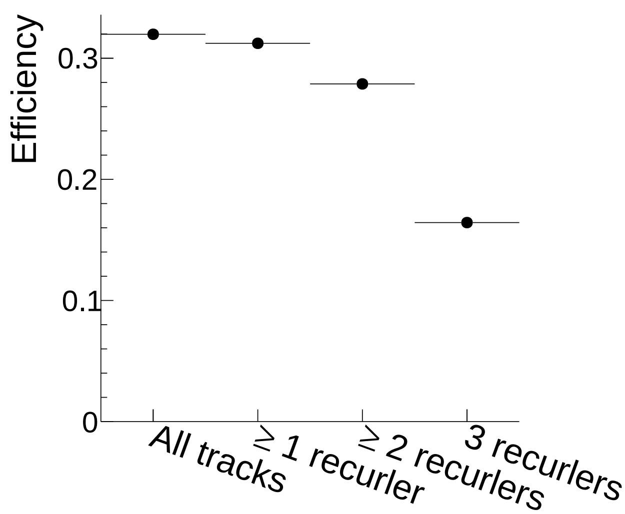


Figure 155: Total efficiency for reconstructing phase-space signal events as a function of the required number of recurling tracks. This includes the geometrical detector acceptance, track and vertex reconstruction and selection inefficiencies.

the pixel time resolution is smaller than the length of the reconstruction frames this can already be used to suppress accidental background. The additional suppression by the dedicated timing detectors depends on the size of the pixel timing window, which here was taken to be $50\,\mathrm{ns}$.

The precise timing of a track is determined by the number of assignable hits in the fibre detector and the existence of a matched tile hit. If a track reaches the tile detector in the recurl stations, the timing is dominated by this much more accurate detector. Detailed studies of the signal efficiency and background suppression of the timing detectors are described in [41] and summarised in chapter 10. Using this we have a working point of $90\,\%$ efficiency for coincident tracks (signal), a timing suppression of 71 for the dominant accidental background with two tracks correlated and one uncorrelated in time and a suppression of more than three orders of magnitude for three uncorrelated tracks is expected.

Kinematic Suppression The largest suppression factors for accidental background come from kinematics, i.e. the requirement that the three momenta sum up to zero (enforced by the total momentum selection) and a mass window around the muon mass. Typical suppression factors are of the order of one million. The kinematics of the event however also strongly affect the suppression power of the vertex fit; the corresponding requirements do unfortunately not factorise and large simulated samples are required.

Vertex Suppression The suppression of accidental background due to the common vertex of three tracks is highly dependent on the kinematics. In the interesting cases of Bhabha scattering or photon conversion, there will be an electron-positron pair with a small opening angle balanced with a positron close to the Michel edge going in the opposite direction. This case is favourable for vertex based background suppression, which is much higher than in a generic three-track arrangement. As vertex and kinematic suppression do not factorise, we have simulated the most common accidental background, Bhabha scattering plus a Michel electron with full statistics, using only mild assumptions. We start with a fraction of $7.7 \cdot 10^{-5}$ of all muon stops that produce Bhabha scattering in the target with both products visible. We then simulate normal frames at $1 \cdot 10^{8}$ muon stops per second overlaid with a muon, where, immediately after the decay, the Michel positron undergoes Bhabha scattering (here we make the further assumption, that the distribution of muon stops corresponds to the distribution of Bhabha scatters). Assuming a timing suppression factor of 70, the $2.4 \cdot 10^{9}$ simulated frames then correspond to Bhabha scattering from $1.1 \cdot 10^{16}$ muon stops. After reconstruction and applying all cuts, three simulated Bhabha events

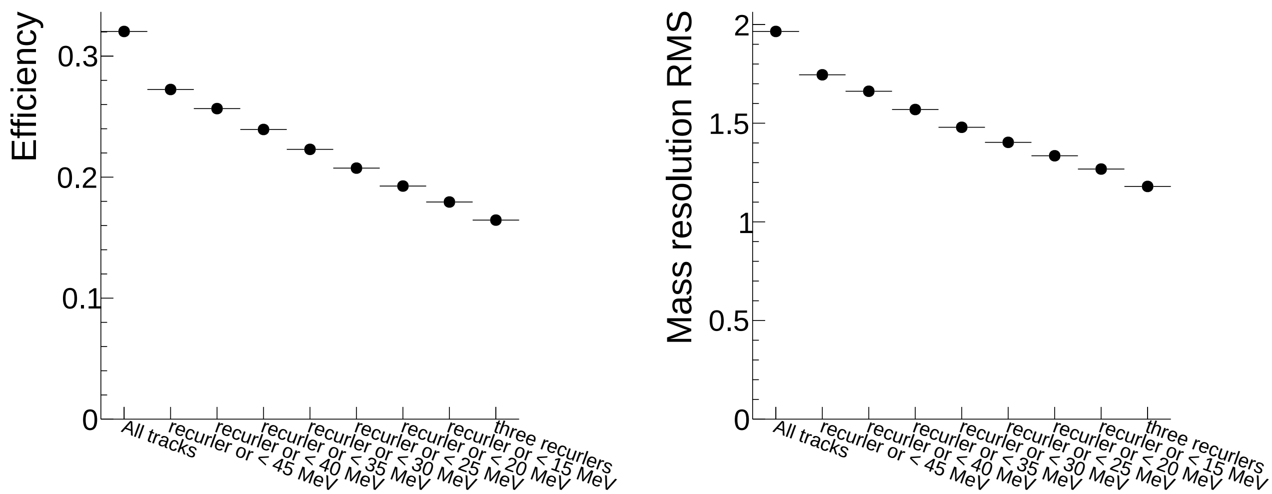


Figure 156: Efficiency before timing selection for reconstructing phase-space signal events (left) and the RMS of the corresponding three-particle invariant mass distribution. Both use the same selection criteria.

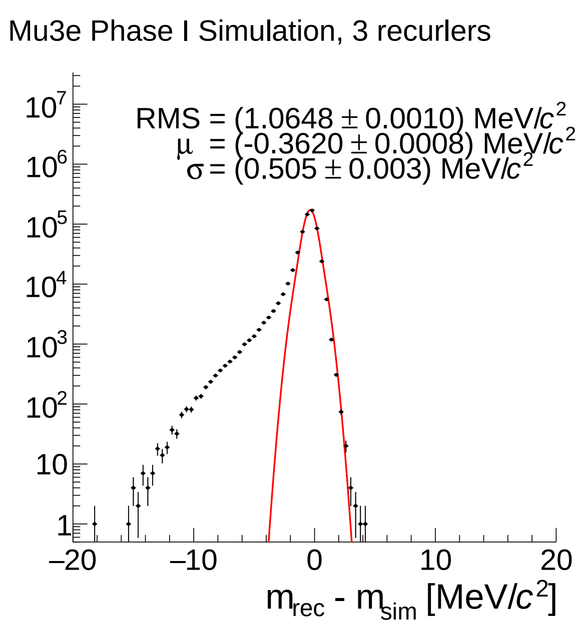


Figure 157: Resolution of the mass reconstruction for internal conversion events with a visible mass above $90\,\mathrm{MeV/c^2}$ for three recurling tracks and a momentum of the three particle system of less than $8\,\mathrm{MeV}$.

with reconstructed masses above $95\,\mathrm{MeV}$ are left. [29]

Here we have not yet used the fact that the Bhabha events almost all have the same e^+e^- invariant mass of around $7\,\mathrm{MeV/c^2}$, see Figure 161. This comes from the minimum momentum transfer required to kick the electron (initially at rest) into the detector acceptance folded with the strongly forward peaked Bhabha cross section. A requirement on this mass can further reduce the Bhabha background, will however also remove a specific part of the signal kinematics.

A similar simulation study for accidental background from combinations of internal conversion decays and Michel decays indicated that this background contributes an expectation of less than 0.1 events in the signal region.

[29] The simulation of the Bhabha background was performed with a slightly older version of the detector geometry with smaller gaps between the pixel sensors in longitudinal direction, leading to an overall efficiency about 1.5% higher than in the current version.

22.3. Sensitivity

With the phase I Mu3e detector we have the capability of suppressing both accidental backgrounds and internal conversion events to a level that allows for a background free measurement for at least $2.5 \cdot 10^{15}$ muon stops. This corresponds to about 300 days of continuous running at $1 \cdot 10^8$ stops per second. The simulated invariant mass distribution is shown in Figure 162, the 2D distribution of invariant mass and CMS momentum is shown in Figure 163 and the sensitivity versus running time is shown in Figure 164.

A. Acknowledgements

We gratefully acknowledge the beamtimes provided by the following facilities: test beam facility at DESY Hamburg (Germany), a member of the Helmholtz Association (HGF), πM1 at Paul Scherrer Institut, Villigen (Switzerland), PS and SPS at CERN, Geneva,

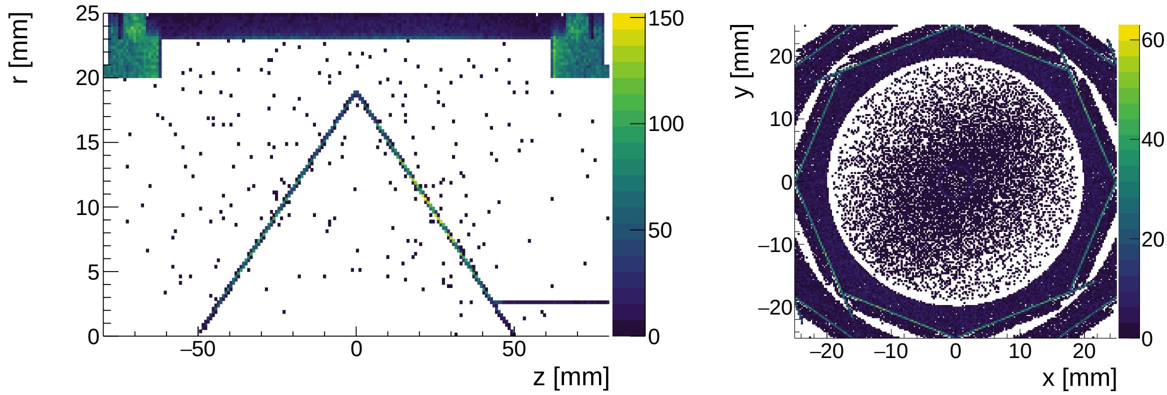


Figure 158: Longitudinal view (left) and transverse view (right) of the loci of Bhabha scattering producing an electron and a positron both in the detector acceptance in the target region for $1.9\,\mathrm{s}$ of running at $1 \cdot 10^8$ muon stops per second.

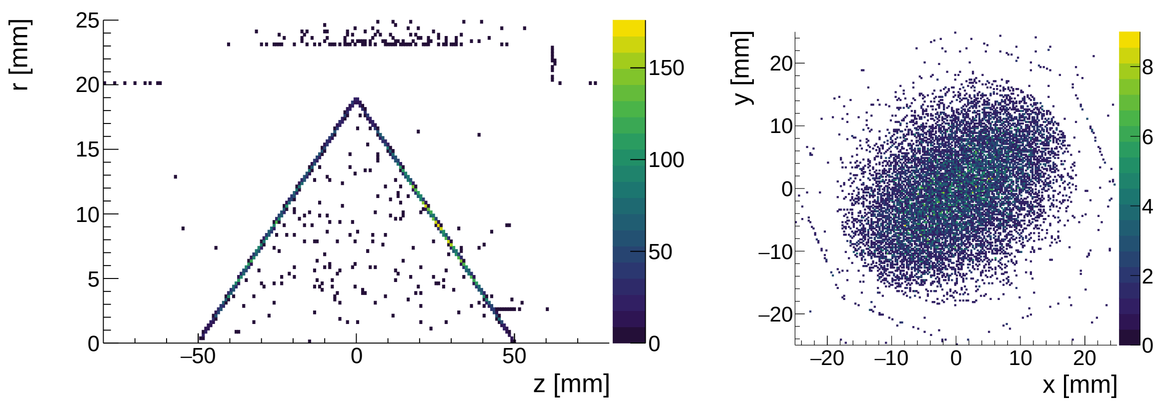


Figure 159: Longitudinal view (left) and transverse view (right) of muon decay vertices leading to a positron then undergoing Bhabha scattering in the target resulting in an electron and a positron both in the detector acceptance in the target region for $1.9\,\mathrm{s}$ of running at $1 \cdot 10^8$ muon stops per second.

(Switzerland), and MAMI A2 and X1 at Institut für Kernphysik at the JGU Mainz (Germany).

The Heidelberg groups acknowledge the support by the German Research Foundation (DFG) funded Research Training Groups HighRR (GK 2058) and “Particle Physics beyond the Standard Model” (GK 1994), by the EU International Training Network PicoSec (grant no. PITN-GA-2011-289355-PicoSEC-MCNet), by the International Max Planck Research School for Precision Tests of Fundamental Symmetries (IMPRS-PTFS) and the Heinz-Götze-Stiftung.

The work of the Mainz group has also been supported by the Cluster of Excellence “Precision Physics, Fundamental Interactions, and Structure of Matter” (PRISMA EXC 1098 and PRISMA+ EXC 2118/1) funded by the German Research Foundation (DFG) within the German Excellence Strategy (Project ID 39083149). We acknowledge the contributions of the PRISMA Detector laboratory to the development of the DC-DC converters.

The Swiss institutes acknowledge the funding support from the Swiss National Science Foundation grants no. 200021_137738, 200021_165568, 200021_172519, 200021_182031 and 20020_172706.

The Particle Physics Department of (DPNC) the University of Geneva gratefully acknowledges support from from the Ernest Boninchi Foundation in Geneva.

The UK institutes thank the Science and Technology Facilities Council for funding their work through the Large Projects scheme, under grant numbers: ST/P00282X/1, ST/P002765/1, ST/P002730/1, ST/P002870/1.

N. Berger, S. Shrestha, A. Kozlinskiy, A.-K. Perrevoort, D. vom Bruch, Q. H. Huang, U. Hartenstein and F. Wauters thank the DFG for funding their work on the Mu3e experiment through the Emmy Noether programme.

F. Meier Aeschbacher, I. Perić, A. Schöning and D. Wiedner thank the DFG for funding their work un-

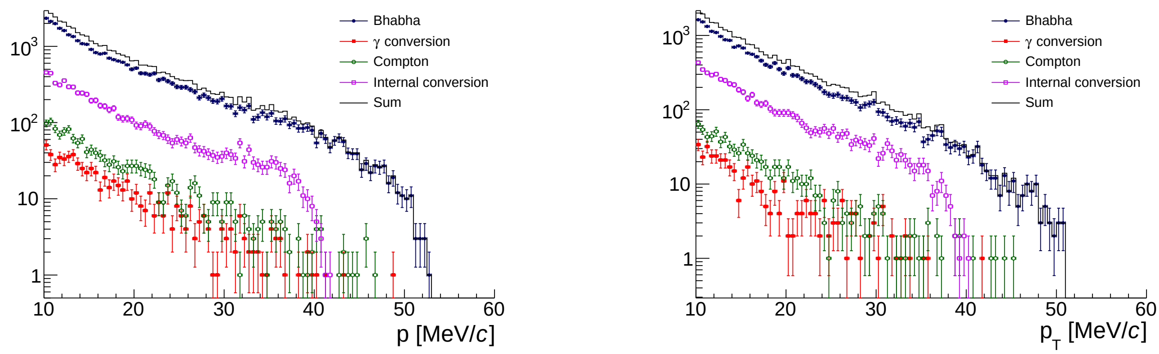


Figure 160: Momentum spectrum (left) and transverse momentum spectrum (right) of electrons produced in the target region.

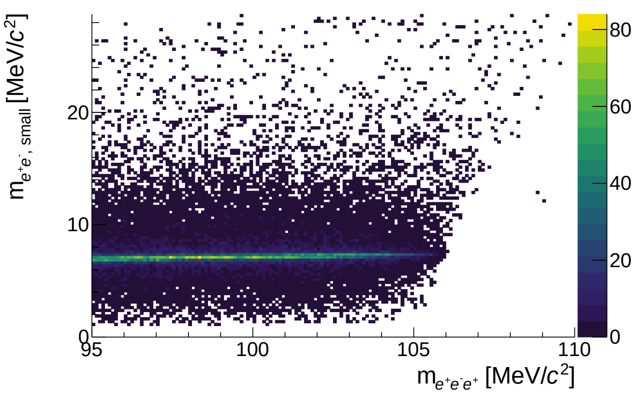


Figure 161: Small invariant mass of e^+e^- pairs versus $e^+e^-e^+$ invariant mass for accidental combinations of a Bhabha e^+e^- pair with a Michel positron.

der grant no. SCHO 1443/2-1.

G. Hesketh gratefully acknowledges the support of the Royal Society through grant numbers UF140598 and RGF\EA\180081.

Bibliography

[1] A. Blondel, S. Bravar, M. Pohl, S. Bachmann, N. Berger, A. Schöning, D. Wiedner, P. Fischer, I. Peric, M. Hildebrandt, P.-R. Kettle, A. Papa, S. Ritt, G. Dissertori, Ch. Grab, R. Wallny, P. Robmann and U. Straumann, *"Letter of Intent for an Experiment to Search for the Decay $\mu \to eee$"*, 2012, Available from `https://www.psi.ch/mu3e/documents`.

[2] A. Blondel et al., *"Research Proposal for an Experiment to Search for the Decay $\mu \to eee$"*, ArXiv e-prints, January 2013, (`arXiv:1301.6113 [physics.ins-det]`).

[3] Y. Kuno and Y. Okada, *"Muon decay and physics beyond the standard model"*, Rev. Mod. Phys., **73** 151–202, 2001, (`arXiv:hep-ph/9909265`).

[4] R. M. Djilkibaev and R. V. Konoplich, *"Rare Muon Decay $\mu^+ \to e^+e^-e^+\nu_e\bar{\nu_\mu}$"*, Phys.Rev., **D79** 073004, 2009, (`arXiv:0812.1355 [hep-ph]`).

[5] P. Blackstone, M. Fael and E. Passemar, *"$\tau \to \mu\mu\mu$ at a rate of one out of 10^{14} tau decays?"*, Eur. Phys. J. C, **80**(6) 506, 2020, (`arXiv:1912.09862 [hep-ph]`).

[6] W. Bertl et al., [SINDRUM Collaboration], *"Search for the decay $\mu^+ \to e^+e^+e^-$"*, Nucl. P, **B 260**(1) 1 – 31, 1985.

[7] G.M. Pruna, A. Signer and Y. Ulrich, *"Fully differential NLO predictions for the rare muon decay"*, Phys. Lett. B, **765** 280–284, 2017, (`arXiv:1611.03617 [hep-ph]`).

[8] M. Fael and C. Greub, *"Next-to-leading order prediction for the decay $\mu \to e\ \left(e^+e^-\right)\ \nu\bar{\nu}$"*, JHEP, **01** 084, 2017, (`arXiv:1611.03726 [hep-ph]`).

[9] R. R. Crittenden, W. D. Walker and J. Ballam, *"Radiative Decay Modes of the Muon"*, Phys. Rev., **121** 1823–1832, Mar 1961.

[10] M. Fael, L. Mercolli and M. Passera, *"Radiative μ and τ leptonic decays at NLO"*, JHEP, **07** 153, 2015, (`arXiv:1506.03416 [hep-ph]`).

[11] S. Egli et al., [SINDRUM Collaboration], *"Measurement of the Decay $\pi^+ \to e^+\nu_e e^+e^-$ and Search for a Light Higgs Boson"*, Phys. Lett., **B222** 533, 1989.

[12] G. Bressi, G. Carugno, S. Cerdonio, E. Conti, A.T. Meneguzzo and D. Zanello, *"New measurement of the $\pi \to \mu\nu\gamma$ decay"*, Nuclear Physics B, **513**(3) 555 – 572, 1998.

[13] A.E. Pifer, T. Bowen and K.R. Kendall, *"A High Stopping Density μ^+ Beam"*, Nucl.Instrum.Meth., **135** 39–46, 1976.

[14] A.M. Baldini, F. Cei, C. Cerri, S. Dussoni, L. Galli et al., *"MEG Upgrade Proposal"*, ArXiv e-prints, January 2013, (`arXiv:1301.7225 [physics.ins-det]`).

[15] U. Rohrer, PSI Graphic Transport Framework, based on a CERN-SLAC-FERMILAB version by K. L. Brown *et al.*, `http://aea.web.psi.ch/Urs_Rohrer/MyWeb/trancomp.htm`.

[16] U. Rohrer, PSI Graphic Turtle Framework by U. Rohrer based on a CERN-SLAC-FERMILAB version by K. L. Brown *et al.*, `http://aea.web.psi.ch/Urs_Rohrer/MyWeb/turtcomp.htm`.

[17] T. Roberts, G4Beamline, `http://g4beamline.muonsinc.com`.

[18] F. Berg, *CMBL - A High-intensity Muon Beam Line & Scintillation Target with Monitoring System for Next-generation Charged Lepton Flavour Violation Experiments*, PhD thesis, ETH Zürich, 2017.

[19] Z. D. Hodge, *Production, Characterization, and Monitoring of Surface Muon Beams for Charged Lepton Flavor Violation Experiments*, PhD thesis, ETH Zürich, 2018.

[20] U. Bellgardt et al., [SINDRUM Collaboration], *"Search for the Decay $\mu^+ \to e^+e^+e^-$"*, Nucl.Phys., **B299** 1, 1988.

[21] W. Bertl, SINDRUM I, Presentation at PSI, 2008.

[22] I. Perić, *"A novel monolithic pixelated particle detector implemented in high-voltage CMOS technology"*, Nucl.Instrum.Meth., **A582** 876, 2007.

[23] M. Oinonen et al., ALICE Silicon Strip Detector module

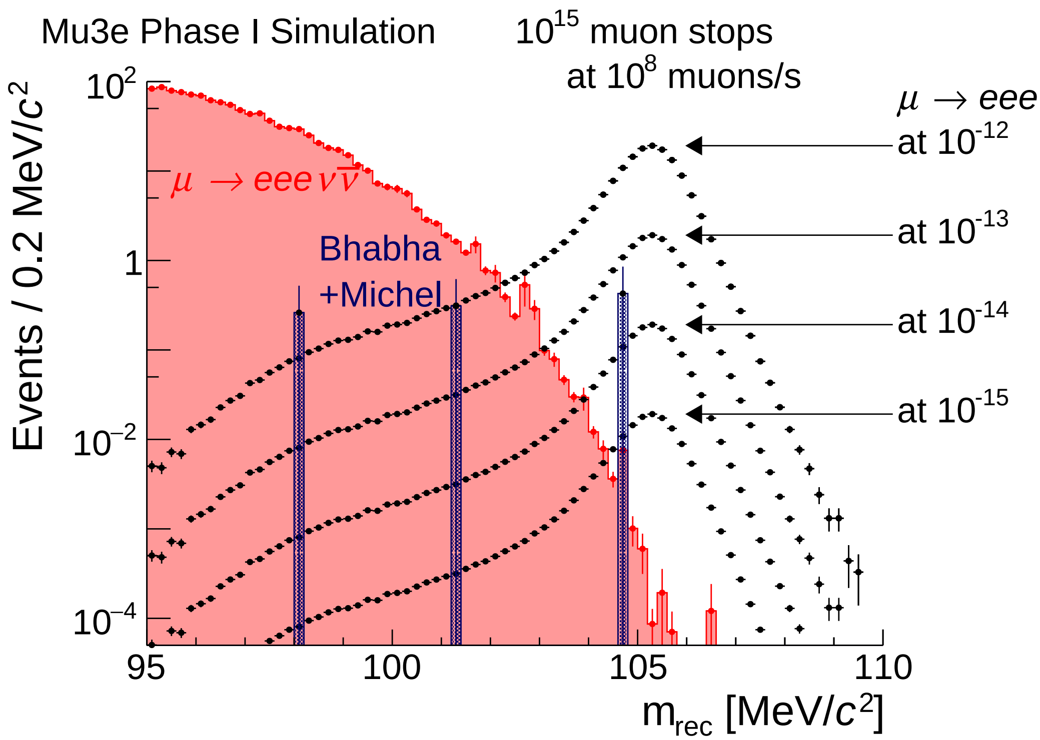


Figure 162: Reconstructed invariant mass for signal events at various branching fractions and internal conversion events. Accidental background from combinations of Bhabha pairs and Michel electrons is also shown. The CMS momentum is required to be less than $8\,\mathrm{MeV/c}$. Note that both the internal conversion and Michel and Bhabha simulation uses weighted events.

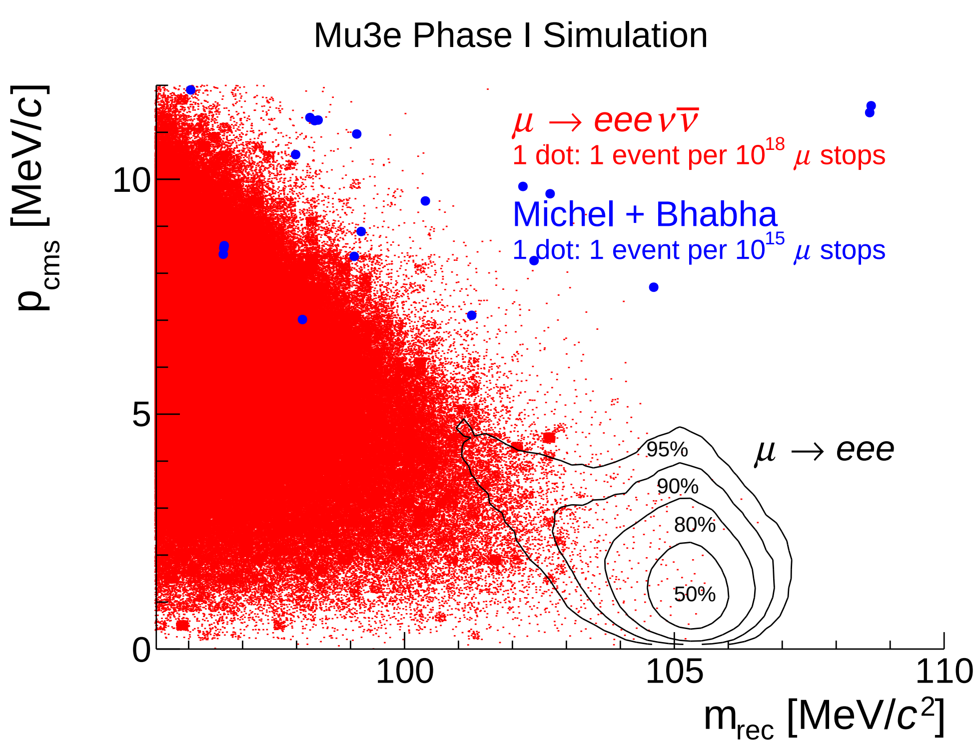


Figure 163: Reconstructed invariant mass versus the CMS momentum for signal events, internal conversion events and accidental background from combinations of Bhabha pairs and Michel electrons. Note that both the internal conversion and Michel and Bhabha simulation uses weighted events.

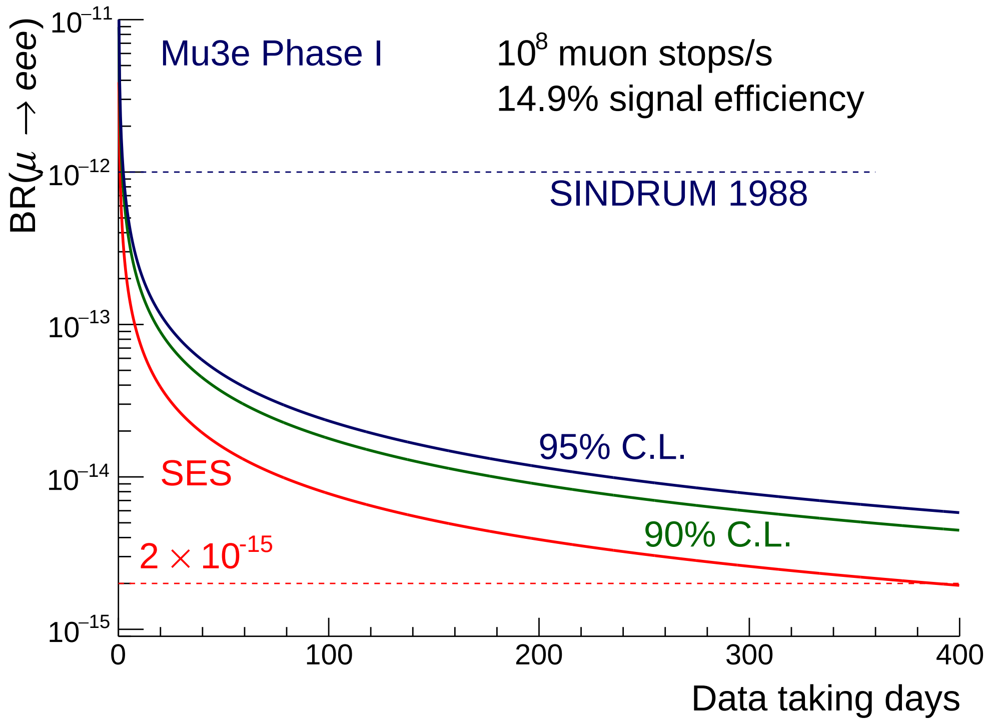


Figure 164: Single event sensitivity (SES) and the corresponding 90% and 95% C.L. upper limits versus data taking days for the phase I Mu3e detector.

assembly with single-point TAB interconnections, In *Proceedings, eleventh Workshop on Electronics for LHC and Future Experiments, Heidelberg, Germany, 12-16 September 2005*, pages 92–98, 2005.

[24] LTU, LED Technologies of Ukraine – `http://ltu.ua/en/index/`.

[25] L. Noethe, *Flexprint design and characterization for the Mu3e experiment*, Bachelor thesis, Heidelberg University, 2016.

[26] Samtec Corporation, Ultra-low profile micro arrays, Catalogue.

[27] M. Zimmermann, *Cooling with Gaseous Helium for the Mu3e Experiment*, Bachelor thesis, Heidelberg University, 2012.

[28] L. Huxold, *Cooling of the Mu3e Pixel Detector*, Bachelor thesis, Heidelberg University, 2014.

[29] Y.W. Ng, *Finite Element Analysis of the Cooling System for the Mu3e Experiment*, Master thesis, University of Applied Science Jena, 2015.

[30] A. Herkert, *Gaseous Helium Cooling of a Thin Silicon Pixel Detector for the Mu3e Experiment*, Master thesis, Heidelberg University, 2015.

[31] K. Tormann, *Thermal Analysis of the Silicon Pixel Detector for the Mu3e Experiment*, Master thesis, Heidelberg University, 2018.

[32] M. Deflorin, *Helium cooling of Silicon Pixel Detector for Mu3e Experiment*, Master thesis, University of Applied Sciences and Arts Northwestern Switzerland, Brugg-Windisch, 2019, in preparation.

[33] F. Meier Aeschbacher, M. Deflorin and L. Noehte, Mechanics, readout and cooling systems of the Mu3e experiment, In *28th International Workshop on Vertex Detectors*, 3 2020.

[34] R. P. Austermühl, *Analyse von Michelson-Interferometriedaten von Vibrationsmessungen eines dünnen gasgekühlten Pixeldetektors*, Bachelor thesis, Heidelberg University, 2015.

[35] L. Henkelmann, *Optical Measurements of Vibration and Deformation of the Mu3e Silicon Pixel Tracker*, Bachelor thesis, Heidelberg University, 2015.

[36] EndoTOFPET-US Proposal, *"Novel Multimodal Endoscopic Probes for Simultaneous PET/Ultrasound Imaging for Image-Guided Interventions"*, European Union 7th Framework Programme, **186** 2007–2013.

[37] H. Chen, *A Silicon Photomultiplier Readout ASIC for the Mu3e Experiment* , PhD thesis, Heidelberg University, 2018.

[38] R. Gredig, *Scintillating Fiber Detector for the Mu3e Experiment*, PhD thesis, University of Zurich, 2016.

[39] G. Rutar, *In Search of Charged Lepton Flavor Violating Decays at PSI*, PhD thesis, ETH Zurich, 2017.

[40] A. Damyanova, *Development of the Scintillating Fiber Detector for Timing Measurements in the Mu3e Experiment*, Phd thesis, University of Geneva, 2019.

[41] S. Corrodi, *A Timing Detector based on Scintillating Fibres for the Mu3e Experiment*, Ph.d. thesis, ETH Zürich, 2018.

[42] A. Kuonena, G. Haefeli, M. E. Stramagliaa and O. Girard, Characterisation of the Hamamatsu MPPC multichannel array for the LHCb SciFi Tracker v.11.2016, Technical report, EPFL, 2017.

[43] P. Eckert, *The Mu3e Tile Detector*, Phd thesis, Heidelberg University, 2015.

[44] Y. Qiang, C. Zorn, F. Barbosa and E. Smith, *"Radiation Hardness Tests of SiPMs for the JLab Hall D Barrel Calorimeter"*, Nucl. Instrum. Meth., **A698** 234–241, 2013, (`arXiv:1207.3743 [physics.ins-det]`).

[45] S. Sanchez Majos et al., *"Noise and radiation damage in silicon photomultipliers exposed to electromagnetic and hadronic radiation"*, Nucl. Instrum. Meth., **A602** 506–510, 2009.

[46] H. Klingenmeyer, Y. Munwes, K. Briggl, T. Zhong, H. Chen, W. Shen and H. Schultz-Coulon, *"Measurements with the technical prototype for the Mu3e tile detector"*, Nucl. Instr. Meth., **A958** 162852, 2019.

[47] M. Hesping, *Air Coils for Powering the Mu3e Experiment*, Bachelor thesis, Mainz University, 2019.

[48] S. Gagneur, *Developement of a DC-DC Converter for the Mu3e Detector*, Master thesis, Mainz University, 2020.

[49] A. X. Widmer and P. A. Franaszek, *"A DC-Balanced, Partitioned-Block, 8B/10B Transmission Code"*, IBM Journal of Research and Development, **27** 440, 1983.

[50] P.A. Franaszek and A.X. Widmer, Byte oriented DC balanced (0,4) 8B/10B partitioned block transmission code, December 4 1984, US Patent 4,486,739.

[51] *"FPGA Mezzanine Card (FMC) standard – `https://cds.cern.ch/record/1172409`"*, 2008, Approved in 2008, revised in 2010.

[52] The Digilent Genesys 2 Kintex-7 FPGA development board, `https://reference.digilentinc.com/reference/programmable-logic/genesys-2/`.

[53] C. Ghabrous Larrea, K. Harder, D. Newbold, D. Sankey, A. Rose, A. Thea and T. Williams, *"IPbus: a flexible Ethernet-based control system for xTCA hardware"*, Journal of Instrumentation, **10**(02) C02019, 2015.

[54] R. Schmidt S. Ritt, MSCB (MIDAS Slow Control Bus), 2001, `http://midas.psi.ch/mscb`.

[55] EPICS (Experimental Physics and Industrial Control System), `http://www.aps.anl.gov/epics`.

[56] Altera, Nios II Classic Processor Reference Guide, Technical report, 2015.

[57] M. Müller, *A Control System for the Mu3e Data Acquisition*, Master thesis, Mainz University, 2019.

[58] M. Köppel, *Data Flow in the Mu3e Filter Farm*, Master thesis, Mainz University, 2019.

[59] S. Corrodi, *Fast Optical Readout of the Mu3e Pixel Detector*, Master thesis, ETH Zürich and Heidelberg University, 2014.

[60] C. Grzesik, *Fast Optical Readout for the Mu3e Experiment*, Bachelor thesis, Heidelberg University, 2014.

[61] P. Durante, N. Neufeld, R. Schwemmer, U. Marconi, G. Balbi and I. Lax, 100Gbps PCI-express readout for the LHCb upgrade, In *Proceedings, 19th Real Time Conference (RT2014)*, 2014.

[62] P. Durante, N. Neufeld, R. Schwemmer, G. Balbi and U. Marconi, *"100 Gbps PCI-Express readout for the LHCb upgrade"*, JINST, **10**(04) C04018, 2015.

[63] J.P. Cachemiche, Pcie40 status report, In *LHCb Electronics upgrade WG*, 2015.

[64] NVIDIA Corporation, *NVIDIA GeForce GTX 680*, 2012, Whitepaper.

[65] AMD Corporation, *AMD Graphics Cores Next (GCN) Architecture*, 2012, Whitepaper.

[66] D. vom Bruch, *Pixel Sensor Evaluation and Online Event Selection for the Mu3e Experiment*, PhD thesis, Heidelberg University, 2017.

[67] T. Wagner, *Clock transmission for the Mu3e data acquisition*, Bachelor thesis, Mainz University, 2018.

[68] K. Olchanski S. Ritt, P. Amaudruz, Maximum Integration Data Acquisition System, 2001, `http://midas.psi.ch`.

[69] K. Abe et al., [T2K Collaboration], *"The T2K Experiment"*, Nucl.Instrum.Meth., **A659** 106–135, 2011, (`arXiv:1106.1238 [physics.ins-det]`).

[70] L.M. Brarkov et al., *"Search for $\mu^+ \to e^+\gamma$ down to 10^{-14} branching ratio"*, Research Proposal to PSI, 1999.

[71] J. Allison, K. Amako, J. Apostolakis, H. Araujo, P.A. Dubois et al., *"Geant4 developments and applications"*, IEEE Trans. Nucl. Sci., **53** 270, 2006.

[72] S. Agostinelli et al., *"Geant4–a simulation toolkit"*, Nucl. Instr. Meth., **A 506**(3) 250 – 303, 2003.

[73] R. Brun and F. Rademakers, *"ROOT: An object oriented data analysis framework"*, Nucl. Instrum. Meth., **A389** 81–86, 1997.

[74] L. Urban, *"A model for multiple scattering in Geant4"*, 2006, CERN-OPEN-2006-077.

[75] V. Ivanchenko et al., [Geant4 Collaboration], *"Progress of Geant4 electromagnetic physics developments and applications"*, EPJ Web Conf., **214** 02046, 2019.

[76] N. Berger et al., *"Multiple Coulomb Scattering in Thin Silicon"*, JINST, **9** P07007, 2014, (`arXiv:1405.2759 [physics.ins-det]`).

[77] F. Scheck, *"Muon Physics"*, Phys.Rept., **44** 187, 1978.

[78] W.E. Fischer and F. Scheck, *"Electron Polarization in Polarized Muon Decay: Radiative Corrections"*, Nucl.Phys., **B83** 25, 1974.

[79] P. Depommier and A. Vacheret, Radiative muon decay, Technical report, TWIST Technote No 55, 2001.

[80] C. Fronsdal and H. Uberall, *"mu-Meson Decay with Inner Bremsstrahlung"*, Phys.Rev., **113** 654–657, 1959.

[81] P. Biallass and T. Hebbeker, *"Parametrization of the Cosmic Muon Flux for the Generator CMSCGEN"*, 2009, (`arXiv:0907.5514 [astro-ph.IM]`).

[82] N. Berger, M. Kiehn, A. Kozlinskiy and A. Schöning, *"A New Three-Dimensional Track Fit with Multiple Scattering"*, Nucl. Instr. Meth A., **844 C** 135–140, 2017, (`arXiv:1606.04990 [physics.ins-det]`).

[83] V. Blobel, C. Kleinwort and F. Meier, *"Fast alignment of a complex tracking detector using advanced track models"*, Comput.Phys.Commun., **182** 1760–1763, 2011, (`arXiv:1103.3909 [physics.ins-det]`).

[84] C. Kleinwort, *"General Broken Lines as advanced track fitting method"*, Nucl.Instrum.Meth., **A673** 107–110, 2012, (`arXiv:1201.4320 [physics.ins-det]`).

[85] S. Liechti, *Particle Track reconstruction using a recurrent neural network at the Mu3e experiment*, Bachelor thesis, Zürich University, 2018.

[86] A.-K. Perrevoort, *Sensitivity Studies on New Physics in the Mu3e Experiment and Development of Firmware for the Front-End of the Mu3e Pixel Detector*, PhD thesis, Heidelberg University, 2018.

[87] U. Hartenstein, *Track Based Alignment for the Mu3e Pixel Detector*, PhD thesis, Mainz University, 2019.

[88] C. Kleinwort, H1 alignment experience, In *Proceedings, first LHC Detector Alignment Workshop, CERN, Geneva, Switzerland, 4-6 September 2006*, pages 41–50, 2006.

[89] T. Lampén, [CMS Collaboration], *"Alignment of the CMS silicon tracker"*, J. Phys. Conf. Ser., **523** 012024, 2014.

[90] S. Chatrchyan et al., [CMS Collaboration], *"Alignment of the CMS tracker with LHC and cosmic ray data"*, JINST, **9** P06009, 2014, (`arXiv:1403.2286 [physics.ins-det]`).

[91] N. Bartosik, [CMS Collaboration], *"Simultaneous alignment and Lorentz angle calibration in the CMS silicon tracker using Millepede II"*, PoS, **EPS-HEP2013** 074, 2013.

[92] G. Flucke, [CMS Collaboration], *"Alignment of the CMS silicon tracker"*, J. Phys. Conf. Ser., **368** 012036, 2012.

[93] J. Behr, [CMS Collaboration], *"Alignment procedures for the CMS silicon tracker"*, J. Phys. Conf. Ser., **396** 022005, 2012.

[94] J. Draeger, *Track based alignment of the CMS silicon tracker and its implication on physics performance*, PhD thesis, Hamburg U., 2011.

[95] M. Weber, [CMS Collaboration], *"Calibration, alignment and tracking performance of the CMS silicon strip tracker"*, Nucl. Instrum. Meth., **A628** 59–63, 2011.

[96] V. Blobel, *"Software alignment for tracking detectors"*, Nucl. Instrum. Meth., **A566** 5–13, 2006.

[97] S. Schenk, *A Vertex Fit for Low Momentum Particles in*

a Solenoidal Magnetic Field with Multiple Scattering, Master's thesis, Heidelberg University, 2013.