17 September 2024
amine.benmahmoud@student.uantwerpen.be

Exploring Detection Asymmetries in the FORMOSA Demonstrator

Amine Ben Mahmoud
FORMOSA, CERN, Geneva, Switzerland

Summary

This report describes the project I was involved in during my internship at CERN, working on the FORMOSA experiment. I contributed to developing Uproot code that uses datasets from the FORMOSA demonstrator to compare different regions and create heatmaps of the channels to understand how radiation and background rates are distributed. In this report, we also briefly explain the underlying theory and the main concepts behind the detectors from which the data was obtained.

CERN-THESIS-2024-182
11/10/2024

Name of University: University of Antwerp

CERN Supervisor: Matthew Daniel Citron

University Supervisor: Pierre Van Mechelen

Contents

1 Introduction

The quest to understand the nature of dark matter remains one of the most profound challenges in modern physics. Among the various hypotheses, the concept of millicharged particles (mCPs) has emerged as a promising avenue to bridge the gap between ordinary matter and dark matter. Millicharged particles are hypothesized to carry a tiny charge—much smaller than that of the electron—allowing them to interact weakly with electromagnetic fields. A compelling theoretical framework suggests that millicharged particles arise through kinetic mixing between the standard photon and a hypothetical dark photon, which is the gauge boson of a hidden U(1) symmetry [1, 2].

The minimal extension to the Standard Model that includes kinetic mixing between the standard photon A_μ and a dark photon A'_μ can be described by the following lagrangian,

$$\mathcal{L} = \mathcal{L}_{\text{SM}} - \frac{1}{2}\epsilon A_{\mu\nu} A'^{\mu\nu} - \frac{1}{4} A'_\mu A'^\mu + i\bar{\chi}\left(\not{\partial} + ie'\not{A} + iM_\chi\right)\chi \qquad (1)$$

where:

- $\mathcal{L}_{\text{SM}}$ is the Standard Model Lagrangian,
- $A_{\mu\nu}$ and $A'_{\mu\nu}$ are the field strength tensors of the photon and dark photon, respectively,
- ϵ is the kinetic mixing parameter, governing the interaction strength between the two photons,
- $ie' \not{A}'$ term describes the interaction between the millicharged particle and the dark photon field A'_μ.
- iM_χ is the mass term for the millicharged particle.

The kinetic mixing term $\epsilon A_{\mu\nu} A'^{\mu\nu}$ induces a small coupling between the Standard Model photon and the dark photon, allowing millicharged particles to interact with electromagnetic fields despite their small charge.[1, 2]

While this explanation of millicharged particles is brief, we do not intend to delve deeper into the theoretical details, as it is not the central focus of this research. However, understanding the significance of these particles and their potential connection to the dark matter realm provides a strong motivation for their detection. The possibility that millicharged particles could serve as a bridge between the visible and dark sectors makes their discovery an exciting prospect in the quest to unravel the mysteries of dark matter.

Due to the lack of sensitivity of LHC detectors in searching for millicharged particles, dedicated detectors like MilliQan have been designed [2]. The FORMOSA detector is a new proposed detector designed specifically to improve the sensitivity for detecting millicharged particles in the forward region of the LHC. The FORMOSA detector is inspired by the MilliQan detector, and while MilliQan is sensitive to mCPs in the central region [3], the forward region has an mCP production cross-section 250 times larger than that of the central region, where the FORMOSA detector is intended to be placed. FORMOSA employs a combination of scintillators and photomultipliers, similar to the MilliQan detector, to detect the weak signals from these particles. The weak signals refer to the minimal energy deposits expected from millicharged particles, which have a very small electric charge, thus leading to smaller interactions with the detector materials.

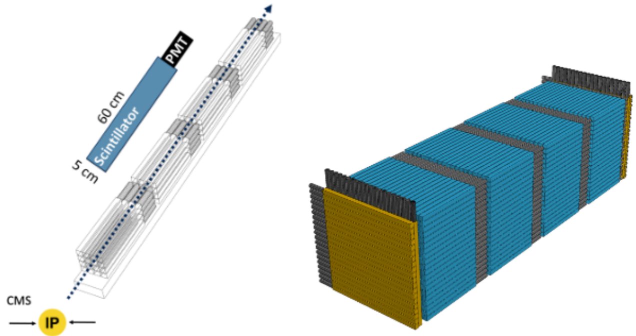

Figure 1: Left: Diagram of the MilliQan detector, consisting of four layers of scintillators with PMTs (photomultiplier tubes) attached to them. "IP" stands for the "interaction point" at the CMS. Right: Diagram of the full-sized FORMOSA detector, including the front and back muon panels (used to detect muons).[4]

Given its high sensitivity, FORMOSA provides a unique opportunity to probe theories which predict mCPs. However, this sensitivity, along with its position in the forward region, where muons are produced more frequently, also introduces a significant challenge: the presence of substantial background. Therefore, the first critical step in the search for millicharged particles is to address a significant challenge: distinguishing genuine particle signals from high background rates. This is the primary focus of our research in this report. We aim to analyze how to tackle these high rates in order to make an mCP search feasible.

2 The Formosa Demonstrator

The FORMOSA Demonstrator is a smaller prototype of the full FORMOSA detector, designed to test and refine the core detection techniques that will be used in the final version. It follows the design principles of the MilliQan detector, but with optimizations for its specific goals.

As mentioned earlier, the weak electromagnetic coupling of millicharged particles means their signals are expected to be characterized by low-energy deposits in the detector's layers. They are more detected by the PMTs, with the FORMOSA Demonstrator consisting of four layers of scintillators, each containing four channels (as shown in Figure 2). To identify an mCP, the detector requires small, energy deposits across all four layers within a narrow time window. This approach effectively filters out random noise, which is unlikely to produce consistent hits in all layers within the same time frame.

Now, When a particle passes through a scintillator bar, it emits light, which is then converted into an electronic signal by photomultiplier tubes (PMTs). These PMTs are sensitive enough to detect even single photons, converting the detected light into electrical pulses. Ensuring the PMTs are responsive to such faint signals is critical, given the low-energy nature of mCP interactions. Additionally, the scintillator bars are designed with a substantial volume (in the shape of bars), which increases the probability of particle interactions and improves the overall sensitivity of the detector.

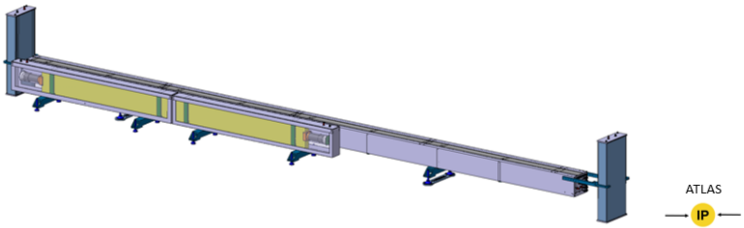

Figure 2: The FORMOSA Demonstrator, consisting of four layers of scintillators, with front and back panels designed to detect and filter out background muons.[4]

Additionally, the front and back muon panels are designed to identify and filter out beam muons. One major challenge posed by the abundance of muons is an effect called afterpulsing. After large pulses, such as those caused by high-energy muons, positive ions drift slowly back to the cathode. When these ions reach the cathode, they can release additional electrons, resulting in a secondary pulse known as an afterpulse. The issue with this effect is that these afterpulses can mimic mCP deposits in the scintillators, complicating the identification of genuine mCP events. This effect is visualized clearly in Figure 3.

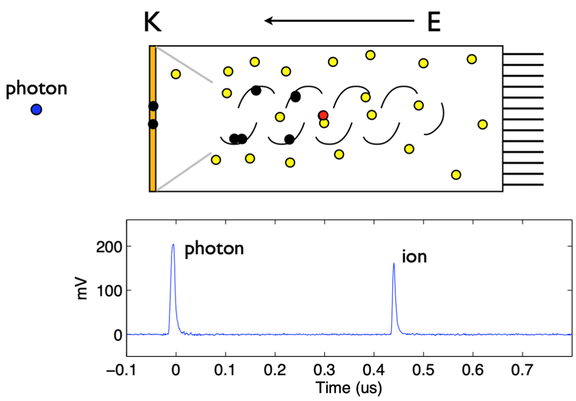

Figure 3: Schematic illustrating the process of afterpulsing. The graph below the diagram shows the main photon pulse (left spike) followed by the ion pulse (right spike), highlighting the afterpulsing effect.[5]

This further emphasizes the need to understand the backgrounds that may mimic the signal of an mCP.

For the FORMOSA demonstrator, radiation from the LHC beam must also be considered. To mitigate this, concrete shielding is placed next to the demonstrator, although part of the detector—the back—remains exposed. In the case of the full FORMOSA detector, which will be located in a separate facility away from the LHC, radiation from the beam will not be a concern, though muons will still reach the detector. Thus, radiation is only an issue for the current location of the demonstrator.

2.1 Motivation for the Location

Recall that the primary difference between the MilliQan and FORMOSA detectors, lies in their locations. Here, we will briefly explain why the location affects the cross-section of millicharged particles (mCPs). To do so, we need to understand the concept of pseudorapidity (denoted as η), a coordinate system used in collider experiments to describe the angle of a particle relative to the beam axis. The pseudorapidity angle ranges from 0 to infinity.[7]

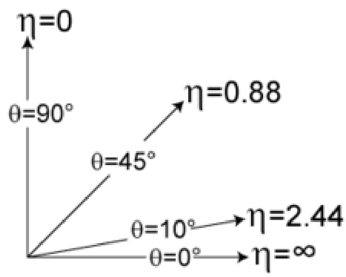

Figure 4: The pseudorapidity angle η illustrated. As particles move closer to the beam axis, the pseudorapidity increases, with interactions occurring at $\eta = \infty$.[7]

The MilliQan detector is located at $\eta \approx 0$, meaning it collects particles emitted more centrally with respect to the beam axis. mCPs are produced evenly across the pseudorapidity range, which limits the number of mCPs collected by MilliQan, as it focuses on a narrower angular range around the beam.

FORMOSA's position in the forward region, where the pseudorapidity is higher, allows it to detect particles across a broader pseudorapidity range compared to MilliQan. The forward region has a higher particle density, increasing the likelihood of producing mCPs by a factor of 250 relative to the central region. Additionally, FORMOSA's placement enables it to subtend a larger solid angle, which describes the detector's coverage around the beam axis in 3D space. This combination of broader pseudorapidity coverage and increased solid angle enhances FORMOSA's ability to capture more particles produced along the beam axis.

2.2 Detector Map

As mentioned before, the FORMOSA demonstrator consists of four layers of scintillators, with each layer containing four separate scintillator bars arranged in a square configuration

(2x2), resulting in a 2x2x4 scintillator setup. Each scintillator functions as a channel through which a particle can pass, giving us a total of 16 separate channels. Additionally, there are 2 extra channels—the front and back panels (there are also 2 side panels, but they are not relevant to my research). In Figure 5, these channels are illustrated, showing an incoming collision muon entering from the left (front panel, channel 16) and moving towards the back panel on the right (channel 18) as it passes through the scintillators.

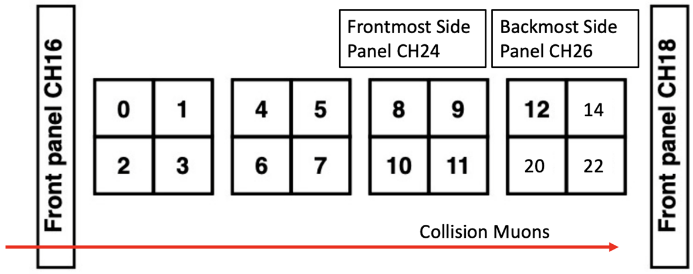

Figure 5: Illustration of the FORMOSA demonstrator's channel layout. The red arrow indicates the direction of incoming muons from collisions.[8]

As previously mentioned, the goal of this research is to study the asymmetry within the detector by comparing different regions and analyzing the number of events detected in each. To facilitate this, we have divided the scintillator channels into four primary regions, defined as follows:

- **Top Region**: Channels 0, 1, 4, 5, 8, 9, 12, and 14
- **Bottom Region**: Channels 2, 3, 6, 7, 10, 11, 20, and 22
- **Left Region**: Channels 0, 2, 4, 6, 8, 10, 12, and 20
- **Right Region**: Channels 1, 3, 5, 7, 9, 11, 14, and 22

In our analysis we will be comparing the top region with the bottom region to assess vertical asymmetry. Following this, we will compare the left and right regions to evaluate any horizontal asymmetry.

Before diving into the analysis, we must first discuss the detector's triggering system. Naturally, we don't want to accept every signal the detector picks up, instead, we need to impose certain restrictions to determine when the detector should register an event. This will be the focus of the next section.

3 Triggers Used to Detect a Signal

The detection system used in the FORMOSA detector relies on a set of triggers. These triggers are essential for determining which events should be recorded by the detector, filtering out irrelevant or unwanted signals to focus on meaningful data:

- **MilliQan Triggers**: These are six different triggers that were originally developed for the MilliQan detector and are still used for FORMOSA, which is why they retain the name:
  - *Four Layers Trigger*: Activated when a particle passes through and hits all four layers of the detector. This trigger ensures the particle has traversed the full length of the detector and is useful for identifying particles with consistent trajectories.
  - *Three in a Row Trigger*: Activated when hits are registered in three consecutive layers. These hits need to be in the same row and must appear in adjacent layers, allowing the detector to track particles passing through multiple layers in a linear path.
  - *Two Separate Layers Trigger*: Activated when hits occur in two non-adjacent layers. This is a trigger with looser requirements than the signal triggers, primarily used to check the performance of signal triggers and to assess background rates.
  - *Two Adjacent Layers Trigger*: Triggered when two adjacent layers register hits. This trigger helps capture particles that only penetrate partially through the detector, hitting consecutive layers.
  - *N Layers Trigger*: A customizable trigger that activates when hits occur in a specific number of layers. For example, if $N = 1$, a single hit in any one layer will activate the trigger.
  - *N Hits Trigger*: This trigger focuses on the total number of hits anywhere in the detector, without regard to which specific layers are hit. It provides a broader detection approach for capturing distributed signals across the detector.
- **Signal Triggers**: These triggers are designed specifically for detecting detecting signal events.
  - *topOnePerLayer Trigger*: This specialized trigger activates when one bar per layer registers a hit, as we would expect from mCP-like deposits.
  - *botOnePerLayer Trigger*: Similar to the topOnePerLayer trigger, this one activates when only one scintillator bar in the bottom layer registers a hit.
- **Zero Bias Trigger**: This trigger fires at a configurable frequency set in the DAQ firmware, based on clock cycles, allowing control over the number of clock ticks between successive triggers.

When collecting data (starting a run), we can customize the set of triggers we want to apply for that specific run. For this research, we used four different runs, each with a unique combination of the triggers explained earlier. The runs are:

- **Run 737**: This run used the MilliQan triggers along with the Zero Bias trigger.
- **Run 1072**: This run employed only the Signal Triggers.
- **Run 1076**: This run combined the MilliQan triggers, the Zero Bias trigger, and the Signal Triggers.

- **Run 1229**: In this run, we set the N-layers trigger to 4, requiring hits in exactly four layers of the detector.

The number of events we obtain depends on the combination of triggers used—certain runs, like Run 1076, will provide a broader range of events, from full detector traversals to isolated low-energy signals. As mentioned in the previous section, we will compare the number of events across different channel regions to assess the asymmetry of the detector.

4 Data Analysis

In this section, we will dive into the data analysis of the events collected during the experiment. To process and analyze the data, we utilized Uproot, a Python library designed for reading and handling ROOT files. One of the advantages of Uproot over ROOT is its faster processing speed, significantly reducing the time spent running our code. However, a downside of Uproot is that it can be more challenging when applying masks, which are cuts on the full data set to isolate only the relevant data for our research.

4.1 Masks

- **Time Cut**: We focus on events occurring between a time window of 400-600 ns. This ensures we are only looking at pulses within the coincidence window, where meaningful detector activity is expected. Pulses outside this window are likely unrelated background noise.

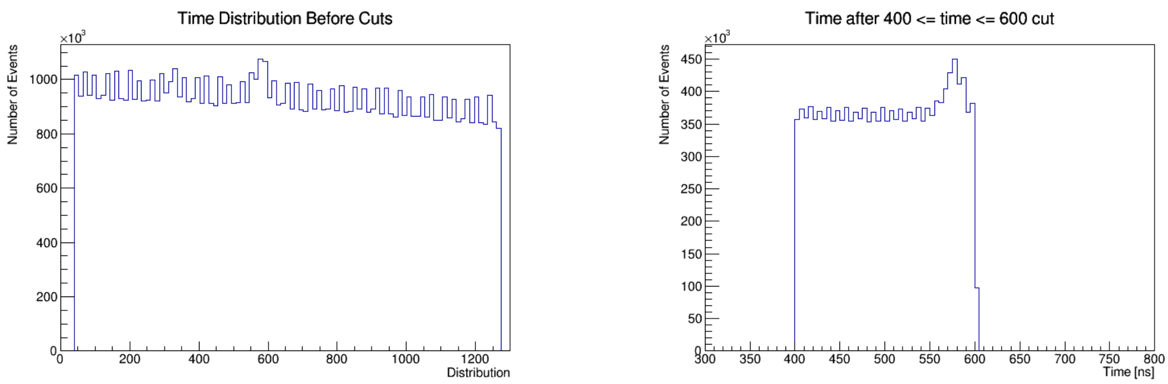

Figure 6: Comparison of initial time distribution (left) and applied time cut (right) for run 1072.

- **Ipulse Cut**: The ipulse variable refers to the pulse index, which indicates the order in which pulses are detected in a specific channel. In many particle detectors, multiple pulses can occur in quick succession due to primary signals or afterpulsing. By setting ‘ipulse == 0‘, we take the primary signals and exclude the secondary stuff in the channels.

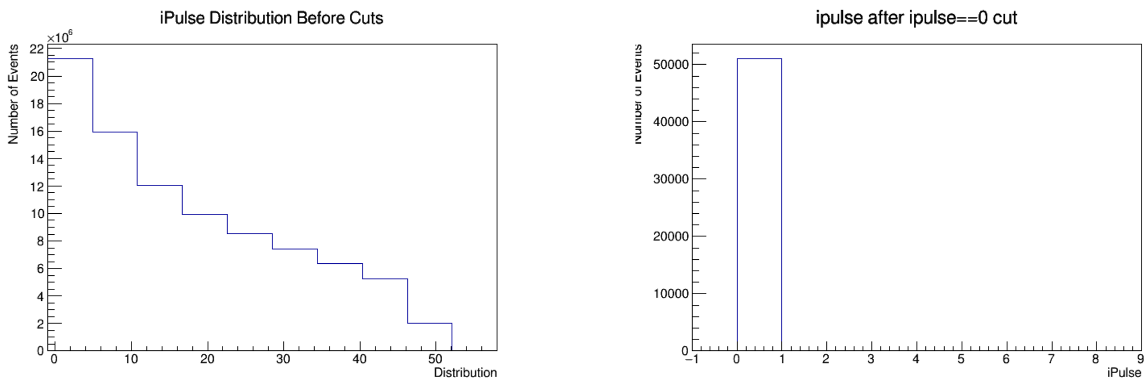

Figure 7: Comparison of initial time distribution (left) and applied iPulse cut (right) for run 1072.

- **Four Layers Cut**: Finally, we select only the events where hits are detected in all four layers of the detector.

These three cuts form the baseline for our analysis. Depending on the additional cuts we apply, we will either focus on muonic or non-muonic activity. The conditions for both are as follows:

- **Muonic Activity**:
  - *Four Layers Area Cut*: In addition to the hits in each layer, we apply thresholds based on the area of the incoming signal (pulse area). By setting a high threshold, we ensure that we are focusing on regions with significant muon activity, as muons are high-energy particles that generate large signal areas. Specifically, we set a minimum area threshold of 95,000 pVs for each layer.

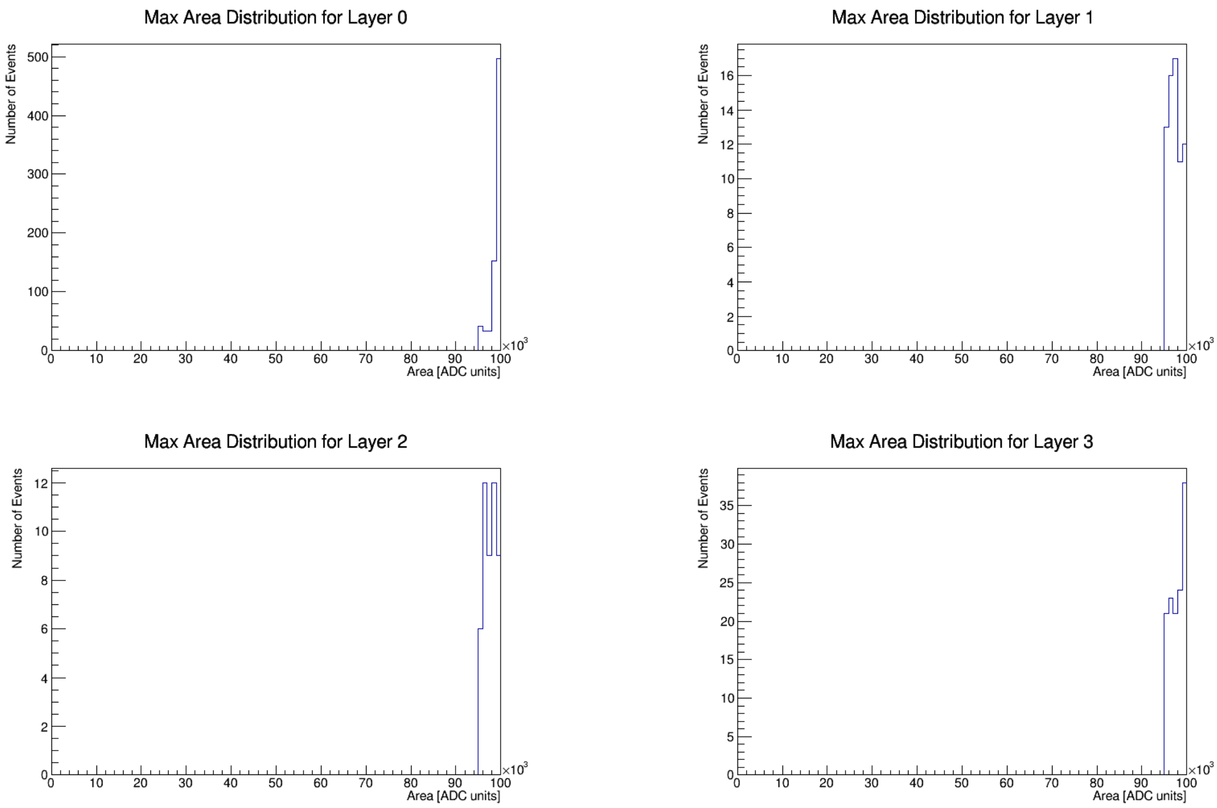

Figure 8: Area cuts are displayed for each layer (where, for example, layer 0 corresponds to channels 0, 1, 2, and 3 as shown in Figure 5). The histogram shows that we only observe events with areas above the threshold of 95,000 pVs.

- *Muon Panels Cut*: We also require hits in both muon panels (channels 16 and 18), with signal areas above a certain threshold. For channel 16, the minimum threshold is 30,000 pVs, and for channel 18, it is 18,000 pVs.

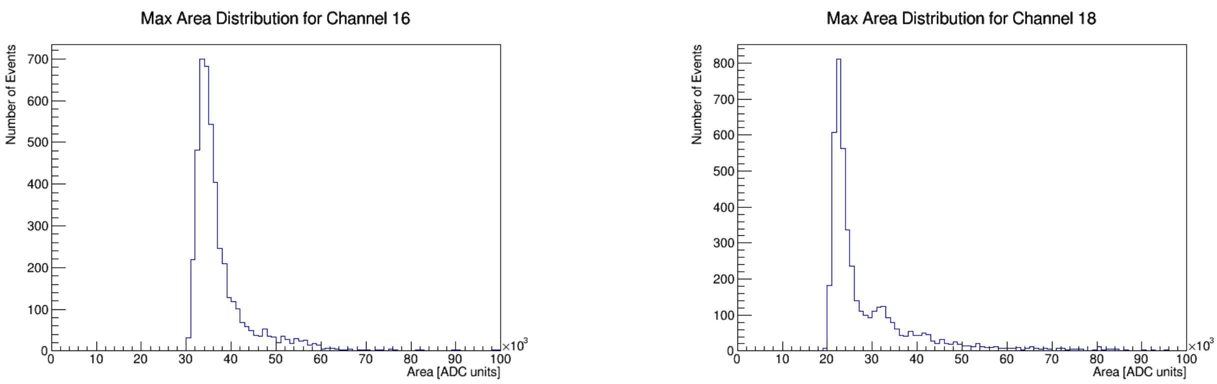

Figure 9: Area cuts are displayed for the muon panels. The histogram shows that we only observe events with areas above the threshold of 30,000 pVs for channel 16 (front panel) and 18,000 pVs for channel 18 (back panel).

- **Non-Muonic Activity**:
  - *No Hits in Muon Panels*:For non-muonic activity, we focus on events where there are no hits in the muon panels. Since muons deposit a significant amount of energy into the detector bars, and we are not looking for muons in this case, we do not apply any area thresholds.

We can now begin comparing the different regions for both the muonic and non-muonic cases.

4.2 Comparing The Different Regions

4.2.1 Non-Muonic Activity

Let's begin by examining the data where we don't consider muonic activity and the histograms we generated, as this analysis was straightforward and presented no complications. Although we previously discussed the different runs we used, the conclusions for each run remain the same. The only noticeable difference in the histograms is the number of events; however, when discussing the results, we can generalize them across all runs. Starting with the comparison of the left and right regions, we have Figure 10.

We can first note that we convert the number of events into an event rate by dividing by the bin width to normalize the data. This ensures that the event rate is comparable across all bins, regardless of their width, and provides a clearer representation of how often events occur within a specific range. Additionally, event rates have physical significance, allowing us to interpret the data in terms of frequency, such as events per second. Looking at Figure 10, we can clearly see that the left region has more events than the right region. In Figure 11, where we compare the top and bottom regions, it is evident that the top region dominates.

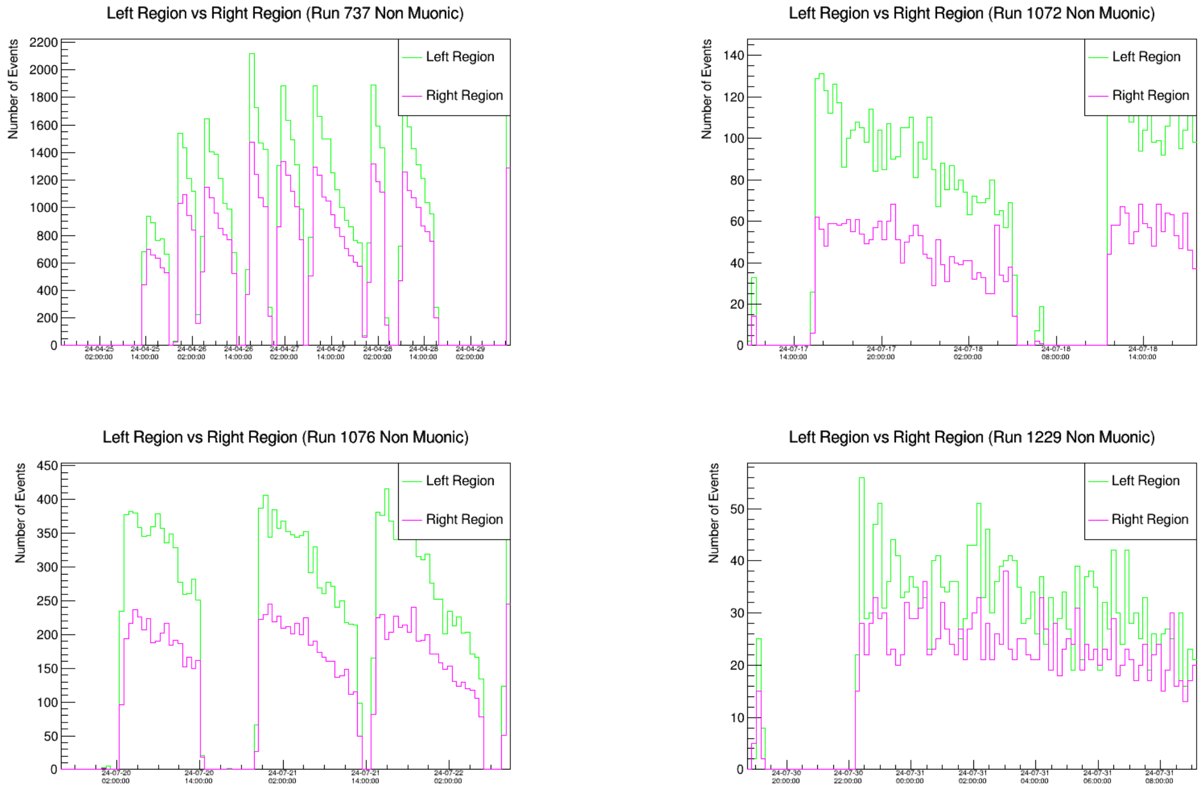

Figure 10: Comparison of the left and right regions for non-muonic activity across four different runs. Each plot shows the event rates (in Hz) for the left and right regions.

It is worth noting that the shape of the histograms closely resembles the well-known LHC beam intensity monitoring plots, as shown in Figure 12. In the case of the non-muonic activity histograms, the shape mirrors that of the LHC beam energy, suggesting that the number of events is directly related to the beam energy, which is a logical conclusion.

Finally, we can make the same comparison by plotting heatmaps for each layer, providing a better visualization of which specific channels have more events. In Figure 13, we show the heatmaps for each run, displaying the 2x2 scintillator channels as seen in Figure 5, arranged in a staircase form with the red line indicating the collision muons. The conclusions drawn from the histograms, namely, that the left and top regions have more events, are consistent with the heatmaps. Note that as we move further into the detector, the number of events increases significantly (as seen clearly in run 737). This is likely due to the radiation discussed earlier, where the detector extends beyond the concrete wall, resulting in a higher number of events.

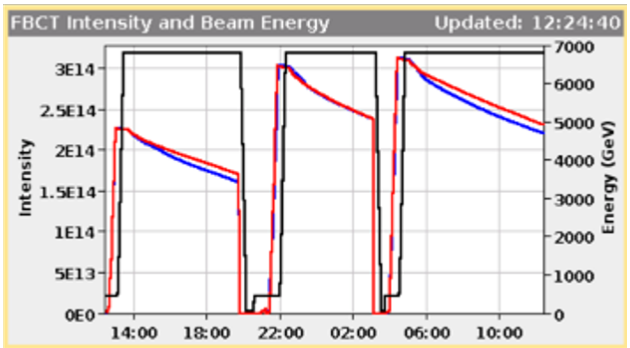

Figure 12: FBCT intensity and beam energy over time. The graph displays the intensity (red and blue lines) and beam energy (black line) of the LHC beam, showing the ramp-up and decay cycles across different time intervals.[9]

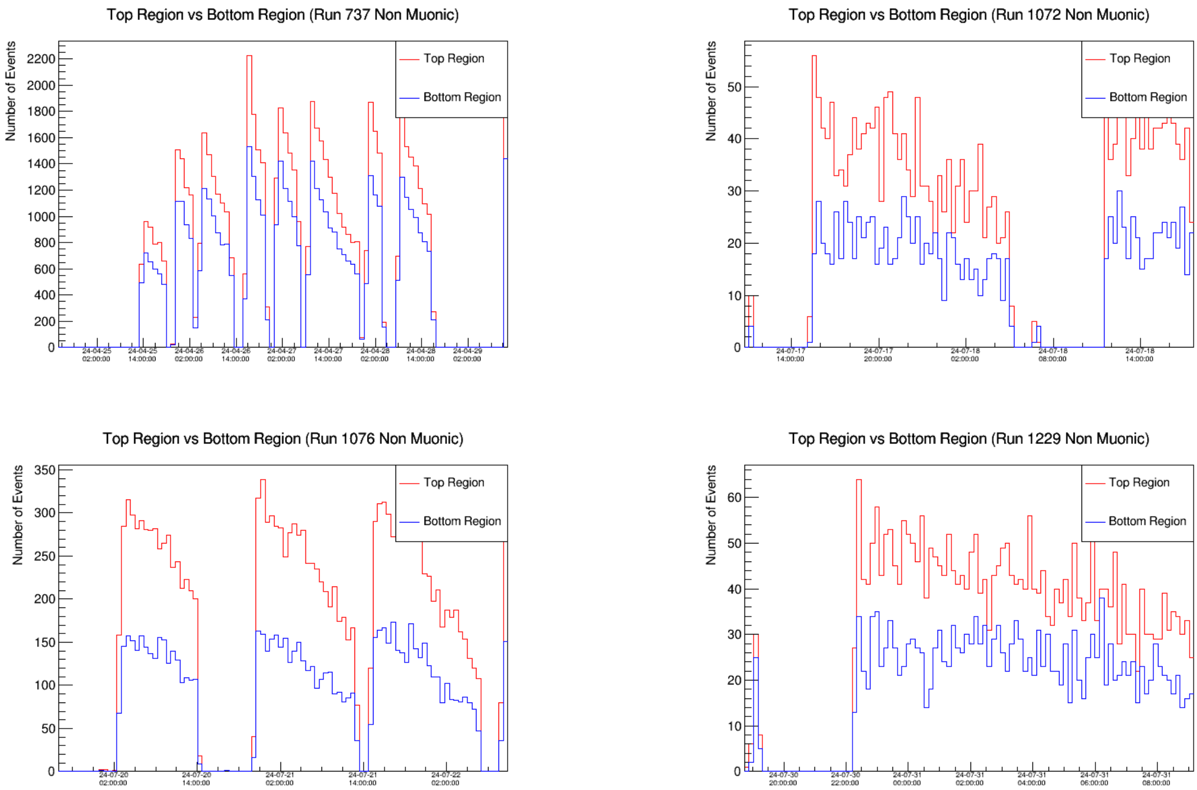

Figure 11: Comparison of the top and bottom regions for non-muonic activity across four different runs. Each plot shows the event rates (in Hz) for the left and right regions.

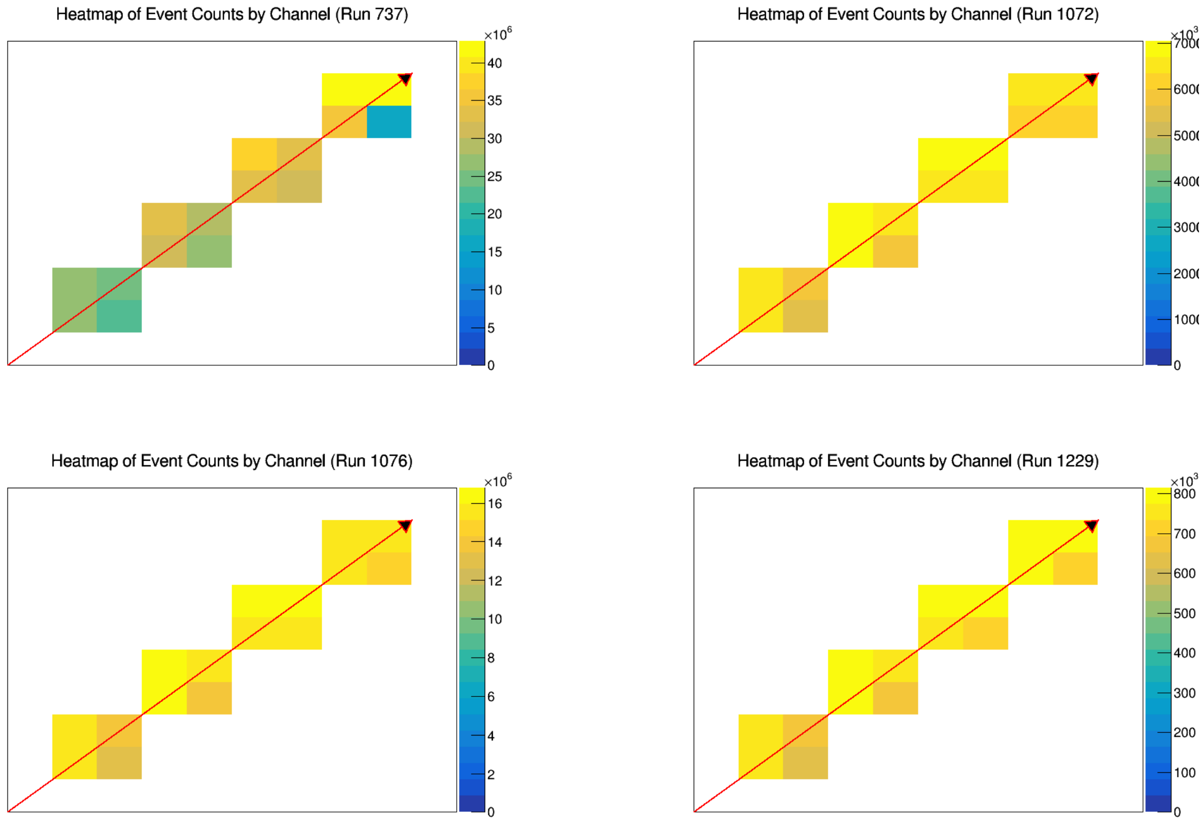

Figure 13: Heatmaps showing the event counts by channel for each run (non-muonic activity). The 2x2 scintillator channels are arranged in a staircase form, with the red arrow indicating the direction of muons coming from collisions at ATLAS. Top left: Run 737; Top right: Run 1072; Bottom left: Run 1076; Bottom right: Run 1229

4.2.2 Muonic Activity

When including the muonic activity using our masks, we can compare the left and right regions in Figure 14. From this figure, we conclude that the left and right regions have approximately the same number of events. Similarly, comparing the bottom and top regions in Figure 16, we observe that the number of events is similar for runs 737 and 1229. However, for runs 1076 and 1072, the top region is more dominant.

The issue with these results is that, when we compare them to the findings from the FORMOSA workshop [10], the results for the top vs. bottom region indicate that the bottom region should have more events. However, when we look at Figure 17, we see that this isn't the case. Additionally, we should expect, that the shapes of these distributions closely resemble those of the ATLAS instantaneous luminosity plots as shown in figure 15. This graph provides the rate of particle collisions, which should align with the scaling of the muonic activity histograms—since the number of muons is directly related to the number of particle collisions.

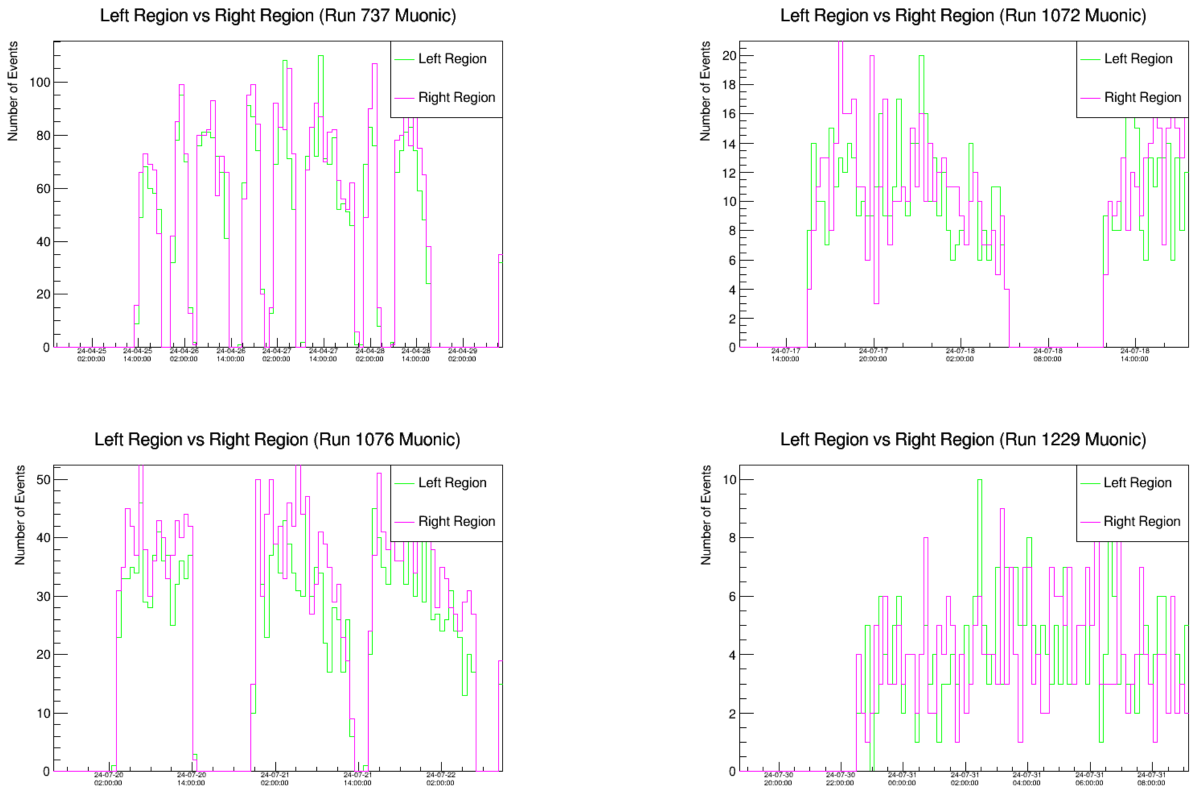

Figure 14: Comparison of the left and right regions for muonic activity across four different runs. Each plot shows the event rates (in Hz) for the left and right regions.

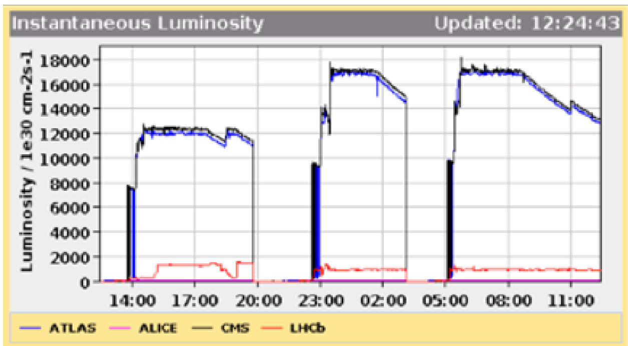

Figure 15: Instantaneous luminosity over time for different experiments at the LHC (ATLAS, ALICE, CMS, LHCb). The plot shows variations in luminosity throughout the day, indicating distinct operational periods for each experiment.[9]

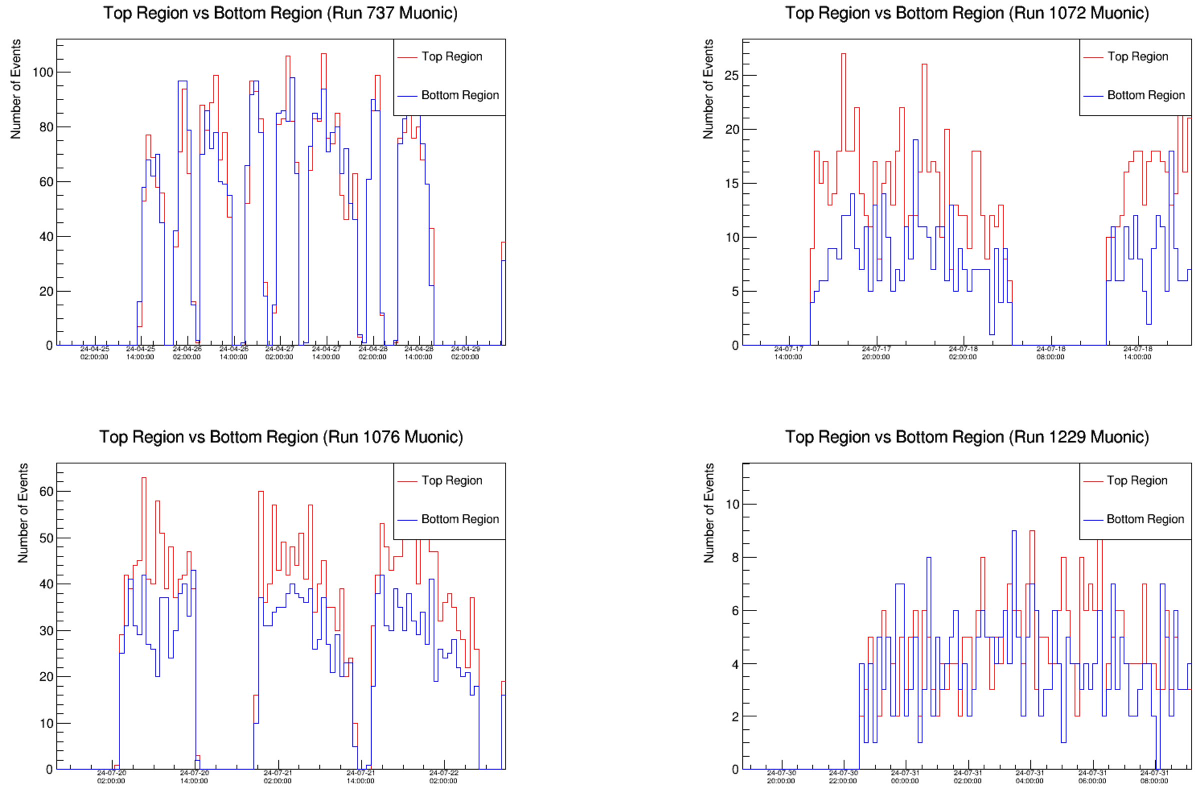

Figure 16: Comparison of the top and bottom regions for muonic activity across four different runs. Each plot shows the event rates (in Hz) for the left and right regions.

Addressing the issue now, we initially suspected a bug in the Uproot script used to generate the histograms. We rewrote the code multiple times in an effort to debug the problem, but this consistently led to the same plots as shown in Figure 16. The main difference between our results and those from the FORMOSA workshop is that their analysis was generated using a ROOT script (without Python). Despite our attempts to replicate the workshop's method by using ROOT instead of Uproot, and expecting the bottom region to show more events, we continued to obtain the same histograms, with the top region showing more events.

The issue seems to arise from radiation originating from the LHC beam. Consider a muon entering the detector: as mentioned earlier, the last section of the detector extends beyond the concrete shielding, resulting in it receiving significantly more radiation compared to the front of the detector. This excess radiation contaminates the muon sample. Specifically, there may be radiation pulses with higher amplitudes than those from muons. This is problematic because, if our code is designed to select events by taking only the maximum pulse, we risk counting non-muonic events, leading to incorrect analysis. While it might seem like an easy fix to simply remove the requirement of selecting the maximum value in our Uproot file, this condition was never explicitly added to the code. Therefore, debugging this issue will likely require revisiting the fundamentals of the code and making adjustments there.

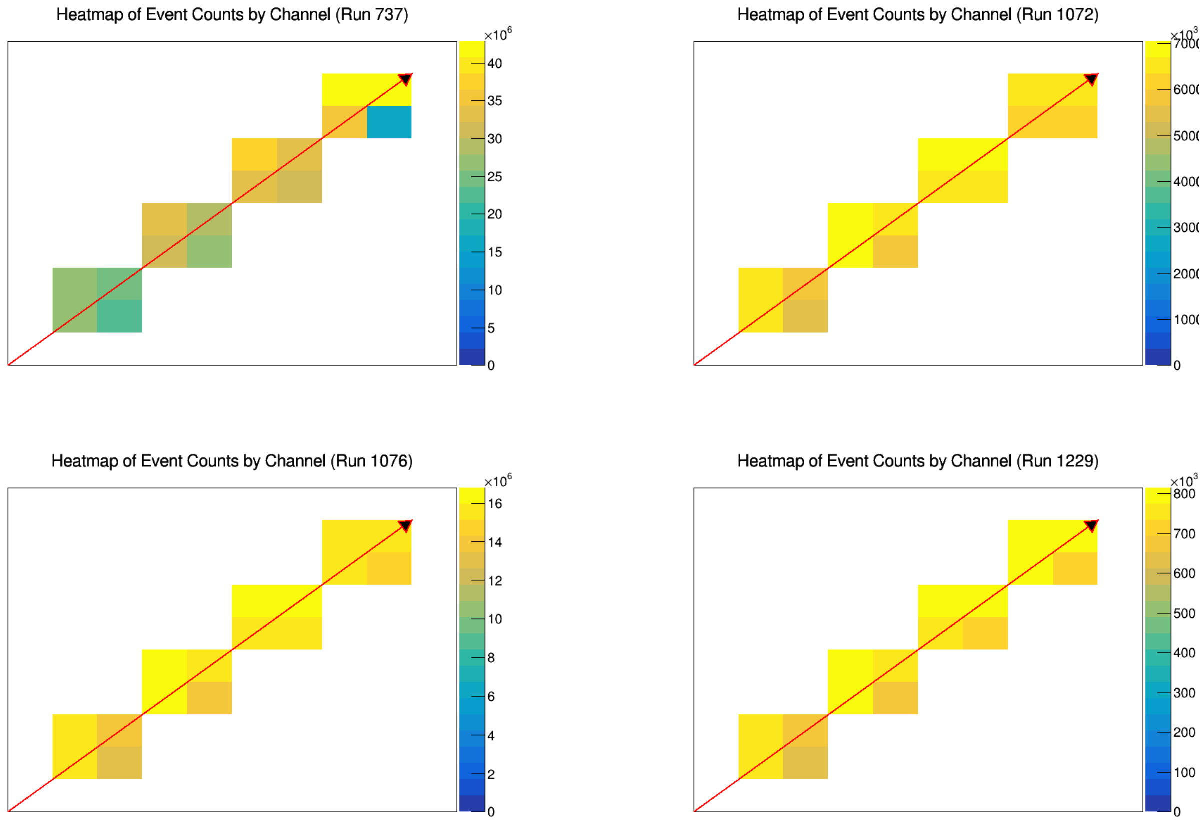

Figure 17: Heatmaps showing the event counts by channel for each run (muonic activity). The 2x2 scintillator channels are arranged in a staircase form, with the red line indicating the collision muons. Top left: Run 737; Top right: Run 1072; Bottom left: Run 1076; Bottom right: Run 1229

5 Conclusion

In this study, we analyzed data from the FORMOSA demonstrator, which is positioned in the forward region of ATLAS, next to the LHC beam. Our focus was on characterizing the asymmetry in particle detection across different regions of the detector and distinguishing between muonic and non-muonic events. To achieve this, we divided the detector into four regions, top, bottom, left, and right, and compared the event rates to evaluate asymmetries. We utilized four datasets from different runs, each with distinct trigger configurations to filter and focus on specific event types.

In our analysis of non-muonic events, we noticed that the left and top regions had more activity. For muonic events, things varied between different runs, some showed an even spread of events, while others had more in the top region. This wasn't what we expected based on earlier results from the FORMOSA workshop, where we thought there would be more events in the bottom region. We also expected a stronger match between the muonic event patterns and the LHC beam collision rate. The differences we saw suggest there might be issues with radiation contamination, especially in the back part of the detector that sticks out beyond the concrete shielding. Radiation pulses in that area could have high amplitudes that interfere with actual muonic signals, making it harder to correctly identify events.

Despite these challenges, this study gives us important insights into how the FORMOSA demonstrator behaves. By looking at event asymmetries and understanding the radiation backgrounds, we're taking crucial steps toward figuring out if a milliQan-like detector could work well in the forward region of the LHC. These results will help us make improvements in the design and data collection process, and ultimately see if a detector in this position can effectively help in the search for millicharged particles.

Acknowledgements

I am grateful to Matthew Daniel Citron for giving me the opportunity to work on this fascinating project and for allowing me to be part of an amazing group during my internship. His guidance and expertise have significantly helped me understand the underlying concepts of this topic much better. For the data analysis and coding aspects of this thesis, I received exceptional guidance from Jacob Henry Steenis, whose support greatly enhanced my understanding of how to use Uproot (and a bit of ROOT), as well as provided me valuable insights into the detectors, an area where I previously had no experience. Under his guidance, everything became clearer, and many (coding) challenges were easily overcome.

I would also like to express my gratitude to Pierre Van Mechelen, who referred me to the team and made it possible for me to attend this internship at CERN. During my time there, I even had the opportunity to visit the CMS tunnels, where I worked with the MilliQan detector for a week, gaining valuable hands-on experience.

I would like to extend my heartfelt thanks once again to all the individuals mentioned above for making this work possible. This has been a wonderful experience, and exploring such an interesting topic has been truly rewarding. My time at CERN was incredible and an enriching experience, providing me with valuable insights into the scientific world and how it truly functions.

References

[1] CMS Collaboration, "Search for fractionally charged particles in proton-proton collisions at $\sqrt{s} = 13$ TeV," arXiv:2402.09932, 2023.

[2] MilliQan Collaboration, "Sensitivity to millicharged particles in future proton-proton collisions at the LHC," arXiv:2104.07151, 2021.

[3] A. Ball *et al.*, "Search for millicharged particles in proton-proton collisions at $\sqrt{s} = 13$ TeV," *Phys. Rev. D*, vol. 102, 2020, p. 032002. arXiv:2005.06518.

[4] UC Davis, "Formosa", https://pex.physics.ucdavis.edu/formosa

[5] C. Carlos Hernandez Faham, "Afterpulsing"

[6] S. Foroughi-Abari, F. Kling, and Y.-D. Tsai, "Looking forward to millicharged dark sectors at the LHC," *Phys. Rev. D*, vol. 104, 2021, p. 035014. arXiv:2010.07941.

[7] Wikipedia, "Pseudorapidity", https://en.wikipedia.org/wiki/Pseudorapidity

[8] UC Davis, "FORMOSA DAQ", Presentation given by Jacob Steenis

[9] CERN, "LHC page 1", From the CERN vistars page (status of the beam), https://op-webtools.web.cern.ch/vistar/

[10] Private Communication