Analysis of LHCb Luminosity Counters for Run 3

Omar Sayeed Saimum [Saimum]

3rd Year Undergraduate Student

Department of Physics, KAIST, Daejeon, Republic of Korea

Intern, Luminosity Group of LHCb, LLR, IN2P3/CNRS, Ecole Polytechnique, France

Contact: omars.saimum@kaist.ac.kr/cern.ch

Table of Contents

Abstract

Luminosity, an important parameter for any physics analysis at LHCb, demands attention to the tiniest detail to ensure a high-precision measurement. As such, detailed study of luminosity counters is of utmost importance. For May and June ’23 Mu Scans in Run 3 of LHC, the luminosity counters of LHCb have been analyzed to determine the set of counters that retains linear behavior to the best order. As per our analysis, VeloTracks and MuonHitsM3R2 provide the best results; however, further analysis may lead to a better result for Tracks vs. Vertices as well. In doing the analysis, a rigorous generalized method to determine empty events and calculate luminosity using LogZero method has been developed to reduce human bias and/or efficiency. To improve the precision of the luminosity measurement, issues such as radioactive activation in SciFi Trackers, pedestal subtraction algorithms in calorimeters in HLT1 and HLT2, presence of LED calibration in bunch crossings were analyzed in-depth and taken care of as best as was possible. It was also found that up to about $\mu\sim 5.7$ measured from VeloTracks, the linearity between a set of counters can be found, signaling a green light to the LHCb collaboration to increase μ at least up to around that value for the future physics runs.

Keywords: Luminosity, Mu Scan, LogZero method, Pedestal, Radioactive Activation, Bunch trains, Relative residuals.

Analysis of LHCb Luminosity Counters for Run 3

For luminosity calculations, the underlying assumption is that the luminosity counters depend linearly on the number of interactions, μ in the event or, at least, we use them only in the linear regime. As such, it is crucial to analyze to what extent each luminosity counter is linear with respect to others. A desired level would be to have the relative residuals of the linearity below 1%. To that end, the goal of my internship was to study each luminosity counter in-depth, revealing potential issues and problems with the detectors. Here to note that the LHCb detectors are still in the commissioning stage after the upgrade and thus there are still issues to be fixed as would be apparent from this report.

The LHCb collaboration also plans to increase the μ value for physics runs, which until Run 2 was set around ~ 1, in the current Run 3. However, from the luminosity side, up to which value it can be increased is still uncertain. By analyzing the May ’23 and June ’23 Mu Scans, one of the goals was to have an estimate of up to which value μ could potentially be increased.

Additionally, I want to note that all the codes for generalized LogZero (Section 11: Linearity of Counters) can be found in my personal publicly available GitHub repository: OSSaimum/LHCb_Lumi (github.com). At the time of writing this report, the repository does not include the codes from other parts of the reports but will be uploaded in the coming days. The repository also includes all my presentations and reports presented at the luminosity working group meetings of LHCb during my internship period.

The LHC

The **L**arge **H**adron **C**ollider (LHC) is a high-energy particle collider, built and operated by European Organization for Nuclear Research (CERN), where particles are accelerated, maintained, and then collided in a 27-kilometer ring of superconducting magnets situated between Switzerland and France. Inside the ring, two separate beams of particles run in opposite directions after getting accelerated in several steps, gaining more and more energy in each step. See figure 1. The beams can consist of only protons (pp collision), proton-lead (pPb collision), and lead-lead ($PbPb$ collision).

On its first operational run from 2009 to 2013, LHC reached an energy of 3.5 TeV on each beam making it the highest energy particle collider in the world, leading to the discovery of Higgs-Boson. In the following operational run from 2015 to 2018, it reached up to 13 TeV (combined energy). And in the current run (run 3), it has already reached the milestone of 13.6 TeV collision energy.

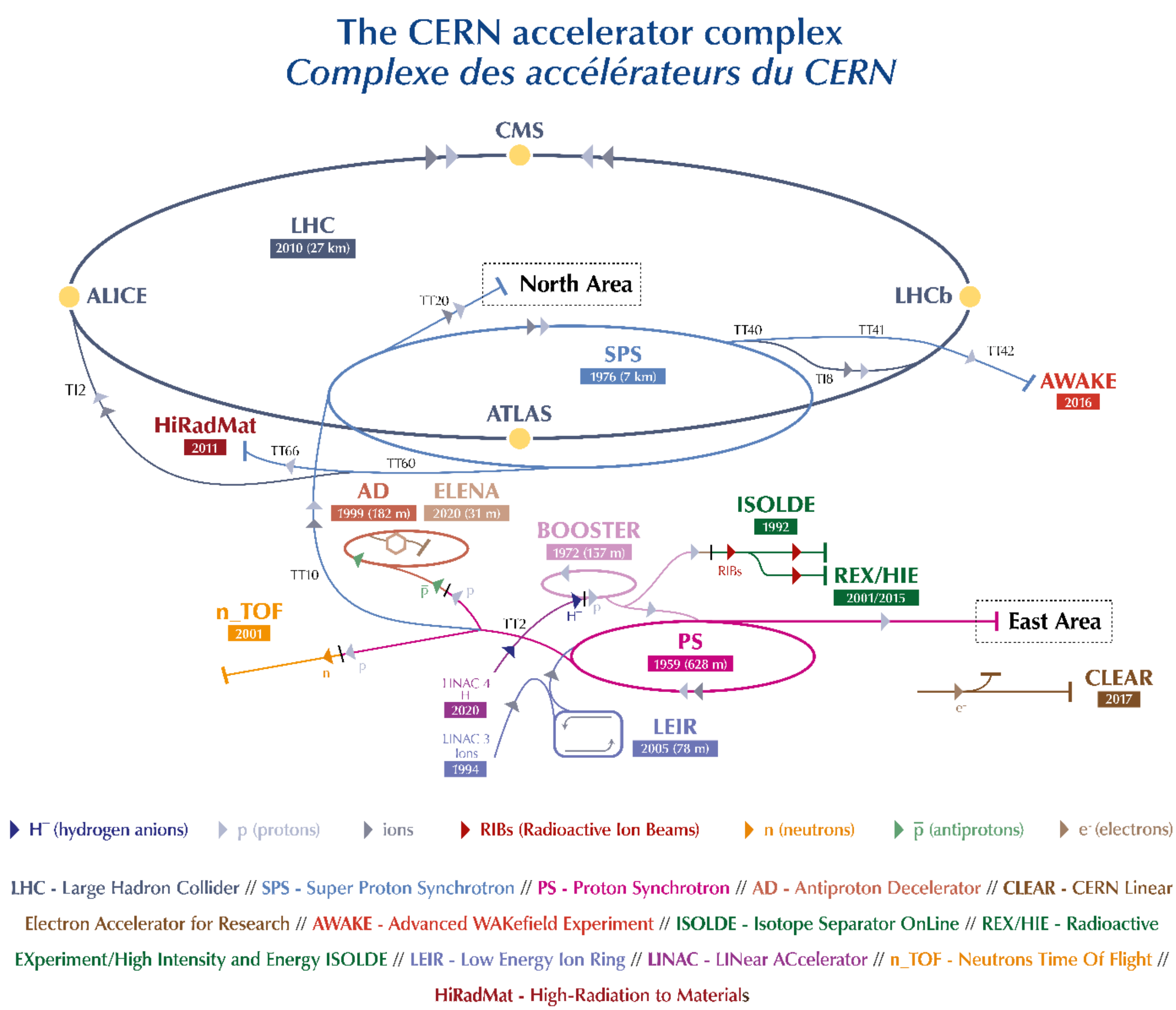

Figure 1: The CERN Accelerator Complex.

Operation

At LHC, the beam size and trajectory are controlled by the quadrupole and dipole magnets, respectively. Inside the beam, there are Radio Frequency (RF) cavities accelerating the particles. The separation between the rf 'buckets' is 2.5 nano sec. For operational stability, only every 10th bucket, called *bunch*, is filled. As such, the LHC *bunches* are separated by 25 nsec. See figure 2.

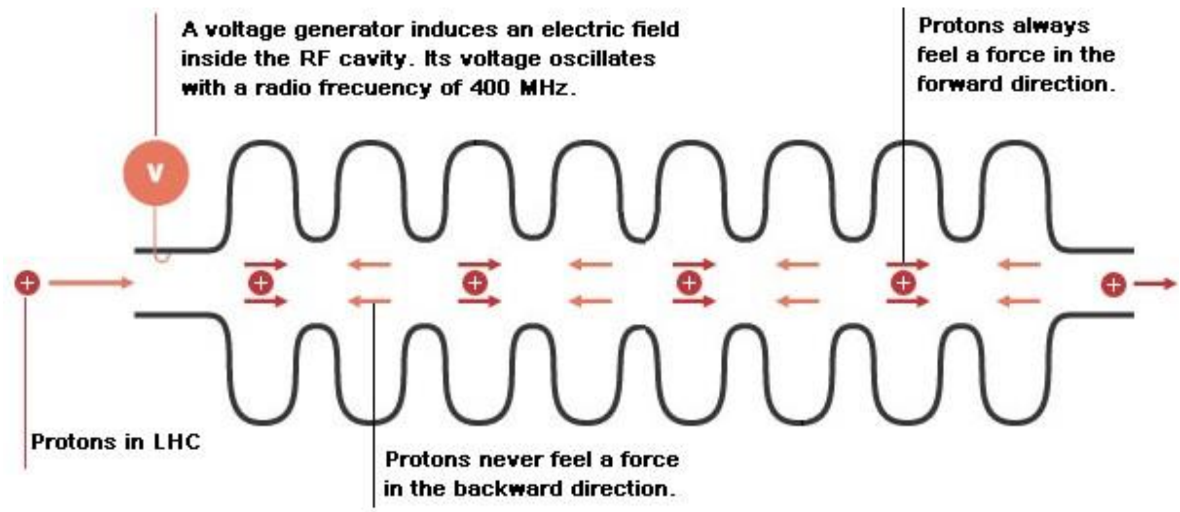

Figure 2: LHC Bunches and Buckets.

The total number of bunches at LHC that can be filled is 3564 and each bunch crossing is commonly referred to as **bxID**; however, for safety and various other reasons [see [1] for some example], not all bunches are filled in a run. Among the filled bunches, at each collision point, the bunches can be characterized by:

1. Beam-Beam (bb): both beams 1 and 2 are present.
2. Beam-Empty (be): beam 1 is present in the bunch and not beam 2.
3. Empty-Beam (eb): beam 2 is present in the bunch and not beam 1.
4. Empty-Empty (ee): both beams 1 and 2 are absent.

Using be, be, and ee, one can estimate the background that might be present in the bb signal.

Of the 3564 LHC bunches, depending on the *filling schemes* in different runs, there can be different numbers of bb, be, eb, and ee bunches. If there are consecutively injected bb bunches, such bunches are called *bunch trains*. For example, if the 500th, 501st, and 502nd bunches are all consecutively injected *bb* bunches, then they are a bunch train. In the analyzed data for my internship, in May ‘23 Mu Scan, there were bunch trains but not for June ‘23 Mu Scan. Bunch trains can be helpful to understand the effect of spill-over, time misalignment, and some detectors’ limit or potential issues that may not be addressed otherwise.

LHCb

At LHC, the two oppositely moving beams collide with each other at 4 different locations where detectors are placed. The four locations equipped with detectors are commonly known by the name of each experiment:

1. **ATLAS: A T**oroidal **L**HC **A**pparatu**s**
2. **CMS: C**ompact **M**uon **S**olenoid
3. **LHCb: LHC-b**eauty
4. **ALICE: A L**arge **I**on **C**ollider **E**xperiment

LHCb specializes in b-physics and studies the parameters of CP-violation and decays of heavy-flavor hadrons to discover new physics. It also works in other aspects of high energy physics such as electroweak physics and quantum chromodynamics. Additionally, the **S**ystem for **M**easuring **O**verlap with **G**as (SMOG) allows only the LHCb collaboration among the LHC experiments to study *fixed target* experiments by injecting gas in some specific region of the beamline.

LHCb Detectors

Unlike ATLAS and CMS which are general purpose detectors with central (low) rapidity, LHCb is a single-arm **forward** spectrometer with a pseudorapidity range of $2 < \eta < 5$. At the front of figure 3, one can locate the **Ve**rtex **Lo**cator (VELO) that reconstructs the tracks and the vertices of particles, followed by **R**ing **I**maging **C**herenkov **D**etectors **1** (RICH1) upstream of the dipole magnets. Since the recent upgrade before Run 3, a dedicated luminosity subdetector called **P**robe **Lu**minosity **Me**asurement (PLUME; not shown in the diagram) has been added upstream of VELO. At the downstream of the dipole magnets, there are three lines of **Sci**ntillating **Fi**ber (SciFi) trackers, followed by RICH2, Muon chamber 1 (M1), the **E**lectromagnetic **Cal**orimeter (ECAL), **H**adronic **Cal**orimeter, and last but not least, 4 more subsections of the Muon chambers (M2-M5).

Among them, RICH1, RICH2, ECAL, HCAL, and Muon chambers (M2-5) provide the particle identification (PID) and the tracking system is handled by VELO, SciFi Trackers, and the silicon-strip **U**pstream **T**rackers (UT; not shown in the diagram).

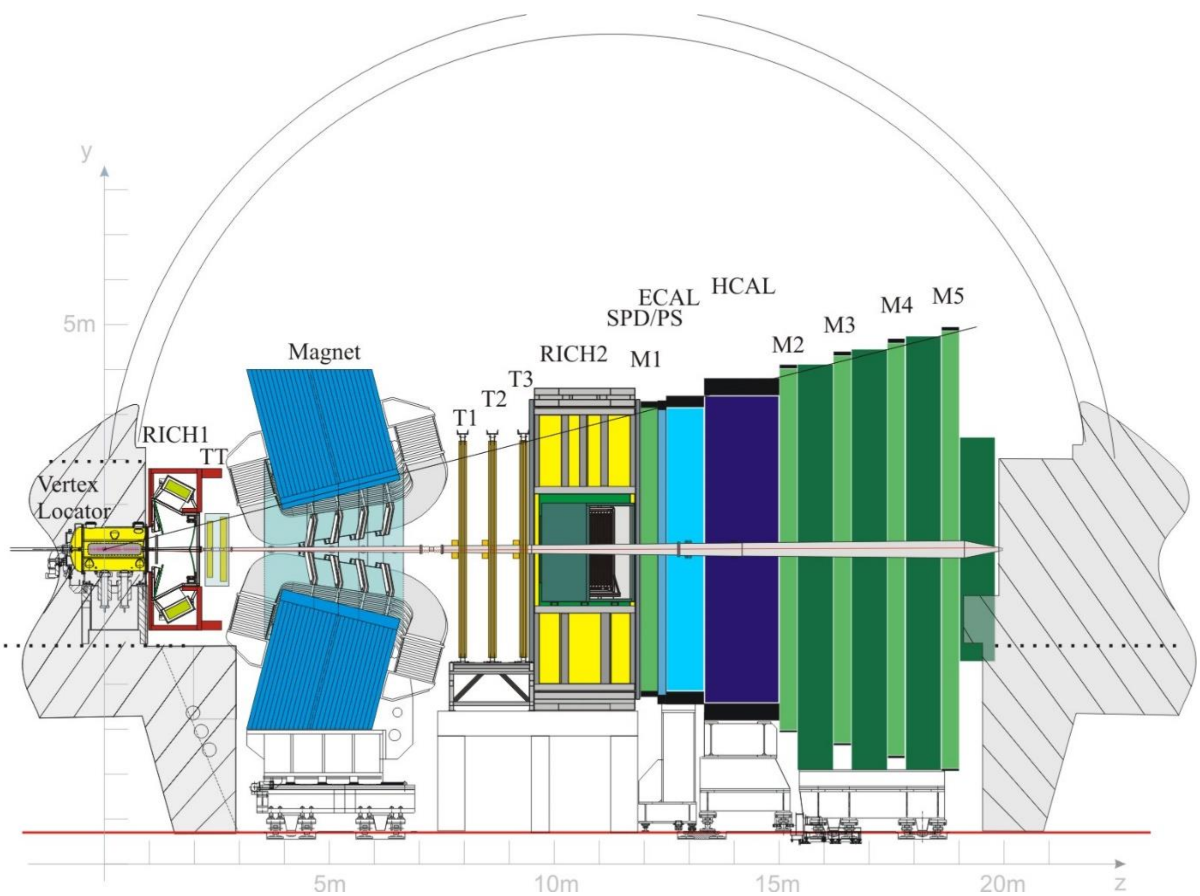

Figure 3: Layout of the Upgraded LHCb detector.

Online Selection at LHCb

At LHCb, online selection of the data is performed in two levels. In **H**igh **L**evel **T**rigger **1** (HLT1) in GPUs where partial detector reconstruction and selections are considered, and in **H**igh **L**evel **T**rigger **2** (HLT2) in offline where full detector reconstruction and selections take place. In our analysis, for example, the 'ECalE's are processed in HLT1 and 'Raw's are processed in HLT2.

Information About the Analyzed Data

The analyses throughout the report were primarily done for the May '23 and June '23 Mu Scans. Some useful information about each is stared below:

- May Mu Scan → Fill: 8782

  Contains 1773 bbs, 601 bes and ebs, and 589 ees

  Bunch trains; Short consecutive ee, be, eb bunches; Both ee and be after bb.

- June Mu Scan → Fill: 9007

  Contains 16 bbs, 124 bes and ees, and 3300 ees

  No bunch trains; Long consecutive ee bunches; Only ee appears after bb.

Why Luminosity?

Cross sectional area, or quantum mechanically, cross section, is one of the most physically meaningful quantities as it provides the probability of a physical process to happen. In particle physics specifically, it abridges the theoretical predictions with the experimental results. In experiments, the data is presented usually as the differential number of events for a given integrated luminosity:

$$dN = Ld\sigma$$

Here, $d\sigma$ is the differential cross section, dN is the differential number of events of the physical process in question, and L is the integrated luminosity. As such, for any particle physics experiment, luminosity is one of the central variables of interest and thus ultra-high precision is desired for luminosity measurement.

Luminosity

Consider two oppositely moving particle bunches with particle numbers N_1 and N_2, respectively, with an effective area of A. Now, if the particle bunches are parallel to the $z - axis$ with particle densities ρ_1 and ρ_2, then the time-integrated luminosity for parallel beams along the $z - axis$ can be found as:

$$L = N_1 N_2 \iint \rho_1(x,y)\rho_2(x,y)dxdy$$

In practice, the parameter that one can consider is μ – the average number of interactions per bunch crossing. As an example, if the two beams do not collide head-on ($(x_1, y_1) \neq (x_2, y_2)$) and beam 1 is displaced by $(\Delta x, \Delta y)$, then:

$$\sigma \iint \rho_1(x_2 - \Delta x, y_2 - \Delta y)\rho_2(x_2, y_2)dx_2dy_2 = \frac{\sigma L}{N_1 N_2} = \frac{\mu(\Delta x, \Delta y)}{N_1 N_2}$$

Then one finds:

$$\sigma = \int \mu_{sp}(\Delta x, \Delta y)d\Delta x d\Delta y$$

Where μ_{sp} is called the specific luminosity.

Methods to Calculate Luminosity

There are two methods that have been used to calculate luminosity in LHC experiments. The first of which is the so-called average method to calculate luminosity (or average number of interactions, μ). The second one is the LogZero method which LHCb collaboration primarily has been using. Below is a short description of the two methods.

Average Method: To calculate luminosity using this method, one calculates the average number of events, giving one the quantity, μ. Although the simplest (and linear), in events where many interactions take place, the so-called *busy* events, this method is prone to any bias associated with those events, i.e., error in reconstructed events gets enhanced proportionally to the number of events since each busy event leads to a higher number of interactions.

LogZero Method: Since the probability of the interactions is independent, if the number of interactions on average is known or a set value, the probability of n interactions should, by Poisson's law, be:

$$P(n) = \frac{\mu^n}{n!} e^{-\mu}$$

Now, if an event yields *no interaction*, the so-called *empty events*, the average number of interactions can be calculated as:

$$\mu = -\log\big(P(0)\big) = -\log\frac{N_0}{N}$$

That is, for any luminosity counters, the challenge is to determine the number of *empty events*, N_0. However, since $\log()$ is a non-linear function, one also needs to consider higher order corrections, typically performed up to the second order. Considering second order correction, one can find (see Appendix 1):

$$\mu = -\log\frac{N_0}{N} - \frac{1}{2}\left(\frac{1}{N_0} - \frac{1}{N}\right)$$

Since the LogZero method does not rely on the reconstruction of busy events, rather concerns only the classification by empty and non-empty events, LHCb has adopted this method for a long time and had been the primary method for calculating luminosity. However, for higher (average) number of interactions, μ, this method does not work very well since the statistics becomes very small for high μ; for example, for a μ value of 8, *only* $\sim .034\%$ of the events have the probability to be empty events. Since LHCb has been operating around $\mu \sim 1$, it hasn't been a concern for the collaboration. But from run 3 and beyond, LHCb collaboration expects to go to higher μ values which calls for two things: a. up to which μ

value it is feasible to use the LogZero method and b. beyond that μ, what method should be adopted. To address the second question, one can, in principle, use the **P**robability **G**enerating **F**unction method (see Appendix 2; to be published in detail by Vladislav Balagura, my internship supervisor). For the first question, one of the ways is to perform so-called *mu scans* where the μ values are changed, and data is collected and analyzed, allowing us, besides determining the applicability of LogZero method, detector behavior and potential issues for different μ values that may not be apparent otherwise.

Determining Empty Events

In a bb signal, there exists both busy events and empty events. To distinguish between them, one must take into account each detector's data-taking method, physical principle or meaning of the data, and other external factors. For example, for VeloVertices or VeloTracks, in principle, if there are no vertex or track, then it should really be an empty event. Therefore, if one obtains the histogram of the bb signal, the number of empty events can be found by calculating the number of events in VeloVertices counter having zero (0) vertex. The fact that empty events correspond to no vertices can also be seen from the ee signal. Below is a schematic example of the histogram of the VeloVertices counter:

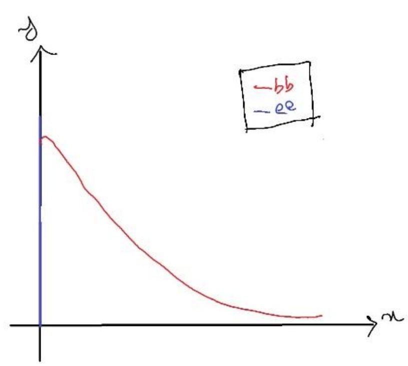

Figure 4: Schematic Example of Number of Events vs. Number of Vertices from VeloVertices.

As such, determining N_0 is the simplest for detectors with such physical principles. However, not all the detectors have this desired property; for example, the ee signals from Calorimeters are not delta function; rather, one obtains a gaussian-like signal. From such a signal, empty events cannot be determined the same way.

Empty Events from Gaussian Like Distributions

A typical signal from an empty-empty bunch crossing may look like the following:

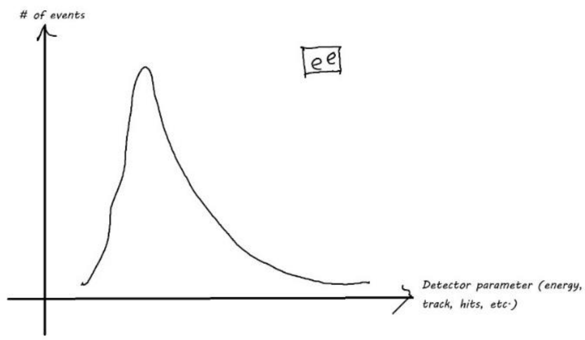

Figure 5: An Example of a Gaussian Shaped Empty-Empty Signal.

And for a bb signal it may look like the following:

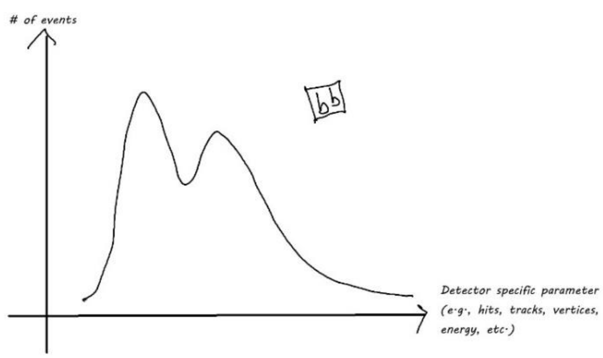

Figure 6: An Example of a Double Gaussian Shaped Beam-Beam Signal.

To extract the empty events from the bb signal, one needs to refer to the ee signal and convey that to the bb signal. For example, if at calorimeter, the energy peaked at a value of 100, and the bb signal had double gaussian with one of the peaks at a value of 100, then that particular peak corresponds to the empty events in the bb signal.

However, as can be understood from the two figures above, when there exists a Gaussian shaped signal, it is not straight-forward to determine from the bb signal what portion of the signal corresponds to the empty events. The commonly used practice was to plot the bb signal and the ee signal of the Gaussian shaped signal and determine the cut value left of which would be considered empty events *by eye*. This method, although worked, poses many potential issues:

a. Not rigorous: there is no proper mathematical formulation behind defining the cut.
b. Subjective: different people can demand different values for the same luminosity counter as the appropriate cut value.

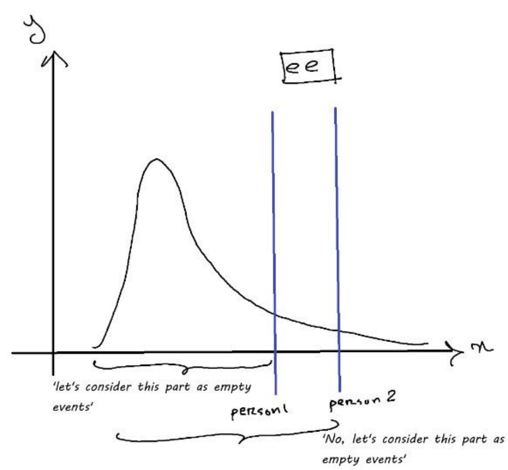

Figure 7: Subjectivity of Determining the Cut by Eye.

c. Inefficient: For different subdetectors, for example, the sub-sections of the transverse component of the **E**lectromagnetic **Cal**orimeter, the cut could be different and for each one, it has to be found and only then relay that information to calculate the luminosity using LogZero method, requiring a lot of hardcoding.

Generalized Cut Function

To address the issues mentioned above, one of the goals of my internship was to develop a rigorous, generalized method that can be used to determine the cut for each counter that takes into account all types of issues that might be encountered. In principle, the cut should be determined and set independently, regardless of the filling scheme.

For the Gaussian shaped histograms, since the Gaussian shaped histogram is usually not symmetric; rather biased to the right side of the histogram, one cannot simply take any percentage of the ee signal and discard the rest. After discussions and considerations, the following algorithm is proposed:

a. Obtain the counter value at the nth percentile of the ee signal.

b. Take the different of the counter values at the nth percentile and the peak of the ee signal.

c. Multiply this difference by m and add to the counter value at the nth percentile to make sure the cut set by the algorithm counts for (almost) all the empty events in the bb signal.

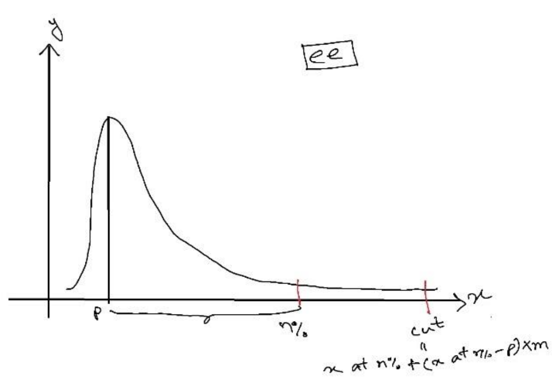

Figure 8: Algorithm to Define the Cut from ee Signal.

After obtaining the cut, we check the ratio of the ee signal that has been labelled as *empty events* to make sure that the cut set by the algorithm was appropriate.

Beyond that, the biggest advantage of developing this algorithm was that if a concerned sub-counter does not have a Gaussian shaped distribution, the other issues such as, shifted histogram or such can be automatically taken into account as needed. For example, the SciFi tracker can have random shifts in its ee signal, or the noise can be triggered at some arbitrary time (See section 8, paragraph 2). Or one of the Muon chambers could be affected by some external noise or high background. All such issues can be considered and cross-checked all at once using this algorithm.

```r
# function to get the cut for the desired counters
# input: a. bin0: initial binning set for ECals in the root file
#        b. bin0: initial binning set for Raws in the root file
#        c. per: set how much of the cumulative sum of the Gaussian to be considered
#        d. jmp: set how far from the distance between the mean and the 'per'th percentile to be added
#        e. names: counters to be considered for cut
cut <- function(bin0,rawbin0,per,jmp,names) {
    # select only ees; note: may seem redundant since it is performed in line 163.
    # however, after running the first time, the in_dt1 data table will be modified.
    # as such, any subsequent execution of the cut function is likely to cause an error.
    # note: line 163 may be removed but kept since it may be needed for analysis.
    in_dt1 <- subset(dt1, bx.type=='ee')
    n1 <- in_dt1[, {
                    # for each counter, obtain the 'combined' histogram
                    . <- in_dt1[,.(y=sum(y)),by=.(name,x)]
                    # for each counter, obtain the cumulative sum
                    .[,ycs:=cumsum(y),by=.(name)]
                    # obtain the median of the histogram
                    .[,peak:=max(y),by=.(name)]
                    # get the x value of the median
                    .[,ipeak:=ifelse(y == peak, x, 0),by=.(name)]
                    .[,peakx:=ifelse(ipeak == 0, max(ipeak), ipeak), by=.(name)]
                    # get the y value up to which threshold the cut is desired
                    .[,ycut:=ceiling(ycs[length(ycs)]*per),by=.(name)]
                    # obtain the first value above that cut
                    .[,cscut:=ycs[ycs>ycut][1],by=.(name)]
                    # get the x value for that cut; note: y value is taken and then the x which may be redundant;
                    # but it's done for easier debugging
                    .[,icut:=ifelse(ycs == cscut, x, 0),by=.(name)]
                    .[,icut:=ifelse(icut == max(icut), icut, -1),by=.(name)]
                    # to make sure the ee signal doesn't affect, we move the 'jmp'*difference between the cut and the peak distance away
                    .[icut!=-1,.(xcut=icut+jmp*max(abs(icut-peakx),1)),by=.(name)]
                },]
    # multiply the initial binning (done when converting from the root files)
    # additionally, if xcut == jmp, set it to zero since it most likely means the zero for that counter is really the empty events
    n1 <- n1[,.(cut=ifelse(name %in% ETots, xcut*bin0, ifelse(name %in% ECals, xcut*bin0, ifelse(name %in% Raws, xcut*rawbin0, ifelse(xcut==jmp, 0, xcut))))),by=name
    # for the non-desired counters, set the cut to simply zero
    n2 <- in_dt2[,.(cut=0),by=name]
    # add the two data table objects and return
    all <- dplyr::bind_rows(n1, n2)
    data.table(all)
}
```

Figure 9: Snippet of the Code for the Generalized Cut() Function.

Further Constraints Determining Empty Events

Throughout my internship time, as we have dug deeper into the counters in order to make our luminosity calculation as precise as possible, we have discovered different issues, problems, and also potentials that can affect the luminosity calculation. In the next several sections, a selected number of issues will be explained in detail.

Radioactive Activation in SciFi Tracker

One of the key findings from my internship was that one of the detectors at LHCb, the SciFi Tracker, suffers from radioactive activation at higher mu values. To explain further, for higher mu values, as the particles are of higher energy and intensity, they may have large enough energy that, when those particles hit the materials of the SciFi, it can lead to activation of the materials of the detector, causing secondary interactions.

Another important finding was that the SciFi tracker also generates arbitrary noises and the histogram of the SciFi tracker sub-counters can be arbitrarily shifted. The following figure shows the two effects:

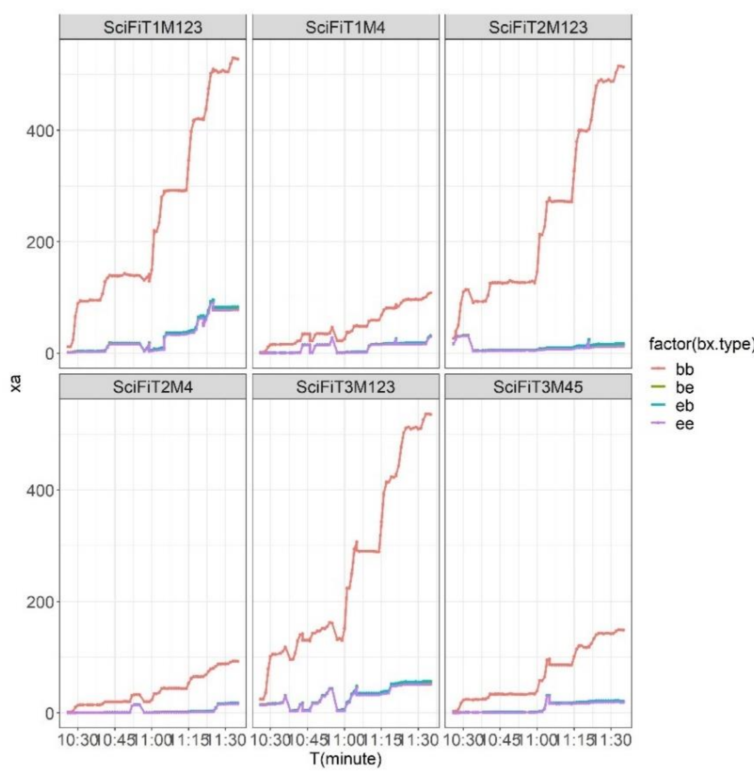

Figure 10: Average hits at SciFi subsections vs. Time (increasing mu values) for May ‘23 Mu Scan.

Now, the ee signal should be constant irrespective of the μ, since otherwise, it means that the ee signal is contaminated by the bb signal to some extent. The fact that the ee signal is contaminated can be confirmed by analyzing the following figures. In the figure, by $bb + x$ means the xth ee bunch crossing *after* the bb bunch crossing.

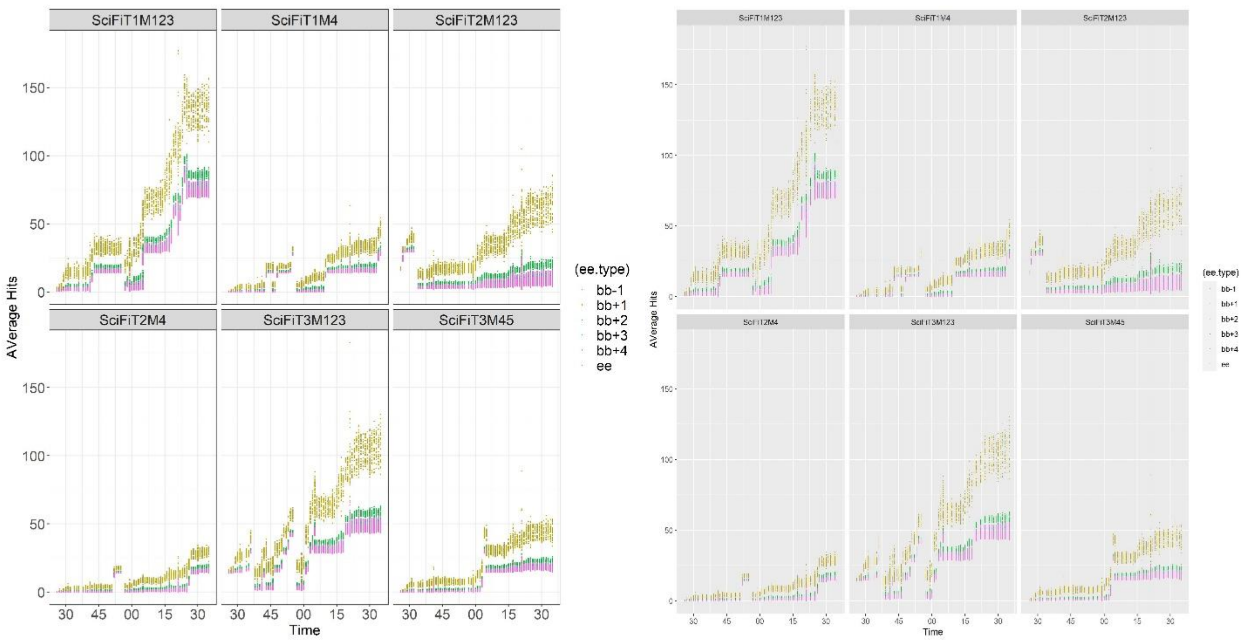

Figure 11: Average Hits vs. Time (increasing μ) of the ee signal for May ’23 Mu Scan (Left) and June ’23 Mu Scan (Right).

Now, for the first issue mentioned above, to determine whether the effect comes from radioactive activation, one can investigate the ee bunch crossings *after* the bb bunch crossings. If the ee bunch crossings are located after the bb bunch crossings, the ee signals drop exponentially, that would indicate radioactive activation. There can also be spill-over effect and time misalignment, but their effect should be limited to the two following bunch crossings at most after the bb bunch crossings. From the figures above, it appears there could be both effects present. As a confirmation, we plot the average signal normalized by a quantity that is proportional to the μ, for example, the (background subtracted) bb signal. If there is radioactive activation, the plots should be linear with decreasing slope as moving away from the bb bunch crossing.

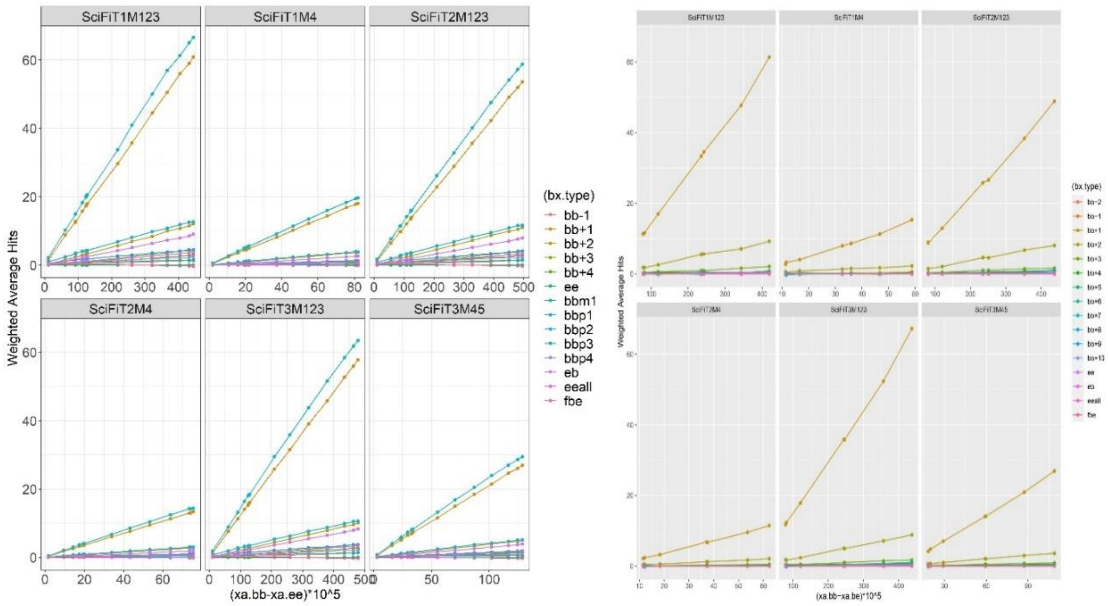

Figure 12: Average ee Signal vs. Background Subtracted bb for May ’23 Mu Scan (Left) and June ’23 Mu Scan (Right). Note for May: $bbpx$ means $be + x$.

And finally, to determine to what extent the activation affects the ee signal, we normalize the average residual ee signal (signal at $bb + x$ subtracted by the ee signal) by the background subtracted bb signal:

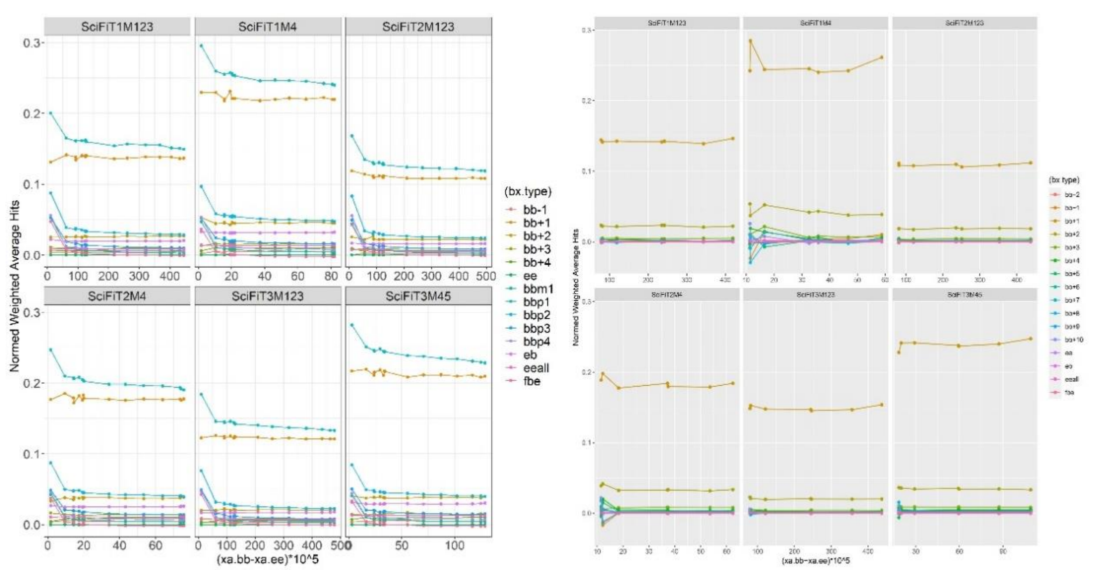

Figure 13: Average ee Residual Signal/Background Subtracted bb Signal vs. Background Subtracted bb for May ’23 Mu Scan (Left) and June ’23 Mu Scan (Right). Note for May: $bbpx$ means $be + x$.

To see that the signal really drops exponentially as moving away from the bb bunch crossings, we plot the $y - axis$ vs. the distance of the ee bunch crossing w.r.t. the bb for June '23 Mu Scan (Recall that in May, there were bunch trains, meaning any $bb + x$ has effect from multiple bunch crossings and the ee bunch crossings were not well-separated.):

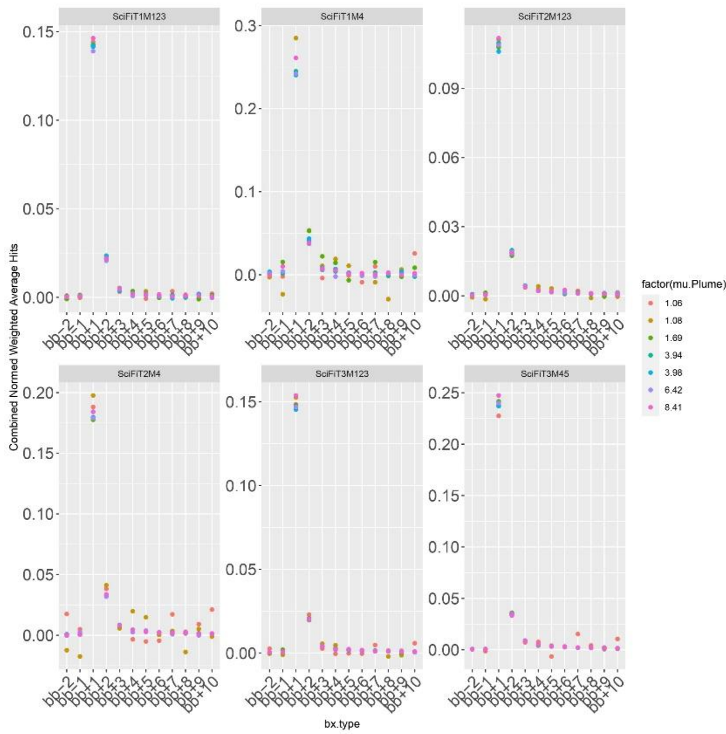

Figure 14: Average ee Residual Signal/Background Subtracted bb Signal vs. Bunch Crossing Types for June '23 Mu Scan.

The radioactive activation, combined with the time misalignment and spill-over effect, it seems to affect the following bb bunch crossing as high as 30%. As the contribution solely from the radioactive activation cannot be estimated due to the other two possible effects, we plot the ee signal vs. the distance from bb bunch crossings from $bb + 2$ onwards after the bb bunch crossings for June '23 Mu Scan:

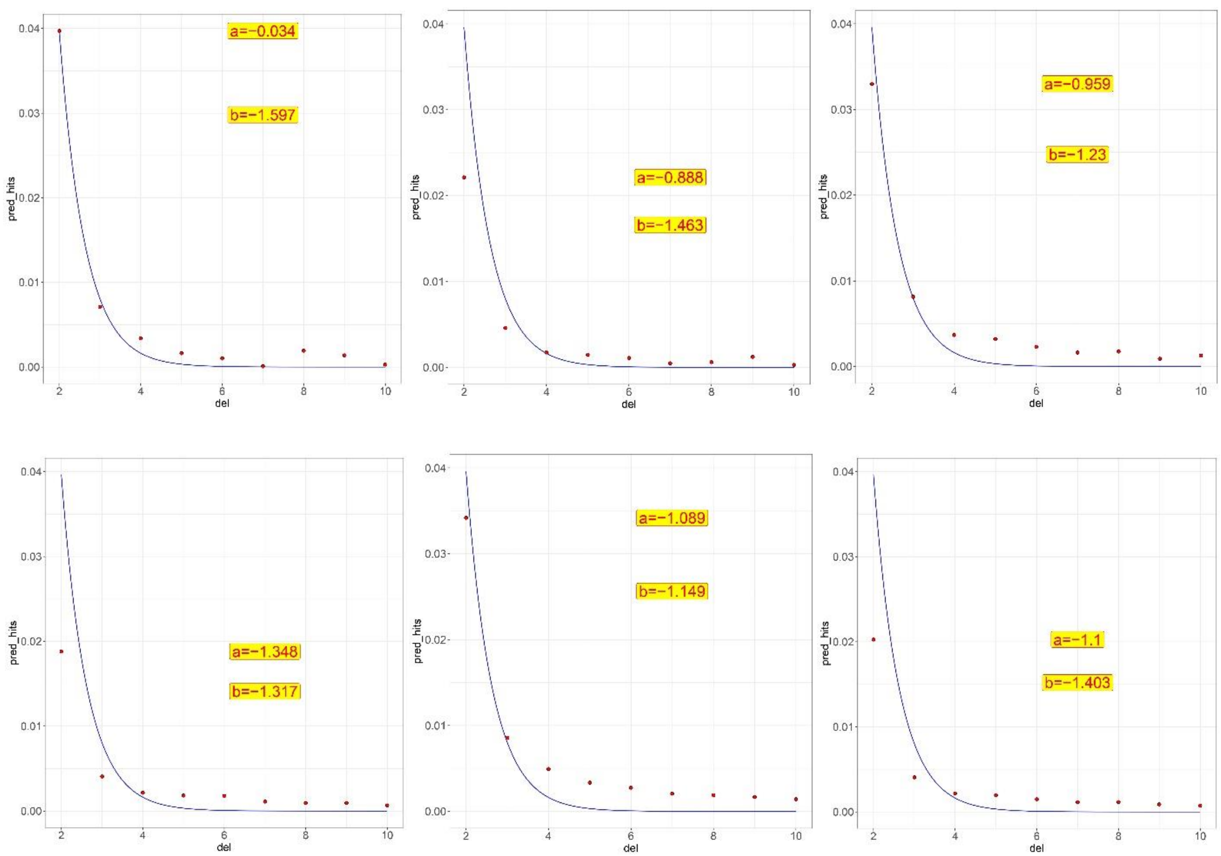

Figure 15: Average ee Residual Signal/Background Subtracted bb Signal vs. Distance from bb Bunch Crossing for June ’23 Mu.

As can be seen, the contribution from the radioactive activation is likely to be in the range of 3% or so.

Negative Pedestal in HLT1 Calorimeter

The Electromagnetic Calorimeter has a total of 8 sub-counters for luminosity calculation. ECalEtot calculates the total energy, ECalET calculates the *transverse energy*, and the six sub-sections of the ECalET counts the transverse energy deposited at each sub-counter.

When the value of the deposited energy in ECals is stored, since there is always a pedestal in the signal, a zero (reference) energy is set via some algorithm and the algorithm always converts the obtained value with respect to the zero energy before storing the value which makes sure the stored value is always positive. There are two algorithms that are used in pedestal subtraction:

a. The pedestal is subtracted by taking the minimum of $bb - 2$ and $bb - 4$.

b. The pedestal us subtracted by taking the minimum of $bb - 2$ and the pedestal at $bb - 2$ (so the effect considers $bb - 4$, which considers $bb - 6$, and so on).

For May '23 Mu Scan, in **H**igh **L**evel **T**rigger 1 (HLT1), the first algorithm was implemented, whereas in June '23 Mu Scan, the second one was implemented in HLT1. For Physics analysis, either of the algorithm may not cause an issue; however, for luminosity calculations, this will lead to a higher estimation of the empty events, leading to incorrect luminosity value.

As per our analysis, using the first algorithm leads to a negative tail to the bb signal as μ is increased since a higher μ value affects the ee signals in the previous bunch crossings. As such, the pedestal is estimated less than the actual values, leading to a negative tail as can be seen in the following figure:

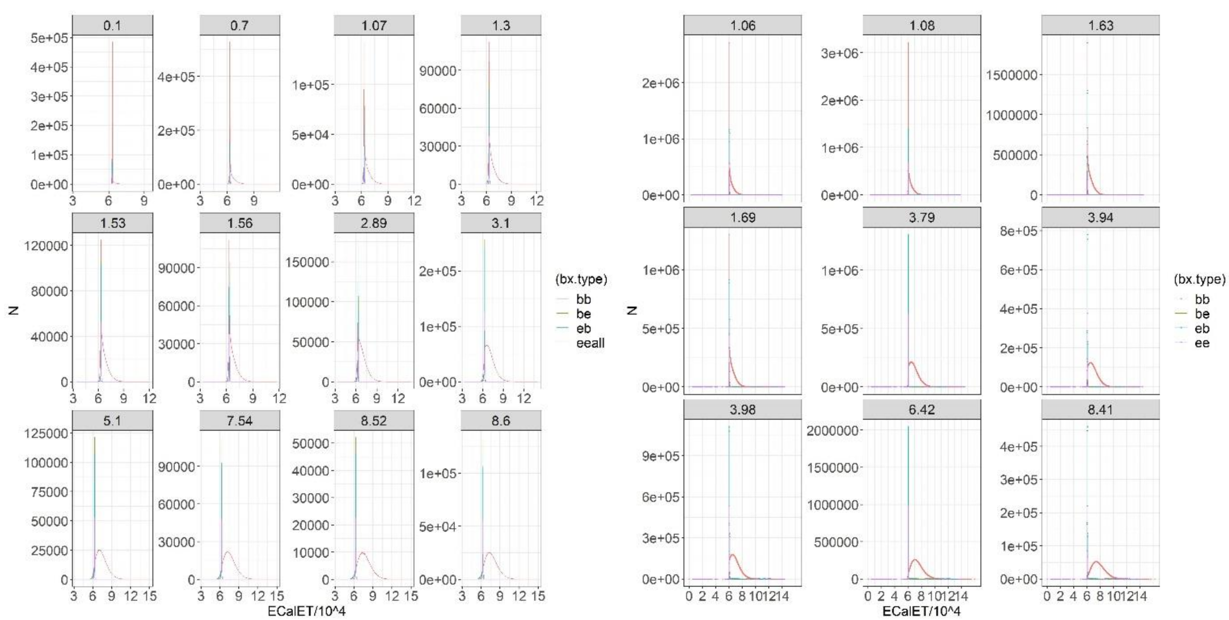

Figure 16: Number of events vs. ECalET deposited energy for May (left) and June (right) '23 Mu Scans.

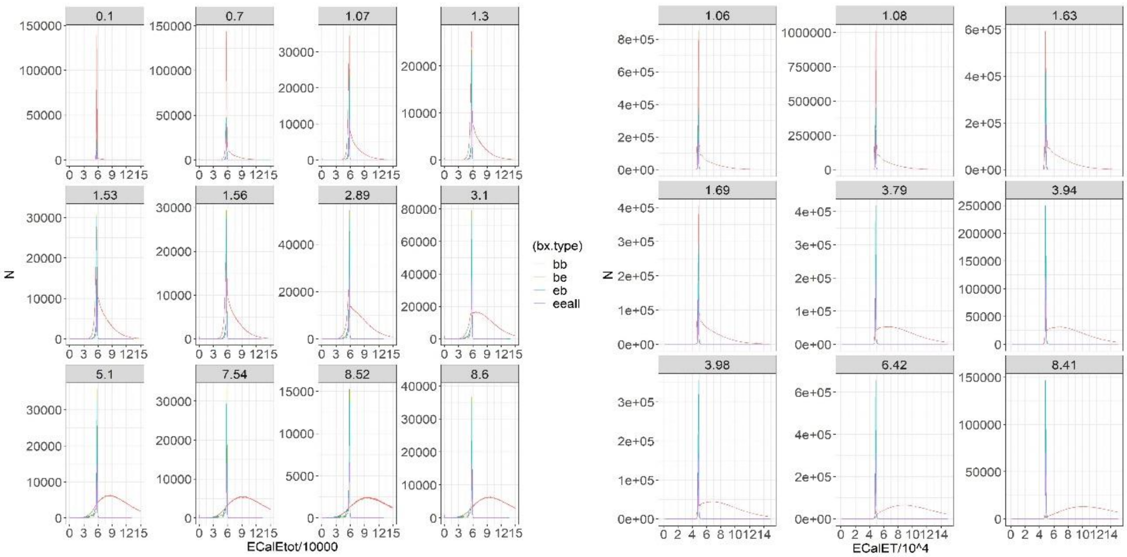

Figure 17: Number of events vs. ECalEtot deposited energy for May (left) and June (right) ’23 Mu Scans.

As such, it is evident that the second algorithm works better for pedestal subtraction. To further strengthen the claim, the ratio of the bb signal to the left and right of the ee signal peak are plotted:

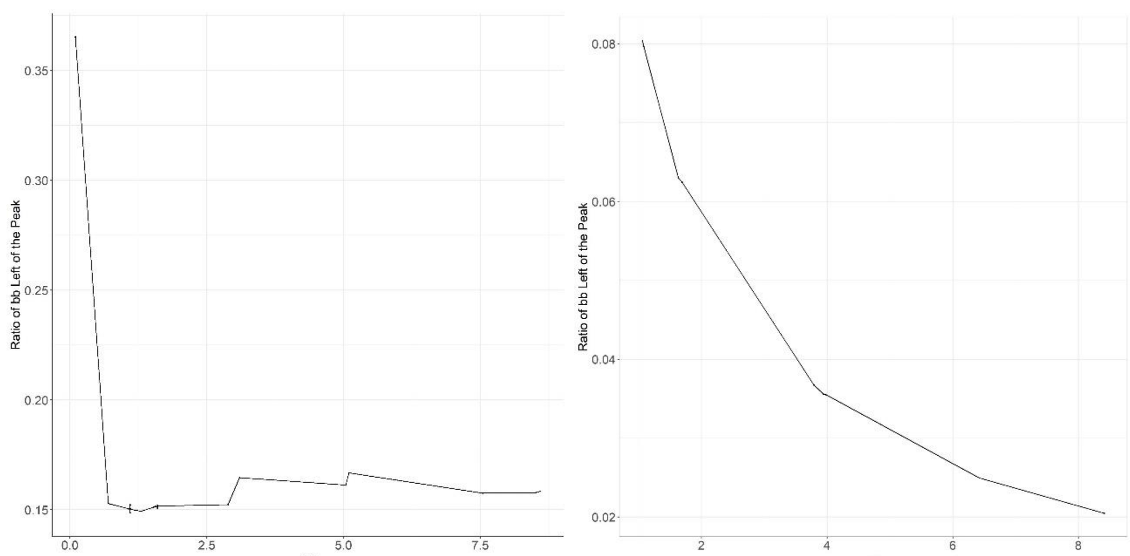

Figure 18: Ratio of the bb signal left of the ee peak for May (left) and June (right) ’23 Mu Scan.

Since as mu increases, the bb signals are supposed to have more and more energy, the ratio of the signals are supposed to decrease as $\sim\frac{1}{x}$ which can roughly be seen in June but not in May Mu Scan. It is because of the incorrect pedestal subtraction, the bb signal gets shifted to the left with increasing μ and keeps the ratio roughly constant.

The negative tail can also be observed if we look into the previous ee bunch crossings subsequent to a bb bunch crossings in the following figure where the ee bunch crossings are plotted that have energy deposition below a threshold. Note that for May mu scan, due to the existence of bunch trains, $bb + 1$ is also affected since it would be the second or fourth bunch crossings for the previous bb bunch crossings.

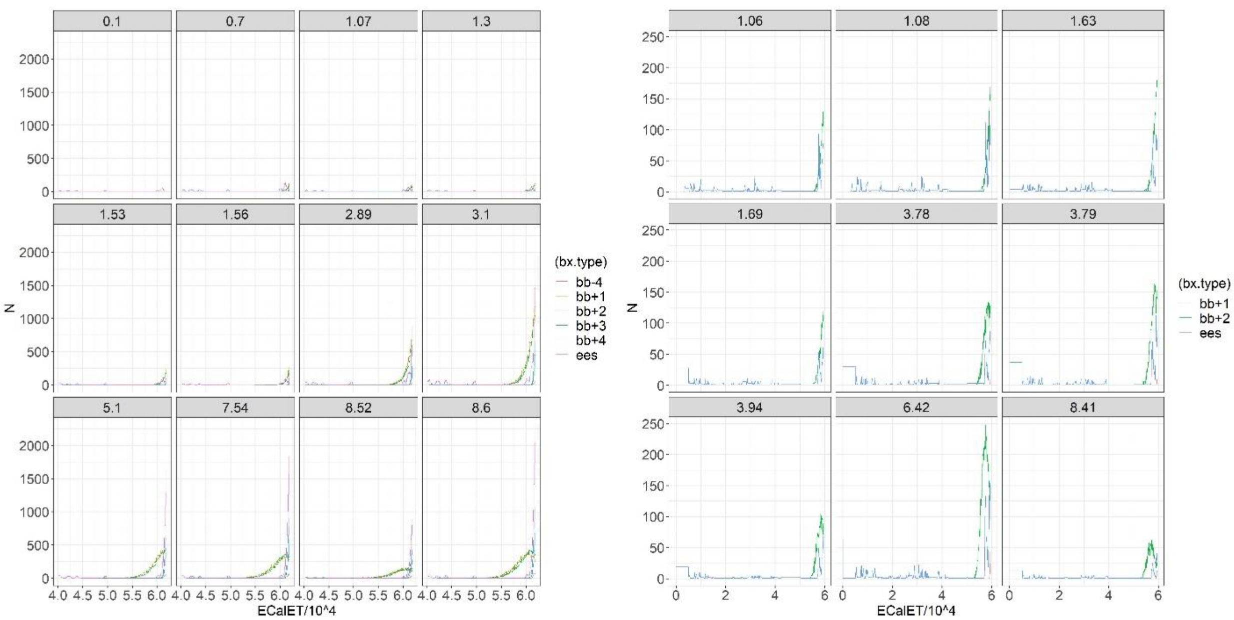

Figure 19: Number of events vs. ECalET deposited energy below a Set Threshold for May (left) and June (right) '23 Mu Scans.

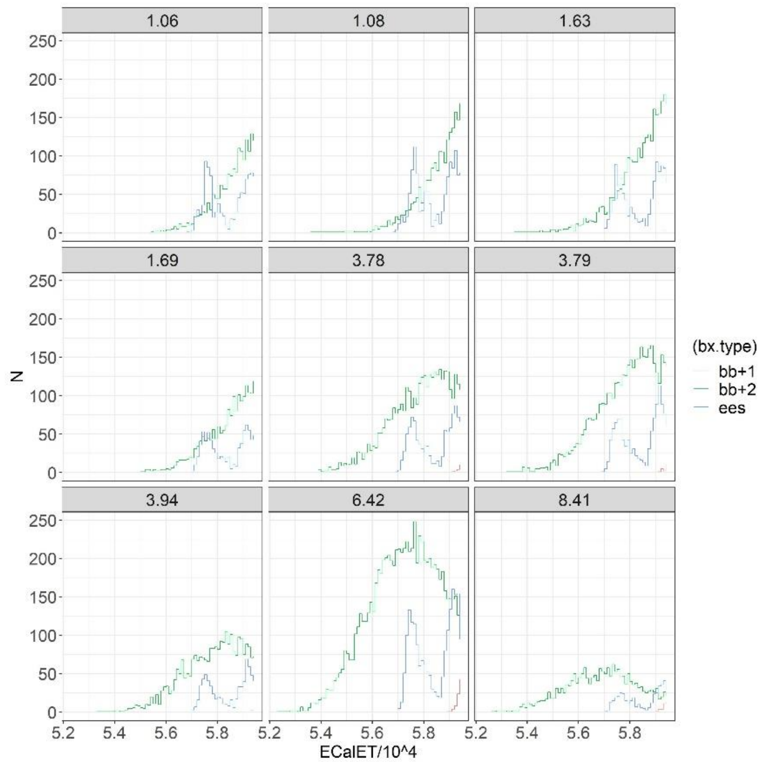

Figure 20: Number of events vs. ECalET deposited energy below a Set Threshold for June ’23 Mu Scan.

The negative pedestal in HLT1 (ECals) can also be seen when comapred with the HLT2 (Raws). After proper scaling and shifting, the following plot can be instructive:

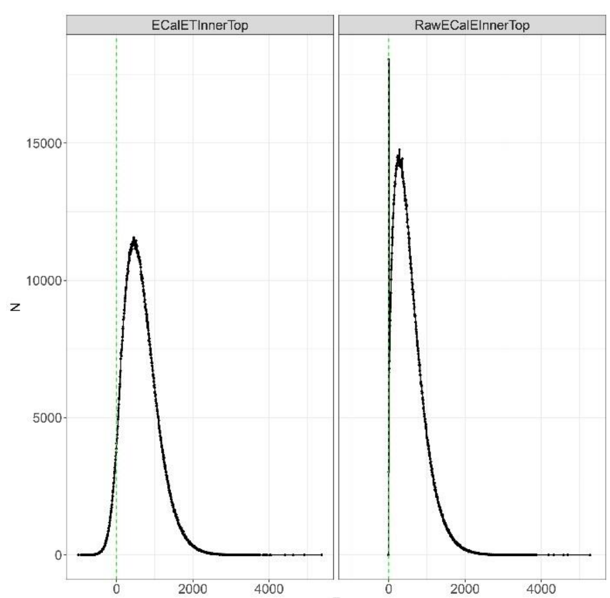

Figure 21: Negative Pedestal for HLT1 but Zero Suppressed in HLT2 in May ‘23 Mu Scan.

Note that, in HLT2, the negative pedestal is zero suppressed (negative values are set to zero) as opposed to the case in HLT1. However, whether the problem is resolved or just the fact that the negative pedestal is just put to zero without addressing the issue appropriately still demands further analysis.

LED Calibration in be Signal in June ’23 Mu Scan

At LHC, to calibrate the Calorimeters, an LED signal is injected. Ideally, this signal should not be present in any bunch crossings such as bb, be, eb, and ee. However, as we found out, for June ’23 Mu Scan, in some be, eb, and ee bunch crossings, there was LED injection. It was confirmed by taking the number of events with energy larger than a set threshold considerably higher than the expected energy.

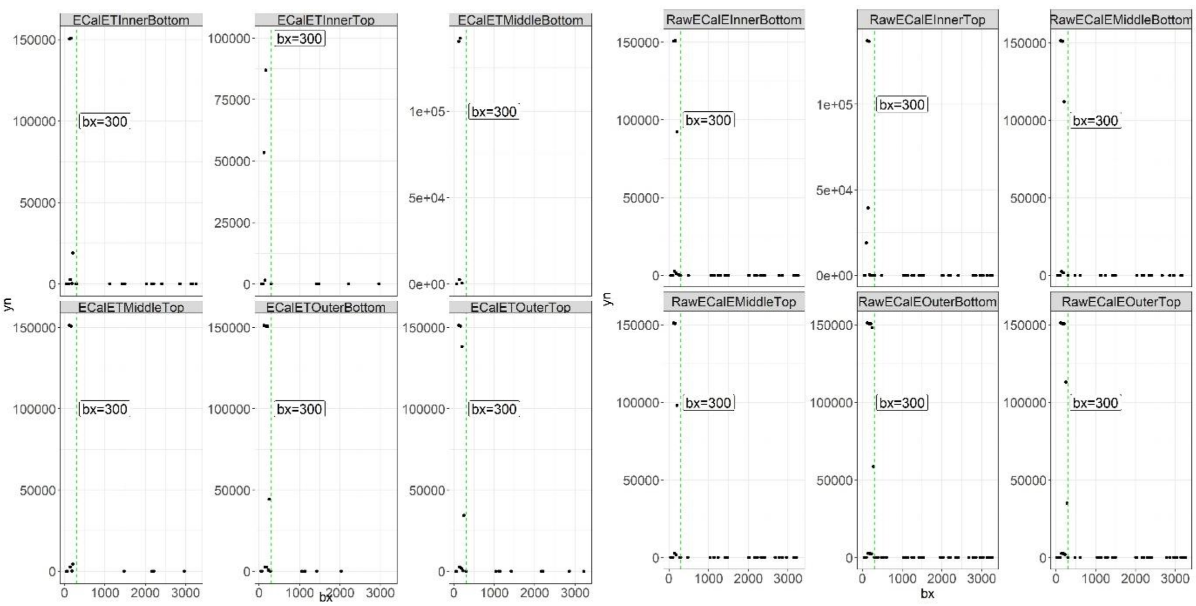

Figure 22: Number of Events Above a Threshold vs. bxID LED Calibration Signal Present in be Signals in June ‘23 Mu Scan.

From the figure, it can be seen that the number of events above a certain energy comes explicitly from a few bunch crossing below bxID 300. As such, it is evident that those bunch crossings had LED injected in them. Due to this reason, for the June ’23 Mu Scan data, the be bunch crossings above bxID 300 are only considered.

Linearity of Counters

Finally, the linearity between the different luminosity counters have been studied in-depth. Luminosity obtained by different counters have been plotted against each other to observe their linear behavior, and the relative residuals of the linear fit have been calculated to show to what extent they behave linearly. Note that the results presented in this section are exclusively from June '23 Mu Scan.

Calorimeters vs. VeloTracks

Between HLT1 and HLT2 calorimeters, it was found that the HLT2 calorimeters (Raws) provide better linearity. The reason can be attributed to the absence of the negative tail by zero suppression algorithm adopted in HLT2.

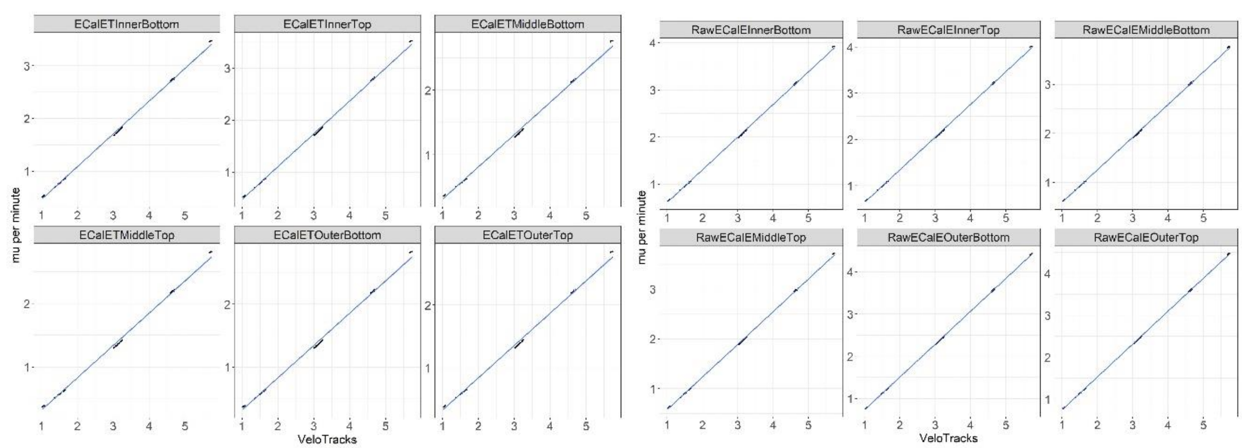

Figure 23: Linearity of μ Obtained from HLT1 (left) and HLT2 (right) Calorimeters vs. VeloTracks.

Note that in these plots, the offset of the linear fit is not set.

To calculate the relative residuals of linearity between the counters, we take the difference between the estimation from the linear fit and the data and divide that quantity by the value obtained from the linear fit.

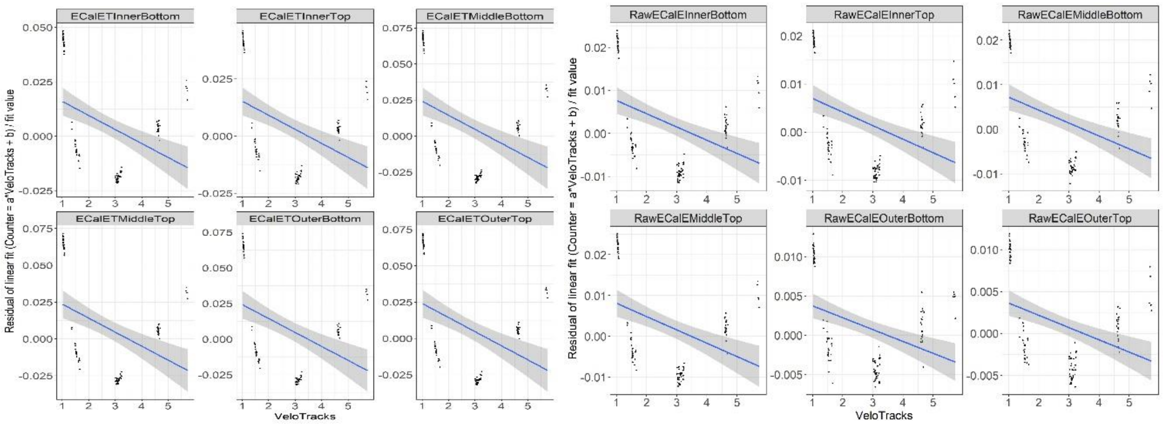

Figure 24: Relative Residual of the Linear Fit of μ Obtained from HLT1 (left) and HLT2 (right) Calorimeters vs. VeloTracks.

When the offset is set to zero, the diversion from the linear fit towards the quadratic fit becomes evident:

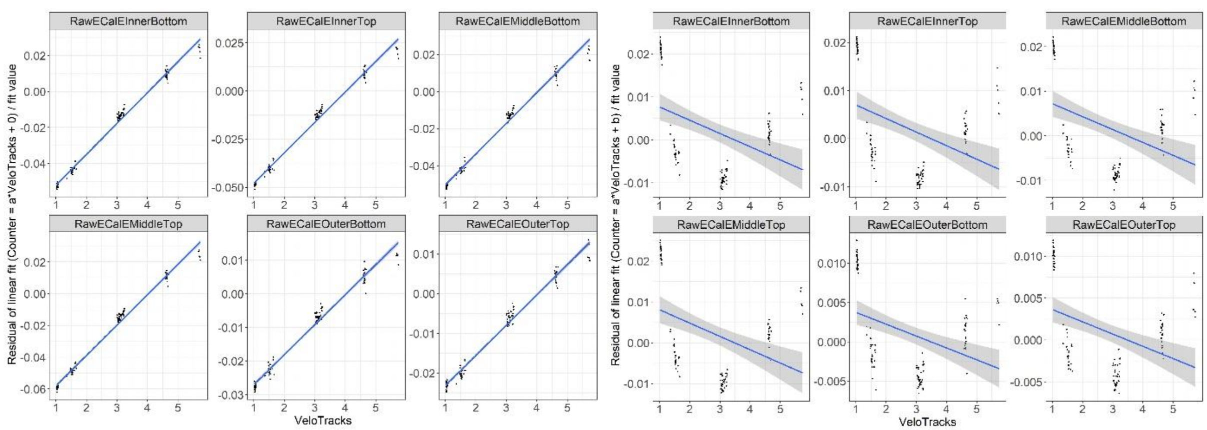

Figure 25: Relative Residual of the Linear Fit of μ Obtained from HLT1 Calorimeters by Setting Offset to Zero (left) and Arbitrary (Right) vs. VeloTracks.

In any case, the relative residuals of the linear fit for the HLT2 calorimeter have the best results for Outer Parts – about 1%.

Muons vs. VeloTracks

As such, to improve the precision, MuonHits and MuonTracks have been analyzed.

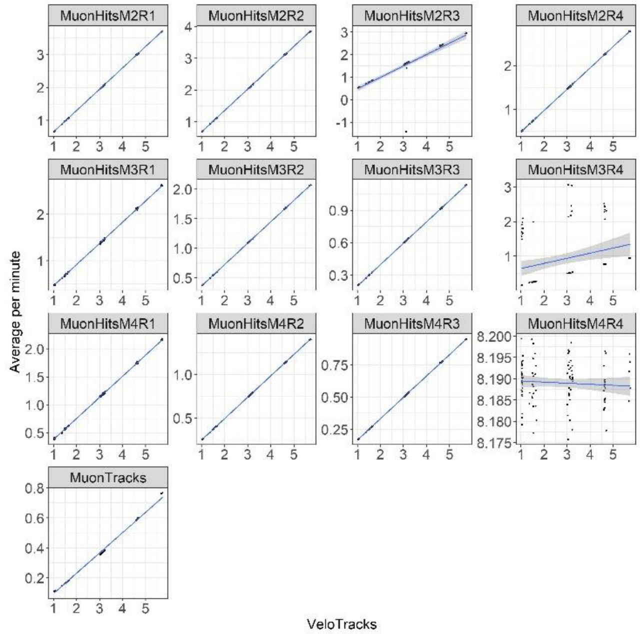

Figure 26: Linearity of μ Obtained from MuonHits and MuonTracks vs. VeloTracks.

Although not all the sub-components behaved linearly with respect to the VeloTracks, the linearity seems to work better than calorimeters with increasing mu for MuonHits sub-counters except for M3R4, M2R3, and M4R4.

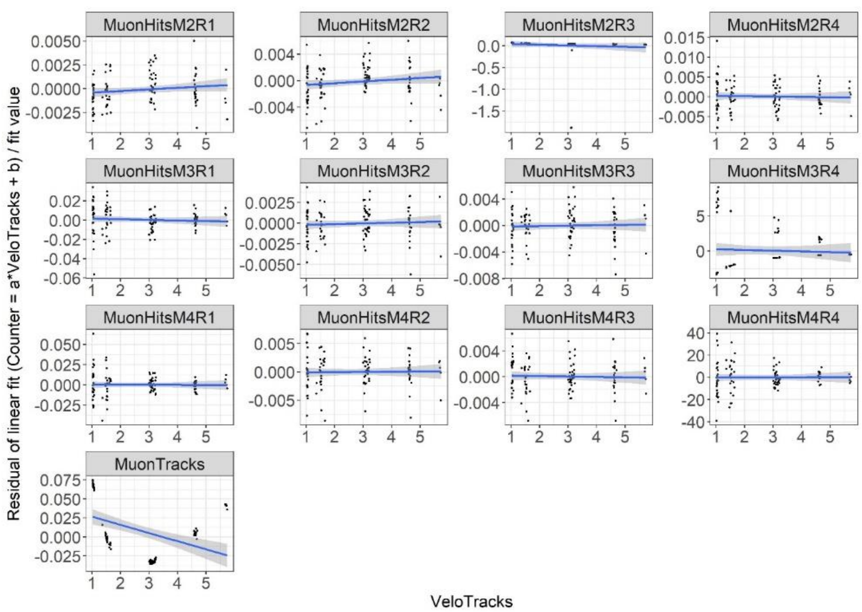

Figure 27: Relative Residuals of the Linear Fit of μ Obtained from MuonHits and MuonTracks vs. VeloTracks.

For MuonHits sub-counters, M2R2, M3R2, M3R3, M4R2, and M4R3, the relative residual of linearity is at most $.4\%$ and as good as around $.25\%$ for MuonHitsM3R2. As such, MuonHitsM3R2 and VeloTracks seem to be the best set of luminosity counters that retain their linear behavior.

Tracks vs. Vertices

We have also analyzed VeloTracks vs. VeloVertices since from the preliminary analysis of the May Mu Scan, those two sets of counters had the best performance.

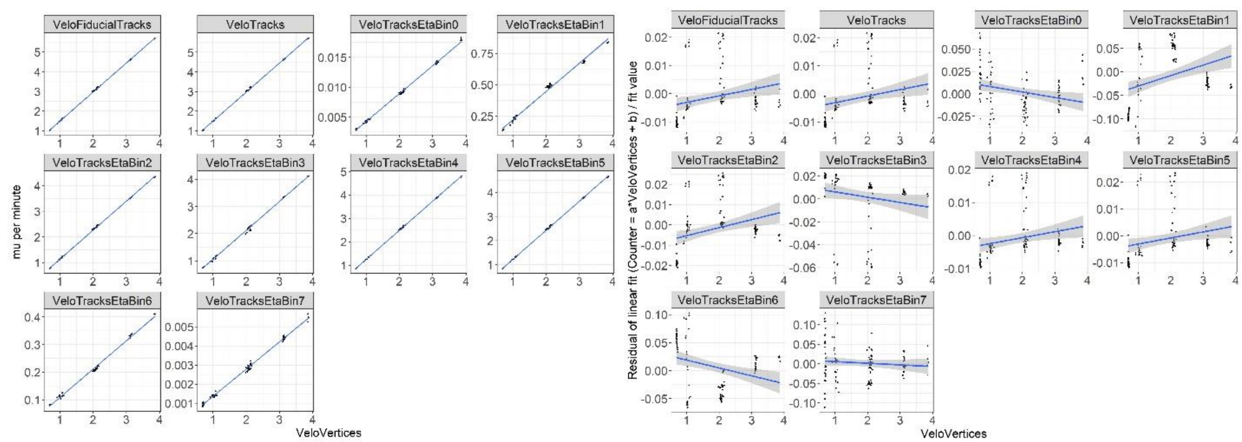

Figure 28: Linearity of μ (left) and Relative Residuals of the Linear Fit of μ (right) Obtained from VeloTrakcs vs. VeloVertices.

As one can see, Trakcs vs. Vertices gives roughly the same level of the relative residuals of the linera fit as the HLT2 calorimeters. However, do note that only some points lead to $\sim 2\%$ relative residula of the linera fit. Further analysis may reveal issues that could potentially improve the precision to a higher level.

Noentheless, the higher relative residual begs the question, since both are sub-parts of the same detector (VELO), between the two, which one is more reliable. To answer that, we plot the MuonHits vs. VeloVertices:

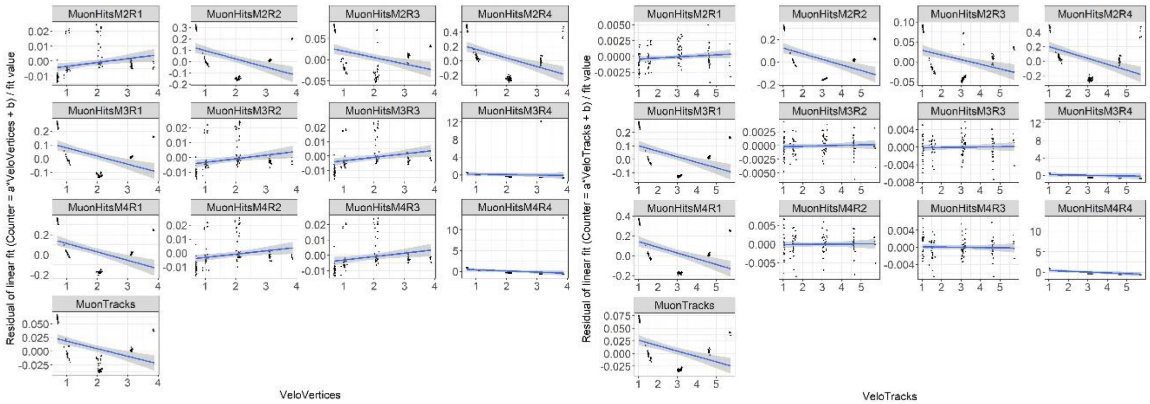

Figure 29: Relative Residuals of the Linear Fit of μ (right) Obtained from MuonHits and MuonTrakcs vs. VeloVertices (left) and VeloTracks (right).

It seems Muons vs. Vertices residuals of linear fit ranges around $\sim 2\%$, indicating VeloTracks work better. So, we see that the less complex counters (Hits < Tracks < Vertices) yield better linearity once the background subtraction and other issues are taken into account appropriately.

Conclusion

Based on the analysis of May '23 and June '23 Mu Scan data, the set of luminosity counter that behave the most linearly is VeloTracks and MuonHitsM3R2; however, further in-depth analysis of VeloVertices may have the potential for a better result. It was also shown that the SciFi Tracker at LHCb suffers from radioactive activation to some extent and the pedestal subtraction algorithm implemented in HLT1 should be improved and the HLT2 algorithm that is zero suppressed has better performance, at least in the scope of the data. Finally, up to at least $\mu \sim 5.7$ counted by VeloTracks, a set of counters can be found that behaves linearly at the order of $\sim .25\%$ with respect to each other.

Appendix 1: Second Order Correction to LogZero Formula

If both μ and N are fixed and the number of empty events, n_0, fluctuates around its average, N_0, due to the non-linearity of $\log()$ function,

$$\mu = -\log\frac{N_0}{N} \neq < -\log\frac{n_0}{N} >$$

Since the available parameter from an experiment is n_0 and not N_0, the average bias:

$$\mu - < -\log\frac{n_0}{N} > = < -\log\frac{n_0}{N_0} > = < \log\left(1+\frac{n_0-N_0}{N_0}\right) >$$

$$\approx < \frac{n_0-N_0}{N_0} > - < \frac{(n_0-N_0)^2}{2N_0^2} > = - < \frac{(n_0-N_0)^2}{2N_0^2} >$$

Since $< (n_0 - N_0)^2 >$ is the variance of the binomial distribution, the second order correction to the LogZero method is:

$$\mu = -\log\frac{n_0}{N} - \frac{1}{2}\left(\frac{1}{N_0} - \frac{1}{N}\right)$$

Appendix 2: PGF Method

If $p(N)$ is the probability of a luminosity counter to have a value of N, then the response to N interactions equals N convolutions of $p(N)$ with itself. Now, if the Fourier transform of $p(N)$ is p^F, then the Fourier transform of response to μ Poisson-distributed interactions is:

$$P(p)^F = \sum_n \frac{(p^F)^n \mu^n}{n!} e^{-\mu} = e^{\mu(p^F-1)}$$

Now, consider a Probability Generating Function, $G_p(z)$ where $z \in \mathcal{Z}$:

$$G_p(z) = \overline{z^n} = \sum_{n=0}^{m-1} p(n) z^n$$

Since the generating function of discrete convolutions $p(N)$ and $q(N)$ is $G(p) \cdot G(q)$, we have:

$$\sum_{n=0}^{m-1} p(n)z^n \sum_{a=0}^{m-1} q(l)z^l = \sum_{l=0}^{m-1}\left(\sum_{n,m,n+m=l} p(n)q(n)\right)z^l = \sum_{n=0}^{m-1}(p\cdot q)(n)\cdot z^n$$

Therefore, the Probability Generating Function to μ Poisson-distributed interactions:

$$G_{P(p)}(z) = \sum_n \frac{\left(G_p\right)^n \mu^n}{n!} e^{-\mu} = e^{\mu\left(G_p(z)-1\right)}$$

In practice, if a detector produces values n_i for an event i,

$$G_{P(p)}(z) = \overline{z^{n_i}} = \frac{\sum_i z^{n_i}}{N}$$

Where N is the total number of events. This leads to luminosity:

$$\log(G_{P(p)}(z)) = -\mu \log\left(G_p(z) - 1\right) \propto \mu \propto Luminosity$$

In fact, the average method and the LogZero methods are a special case of the Probability Generating Function. If one substitutes $z = 1 + \epsilon$ and takes the limit $\epsilon \to 0$ in PGF, one obtains the Average method, whereas taking $z = \epsilon$ and then $\epsilon \to 0$ in PGF leads to the LogZero method.

Acknowledgement

I want to sincerely thank my supervisor Vladislav Balagura and co-supervisor Rita Sadek for the invaluable discussions, suggestions, feedback, advice, and without whom this work would not have been possible; it has truly been an honor for me to share the office with them for the last 6 months. My sincere gratitude extends to all the members of LHCb group at LLR, and the Luminosity Working Group of LHCb collaboration.

References

1. Balagura, Vladislav (2021) *StarterTalk on Luminosity*

2. Alessio, Frederico & LHCb Collaboration (2022) *LHCb Upgrades*

3. LHCb Collaboration (2023) *The LHCb Upgrade I*

4. Balagura, Vladislav (2020) *Van der Meer Scan Luminosity Measurement and Beam–Beam Correction*, Eur. Phys. J.