SISSA

Muon identification for LHCb Run 3

View the article online for updates and enhancements.

IOP ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

TECHNICAL REPORT

Muon identification for LHCb Run 3

L. Anderlini,[a] F. Archilli,[b] A. Cardini,[c] V. Cogoni,[c] M. Fontana,[d] G. Graziani,[a] N. Kazeev,[e,f] H. Kuindersma,[g] R. Oldeman,[c] M. Palutan,[h] M. Santimaria,[h,1] B. Sciascia,[h] P. De Simone[h] and R. Vazquez Gomez[h,i]

[a] *INFN, Sezione di Firenze,*
 Via Giovanni Sansone 1, Sesto Fiorentino, Italy
[b] *Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg,*
 Grabengasse 1, Heidelberg, Germany
[c] *INFN, Sezione di Cagliari,*
 Strada provinciale per Sestu Km 1, Monserrato, Italy
[d] *CERN,*
 Esplanade des Particules 1, Meyrin, Switzerland
[e] *HSE University,*
 Pokrovsky Boulevard 11, Moscow, Russia
[f] *Sapienza University of Rome,*
 Piazzale Aldo Moro 5, Roma, Italy
[g] *Nikhef, National Institute for Subatomic Physics,*
 Science Park 105, Amsterdam, Netherlands
[h] *INFN, Laboratori Nazionali di Frascati,*
 via Enrico Fermi 54, Frascati, Italy
[i] *IGFAE, Universidade de Santiago de Compostela,*
 Rúa Xoaquín Díaz de Rábago, Santiago de Compostela, Spain

 E-mail: marco.santimaria@cern.ch

ABSTRACT: Muon identification is of paramount importance for the physics programme of LHCb. In the upgrade phase, starting from Run 3 of the LHC, the trigger of the experiment will be solely based on software. The luminosity increase to $2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ will require an improvement of the muon identification criteria, aiming at performances equal or better than those of Run 2, but in a much more challenging environment. In this paper, two new muon identification algorithms developed in view of the LHCb upgrade are presented, and their performance in terms of signal efficiency versus background reduction is shown.

[1]Corresponding author.

Contents

1 Introduction

The LHCb experiment [1] at the LHC is a single-arm forward spectrometer specialised in studying particles containing b or c quarks. Thanks to a versatile reconstruction and trigger system, the LHCb physics programme has been extended to electroweak, soft QCD and even heavy-ion physics. Many of the physics channels are identified by their very clean muon signatures, therefore muon identification and trigger are crucial for the success of the experiment.

A brief description of the Run 2 muon detector and reconstruction techniques follows, which sets the basis for the improvements later discussed in view of Run 3. A comprehensive description of the LHCb trigger for Run 2 can be found in [2]. During Run 1 and Run 2, the tracking system of LHCb provided a measurement of the momentum (p) of charged particles with a relative uncertainty that varied from 0.5% at low momentum to 1.0% at 200 GeV/c [3, 4]. The minimum distance of a track to a primary pp collision vertex (PV), the impact parameter (IP), was measured with a resolution of $(15 + 29/p_{\mathrm{T}})$ μm, where p_{T} is the component of the momentum transverse to the beam, in GeV/c [5]. Muons were identified and triggered by a system composed of five stations, M1-M5, of rectangular shape, placed along the beam axis as shown in figure 1. Stations M2 to M5 were placed downstream the calorimeters, and were interleaved with 80 cm-thick iron absorbers to select penetrating muons. The M1 station was placed in front of the calorimeters and used to improve the p_{T} measurement in the trigger [7].

Each muon station is subdivided into four regions, as shown in figure 2, with different read-out schemes defining the x, y resolutions. The dimensions of the logical pads were chosen such that their contribution to the p_{T} resolution was approximately equal to the multiple scattering contribution [6]. These logical pads were obtained from the crossing of horizontal and vertical strips (either cathodic pads or group of wires), with the exception of the full M1 station and the innermost regions of stations M4 and M5, where the logical pads corresponded to physical channels on the detector and were readout directly.

A schematic diagram showing the trigger data flow in Run 2 is depicted in figure 3. The trigger and reconstruction scheme followed three basic steps:
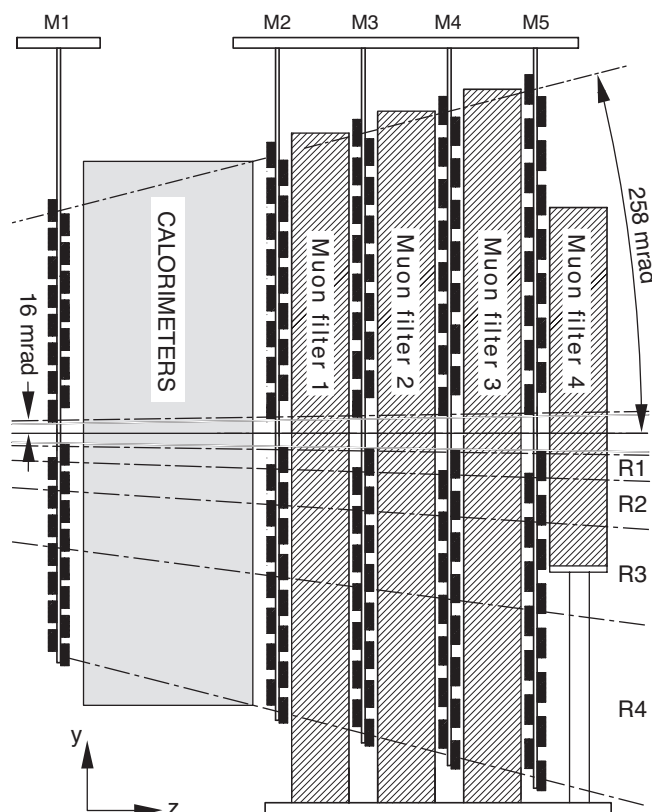
Figure 1. Side view of the muon system in the y-z plane [6].

- A hardware trigger (L0), based on selected calorimeter and muon information, to reduce the interaction rate of 20 MHz[1] to 1 MHz, which corresponded to the readout bandwidth of the detector. The L0 muon trigger was based on the coincidence of one hit in each of the five stations, selected in a projective Field Of Interest (FOI) defined in the x-y plane, from which a muon standalone p_T reconstruction was performed with $\sim 20\%$ resolution [6]. Candidate tracks above a p_T threshold of about 1.5 GeV/c were then used to build single and dimuon topologies.

- A first software stage (HLT1) based on partial reconstruction of tracks from the spectrometer, which allowed to put more strict constraints on the candidate p_T and IP. Concerning muons, the candidates from L0 were not used. Instead, a loose and efficient selection was performed, called `IsMuon`, based on the coincidence of hits in M2 to M5 stations, and combined with the information of the spectrometer. The muon hits were selected in a FOI centered around the track extrapolation position on the muon stations: the number of hits required was two, three or four in the momentum ranges 3–6, 6–10 and above 10 GeV/c, respectively, as expected from the muon penetration power in the iron absorbers [8].

---

[1]Out of the total LHC bunch crossing rate of 40 MHz, there are about 30 MHz of inelastic collisions, of which around 2/3 are visible in the LHCb detector.
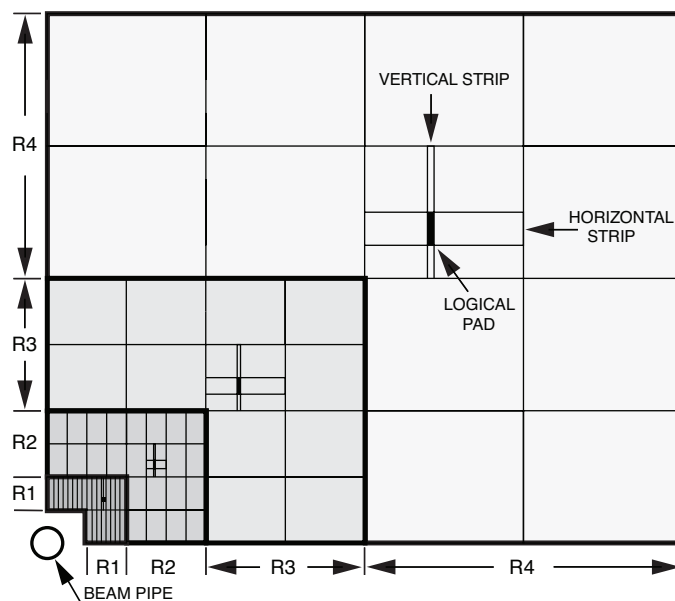
Figure 2. Front view of one quadrant of M2 showing the four regions. The intersection of a horizontal and a vertical strip defines a logical pad. The region and channel dimensions scale by a factor of two from one region to the following.
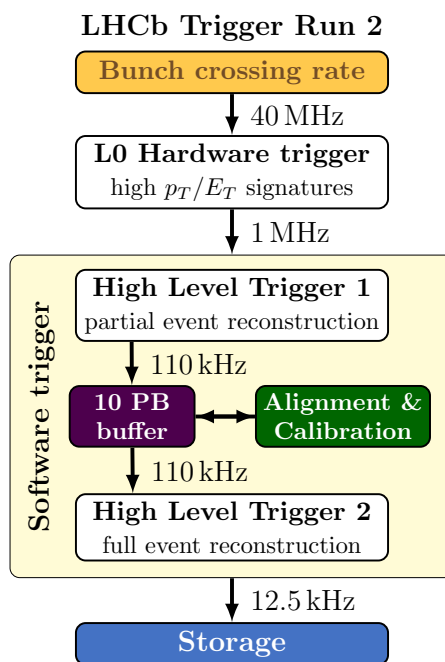

Figure 3. The LHCb trigger scheme in Run 2 [2].

- A more refined software trigger (HLT2), exploiting the full reconstruction of the detector information to reconstruct more complex signal topologies. Concerning muons, a better discrimination than `IsMuon` was achieved by using a likelihood variable (`MuonDLL`), built

upon the uncorrelated sum of the spatial residuals of the muon hits with respect to the track extrapolation position in each station [8], defined as:

$$D^2 = \frac{1}{N} \sum_{i=1}^{N} \left[\left(\frac{x_{\text{closest}}^i - x_{\text{track}}^i}{\text{pad}_x^i} \right)^2 + \left(\frac{y_{\text{closest}}^i - y_{\text{track}}^i}{\text{pad}_y^i} \right)^2 \right], \qquad (1.1)$$

where the index i runs over the N stations containing hits inside the FOI, and the *closest* coordinates represent the position of the hit which is closest to the track extrapolation point. The hit residuals were normalised to the logical pad size in the x and y directions, pad_x and pad_y respectively. The D^2 distribution for muons exhibits a narrow peak at 0, while hadrons satisfying the `IsMuon` criterion have a broader distribution. Using the D^2 spectra of muons as a signal proxy and of protons as a background proxy (pions are instead contaminated by real muons from decays in flight), the `MuonDLL` likelihood was defined, which measured the difference in probability for a candidate track to match the signal or background hypotheses. Using the above variable, on top of `IsMuon`, the misidentification probability for protons was kept at the 2–3 per mille level on the full momentum spectrum, with a muon efficiency above 90%. For pions, similar misidentification probabilities were obtained only for momenta higher than 50 GeV/c, with decays in flight contributing for another few per mille at low momenta [9, 10].

The LHCb detector will be upgraded for Run 3 to sustain a factor of five increase in the instantaneous luminosity, up to 2×10^{33} cm^{-2}s^{-1}. The 1 MHz readout limitation of the current detector will be removed, allowing for the full event rate to be processed in software without the need for a hardware stage [11]. For this reason, a full software trigger has been implemented, which will allow to select signal events with higher efficiency, and with the goal of achieving an order of magnitude increase in the physics bandwidth with respect to Run 2.

In preparation of Run 3, the M1 station has been removed due to the much higher occupancy which will be reached in front of the calorimeter, where the station is located (figure 1). In addition, its main contribution, consisting in the improvement of the standalone muon p_T determination in the hardware L0 trigger is no longer relevant. When working on the implementation of the future software muon trigger lines, two aspects have to be taken into account:

- The need to keep a high efficiency at HLT1, with a smooth dependence on the running conditions and on the phase space, and with a fast execution time. Concerning the bandwidth, a high rejection power must be guaranteed against combinatorial background, especially important at low momentum. This background originates from tracks extrapolated to the muon detector and paired to accidental hits in the muon chambers due for example to other muons in the event or to electronic noise. Accidental hits are expected to increase almost linearly with the luminosity.

- The possibility to tune highly selective cuts in order to achieve very low mis-identification levels at HLT2, especially useful for example in rare decay searches. The full information from the LHCb particle identification detectors may conveniently be used in this case as the constraints on the execution time are less stringent.

These goals can be achieved by following different approaches. In this paper we discuss a baseline strategy for the HLT1 which is an evolution of the present scheme. This assumes that

tracks in the spectrometer are reconstructed upfront, and that the muon identification is applied in two steps: a first step based on `IsMuon` as it is now, plus a second step based on a correlated χ^2 variable (section 2), which represents an improvement of the `MuonDLL`. At the HLT2 stage, the muon identification performance is further refined by means of a multivariate classifier (section 3).

2 Correlated χ^2

The misidentification of charged pions and kaons to muons has an almost irreducible component due to decays in flight, together with a combinatorial component that is relevant especially at $p < 10\,\mathrm{GeV}/c$, where the hit coincidence is less stringent. The present muon identification algorithm was optimised for a low occupancy scenario, without prioritising the rejection of the combinatorial background. The higher luminosity of Run 3 will require instead to suppress this background more effectively, especially in the central detector regions where the occupancies are higher.

An obvious limitation of the present approach based on the D^2 variable is that it does not include the information from the multiple scattering experienced by charged particles traversing the calorimeter and the iron absorbers, as well as the correlation between the hit positions across the muon system. The importance of taking into account these correlations is evident in figure 4, where two very different hit combinations are shown, yet giving a similar `MuonDLL` value. On the left, a random combination of hits scattered around the track extrapolation is shown, receiving contributions from uncrossed logical pads, indicated by the larger error bars. Such events are more affected by electronic noise and spillover hits. On the right-hand side, a clear pattern of hits is visible, which are displaced with respect to the extrapolated track due to multiple scattering.

These two topologies can be discriminated by using a χ^2 variable, expressed as

$$\chi^2_{\mathrm{CORR}} = \vec{\delta x}^T \mathrm{V}_x^{-1} \vec{\delta x} + \vec{\delta y}^T \mathrm{V}_y^{-1} \vec{\delta y}, \tag{2.1}$$

where $\vec{\delta x}$ and $\vec{\delta y}$ are the distances, in the x and y directions, between the track extrapolation points and the closest hit positions, with indices running over the stations M2 to M5. The covariance matrices V_x and V_y both have a diagonal contribution from the detector resolution and
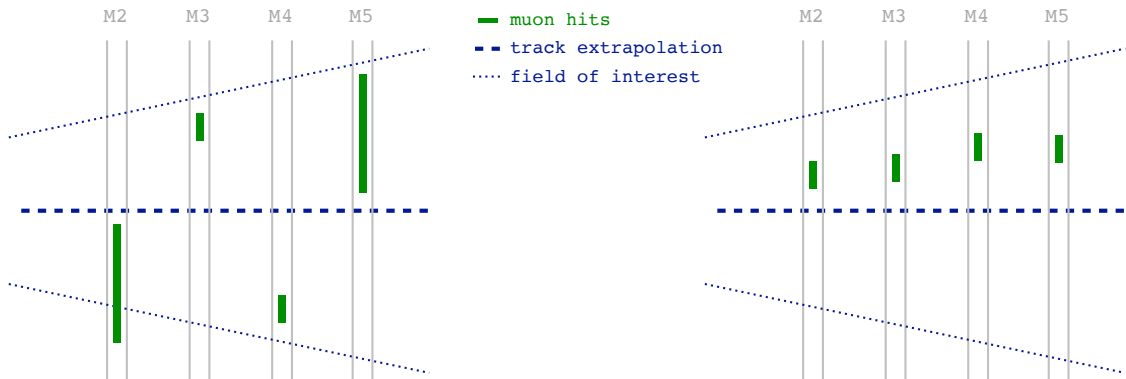

Figure 4. Two different combination of muon hits having a similar value of `MuonDLL`: a combinatorial background event (left) and a clear muon pattern (right).

off-diagonal terms quantifying the correlations induced by the multiple scattering. The diagonal terms corresponding to the hit position resolutions are of the form

$$V_{jj}^{\text{RES}} = \left(\text{pad}_{x,y}^{j}/\sqrt{12}\right)^2,\tag{2.2}$$

where the pad size along x and y corresponding to the muon hit in the given station are used. The multiple scattering (MS) contributions have been modelled as

$$V_{jk}^{\text{MS}} = \sum_{z_i < z_j, z_k} (z_j - z_i)(z_k - z_i)\sigma_{\text{MS},i}^2,\tag{2.3}$$

where $z_{j,k}$ represent the coordinates of stations M2 to M5 along the beam axis, z_i represents the coordinates of the main absorbers, namely the calorimeters and the muon iron filters, as listed in table 1, and $\sigma_{\text{MS},i}$ represents the MS deviation. This term takes the expression [12]

$$\sigma_{\text{MS},i} = \frac{13.6 \text{ MeV}}{\beta c p}\sqrt{\Delta z_i / X_0},\tag{2.4}$$

where p and βc are the momentum and the velocity of the incident particle, respectively, and $\Delta z_i / X_0$ is the thickness of the absorber at the given position z_i in units of radiation length, also listed in table 1. The number of degrees of freedom (ndof), i.e. the order of the covariance matrices, corresponds to the number of stations in which a hit is found.

Absorber	z position (m)	$\Delta z_i / X_0$
ECAL	12.8	25
HCAL	14.3	53
Muon filter 1	15.8	47.5
Muon filter 2	17.1	47.5
Muon filter 3	18.3	47.5

Table 1. Position along the beam axis and thickness in units of radiation length for the main scattering media contributing to the multiple scattering experienced by particles reaching the muon detector.

The probability for a muon to penetrate the iron absorbers and reach a given muon station depends on its momentum. In particular, below 6 GeV/c the probability to reach M4 and M5 stations can be substantially smaller than one, so that hits falling in the FOI of the track are in this case often due to accidental background. For this reason, in that momentum interval only the hits on M2 and M3 stations are included in the χ_{CORR}^2 computation.

The performance of the χ_{CORR}^2 variable is evaluated on muons and protons from data control samples collected in 2016. Data samples are preferred over simulation as the performance of the algorithms is very sensitive to the occupancy of the detector. The number of hits in the detector is dominated by low energy background from particles with energies below the simulation thresholds. Hence, data calibration samples are a better proxy for the background expected in Run 3. An abundant source of muons is provided by $J/\psi \to \mu^+\mu^-$ decays: by requiring the reconstructed J/ψ to have a large flight distance significance and good decay vertex quality, most of the combinatorial background from the tracks originating from the primary vertex is removed, and the sample gets
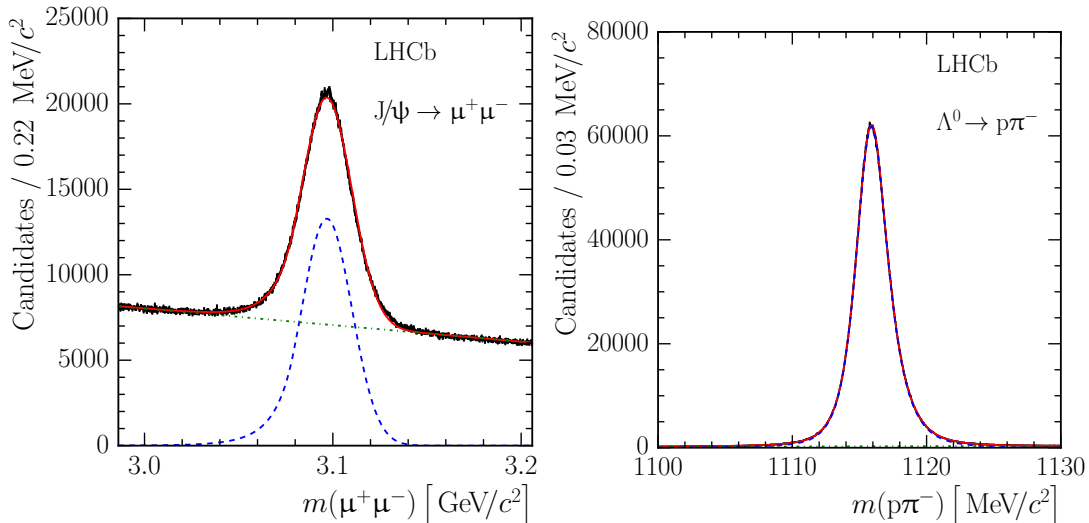
Figure 5. Typical invariant mass distributions for $J/\psi \rightarrow \mu^+\mu^-$ (left) and $\Lambda \rightarrow p\pi^-$ (right) calibration samples. The superimposed fit (red line) is composed of a signal (dashed blue) and background (dotted dash green) component [13].

enriched by $B \rightarrow J/\psi X$ candidates. To further reduce the background, one of the decay tracks, the *tag* muon, is required to be positively identified in the muon detector; the other track, the *probe* muon, is unbiased with respect to particle identification and trigger requirements and it is therefore used to measure the algorithm performances. Protons are selected from $\Lambda \rightarrow p\pi^-$ decays with vertex quality criteria and detachment of the decay vertex from the primary one. In addition, the invariant mass obtained by assigning the π mass to the two daughters is required to be outside the nominal K_S^0 mass window. Examples of mass spectra for muon and proton calibration samples are shown in figure 5.

For both samples, the residual background contribution is subtracted by using the *sPlot* method [14]. To perform unbiased studies, the muon and proton samples have been weighted in order to equalise their momentum, transverse momentum, and track multiplicity spectra. In addition, since the main challenge for Run 3 is the fivefold luminosity increase with respect to Run 2, a weighting procedure that adds more emphasis on high multiplicity events is applied to each calibration sample. Since there is not enough data to accurately emulate the upgrade conditions, the samples have been weighted in such a way to reproduce an occupancy spectrum which is in-between the two actual running conditions.

As a result, in figure 6 the χ^2_{CORR} spectrum for muons and protons satisfying the `IsMuon` requirement is shown, demonstrating a good separation between signal and background. A quantitative comparison between the performance of the χ^2_{CORR} and `MuonDLL` variables is shown in figure 7, where the proton rejection as a function of the muon efficiency, ROC curve in the following, is displayed for tracks satisfying the `IsMuon` requirement. The ROCs are shown in different momentum and transverse momentum intervals, which allow to probe the response of the muon identification algorithms in different regions of the detector and in different momentum regimes. The performance of the χ^2_{CORR} variable is definitely better than the `MuonDLL` in all regions of the phase space, and especially at low momenta. In particular, at muon efficiency of $\sim 98\%$, which is a
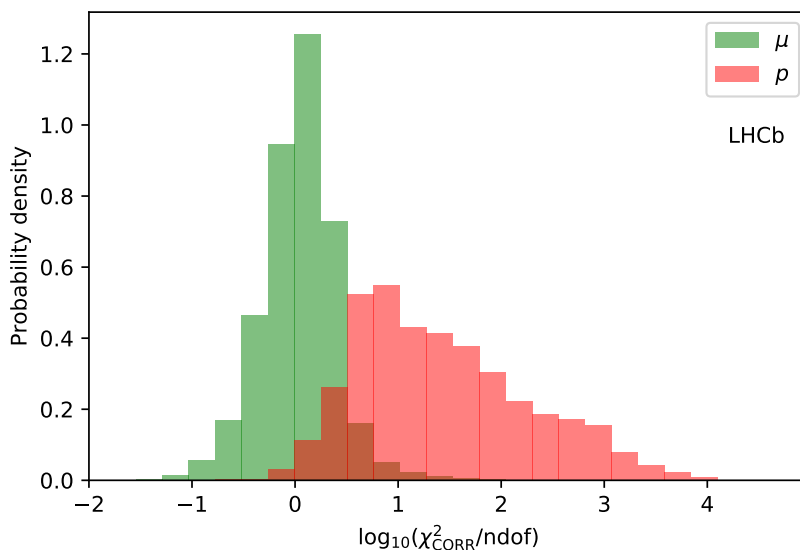
Figure 6. Spectrum of the χ^2_{CORR}, normalised to the degrees of freedom, for muons and protons as evaluated on 2016 calibration samples.

good working point for efficient trigger selections, the gain in background rejection is a factor ~ 1.4 in the region $p > 10$ GeV/c, $p_{\text{T}} < 2$ GeV/c, and exceeds a factor of 2 in the rest of the phase space.

2.1 Performance in HLT1

As discussed in section 1, for the HLT1 in Run 3 it will be crucial to guarantee a high efficiency for muons and a fast execution time of the algorithms. Moreover, a tighter rejection against combinatorial background with respect to the present `IsMuon` selection will be certainly needed.

Given the good performances of the χ^2_{CORR} variable in rejecting protons, which constitute pure combinatorial background to the muon detector, we consider as interesting to provide the rejection estimates on trigger unbiased events, which are mostly populated by pions, selected from a Run 2 data sample without any trigger requirement. As a preliminary selection for this benchmark, the events are filtered by requiring at least one track to satisfy `IsMuon` and the cuts $p_T > 800$ MeV/c and IP$\chi^2 > 35$,[2] which represent the main requirements of the Run 2 HLT1 single muon line. The rejection is therefore computed relatively to the above selection, and thus represents the improvement with respect to the present HLT1, and by removing the L0 trigger. To select high multiplicity events, only those having at least 3 primary vertices (nPVs) are used, whereas average Run 2 events have one primary vertex.

This study is done in three momentum intervals, 3–6, 6–10 and $p > 10$ GeV/c, since the number of hits selected by `IsMuon` is different in each one, as described in section 1. In each interval, a χ^2_{CORR} cut with $\sim 98\%$ muon efficiency, as evaluated on muon calibration data, is chosen. The results are shown in table 2 and demonstrate the effectiveness of this variable in rejecting about half of the trigger unbiased events, with a very small efficiency loss, on top of the Run 2 HLT1 muon

---

[2]The impact parameter χ^2, IPχ^2, is defined as the difference in the primary vertex fit χ^2 with and without the given track. Detached tracks, for example those coming from a B decay, have larger IPχ^2 values.
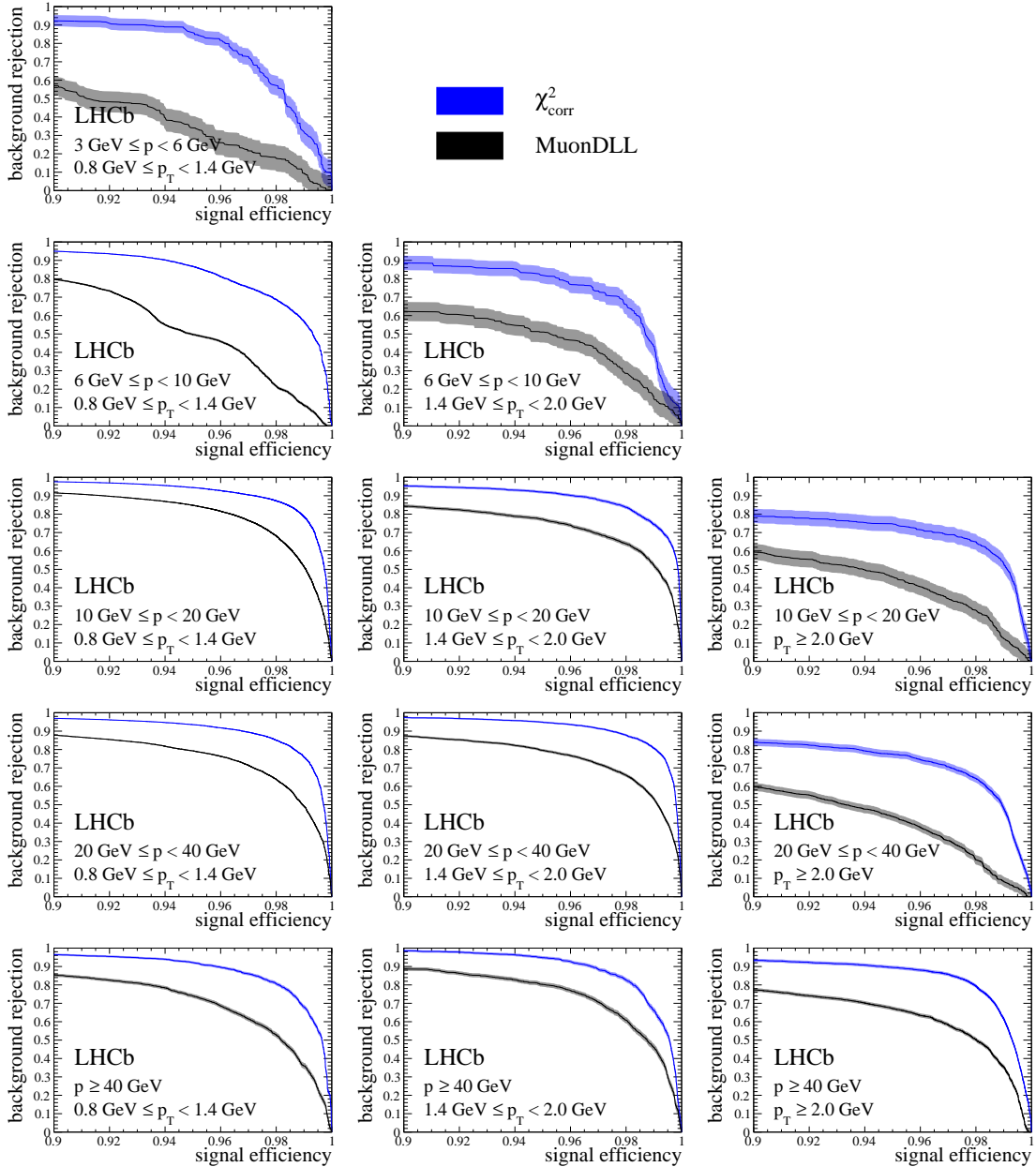
Figure 7. Proton rejection as a function of muon efficiency for tracks satisfying `IsMuon` obtained with the χ^2_{CORR} (blue) and `MuonDLL` (black) variables on 2016 calibration data. Low momentum bins, which are not covered by the calibration samples, are not shown. The bands represent statistical uncertainties.

selection. In particular, the highest rejection is achieved for $6 < p < 10\,\mathrm{GeV}/c$, where the fraction of pion decays in flight is lower with respect to $3 < p < 6\,\mathrm{GeV}/c$, and the multiple scattering correlations provide sensible discrimination as the momentum is not too high.

Finally, the χ^2_{CORR} execution time is tested within the HLT1 upgrade sequence. Throughput tests[3] are performed on simulated Run 3 data and show a χ^2_{CORR} resource usage of about 0.4% out

---

[3]On nodes mounting two Intel Xeon E5-2630 v4 CPUs at 2.20 GHz (40 threads/node).

Momentum range	Rejection factor
$3 < p < 6$ GeV/c	1.8
$6 < p < 10$ GeV/c	3.2
$p > 10$ GeV/c	2.2

Table 2. Rejection factors of the χ^2_{CORR} variable on trigger unbiased events, for a muon efficiency of $\sim 98\%$. The rejection is evaluated on top of the $p_T > 800$ MeV/c, IP$\chi^2 > 35$, `IsMuon` and nPVs ≥ 3 requirements.

of a total HLT1 throughput rate of ~ 36 MHz, and in view of a data taking rate of 30 MHz. This result makes the χ^2_{CORR} well suited for a usage in the upgraded HLT1 trigger of the experiment.

3 Multivariate algorithm

At the second stage of the trigger, the timing budget allows to use more complex algorithms. Besides the spatial information, each muon hit also carries two different time counters, one for each view, x and y. The number of views, i.e. the fact that the hit is crossed or uncrossed, also provides valuable information, as noise or spillover hits typically have one view only. This information, along with its correlations, can be exploited in a multivariate operator. To this purpose, a recent variant of gradient tree boosting available in the CatBoost library from Yandex [15] has been implemented [16, 17]. It uses oblivious decision trees as weak learners, as explained in the following.

A regular decision tree selects each split independently, while an oblivious decision tree has the same split on each level. The difference is illustrated in figure 8. An oblivious decision tree is less expressive but is much faster to evaluate, as it makes possible to unwrap the tree into a table and look up the correct leaf in one operation, instead of the multiple conditional jumps of a regular tree. According to a benchmark study by the CatBoost authors, which we were able to reproduce, this provides 30–100 faster prediction compared to the competing state-of-the-art gradient boosting libraries [18].
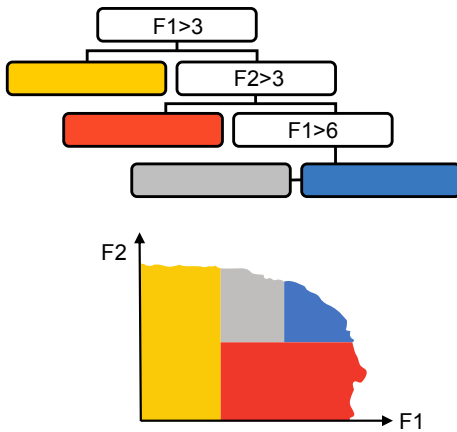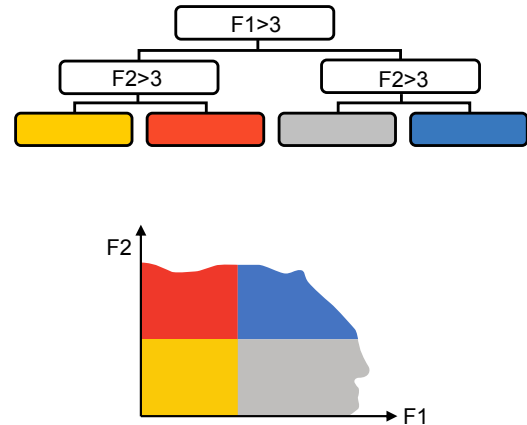

Figure 8. Classic versus oblivious decision trees. Reproduced from [19].

For the muon identification, five variables for each muon station M2 to M5 are used as input to the CatBoost algorithm:

- x_{res}: the difference between the closest hit x position and the track extrapolation, normalised to the total uncertainty;

- y_{res}: the difference between the closest hit y position and the track extrapolation, normalised to the total uncertainty;

- t_x: the time of the x view;

- $dt = t_x - t_y$: the temporal difference between the x and y views;

- N_{views}: the number of views.

The uncertainty in the residuals x_{res} and y_{res} contains the pad size and the contribution from the multiple scattering (eq. 2.4), summed in quadrature. In addition to the hit information, for each event the track extrapolation x and y coordinates on M2 are used, to allow the algorithm to discriminate between different detector regions. Finally, the aforementioned χ^2_{CORR} variable (section 2) of the track is added. This set of up to 23 variables per event has been found to be the smallest one containing the maximum associated information, without introducing excessive correlations. This feature is very important in order to decrease the complexity and hence the computation time of the operator in the trigger.

The classifier is trained using samples from 2016 data, to which the IsMuon requirement is applied. Since the IsMuon algorithm is very fast to execute and already rejects around 99% of background, evaluating the classifier only for events that pass the IsMuon requirement allows to significantly reduce the computational cost and to focus on reducing the remaining background. The data samples used in the training are the same calibration samples used for the χ^2_{CORR} evaluation:

- muons from $J/\psi \rightarrow \mu^+\mu^-$ decays,

- protons from $\Lambda \rightarrow p\pi^-$ decays,

- pions from $D^{*-} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^-$ decays.

While protons represent pure combinatorial events, the pion sample is added to boost the training statistics and accounts for another source of classification error due to particles that decay in flight to muons before reaching the muon stations. These samples have been treated as described in section 2, including kinematic reweighting, background subtraction and multiplicity weights.

To deal with negative sWeights, the solution proposed in ref. [16] is used, which consists of first using a machine learning regression to estimate the expected sWeight in each point of the training variable phase space and, second, using the expected sWeight as event weight during classification.

Finally, since the classifier is trained on the same 2016 calibration data which are used to evaluate its performance, a cross-validation method is used to obtain unbiased predictions. The dataset is split into 5 subsets of equal size, and the model is independently trained on all subsets but the i-th, for which predictions are made. The ROC curves of the CatBoost algorithm are shown in
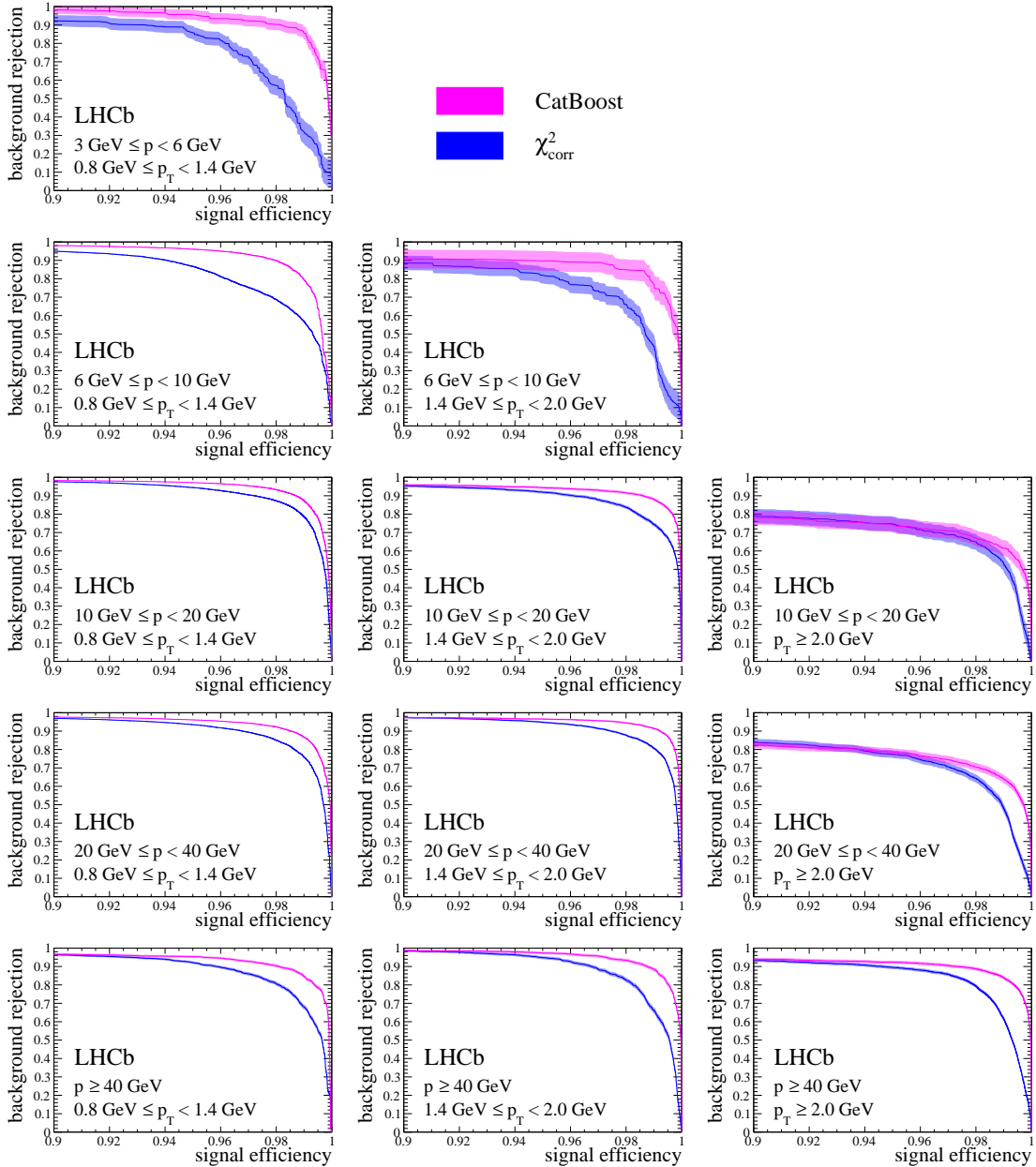
Figure 9. Proton rejection as a function of muon efficiency for tracks satisfying `IsMuon` obtained with the CatBoost algorithm (magenta) and $\chi^2_{\rm CORR}$ (blue) on 2016 calibration data. Low momentum bins, which are not covered by the calibration samples, are not shown. The bands represent statistical uncertainties.

figure 9 for muon efficiencies above 90%. For comparison, the ROC curves for the $\chi^2_{\rm CORR}$ variable are superimposed.

As a result, for high muon efficiency the CatBoost algorithm has better discriminating power than $\chi^2_{\rm CORR}$ in all the momentum bins. The greatest difference in background rejection is observed around 98% muon efficiency, where for $p < 10$ GeV/c it lies in $20 - 40\%$ range and $4 - 10\%$ for $p \geq 10$ GeV/c.

With a similar setup as the one for HLT1 (section 2.1), throughput tests are performed on simulated Run 3 data and show a resource usage of about 0.4% out of a total HLT2 throughput rate of ~ 129 Hz. Therefore, the CatBoost operator is fast enough to be employed in the upgraded HLT2 trigger of the experiment.[4]

4 Conclusions

Two new muon identification algorithms have been developed in view of the LHCb Run 3 upgrade. The first one, χ^2_{CORR}, expands on the muon likelihood variable developed in Run 1 by including the correlation among the muon hits. The second one features a multivariate algorithm based on the CatBoost machine learning toolkit. The performances of both algorithms in terms of background rejection versus signal efficiency are characterised on 2016 proton calibration data, and in both cases are found to improve considerably those of the muon likelihood used during Run 1 and Run 2, with the CatBoost classifier offering a slightly better performance. As far as the computational times are concerned, the χ^2_{CORR} has been proven to be fast enough to be included in the upgrade HLT1 muon trigger lines. The CatBoost algorithm, while improving the competing state-of-art gradient boosting libraries, can be computed in the HLT2, where the time constraints are less stringent.

---

[4]Being around 10 times slower than χ^2_{CORR}, the CatBoost implementation is not currently cost effective for HLT1.

References

[1] LHCb collaboration, *LHCb Detector Performance*, *Int. J. Mod. Phys. A* **30** (2015) 1530022 [arXiv:1412.6352].

[2] LHCb collaboration, *Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, 2019 *JINST* **14** P04013 [arXiv:1812.10790].

[3] LHCb collaboration, *Measurement of the track reconstruction efficiency at LHCb*, 2015 *JINST* **10** P02007 [arXiv:1408.1251].

[4] LHCb Outer Tracker Group collaboration, *Improved performance of the LHCb Outer Tracker in LHC Run 2*, 2017 *JINST* **12** P11016 [arXiv:1708.00819].

[5] R. Aaij et al., *Performance of the LHCb Vertex Locator*, 2014 *JINST* **9** P09007 [arXiv:1405.7808].

[6] LHCb collaboration, *LHCb muon system: Technical Design Report*, Tech. Rep., CERN-LHCC-2001-010, CERN, Geneva (2001).

[7] R. Aaij, J. Albrecht, F. Dettori, K. Dungs, H. Lopes, D. Martinez Santos et al., *Optimization of the muon reconstruction algorithms for LHCb Run 2*, Tech. Rep., LHCb-PUB-2017-007, CERN-LHCb-PUB-2017-007, CERN, Geneva (2017).

[8] G. Lanfranchi, X. Cid Vidal, S. Furcas, M. Gandelman, J.A. Hernando, J.H. Lopez et al., *The Muon Identification Procedure of the LHCb Experiment for the First Data*, Tech. Rep., LHCb-PUB-2009-013, CERN-LHCb-PUB-2009-013, CERN, Geneva (2009).

[9] M. Gandelman and E. Polycarpo, *The Performance of the LHCb Muon Identification Procedure*, Tech. Rep., LHCb-2007-145, CERN-LHCb-2007-145, CERN, Geneva (2008).

[10] F. Archilli et al., *Performance of the Muon Identification at LHCb*, 2013 *JINST* **8** P10020 [arXiv:1306.0249].

[11] LHCb collaboration, *Framework TDR for the LHCb Upgrade: Technical Design Report*, Tech. Rep., CERN-LHCC-2012-007, LHCb-TDR-12, CERN, Geneva (2012).

[12] Particle Data Group collaboration, *Review of Particle Physics*, *Chin. Phys. C* **40** (2016) 100001.

[13] O. Lupton, L. Anderlini, B. Sciascia and V. Gligorov, *Calibration samples for particle identification at LHCb in Run 2*, Tech. Rep., LHCb-PUB-2016-005, CERN-LHCb-PUB-2016-005, CERN, Geneva (Mar, 2016).

[14] M. Pivk and F.R. Le Diberder, *SPlot: A Statistical tool to unfold data distributions*, *Nucl. Instrum. Meth. A* **555** (2005) 356 [physics/0402083].

[15] L. Prokhorenkova, G. Gusev, A. Vorobev, A.V. Dorogush and A. Gulin, *CatBoost: unbiased boosting with categorical features*, in *Advances in Neural Information Processing Systems*, pp. 6638–6648, 2018 [arXiv:1706.09516].

[16] M. Borisyak and N. Kazeev, *Machine Learning on sWeighted Data*, *J. Phys. Conf. Ser.* **1525** (2020) 012088 [arXiv:1912.02590].

[17] M. Borisyak and N. Kazeev, *Machine Learning on data with sPlot background subtraction*, 2019 *JINST* **14** P08020 [arXiv:1905.11719].

[18] Yandex, *Model evaluation benchmark*, https://github.com/catboost/benchmarks/blob/master/model_evaluation_speed/model_evaluation_benchmark.ipynb (2019).

[19] M. Levin, *MatrixNet Applications at Yandex*, https://www.slideshare.net/mlprague/michael-levin-matrixnet-applications-at-yandex (2016).