Flavours of Physics: the machine learning challenge for the search of $\tau^- \to \mu^-\mu^-\mu^+$ decays at LHCb

Thomas Blake[1], Marc-Olivier Bettler[2], Marcin Chrząszcz[3,4], Francesco Dettori[2], Andrey Ustyuzhanin[5,6,7], Tatiana Likhomanenko[5,6,7]

[1] University of Warwick, Coventry, United Kingdom
[2] CERN, European Organization for Nuclear Research, Geneva, Switzerland
[3] Physik-Institut, Universität Zürich, Zürich, Switzerland
[4] Institute of Nuclear Physics, Polish Academy of Sciences, Krakow, Poland
[5] Yandex School of Data Analysis, Moscow, Russia
[6] NRC "Kurchatov Institute", Moscow, Russia
[7] National Research University Higher School of Economics (HSE), Russia

July 20, 2015

Abstract

The Flavours of Physics machine learning challenge has been proposed to enhance the exchange of knowledge between the particle physics and data science communities. The LHCb collaboration provides simulated and real data used in the search for lepton flavour violationg decay $\tau^- \to \mu^-\mu^-\mu^+$. The challenge will be hosted by Kaggle (`https://www.kaggle.com`). This document describes in details the challenge with physics motivation for studying the $\tau^- \to \mu^-\mu^-\mu^+$ decay and analysis steps. No prior knowledge of high energy physics is required. All information on how to participate in the Challenge is available on the Kaggle web site.

1 Introduction

In this document we present the details of the "Flavours of Physics" challenge for the search of $\tau^- \to \mu^-\mu^-\mu^+$ decays at the LHCb experiment. The document is organized as follows. In the first section we define the physics case and briefly introduce the Large Hadron Collider (LHC) machine at CERN and the LHCb experiment. The second section is devoted to a detailed explanation of the evaluation procedure for the challenge. The third section discusses the physics background relevant for this search.

1.1 Motivation

The Standard Model (SM) of particle physics is the current theory describing fundamental particles as well as strong, weak and electromagnetic interactions. It was formulated in the 1960's and 1970's and since then has passed with flying colours extensive experimental tests. The final missing piece of the SM was the Higgs Boson, which was discovered by the LHC in 2012. Despite these facts, the SM has several severe imperfections. For instance SM does not explain the matter-antimatter asymmetry in the Universe or the structure of generations of elementary particles. Furthermore it does not predict the existence of the dark matter or describes quantum gravity. This leads to extensive studies of extended theories, commonly labelled as physics beyond the Standard Model (BSM).

In the SM, fermions (particles that build matter) can be group into two categories: leptons and quarks. For unknown reasons each category is organised in three generations of particles. Curiously, one can show that a minimum of three generations are needed to allow for matter-antimatter asymmetries to exist. In the lepton family the particles are: (e^-, ν_e), $(\mu^-\nu_\mu)$, $(\tau^-\nu_\tau)$. Lepton Flavour (L_e, L_μ, L_τ) and Lepton Number (L) are properties [1], assigned to each lepton as detailed in Tab. 1. In the SM, these numbers are conserved, *i.e.* all the processes foreseen by the SM do not allow the variation of the total lepton number or of the total lepton flavour numbers.

Notably, lepton flavour violation (LFV) was discovered in the neutrino sector. The phenomena showing this effect is known as neutrino oscillations [1], in which one kind of neutrino (*e.g.* ν_e) transforms into a different flavour (ν_μ or ν_τ). Observation of LFV in the charged lepton sector would be a clear indication of Physics beyond present theory of particle physics.

The search for decays that do not conserve these numbers started in the late 1930's with the discovery of the muon (μ). It was believed that muons were an excited electron state in which case one would expect to observe a decay $\mu^- \to e^-\gamma$, where a photon with predictable energy would be emitted. No such process has ever been observed. The muon decays instead through

[1] Technically these are called "quantum numbers", as they can take only certain discrete values.

Table 1: Assignment of lepton numbers L and $L_f,\ f = e,\ \mu,\ \tau$ to elementary leptons. All quark states assume values of zero for all these quantum numbers.

Particle	L	L_e	L_μ	L_τ
e^- (e^+)	+1 (−1)	+1 (−1)	0	0
μ^- (μ^+)	+1 (−1)	0	+1 (−1)	0
τ^- (τ^+)	+1 (−1)	0	0	+1 (−1)
ν_e ($\overline{\nu}_e$)	+1 (−1)	+1 (−1)	0	0
ν_μ ($\overline{\nu}_\mu$)	+1 (−1)	0	+1 (−1)	0
ν_τ ($\overline{\nu}_\tau$)	+1 (−1)	0	0	+1 (−1)

the process $\mu^- \to e^- \nu_\mu \bar{\nu}_e$, with emission of a muon neutrino and an electron anti-neutrino to preserve the total electronic and muonic lepton numbers.

Similarly, in the 1970s an even heavier lepton was discovered as product of e^+e^- annihilations: the tau (τ) lepton, with a mass equivalent to about 3500 electrons. Typical decays of the τ leptons are $\tau^- \to e^- \nu_\tau \bar{\nu}_e$ and $\tau^- \to \mu^- \nu_\tau \bar{\nu}_\mu$, that conserve the various lepton numbers involved.

However, if lepton flavour is not a perfectly conserved quantity in nature, and various explanations of the matter asymmetry in the universe require this, then the τ lepton can also decay into three muons though the reaction $\tau^- \to \mu^- \mu^- \mu^+$, forbidden instead in the Standard Model. The discovery of such a reaction would therefore be a major breakthrough on the laws of nature. Recently, the LHCb collaboration performed a search for the $\tau^- \to \mu^- \mu^- \mu^+$ decay [2], in τ leptons produced by the LHC proton-proton collisions.

1.2 Detector and data analysis

The Large Hadron Collider (LHC) [3], located at CERN, in Geneva, Switzerland, is the world's largest particle accelerator providing collisions with the highest available energy. It accelerates protons and brings them to collide, at 8 TeV (as of for the 2012 running period) centre-of-mass energy.

There exist four interaction points, where protons collide, where seven experiments (ATLAS [4], CMS [5], LHCb [6], ALICE [7], TOTEM [8], LHCf [9] and MoEDAL [10]) are located. The protons are accelerated using 400 MHz radio frequency (RF) cavities and their trajectory is bent by superconducting dipole magnets.

The LHCb experiment was designed to study the properties of particles containing b and c quarks, but thanks to outstanding performance it allows also to study other particles produced at LHC collisions.

The LHCb detector is a single-arm forward spectrometer. The detector covers the pseudo-rapidity (η) range $2 < \eta < 5$, which corresponds to the geometric acceptance defined by a solid angle of 10–250 mrad (10 – 300 mrad) in the vertical (horizontal) plane[2].

In the first run of the LHC, the LHCb detector saw on average 1.4 interactions per crossing of proton bunches occurring every 50 ns. Usually each $p-p$ interaction creates about 80 charged particles visible in the LHCb detector that originate from the interaction point (collision point of two protons), also called primary vertex (PV).

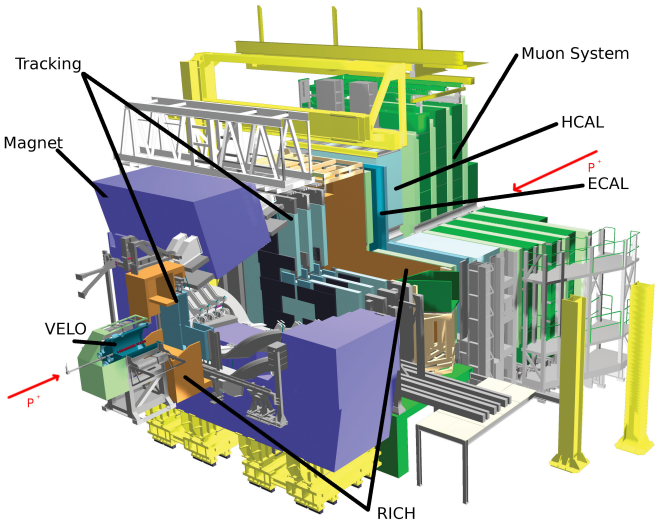

Figure 1: Schematic drawing of LHCb detector.

The LHCb spectrometer is composed of several sub-detectors, serving different and complementary purposes. The full layout of the LHCb detector is presented in Fig. 1. The reconstruction of electrically-charged particles is performed using the information collected by the Vertex Locator (VELO), fine granularity silicon tracker and the magnet. Momentum of the charged particles is measured thanks to the bending power of the dipole magnet. The particle identification is performed by two Ring Imaging Cherenkov detectors (RICH1 and RICH2) and the muon system (M1-M5). If a charged particle passes through a material faster than the local speed of light in that material ($\frac{c}{n}$, where c is the speed of light and n is the refractive index of the material) the particle loses energy by emitting light, known as Cherenkov radiation. This is the same process that makes the water in nuclear reactors appear to glow. The Cherenkov photons are emitted at a fixed angle depending on

[2] $\eta = -\ln\tan\frac{\theta}{2}$, where θ is the polar angle w.r.t. the beam axis.

the speed of the particle, which when combined with the momentum of the particle allows LHCb to tell different charged particles apart (based on their different masses). The Scintillating Pad Detector (SPD), Pre-Shover (PS), Electromagnetic Calorimeter (ECAL) and Hadronic Calorimeter (HCAL) measure the energy of neutral particles.

1.3 The challenge goal

The main goal of this challenge is to gain sensitivity in the search for $\tau^- \to \mu^-\mu^-\mu^+$ decays. That is achieved by improving the discriminating power between signal events (where the decay did occur) and background events (where it did not). We provide signal and background samples for training and testing. The evaluation is done in two steps: firstly, the classifier is checked not to depend too strongly on the discrepancies between real data and simulation and the classifier is checked not to be too correlated with the τ mass. The second step is the evaluation of the classifier using the weigthed area under the ROC curve. The details of the evaluation procedure will be given in Sect. 3.

2 Data description

In this section we present the signal production processes and possible background contributions.

2.1 Tau lepton production channels

Tau leptons are produced by different mechanisms. At LHCb, taus are produced in the decay of heavy flavoured particles (containing a c or a b quark), which are listed in Table 2. They are mainly produced in the decays of D_s^- or D_s^- particles, such as $D_s^- \to \tau\overline{\nu}$. In the simulation samples provided, the correct proportions of the different tau production mechanisms are respected, and the production mechanism is identified by the label `production`.

2.2 Physics backgrounds

The background for $\tau^- \to \mu^-\mu^-\mu^+$ decay can be divided in two categories. The first one consists of cases where one or more light hadrons (pion or kaon) is wrongly identified as a muon. The main process in this category is $D^+ \to K^-\pi^+\pi^+$. The invariant mass distributions for this process are shown in Fig. 2. These two mass distributions differ because of the mass assigned to the final states and thus used when computing the mass of the initial state. On the left-hand side, the muon mass ($105.66\,\mathrm{MeV}/c^2$) is assigned to all final states, while, on the right-hand side, the correct masses for kaons and pions ($139.57\,\mathrm{MeV}/c^2$ for $\pi^\pm$ and 493.68 for $K^\pm$) are used.

Table 2: Production mechanisms and their proportions for tau leptons at LHCb, according the centre-of-mass energy. X_b denotes any particle containing a beauty (b) quark. `production` is a label that, in the simulation, specifies the production mechanism of τ. In data, this label is set to -99.

Mode	7 TeV	8 TeV	`production`
Prompt $D_s^- \to \tau$	$71.1 \pm 3.0\,\%$	$72.4 \pm 2.7\,\%$	1
Prompt $D^- \to \tau$	$4.1 \pm 0.8\,\%$	$4.2 \pm 0.7\,\%$	2
Non-prompt $D_s^- \to \tau$	$9.0 \pm 2.0\,\%$	$8.5 \pm 1.7\,\%$	5
Non-prompt $D^- \to \tau$	$0.18 \pm 0.04\,\%$	$0.17 \pm 0.04\,\%$	6
$X_b \to \tau$	$15.5 \pm 2.7\,\%$	$14.7 \pm 2.3\,\%$	4

Hence the mis-identification results in a shift in the mass of the initial state. This background can be removed — and we suggest the participants to do so with a requirement on the probability for the final states to be muons, `min_ANNmuon`> 0.4 (the meaning of this variable will be explained in the appendix).

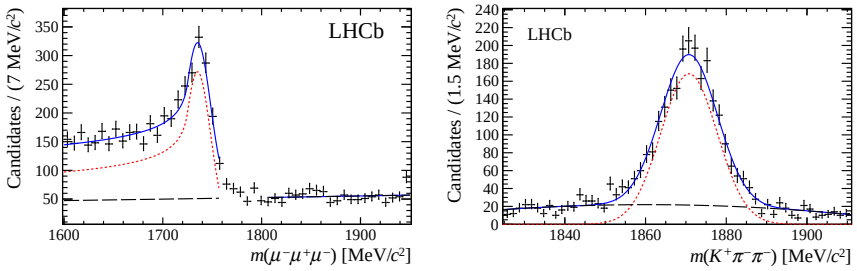

Figure 2: D meson invariant mass distribution in $D^+ \to K^-\pi^+\pi^+$ decays as observed in data. On the left-hand side, all hadrons have been assigned muon mass hypothesis before computing the mass of the mother, on the right-hand side, the correct mass hypotheses have been used.

A second dangerous background $D_s \to \eta(\to \mu^+\mu^-\gamma)\mu^-\nu_\mu$ originates from the decay in which there are three real muons that can mimic the signal signature. This background can be effectively removed requiring all mass combinations of two muons of opposite sign to be greater than $450\,\mathrm{MeV}/c^2$: $m_{\mu^-\mu^+} > 450\,\mathrm{MeV}/c^2$.

We provide four data sets:

- `training.csv` is a labelled data set (the `signal` being 1 for signal events, 0 for background events) to use for training the classifier. Background events come from real data mass side-bands[3] and from

[3] The mass side-bands are mass regions around the τ mass region in which the signal

simulation.

- `check_agreement.csv` is a labelled data set (the `signal` being 1 for simulated data, 0 for real data) with the same features as in the `training.csv`. This data set is used to check yourself the agreement between simulated and real data as described in Sec. 3.1.

- `check_correlation.csv` is a data set with the same features as the `training.csv`, to check yourself correlation of the classifier with the τ mass as described in Sec. 3.2.

- `test.csv` is a non-labelled (signal and background are mixed and indistinguishable for the participants) data set, containing simulated signal events and real background data, simulated events and real data for the control channel (latter form the main part of the test data set).

All data sets have a format along the following example, where `LifeTime` and `dira` are examples of the features:

id	signal	LifeTime	dira	...
1	0	0.1	1.0	...
2	1	0.99	3.2	...
3	1	0.4	2.4	...

3 Evaluation procedure

In this section we discuss in detail the evaluation procedure and all requirements that are requested from the final classifier.

3.1 Verification of the agreement between data and simulation

Since the classifier is trained on simulation data for the signal and real data side-bands for the background, it is possible to reach a high performance by picking features that are not perfectly modeled in the simulation (`min_ANNmuon` is an example of such feature). We require for the classifiers not to have large discrepancy when applied to data and to simulation. To check this we use a control channel, $D_s^+ \to \phi(\to \mu^-\mu^+)\pi^+$, that has a similar topology as the signal decay, $\tau^- \to \mu^-\mu^-\mu^+$. We provide both data and simulation samples for this decay, as the `check_agreement.csv` data set, to which the classifier can be verified. The Kolmogorov–Smirnov (KS) test is used to evaluate the differences between the classifier distributions on both samples. We require the KS-value of the test to be smaller than 0.09. The

is ultimately searched. The mass side-bands are used as a proxy for the background.

test is performed by the function `compute_ks` in the script `evaluation.py`, which returns the KS-value. **If KS-value is** $\geqslant 0.09$ **then the figure of merit will be equal to** -1**.**

The cumulative distribution (CDF) functions are computed for simulated data predictions and real data predictions and Kolmogorov-Smirnov metric is calculated:

$$KS = \max |F_{\text{simulation}} - F_{\text{real}}|,$$

where $F_{\text{simulation}}$ and F_{real} are cumulative distribution functions for Monte Carlo data and real data corresponding. KS is appropriate for agreement check because it reflects the CDF point with disagreement.

3.2 Checking the correlation with the τ mass

We further require that the classifier be uncorrelated with the τ mass, as such correlations can cause an artificial signal-like mass peak or lead to incorrect background estimations. We perform the Cramer–von Mises (CvM) test [11] to that end. We require the CvM-value of the test to be smaller than 0.002. The test is performed by the function `compute_cvm` in the script `evaluation.py` which returns the CvM-value. Participants can verify classifiers on the `check_correlation.csv` data set. **If CvM-value is** $\geqslant 0.002$ **then the figure of merit will be equal to** -2**.**

For computing CvM metric the global predictions CDF is compared to a local (in some mass interval) predictions CDF. After that all intervals are averaged:

$$CvM_{\text{interval}} = \int (F_{\text{global}} - F_{\text{interval}})^2 dF_{\text{global}},$$

$$CvM = < CvM_{\text{interval}} >_{\text{interval}},$$

where F_{global} and F_{interval} are predictions cumulative distribution functions for all data and data in some mass interval corresponding.

3.3 Final Evaluation

The calculation of the final figure of merit is performed only if the two above tests are passed on `test.csv` with success and is calculated only using events with `min_ANNmuon`> 0.4.

Originally, the LHCb used the CLS [12] method to determine the upper limit. This method unfortunately is computationally expensive and can't be used in this challenge. Instead we proposed a much simpler metric, which is the weighted area under the ROC curve. The areas and their weights are defined in Fig. 3.

The reason to assign different weights to different bins of signal efficiency is that the sensitivity to a given process is not a linear function of expected background events. Most of the sensitivity is obtained when number of expected background events is $\mathcal{O}(1)$. For example see Table 208 in Ref. [13].

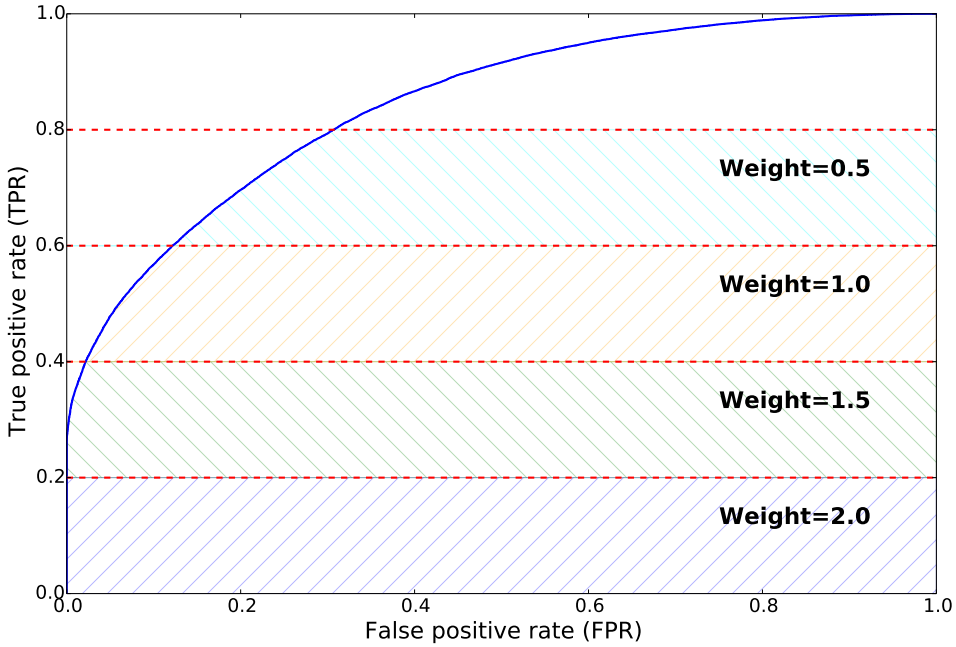

Figure 3: Weights assigned to the different segments of the ROC curve for the purpose of submission evaluation. The x axis is the False Positive Rate (FPR), while the y axis is True Positive Rate (TPR).

A Special Relativity for non-physicists

In this appendix we give a short introduction to Special Relativity for non-physicists. This appendix can be skipped if you are familiar with the subject.

Special Relativity combines the space-related quantities (as the momentum $\overrightarrow{p}$) to time-related quantities (as the energy E) in objects called four-vectors. A prime example of such four-vector is four-momentum ($\mathbf{p}$) defined as:

$$\mathbf{p} = \begin{pmatrix} E \\ p_x \\ p_y \\ p_z \end{pmatrix} = \begin{pmatrix} E \\ \overrightarrow{p} \end{pmatrix}. \tag{1}$$

The above equation assumes the *natural system of units* in which the speed of light, c, equals 1. Fundamental principle of Special Relativity for a particle four-momentum reads:

$$m^2 = \mathbf{p}^T \eta \mathbf{p} = \begin{pmatrix} E & \overrightarrow{p}^T \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \begin{pmatrix} E \\ \overrightarrow{p} \end{pmatrix} = E^2 - p^2, \tag{2}$$

where m is the mass of the particle, $p^2 = |\overrightarrow{p}|^2$ and η is called the Minkowsky metric. The Eq. 2 allows to calculate the mass of a decaying particle like the

τ in a $\tau^- \to \mu^-\mu^-\mu^+$ decay using the energy and momentum of the three muons by adding their four-momenta:

$$\mathbf{p}_\tau = \sum_{i=1}^{3} \mathbf{p}_{\mu,i} = \sum_{i=1}^{3} \begin{pmatrix} E_{\mu,i} \\ \overrightarrow{p}_{\mu,i} \end{pmatrix} \tag{3}$$

Knowing the $\mathbf{p}_\tau$ for a τ lepton one can calculate its mass using Eq. 2 (it is worth pointing out that real particles have only positive masses so there is no ambiguity in Eq. 2, as only one positive solution exists).

Useful quantity used by physicist is the *transverse momentum*, p_T. It is defined as projection of the total momentum ($|\overrightarrow{p}|$) on the plane perpendicular to the colliding proton beams. In LHCb, the z axis is choosen to be the direction of the beam, as shown with red arrows in Fig. 1 and the component transverse to the beam-line $p_T = \sqrt{p^2 - p_z^2} = \sqrt{p_x^2 + p_y^2}$.

Another useful kinematic variable used is the pseudorapidity, η. It is defined as

$$\eta = -\ln\left[\tan\frac{\theta}{2}\right] = \operatorname{arctanh}\left(\frac{p_z}{|\overrightarrow{p}|}\right), \tag{4}$$

where θ is the angle between the beam axis and the track. Rewriting Eq. 2 in terms of η, $|\overrightarrow{p}|$ and p_T reads:

$$m^2 = E^2 - p_T^2 - |\overrightarrow{p}|^2 \tanh^2\eta \tag{5}$$

B Geometric variables in LHCb detector

LHCb detector uses geometric properties to discriminate between signal and background events. In this chapter we will review the variables that are relevant for $\tau^- \to \mu^-\mu^-\mu^+$ search.

Impact parameter (IP) of a particle is defined as the closes distance between PV and the track. This quantity is shown in Fig. 4. Throughout "significance" is defined as $\dfrac{\text{IP}}{\sigma\text{IP}}$, where σIP is the uncertainty of the measured IP.

Another useful quantity used to discriminate between signal and background events is the Vertex χ^2. The full definition of this quantity is beyond the scope of this document and can be found in [14]. Important thing is to notice that the higher this quantity is the poorer the vertex was reconstructed making it more background like.

A crucial variable in searches for rare decays like $\tau^- \to \mu^-\mu^-\mu^+$ are isolation variables. There are two types of "isolations": track isolation and cone isolation. The track isolation gives information if a track of one of the thee muons in the decay in question forms a good vertex with another track in the event. Cone isolation describes the how well the τ decay vertex

is separated from other tracks in the decay. For full derivation of those variables see [15].

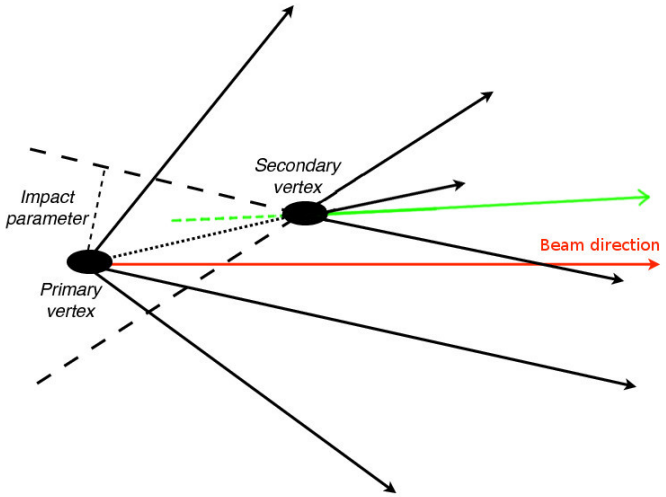

Figure 4: Schematic drawing of event in LHCb detector.

C Data features explanation

The list of the features (variables) available follows. The notation `p0, p1, p2` refer to the final states tracks.

- FlightDistance - Distance between τ and PV (primary vertex, the orginal protons collision point).
- FlightDistanceError - Error on FlightDistance.
- mass - τ candidate invariant mass, which is absent in the test samples.
- LifeTime - Life time of tau candidate.
- IP - Impact Parameter of tau candidate.
- IPSig - Significance of Impact Parameter .
- VertexChi2 - χ^2 of τ vertex.
- dira - Cosine of the angle between the τ momentum and line between PV and tau vertex. The above mentioned angle is the angle between the green dashed line and the dashed line connecting the Secondary vertex with the Primary vertex on Fig. 4.
- pt - transverse momentum of τ.

- DOCAone - Distance of Closest Approach between `p0` and `p1`.
- DOCAtwo - Distance of Closest Approach between `p1` and `p2`.
- DOCAthree - Distance of Closest Approach between `p0` and `p2`.
- IP_p0p2 - Impact parameter of the `p0` and `p2` pair.
- IP_p1p2 - Impact parameter of the `p1` and `p2` pair.
- isolationa - Track isolation variable.
- isolationb - Track isolation variable.
- isolationc - Track isolation variable.
- isolationd - Track isolation variable.
- isolatione - Track isolation variable.
- isolationf - Track isolation variable.
- iso - Track isolation variable.
- CDF1 - Cone isolation variable.
- CDF2 - Cone isolation variable.
- CDF3 - Cone isolation variable.
- production - source of τ, Explained in detail in Tab. 1. This variable is absent in the test samples.
- ISO_SumBDT - Track isolation variable.
- p0_IsoBDT - Track isolation variable.
- p1_IsoBDT - Track isolation variable.
- p2_IsoBDT - Track isolation variable.
- p0_track_Chi2Dof - Quality of `p0` muon track.
- p1_track_Chi2Dof - Quality of `p1` muon track.
- p2_track_Chi2Dof - Quality of `p2` muon track.
- min_ANNmuon - Muon identification. LHCb collaboration trains Artificial Neural Networks (ANN) from informations from RICH, ECAL, HCAL, Muon system to distinguish muons from other particles. This variables denotes the minimum of the three muons ANN. min_ANNmuon should not be used for training. This variable is absent in the test samples.

- p0_pt - Transverse momentum of `p0` muon.
- p0_p - Momentum of `p0` muon.
- p0_eta - Pseudorapidity of `p0` muon.
- p0_IP - Impact parameter of `p0` muon.
- p0_IPSig - Impact Parameter Significance of `p0` muon.
- p1_pt - Transverse momentum of `p1` muon.
- p1_p - Momentum of `p1` muon.
- p1_eta - Pseudorapidity of `p1` muon.
- p1_IP - Impact parameter of `p1` muon.
- p1_IPSig - Impact Parameter Significance of `p1` muon.
- p2_pt - Transverse momentum of `p2` muon.
- p2_p - Momentum of `p2` muon.
- p2_eta - Pseudorapidity of `p2` muon.
- p2_IP - Impact parameter of `p2` muon.
- p2_IPSig - Impact Parameter Significance of `p2` muon.
- SPDhits- Number of hits in the SPD detector.

Further improvements of the metric regarding classifier uncertainties are postponed until further competition development.

References

[1] A. B. Balantekin and W. C. Haxton. Neutrino oscillations. *Prog.Part.Nucl.Phys.*, 71:150–161, 2013.

[2] Roel Aaij et al. Search for the lepton flavour violating decay $\tau^- \to \mu^-\mu^+\mu^-$. *JHEP*, 1502:121, 2015.

[3] `http://home.web.cern.ch/topics/large-hadron-collider`.

[4] `http://home.web.cern.ch/about/experiments/atlas`.

[5] `http://home.web.cern.ch/about/experiments/cms`.

[6] `http://home.web.cern.ch/about/experiments/lhcb`.

[7] `http://home.web.cern.ch/about/experiments/alice`.

[8] `http://home.web.cern.ch/about/experiments/totem`.

[9] `home.web.cern.ch/about/experiments/lhcf`.

[10] `home.web.cern.ch/about/experiments/moedal`.

[11] H. Cramér. On the composition of elementary errors. *Scandinavian Actuarial Journal*, 1928.

[12] A L Read. Modified frequentist analysis of search results (the CL_s method). (CERN-OPEN-2000-205), 2000.

[13] Y. Amhis et al. Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014. 2014.

[14] `https://twiki.cern.ch/twiki/bin/view/LHCb/VertexFitters`.

[15] M. Chrzaszcz, T. Lesiak, and A. Lusiani. *Searches for charged Lepton Flavour Violation*. Scholars' Press, 2015.