Universität Zürich[UZH]

July 28, 2018

Particle track reconstruction using a recurrent neural network at the $\mu \to eee$ experiment

Bachelor thesis of
Sascha Liechti

Supervised by
Prof. Nicola Serra
Dr. Patrick Owen

Abstract

During the $\mu - 3e$ experiment we faced the challenge of reconstructing the paths of certain low momentum particles that curled back into the detector and cause additional hits. To face this, a recurrent neural network was used which found the right track for 87% of these particles.

Contents

1 Standard Model **1**
1.1 Elementary particles and forces 1
1.2 Interaction rules 3

2 LHCb Experiment **5**
2.1 LHC 5
2.2 Detector 5
2.2.1 Vertex locator 6
2.2.2 Tracking system 6
2.2.3 RICH 6
2.2.4 Calorimeter 6
2.2.5 Muon system 7
2.3 Trigger 7
2.4 Software 7
2.4.1 Track reconstruction and fit 7

A Appendix **10**
A.1 Preselection 10
A.2 Reweighting 12
A.3 Selection 13

References **14**

1 Standard Model

1.1 Elementary particles and forces

The Standard Model(SM) describes all known elementary particles as well as three of the four known forces[1].
The elementary particles that make up matter can be split into two categories, namely quarks and leptons. There are 6 types of quarks and six types of leptons. The type of a particle is conventionally called flavour. The six quark flavours and the six lepton flavours are separated over 3 generations (each which two quarks and two leptons in it). Experimental evidence suggests that there exist exactly three generations of particles. Each particle of the first generation has higher energy versions of itself with the same characteristics (e.g. $e^- \rightarrow \mu^- \rightarrow \tau^-$)as in other generations. Contrary, each following generation has a higher mass than the generation before.

Table 1: Quarks in the Standard Model

Quarks				
	Particle		Q[e]	$\frac{mass}{GeV}$
1. Gen.	up	u	$-\frac{1}{3}$	0.003
	down	d	$\frac{2}{3}$	0.005
2. Gen.	strange	s	$-\frac{1}{3}$	0.1
	charm	c	$\frac{2}{3}$	1.3
3. Gen.	bottom	b	$-\frac{1}{3}$	4.5
	top	t	$\frac{2}{3}$	174

One category consists of quarks(q)(see Table 1). In this, we differentiate between up-type quarks, with charge $-\frac{1}{3}e$, and down-type, quarks with charge $\frac{2}{3}e$. Quarks interact with all fundamental forces.
Each quark carries a property called colour-charge. The possible color charges are red(r), green(gr), blue(bl) in which anti-quarks carry anti-colour. Quarks can only carry one colour, whilst every free particle has to be colorless[2]. In conclusion we cannot observe a single quark.
Free particles can achieve being colourless in two ways. Either by having all three colors present in the same amount (one quark of each color), which creates the characteristic group of baryons(qqq) and anti-baryons($\bar{q}\bar{q}\bar{q}$) or by having a color and its anticolor present, which creates the group of mesons($q\bar{q}$).

[1] Strong, weak and electromagnetic forces
[2] Colour confinement

Table 2: Leptons in the standard model

Leptons				
	Particle		Q[e]	$\frac{mass}{GeV}$
1. Gen.	electron	e^-	-1	0.005
	neutrino	ν_e	0	$< 10^{-9}$
2. Gen.	muon	μ^-	-1	0.106
	neutrino	ν_μ	0	$< 10^{-9}$
3. Gen.	tau	τ^-	-1	1.78
	neutrino	ν_τ	0	$< 10^{-9}$

The other group consists of leptons(l)(see Table 2). They only interact through the weak and the electromagnetic force. Each generation consists of a lepton of charge -1 and a corresponding EM neutrally charged neutrino. The electron has the lowest energy of all charged leptons. This makes the electron stable while the higher generation particles decay to lower energy particles.

The leptons of one generation, namely the charged lepton and its corresponding neutrino are called a lepton family. A lepton of a family counts as 1 to its corresponding lepton family number whilst a anti-lepton counts as -1.

Table 3: Fundamental forces

Force	Strengh	Boson		Spin	Charge	$\frac{mass}{GeV}$
Strong	1	gluon	g	1	0	0
Electromagnetism	10^{-3}	photon	γ	1	0	0
Weak	10^{-8}	Z boson	Z	1	0	80.4
	10^{-8}	W boson	$W^\pm$	1	± 1	91.2

The particles of the SM interact through the 3 fundamental forces of the SM. In these interactions, particles called bosons are being exchanged which are the carriers of their respective force (see Table 3).
As mentioned above, only quarks can interact through the strong force, in which they exchange gluons. Gluons are massless and EM neutrally charged. The strong force has the biggest coupling strengh of 1 (though it decreases with higher energies as a result of gluon-gluon self interaction loops, which interfere negatively in perturbation theory)[3]. A gluon carries colour charge and hence can change the colour of a quark but it conserves its flavour. The strong interaction has an underlying gauge symmetry of SU(3). Therefore, it can be derived that color charge is conserved through the strong interaction[4].
The electromagnetic(EM) force is propagated through the photon. It carries zero charge and no invariant mass. Exclusively charged particles can interact through the electromagnetic force. The coupling strength is $\alpha \approx \frac{1}{137}$, contrary to the strong force the

[3]Mark Thomson - Modern Particle physics - 10.5.2
[4]E.g. through Gell-Mann matrices

coupling constant increases with higher energies[5]. This difference stems from the fact that photon-photon interaction loops are not allowed whereas gluon-gluon interaction loops are. In perturbation theory this results in only positive terms being added to the coupling strength. The underlying gauge symmetry is of SU(1). The electromagnetic force also conserves flavour.
The weak force has two types of bosons. The bosons of the weak force are the only bosons to have an inertial mass.
First we will discuss the EM neutrally charged Z boson[6]. Even though the Z boson belongs to the weak force it, it also has an electromagnetic part additionally to the weak force part[7]. It follows directly, that the Z boson couples weaker to uncharged particles.
The other boson of the weak force is the W boson[8]. In the classical SM, the only way particles can change flavour is through the weak force by emitting or absorbing W boson. It is important to notice that, besides of having an invariant mass, the W boson is the only boson with a non zero charge ($Q_{W^\pm} = \pm 1e$). In the gauge symmetry of the weak force the $W^\pm$ are actually the creation and annihilation operators of said symmetry[9].
An important characteristic of the weak force is that it exclusively couples to left-handed(LH) particles and righthanded(RH) antiparticles (describing chirality states)[10]. The chirality operators for left- and righthandedness are:

LH: $\frac{1}{2}(1-\gamma^5)$, RH: $\frac{1}{2}(1+\gamma^5)$

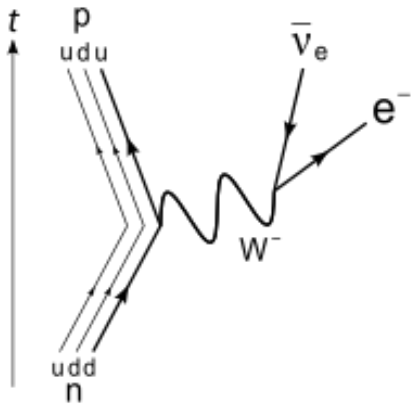
As a consequence RH particles and LH anti-particles cant couple to the W boson at all. This also results in charged RH particles and LH anti-particles to couple to the Z boson only through the electromagnetic part of the itself, while uncharged RH particles and LH anti particles (e.g. RH ν, LH $\bar{\nu}$) don't couple with the EM force nor the weak force.

1.2 Interaction rules

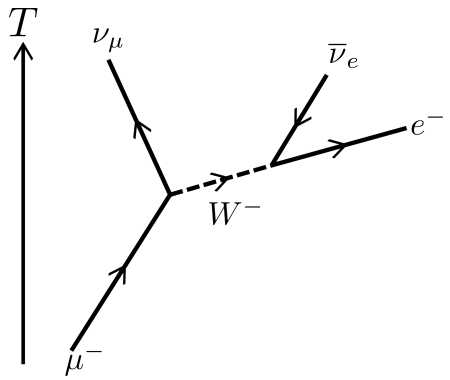
Now we will establish the general rules for interactions in the SM.

Baryon number is conserved
As we already established before, the only interaction that can change flavour is the weak force through the W boson. We directly see that all other interactions baryon number has to be conserved. So any up-type quark can be changed to a down-type quark and backwards by emitting or absorbing a W boson. In the end however, there are still 3 quarks which form a baryon[11], even though it changed its type and charge. A well known example is the beta decay, where a down quark in a neutron decays into a an up quark to form now a proton(e.g. see Figure 1a). We easily see that the baryon number is conserved.

[5]Mark Thomson - Modern Particle physics - 10.5.1
[6]Discovered at Super Proton Synchrotron accelerator - Cern - 1983
[7]$Z \rightarrow EM_{part} + W^3$, Modern Particle Physics - Blababla
[8]Discovered at Super Proton Synchrotron accelerator - Cern - 1983
[9]$W^\pm = W_1 \pm iW_2$
[10]In the ultrarelativistic limit helicity and chirality eigenstates are the same
[11]I exclude $q\bar{q}$ pair-antipair production to form a pentaquark($qqqq\bar{q}$) and other exotic states



(a) Feynman diagram of the β-decay



(b) Feynman diagram of a μ-decay

Lepton family number is conserved

According to the SM lepton family number is conserved. As all interactions beside the W conserve particle flavour, it is easy to see that lepton family number is conserved. Whenever a lepton interaction with a W boson, it just changes a lepton to its corresponding lepton neutrino and or the other way around (e.g. see Figure 1b).

2 LHCb Experiment

2.1 LHC

The Large Hadron Collider (LHC) is a proton-proton synchrotron situated nearly 200 m below the surface in a tunnel. The ring of superconducting magnets has a total length of 27 km containing two beam pipes filled with protons that are brought to collision at several points. Those collisions occurred at a total centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ in 2011 and 8 TeV in 2012. After an upgrade, the energy has been increased to the centre-of-mass energy of 13 TeV in 2015 and 2016. The proton beams interact simultaneously in four detector points in the LHC ring which experiments built around, ATLAS, CMS, LHCb and ALICE. Two of them, CMS and ATLAS, are more general-purpose experiments with a toroidal structure covering the whole space around the interaction point and operating at the full collision rate. With another goal in mind, there is also ALICE, an experiment designed to study gluon-plasma and high-density events. For a fraction of the running time, the LHC is filled with lead-ions in order to create lead-proton or lead-lead interactions.

2.2 Detector

The Large Hadron Collider beauty (LHCb) is one out of four experiments situated at the LHC at CERN [6]. The LHCb is designed to perform high-precision measurements of particles containing b and c quarks to study rare decays and CP violation. In contrast to the other experiments located at the LHC, the LHCb is a single-arm forward spectrometer. This allows for measurements in the region of the pseudorapidity range $2 < \eta < 5$, the predominant flight direction of $b\bar{b}$-production.

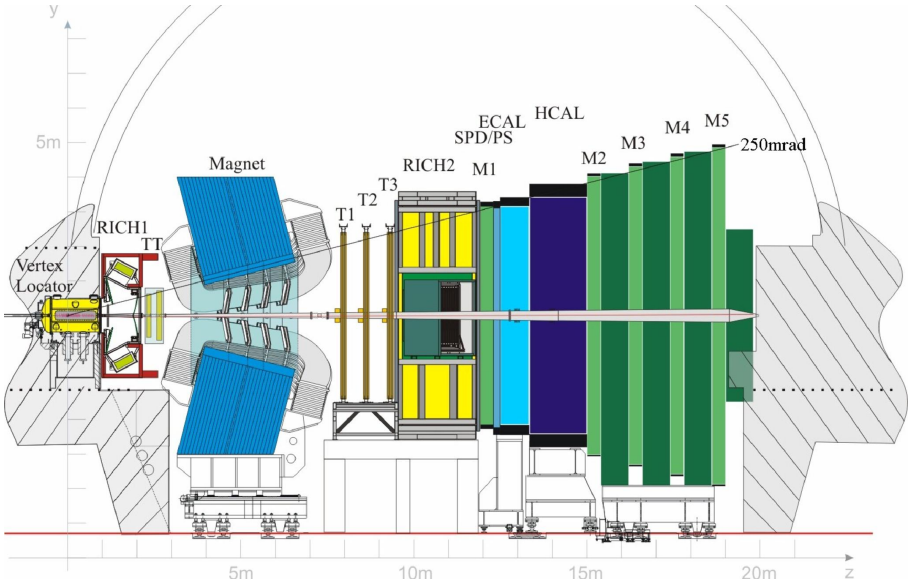


Figure 2: A schematic view of the non-bending plane of the LHCb detector. Particles are produced in the collision point on the left side inside the vertex locator and are bent by the magnet afterwards.

2.2.1 Vertex locator

An important aspect of the b and c physics is the tracking of the particles close to the interaction point in order to precisely determine the primary and secondary vertices of heavy mesons. At the LHCb this is achieved with the vertex locator (VELO), a series of modules with sensors made up of lightweight, radiation-hard silicon-strips. Each sensor is able to either measure the azimuthal coordinate or the radial distance to the beam axis, a single module contains both complementary sensors. The tracker is located about 8 mm from the aligned beam and is placed inside a beam-pipe independent vacuum system.

2.2.2 Tracking system

In addition to the VELO, several other tracking stations measure the tracks and bending of the particles. In front of the 4 Tm dipole magnets, the Tracker Turicensis is installed. It consists of four layers of silicon-strip detectors and allows for the detection of low-momenta particles which will be bent away in the magnetic field.

After the dipole magnet, the three tracking stations T1, T2 and T3 are placed. Each of them consists of an inner tracker situated close to the beam pipe and an outer tracker, covering the largest area of the tracker plane. The inner tracker is a silicon-strip detector covering the area with a high density of tracks. Another detector technique is used in the outer tracker as it covers a greater area without the need for the same precision as required in the inner tracker. In this case four layers of straw tubes filled with gas are used as drift chambers.

2.2.3 RICH

In b physics it is important to have a good discrimination between charged particles, *e.g.* K and π. In order to achieve a good particle identification, there is a ring imaging Cherenkov detector (RICH) on each side of the magnet, which measure the Cherenkov emission angle θ_c. The Cherenkov radiation is detected by pixel hybrid photon detectors. As the angle of the radiation relative to the particles flight direction depends on the velocity of the passing particle only, using additionally the information about the momentum from the tracker allows to determine the mass of the particle and therefore its identity. The Cherenkov angle also depends on the materials refractive index the charged particle is passing through. In order to cover a large momentum range with a good angle resolution, the RICH detectors are filled with materials of different refractive indices.

2.2.4 Calorimeter

A calorimeter measures the total as well as the differential energy loss by completely absorbing it through interactions with the material. For the LHCb, a classical architecture of an electromagnetic calorimeter (ECAL) in front of a hadronic calorimeter (HCAL) was chosen. Both are optimized for particle identification, mostly for e/π and π^0/γ discrimination, as well as for a fast readout. The information will be used, among others, in the first trigger stage (see Sec. 2.3).

In front of the ECAL, a scintillator pad detector is placed to detect the pass-through of charged particles followed by a pre-shower detector. The ECAL itself is built of a sampling scintillator/lead structure (shashlik technology) and has a total depth of $25X_0$. As the hit density rapidly drops with increasing distance from the beam pipe, the ECAL is split into three different sections with appropriate cell sizes.

The HCAL of the LHCb is a sampling calorimeter with a special structure. It consists of lead/scintillator tiles directed *parallel* to the beam-pipe. Each thin row consisting of several tiles has a neighbour-row with inverted lead/scintillator tiles. The scintillation light is detected by photomultiplier tubes and collected by fibres. The total length equals to 5.6 hadronic interaction lengths.

2.2.5 Muon system

The muon system is responsible for the identification of muons and provides a standalone, fast signal to the trigger in case of muons with high transverse momentum (p_{T}) passing it. The whole system is composed of the five stations M1–M5. M1 is placed in front of the calorimeters to improve the p_{T} resolution whereas the others are located downstream. The stations are separated by iron absorbers to prevent any non-muons from passing through the detectors. All systems provide spatial resolved hit information with decreasing segmentation scale for increasing distance to the beam pipe. M4 and M5 are mostly used for penetration testing and offer only sparsely location informations.

2.3 Trigger

At the nominal LHC conditions, the bunch crossing frequency can reach up to 40 MHz which leaves 25 ns in between two crossings. This high frequency has to be reduced down to 1 kHz in order to be able to store the data for offline analysis. Two trigger-systems, a low-level trigger (L0) and a high-level trigger (HLT) consisting of two stages, HLT1 and HLT2 select which events to keep.

The L0 stage is a hardware implemented trigger and consists of a custom electronics set-up built with FPGA. It takes information from three different sources into account. The first is a pile-up system inside the VELO, estimating the number of events that occurred during the collision. Information from the calorimeter is used to estimates the transverse energy (E_{T}) of certain particles and decides to keep the event in case of high E_{T}. The muon system feeds the trigger with information about the p_{T} of muons in order to trigger on a certain threshold.

The next stage is the HLT1. It reconstructs some parts of the tracks to confirm the L0 decision as well as to further reduce the event rate. This is now low enough to allow the HLT2 to reconstruct b events and make more refined decisions. The events which pass HLT2 with a frequency of around 1 kHz are then stored for offline analysis.

A general distinction is made on whether an event passed the trigger because of the events signature itself (trigger on signal, `TOS`) or because of some other particles signature (trigger independent of signal, `TIS`).

2.4 Software

Once the events are stored, offline tools are used to reconstruct and fit tracks and apply sets of exclusive selections prior to the data manipulation.

2.4.1 Track reconstruction and fit

For the event reconstruction, information from the tracking system (including the VELO) is used. First of all, a clustering algorithm determines *track seeds* by searching for candidates in a low magnetic field region of the spectrometer. A Kalman filter algorithm is then

fitted to the data using the track seeds as initialisation. An advantage of reconstructing and fitting with this algorithm is that the result is equivalent to a least square fit of the tracks to the hits. For the particle propagation with the Kalman filter, the inhomogeneous magnetic field as well as multiple scattering occurring from detector material is taken into account.

Acknowledgements

I would like to to thank the Physics Department of the University of Zurich, the CERN collaboration and especially the LHCb collaboration for providing such an impressive infrastructure of computing power, accelerators and detectors.

I would like to express my gratitude to Prof. Nicola Serra from the University of Zurich for letting me do my bachelor thesis in his research group and who managed to give me a very positive impression on the world of data analysis. I like to especially thank my supervisor Dr. Rafael Silva Coutinho for the support and explanations for all kind of problems and questions I had as well as for the encouragement if things did not go as well as expected. I also like to thank Dr. Albert Piug for his support, advices and discussions. I further want to thank my colleague Alexander Daetwyler for explanations and discussions as well as for providing me several, useful code snippets.

A Appendix

A.1 Preselection

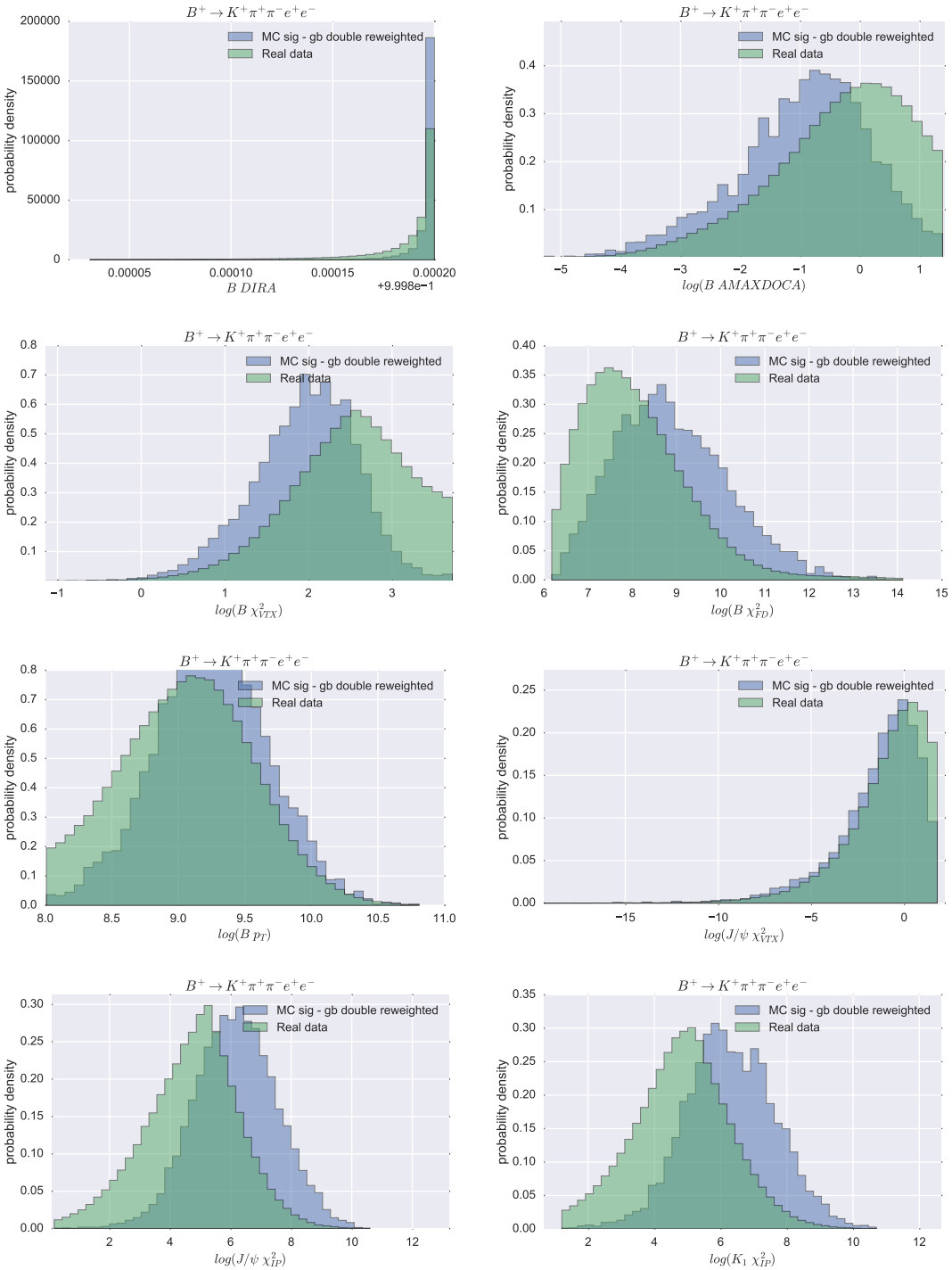


Figure 3: Variables used in the preselection as described in Sect. **??**

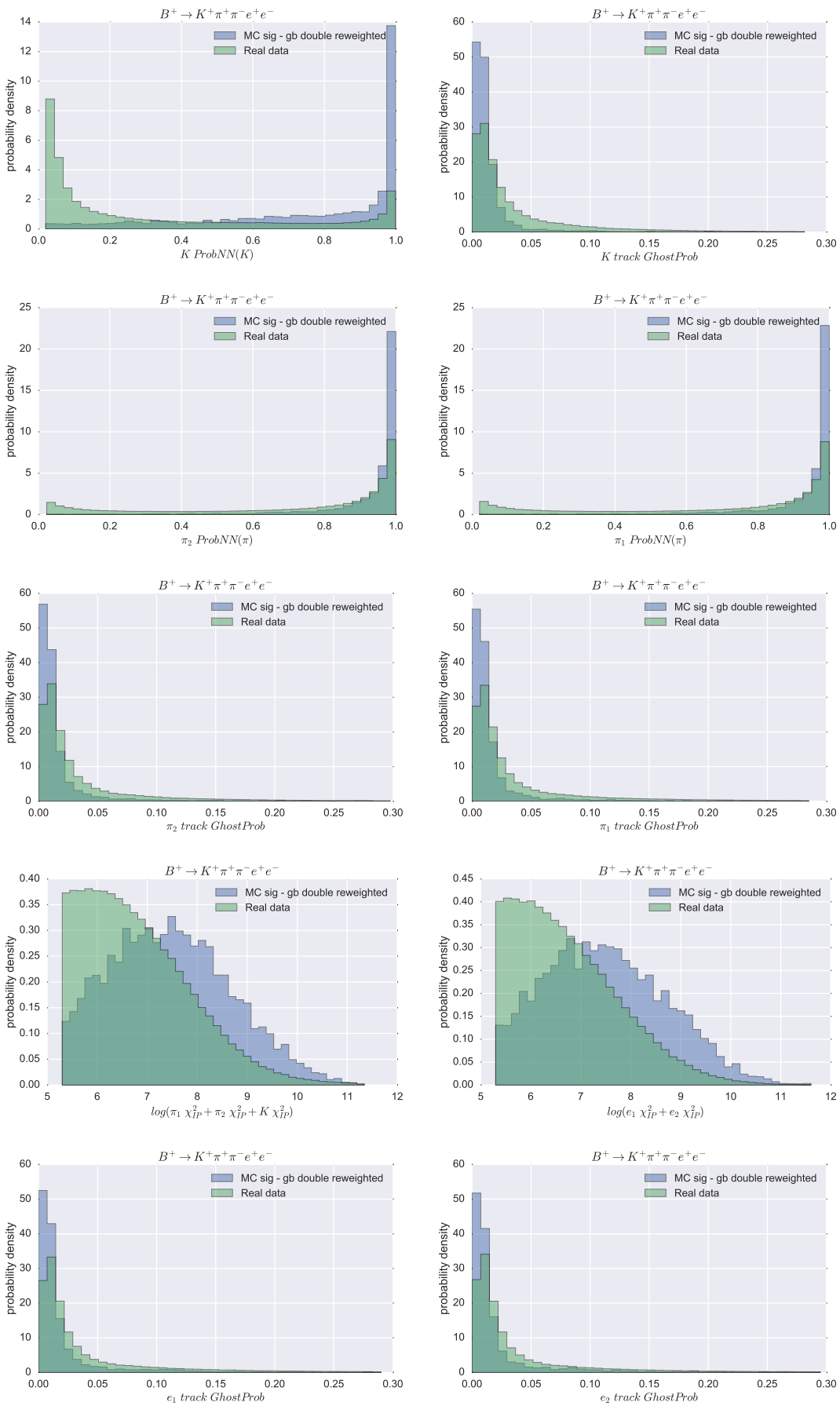


Figure 3: Variables used in the preselection as described in Sect. **??**

A.2 Reweighting

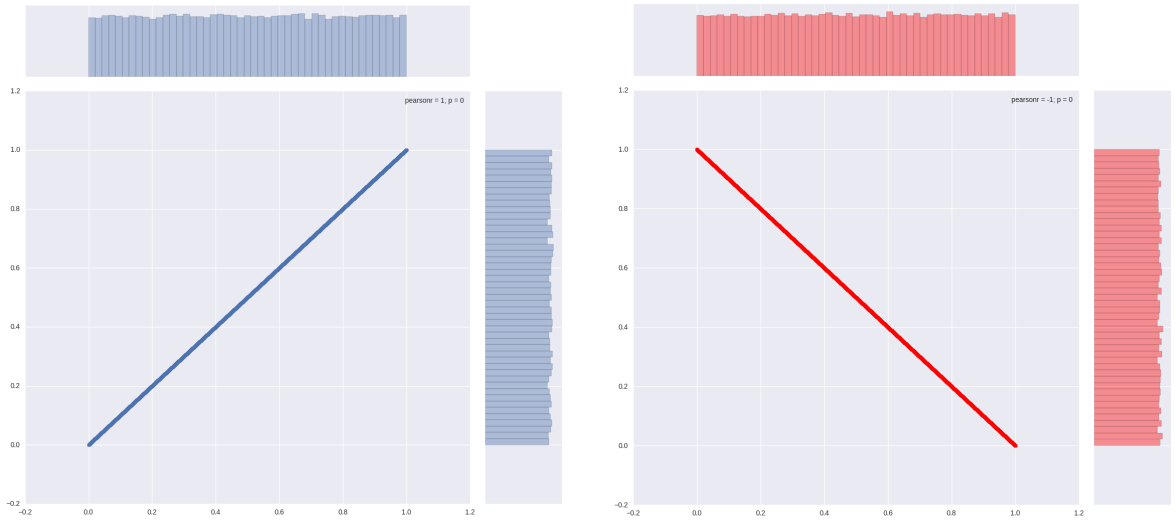


Figure 4: A two dimensional distribution and its projections. Even tough the distributions can be easily discriminated by looking at their higher order – second order here – correlations, there projections do not reveal that.

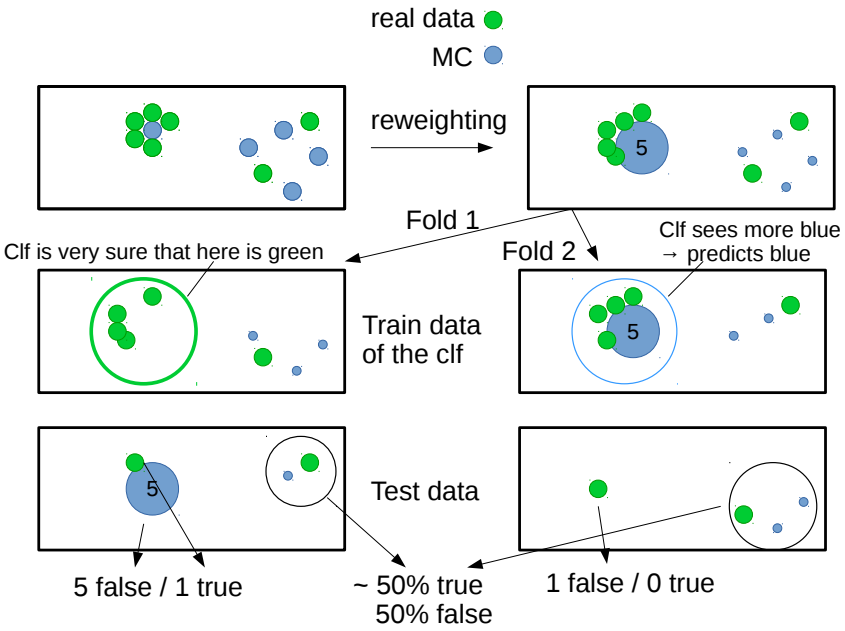


Figure 5: ROC AUC bias with weights visualized. The reweighter works quite well for this example and assigns a weight of 5 to the single blue point. Then the data is split in two different ways (Fold 1 and 2) into training and test data in order to compare two possible outcomes. The total outcome can be thought as an average of both cases.

A.3 Selection

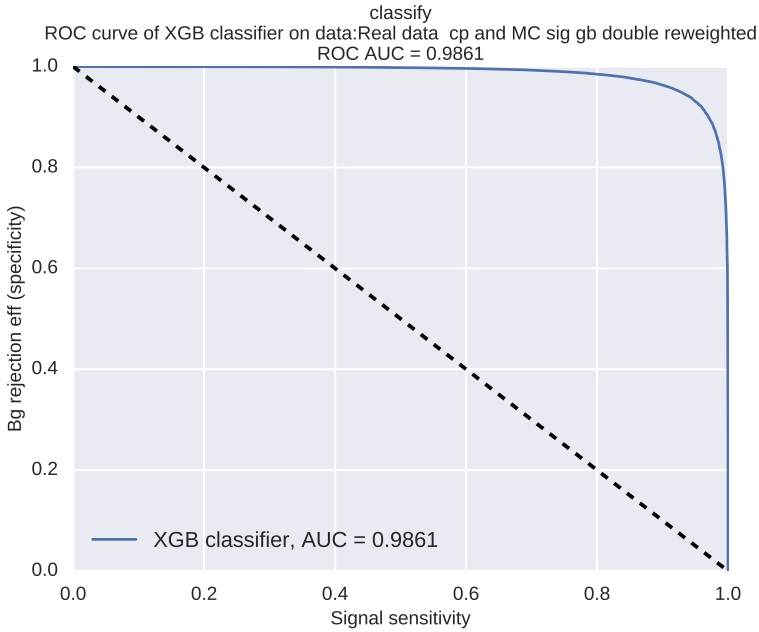


Figure 6: ROC curve of the XGB trained on the MC against the right side band (B mass vertex constrained $> 5600\,\text{MeV}$) of $B^+ \to K^+\pi^+\pi^-e^+e^-$.

References

[1] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Phys. Rev. D **2** (1970) 1285.

[2] LHCb, R. Aaij *et al.*, *Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays*, Phys. Rev. Lett. **113** (2014) 151601, arXiv:1406.6482.

[3] LHCb, R. Aaij *et al.*, *Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays*, arXiv:1705.05802.

[4] LHCb, R. Aaij *et al.*, *First observations of the rare decays $B^+ \to K^+\pi^+\pi^-\mu^+\mu^-$ and $B^+ \to \phi K^+\mu^+\mu^-$*, JHEP **10** (2014) 064, arXiv:1408.1137.

[5] H. Hatanaka and K.-C. Yang, $K_1(1270) - K_1(1400)$, Phys. Rev. D **78** (2008) 074007.

[6] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[7] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.

[8] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[9] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[10] A. P. Bradley, *The use of the area under the roc curve in the evaluation of machine learning algorithms*, Pattern Recogn. **30** (1997) 1145.

[11] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[12] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, J. Comput. Syst. Sci. **55** (1997) 119.

[13] T. Chen and C. Guestrin, *Xgboost: A scalable tree boosting system*, CoRR **abs/1603.02754** (2016).

[14] G. Punzi, *Sensitivity of searches for new signals and its optimization*, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology* (L. Lyons, R. Mount, and R. Reitmeyer, eds.), p. 79, 2003. arXiv:physics/0308063.

[15] LHCb, R. Aaij *et al.*, *Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions*, Phys. Rev. Lett. **118** (2017), no. 5 052002, arXiv:1612.05140.

[16] LHCb collaboration, R. Aaij *et al.*, *Search for the $\Lambda_b^0 \to \Lambda\eta$ and $\Lambda_b^0 \to \Lambda\eta'$ decays with the LHCb detector*, JHEP **09** (2015) 006, arXiv:1505.03295.

[17] LHCb collaboration, R. Aaij *et al.*, *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics*, JHEP **04** (2013) 001, arXiv:1301.5286.

[18] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.