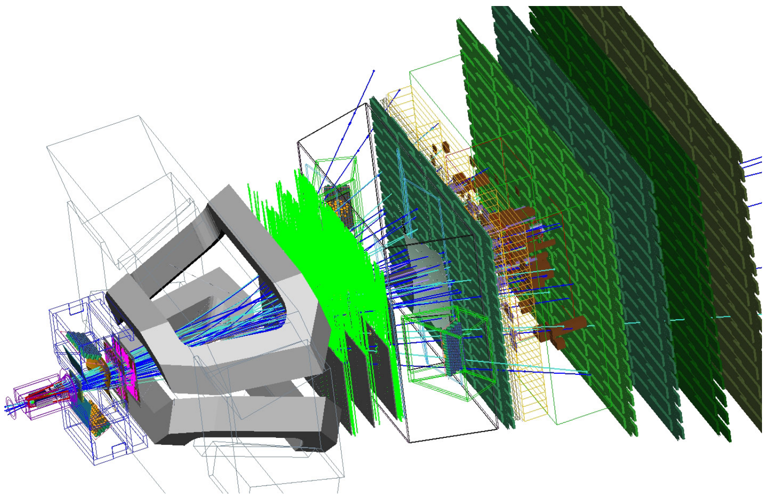
March, 2015

THESIS

SEARCH FOR LONG-LIVED EXOTIC PARTICLES DECAYING TO MUON PAIRS

A STUDY PERFORMED WITH 3 fb^{-1} DATA COLLECTED AT THE LHCb EXPERIMENT IN 2011 AND 2012



E.M. Löbker

Faculty of Science and Technology
Chair of Energy, Materials and Systems
NIKHEF National Institute for Subatomic Physics

Exam committee:
Prof. dr. ir. H.H.J. ten Kate
Dr. W.D. Hulsbergen
Dr. ir. H. Wormeester

Nikhef

LHCb

UNIVERSITY OF TWENTE.

Contents

1 Introduction

The Standard Model (SM) of particle physics has been very successful at predicting and explaining a wide range of experimental results and is the widely accepted theory of interactions and elementary particles [1, 2]. However, some phenomena remain unexplained by it, like the inconsistencies between the gravitational effects of astronomical objects and their mass from observable matter. This indicates the presence of ‘dark matter’ [3]. Gravity is not included in the SM and the dominance over matter over antimatter in the universe cannot be explained by it.

Therefore, searches are performed for Beyond the Standard Model (BSM) physics, based on different theoretical models. Several of those predict the existence of massive long-lived particles that can decay to a muon pair [4–6], as illustrated in Figure 1. Because of the long lifetime, the muons originate from a displaced vertex, which provides a clear signature for a search. The SM background is easily suppressed here, which will become clear in this analysis.

This topology thus has potential for finding new physics and this analysis will search for this signal using data from the LHCb detector at CERN. The Hidden Valley theory is used as a benchmark model, which predicts the existence of neutral π_v^0 (v-pion) particles with long lifetimes, that can decay to SM muons. [6, 7].

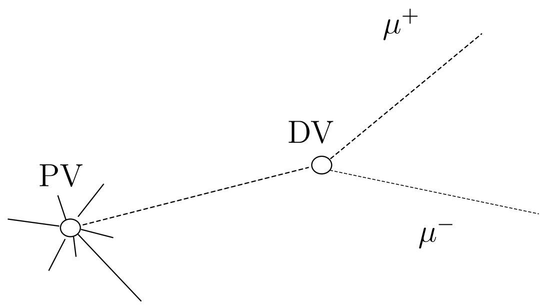


Figure 1: Signal topology, a long-lived particle originating in the primary vertex (PV) and decaying to a muon pair in the displaced vertex (DV).

1.1 Research objective and scope

In this analysis, a search is performed for long-lived exotic (BSM) particles decaying to oppositely charged muons in data collected with the LHCb detector in 2011 and 2012. The search is confined to particles with a mass ranging from 7 GeV/c^2 to 50 GeV/c^2 and with lifetimes ranging from 10 to 100 ps. These values are motivated by detector constraints. As a benchmark model, the Hidden Valley π_v^0 decay is used.

When no evidence is found of BSM physics, a limit will be set on the likelihood of a proton proton collision resulting in a $\pi_v^0 \to \mu^+\mu^-$ decay as a function of the π_v^0 mass. This is represented by the quantity $\sigma(pp \to H) \times \mathcal{B}(H \to \pi_v^0\pi_v^0) \times \mathcal{B}(\pi_v^0 \to \mu^+\mu^-)$. It is assumed that the branching fraction $\mathcal{B}\ (\pi_v^0 \to \mu^+\mu^-) = 1$, i.e. π_v^0 particles always decay to a muon pair.

1.2 Analysis overview

In the first two chapters, a theoretical motivation is presented (2) and the LHCb detector and its relevant features are described (3).

To give direction, signal events are simulated. Events are collisions of the proton beams in the LHCb detector. The formation of a Higgs in the collision is simulated, which decays to two π_v^0 particles, each of which decay to two muons. Other possible processes that result in a muon pair are identified, these background processes are also simulated. In Chapter 4, details of these samples are given.

Data from the detector is ordered according to certain properties. Suitable data has to be selected and a proper preselection needs to be applied. The result is data with 'high-quality' muon pairs. This is described in Chapter 5.

In Chapter 6, properties of the signal and background decays are compared. Based on their differences, selection criteria are defined that keep data with signal properties and discard of data that most likely contains background decays.

The data after the selection gives an indication whether or not the π_v^0 signal is found. When this is not the case, a statistical method is used to set a limit on the occurrence of a signal decay as explained in Chapter 8. The number of expected signal decays μ_{Sig} after the selection, can be calculated by Equation 1:

$$\mu_{\mathrm{Sig}} = L \times \sigma(pp \to H) \times \mathcal{B}(H \to \pi_v^0\pi_v^0) \times \mathcal{B}(\pi_v^0 \to \mu^+\mu^-) \times \epsilon. \tag{1}$$

Here, L is the total integrated luminosity in fb^{-1} which represents the total number of pp collisions, and the $\sigma \times \mathcal{B} \times \mathcal{B}$ quantity represents the likelihood of a pp collision resulting in a signal decay. ϵ is the efficiency of the selection on π_v^0 decays, the fraction of π_v^0 decays that survive the selection. Based on the number of muon pairs after the selection, a limit is set on μ_{Sig}, which directly translates to a limit on $\sigma(pp \to H) \times \mathcal{B}(H \to \pi_v^0\pi_v^0)$ by using the known L and the ϵ calculated on π_v^0 simulation samples.

Systematic errors are discussed in Chapter 7.

Lastly, in Chapter 9 the conclusion and discussion are presented and an outlook on future research is given.

2 Theoretical motivation

In this Chapter, the Standard Model (SM) and its limitations are briefly introduced, which are the motivation for Beyond the Standard Model (BSM) theories and searches. The Hidden Valley (HV) model used in this search is described and the search by the CMS experiment for this signal is discussed.

2.1 Standard Model and limitations

The SM of particle physics [1, 2] is a well-established theory that describes the dynamics of all known subatomic particles and three of the four natural forces. It comprises three generations of fermions (both leptons and quarks), four mediators and the scalar Higgs boson. The lepton generations are the electron, muon and tauon and their associated neutrinos. The quarks carry colour and cannot exist on their own right. This colour confinement makes them combine to form colour neutral hadrons. Either a quark and anti-quark (carrying anti-colour) combine to form a meson, or three quarks form a hadron. The natural forces described by the SM are the strong, electromagnetic and weak forces. The are mediated by gluons, photons and W and Z bosons respectively.

Although the SM has been very successful at explaining many experimental results, several phenomena remain unexplained by it. The most obvious is the omission of the fourth fundamental force, gravity. No model is yet able to unify general relativity and quantum field theory. Next to that, inconsistencies are observed between the gravitational effects of astronomical objects and their mass from observable matter [3]. This could be explained by the existence of ‘dark matter’, a not yet observed new type of matter with at least one stable, neutral particle. ‘Dark energy’ theories try to explain the observed accelerating expansion of the universe and together with dark matter, it is hypothesised to make up 95.1% of our universe [8]. It is therefore a subject of high interest in current research.

2.2 Beyond the Standard Model physics

The limitations of the SM give rise to numeral Beyond the SM (BSM) theories. They predict a wide range of phenomena that are often hard to verify by experiments. Therefore, it is a challenge to design searches that have potential to discover new physics. Approaches often contain distinct topologies in order to make them more likely to be detected.

One of those distinct topologies is a long-lived exotic particle that creates a displaced vertex. Several models predict these particles, among others super symmetry (SUSY) theories [4, 5], dark matter theories and Hidden Valley models [6, 9]. A HV model is used in this search and this class of models is further described here.

2.2.1 Hidden Valley models

The name ‘Hidden Valley’ refers models that describe the presence of a low mass ‘valley’ that couples weakly to the SM, as illustrated in Figure 2. It introduces new particles that are ‘v-charged’ and are neutral under the SM. This hidden sector may communicate with the SM by a heavy mediator at the TeV scale, which could explain why it has not been discovered in previous experiments. Efforts have been made to describe a signature of HV that could be detected in current experiments [6, 10].

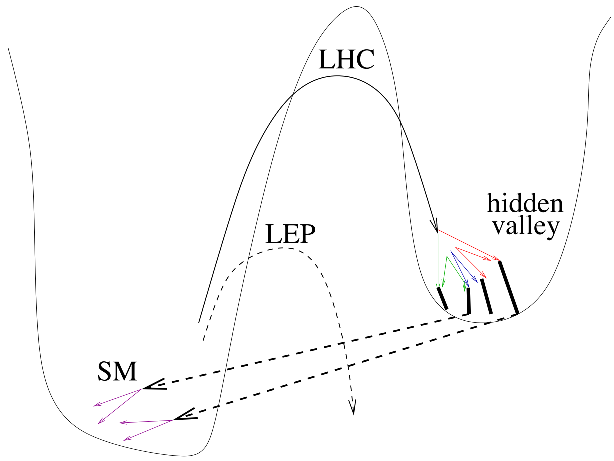


Figure 2: Schematic view of production and decay of v-hadrons. While LEP was unable to penetrate the barrier separating the sectors, LHC may easily produce v-particles. These form v-hadrons, some of which decay to SM particles. From [11].

A possibly detectable scenario is described by making a simple addition to the SM. In this theory, the v-particles combine to form v-hadrons, which are predicted to decay quickly to v-nucleons and v-pions, the lightest v-mesons. The π_v^+ and π_v^- are v-charged and cannot decay to SM particles. These stable particles are thus dark matter candidates. The π_v^0 however is v-charge neutral and can decay to SM fermion pairs via the heavy Z boson. It is therefore long-lived. Multiple production processes are proposed. Two SM quarks formed in a collision can combine to form a virtual Z, which can decay to v-quarks. They are expected to decay to multiple v-pions. Another possible production process is through the resonant decay of a Higgs boson.

2.3 Model for this search

The benchmark model for this search is the described HV model with the Higgs production process. The muon decay channel is chosen as it provides a clear signature for a search. As a second generation lepton, the muon is more massive than the electron. It emits less bremsstrahlung in matter and therefore penetrates far deeper into material.

The HV model is only used to the extend that the exotic particle is required to be long-lived, massive, spinless and decaying to two muons. The search is therefore also sensitive to other BSM theories.

2.4 Previous experiments

A search for long-lived exotic particles to a muon pair has previously been carried out by the CMS experiment [12, 13]. The search in this analysis is complementary, as the CMS detector is sensitive in a different area. They cover exotic particle lifetimes ranging from 0.3 to 3000 ps and masses from 20 to 350 GeV/c^2. This analysis thus reaches lower masses.

The latest CMS result is displayed in Figure 3. No evidence for BSM physics was found and a limit was set on the production rate as a function of the lifetime in cm, where 10^{-2} cm corresponds to 0.3 ps. The CMS detector has a higher luminosity than LHCb, which makes the amount of data larger. Thus more stringent choices are made in which data is stored and lower mass data is discarded more than in LHCb. It is thus challenging for CMS to search for masses lower than 20 GeV/c^2.

The CMS detector is a hermetic detector, it covers the area around the interaction point as fully as possible. This enables the search to extend to higher lifetimes than in LHCb. It is shown in this analysis that much higher lifetimes than 100 ps can not efficiently be detected in the forward single-armed LHCb detector.

It has been proven that LHCb is at least competitive with CMS in detecting the $B_s^0 \to \mu^+\mu^-$ decay [14]. This analysis can show if LHCb is competitive with the shown result for low lifetimes.

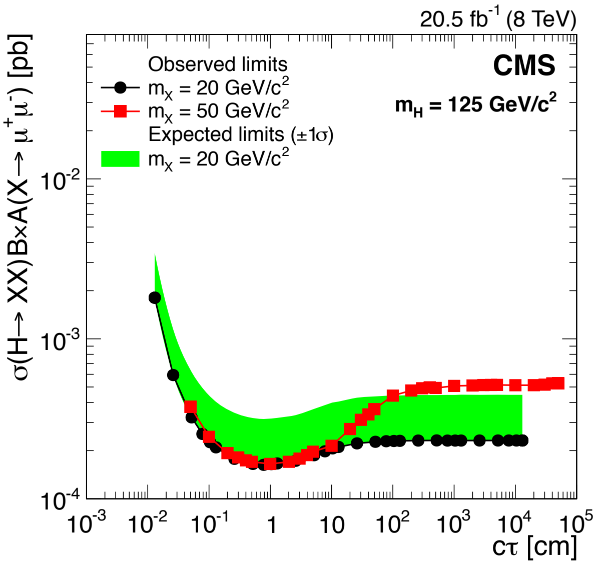


Figure 3: The 95% CL upper limit on $\sigma(H \to XX)\mathcal{B}(X \to \mu^+\mu^-)$, as a function of the mean proper decay length of the X boson. The shaded band shows the $\pm1\sigma$ range of variation of the expected 95% CL limits for the case of a 20 GeV/c^2 X boson mass. From [13].

3 LHCb detector

The LHCb detector [15, 16] is one of the four large detectors of the Large Hadron Collider (LHC). In the LHC, two proton beams circle in opposite direction and are made to collide in order to study interactions. The other detectors are CMS, ATLAS and ALICE, depicted in Figure 4. CMS and ATLAS are general-purpose particle detectors, but LHCb has a more specific goal: it was designed to look for new physics in CP-violation and rare decays of beauty and charm hadrons.

LHCb has a lower luminosity than CMS and ATLAS: $2 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ versus $1 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ for the latter two. This is a measure for the number of proton proton collisions per second. The lower luminosity has the advantage that when the two proton beams are crossed (which is called an event), the interaction is dominated by a single proton proton interaction. This makes the events easier to analyse, which is necessary to study rare decays and oscillations in CP-violation. To serve this goal, LHCb was also designed for a very good momentum and vertex resolution. The former ranges from $\sigma/p = 0.35\%$ for low momentum tracks to 0.55% for high momentum tracks.

This study exploits these properties of the detector: the high resolution enables reconstructing displaced vertices of long-lived particles and calculating the mass and lifetime of these particles very precisely.

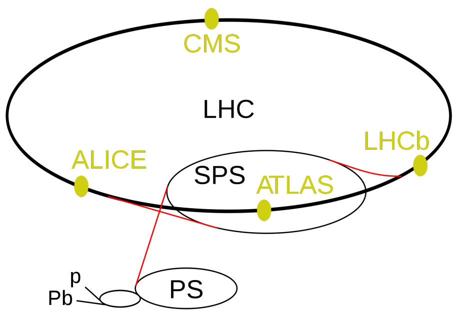


Figure 4: CERN's accelerator complex for the LHC, displaying the four large experiments.

3.1 Layout

The LHCb detector is a single-arm spectrometer, covering a forward angle of about 10 to 250 mrad. Its layout is displayed in Figure 5. The right-handed detector coordinate system is defined with the z-axis along the beam, the x-axis perpendicular to the beam and parallel to the floor, and a vertical y-axis.

The interaction point is in the Vertex Locator system (VELO), described in more detail in Section 3.2. Part of the particle identification is performed by two Ring Imaging Cherenkov counters (RICH 1 and RICH 2), which measure the characteristic Cherenkov radiation emitted by particles. The large dipole magnet is used to bend the charged particles for measuring the momentum. The trajectories of charged particles are determined

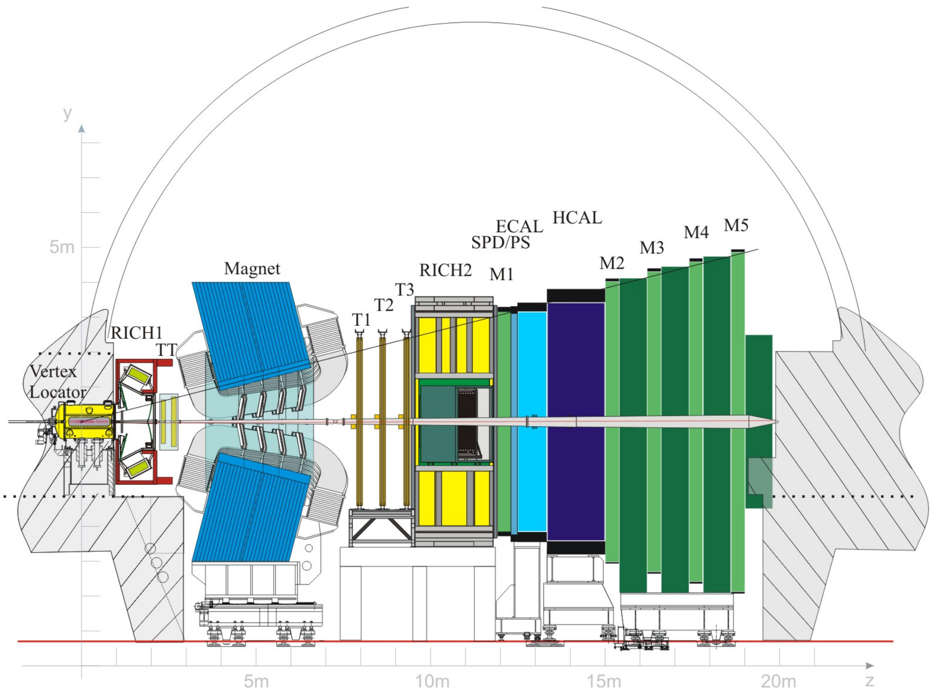


Figure 5: Layout of the LHCb detector, longitudinal section. From [17].

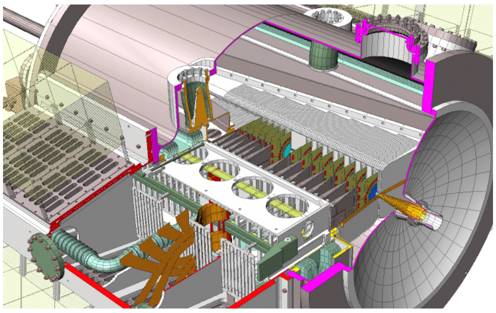
by the tracking system, consisting of the Trigger Tracker (TT) before the magnet and three tracking stations (T1-T3) after the magnet. The calorimeter system consists of a Scintillator Pad Detector (SPD) and Preshower (PS) and an electromagnetic and hadronic calorimeter (ECAL and HCAL). They contribute to the identification of electrons, photons and hadrons and measure their energies and positions. The only particles passing the calorimeters are muons and neutrinos. Only the first one of the two can be detected, by the muon system, which consists of five muon stations (M1-M5).

3.2 VELO

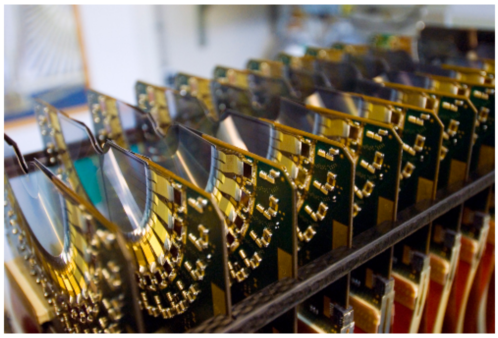
The VELO [18,19] is a silicon microstrip detector designed to provide precise measurements of track coordinates close to the interaction point. Figure 6 shows that it consists of 21 silicon modules (excluding two pile-up detecting modules), with a sensitive radius ranging from 8 to 42 mm from the beam axis. The distance between the first and last module is 925 mm. The π_v^0 particles in the scope of this search have radial decay distances ρ up to 30 mm, determined from simulations. The decay distance along the beamline d_z of the π_v^0 particles ranges up to 200 mm, so it is possible that they leave the VELO before they decay.

The primary vertex (PV) resolution depends on the number of tracks used to calculate its position and ranges in z-direction from 60 (40 tracks) to 260 μm (5 tracks). The resolution in x- and y-direction is even better, ranging from 10 to 35 μm .

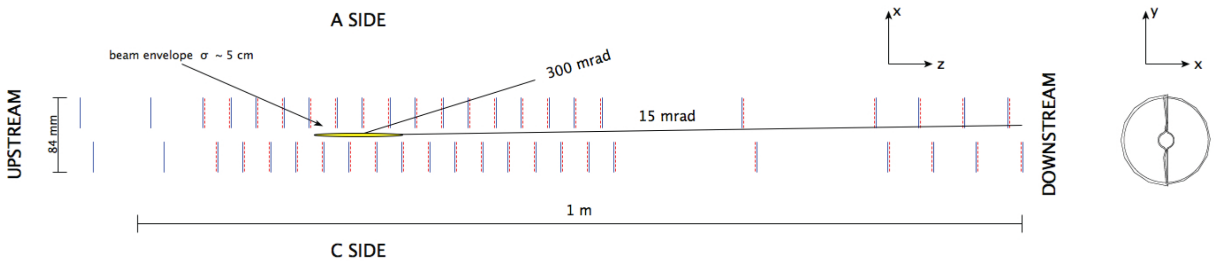
The Impact Parameter (IP) is defined as the shortest distance from a particle track to



(a)



(b)



(c)

Figure 6: (a) The LHCb VELO vacuum tank. (b) One half of the silicon modules and readout system during assembly. (c) Cross-section in the xz plane at $y = 0$ of the silicon modules and a view of a module in the xy plane. From [19].

the PV. The resolution on IP_x is $< 35\ \mu$m for particles with a momentum in the xy-plane $p_{\mathrm{T}} > 1$ GeV/c. The resolution on the decay time depends on the resolution of the displaced vertex, which depends on the topology of the decay. Therefore, it needs to be determined from simulations or data.

4 Data and simulation samples

This analysis has been performed using the data recorded in the LHCb detector in 2011 and 2012. Hidden valley $\pi_v^0 \to \mu^+\mu^-$ events are generated using simulation software in order to see which properties these events have and what effect the selection has on them. In this Chapter, the details of data and π_v^0 simulation samples are described. Samples are also made for possible background decays. See Appendix A for a description of the used software.

4.1 Data processing

The samples used in this search consist of the information recorded for proton proton collisions (events) where a muon pair could be identified. This pair is called a π_v^0 candidate, as the muons might have come from a π_v^0.

In order to go from events in the detector to a sample with π_v^0 candidate events, several steps are taken, as displayed in Figure 7. Firstly, the trigger decides online which of the events in the detector are stored on disk. Data with similar properties is grouped in 'trigger lines', which can be selected for analyses.

From the stored 'raw data' coming from the detector, particle tracks are reconstructed and identified. Then a preselection is applied which in this analysis selects only events that contain a 'high-quality' muon pair. These muon pairs are the starting point of the analysis. In Chapter 5, the 'high-quality' requirements are described.

For the π_v^0 simulation sample, $\pi_v^0 \to \mu^+\mu^-$ events are generated and a detector simulation is applied. From that point, the events are processed in the same way as real data. The trigger is simulated and reconstruction and preselection are applied, leading to samples of π_v^0 candidate events.

The samples are analysed by calculating properties of the events and studying the distributions of these properties.

4.2 Data samples

For this analysis, four data samples are used: data recorded in 2011 and 2012 with the magnetic field in the detector either pointing upwards or downwards (from now on called Up and Down). The center-of-mass energy of the proton beams was $\sqrt{s} = 7$ TeV in 2011 and $\sqrt{s} = 8$ TeV in 2012. The integrated luminosity L of these samples is a measure of the number of proton-proton collisions N_{pp}, which can be calculated by:

$$N_{pp} = \sigma_{pp} \cdot L.$$

Here, σ is the cross-section of the scattering of two protons in pb, a unit of area. The unit of L is pb^{-1}. The four samples and their integrated luminosities are listed in Table 1.

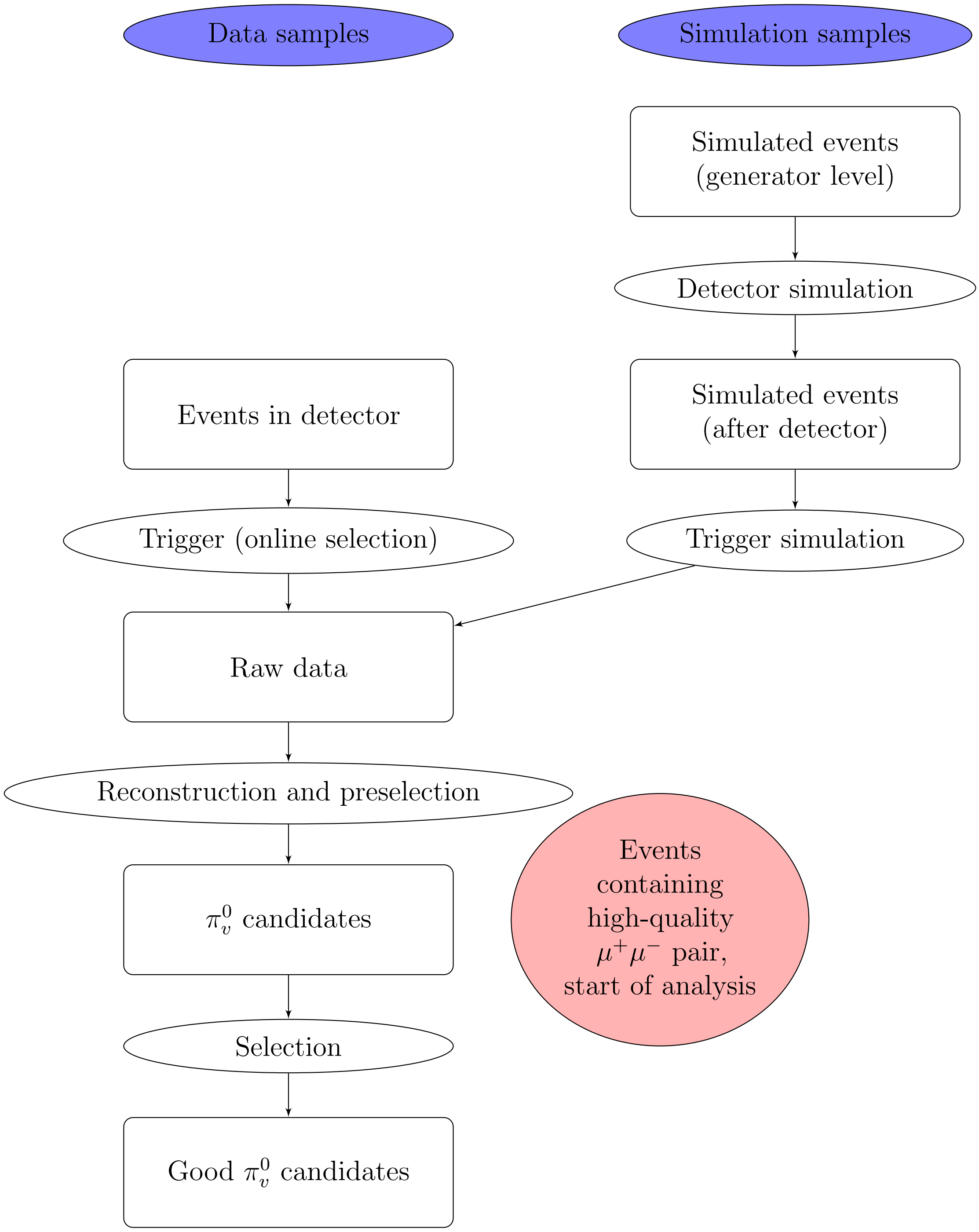
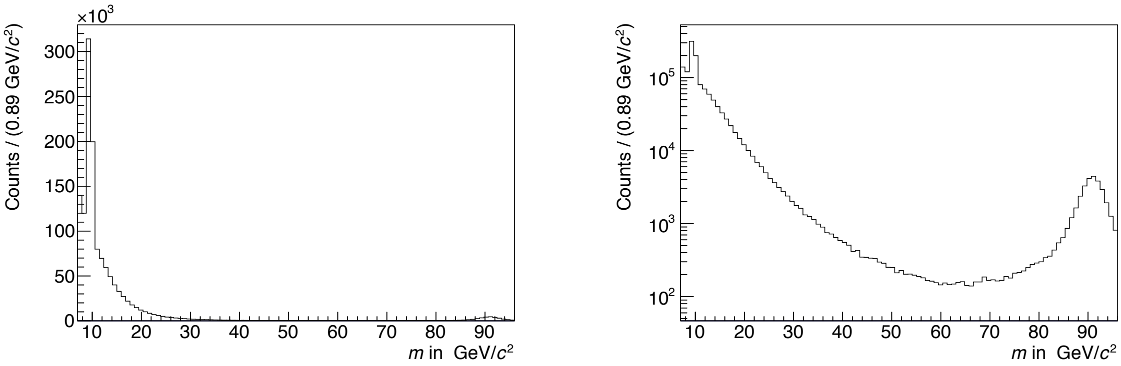


Figure 7: Diagram of data processing.

Table 1: Table of data samples, their integrated luminosities L and systematic error on L in pb^{-1}.

Data sample	L in pb^{-1}		
2011 Up	422.16	$\pm$	7.21
2011 Down	563.61	$\pm$	9.63
2012 Up	999.13	$\pm$	11.58
2012 Down	987.75	$\pm$	11.45
Total	2972.66	$\pm$	20.25

For each event in a data sample, the mass of the π_v^0 candidate is calculated. The mass histogram is displayed in Figure 8, where each entry corresponds to a π_v^0 candidate. A rapidly, approximately exponentially decaying mass distribution is observed, with a peak near a mass of 10 and 90 GeV/c^2.



(a) (b)

Figure 8: Invariant mass distribution of π_v^0 candidates for data recorded in 2012 with magnetic field Up, on (a) linear and (b) logarithmic scale.

4.3 π_v^0 simulation samples

A Standard Model-like scalar Higgs boson is generated with a mass of 125 GeV/c^2, which is forced to decay to two π_v^0 particles. The natural width is set to 4 MeV/c^2. This is a measure of the uncertainty in the mass, as the uncertainty principle $\Delta E \Delta t > \hbar/2$ implies an uncertainty in the energy of short-lived resonances. The π_v^0 particles each decay to two oppositely charged muons.

Different samples are generated with π_v^0 masses of 7, 10, 20, 35 and 50 GeV/c^2 and lifetimes of 10 and 100 ps respectively. Samples with 2011 and 2012 conditions are produced, which differ in beam properties, triggering and detector settings. The center-of-mass energy $\sqrt{s}$ was 7 TeV in 2011 and 8 TeV in 2012. The average number of generated proton-proton interactions per event are 2.0 and 2.5 for the respective years. Samples are generated with the magnetic field in the detector either pointing upwards or downwards. This adds up to a total of 40 π_v^0 samples. In Table 2, only the samples with 2012 simulation conditions and magnetic field Up are shown. The total table is listed in Appendix B.

In order to get a sample with events that can be detected, the muon daughters of at least one of the π_v^0 particles are required to have a momentum p larger than 2 GeV/c and an angle with the beam axis θ smaller than 400 mrad. Muons with a lower momentum do not reach the muon stations and muons with an angle that is too large, will not fly through the detector. In other words, the muons have to fall within the 'acceptance' of the detector. The ratio of the number of events before and after these constraints are applied,

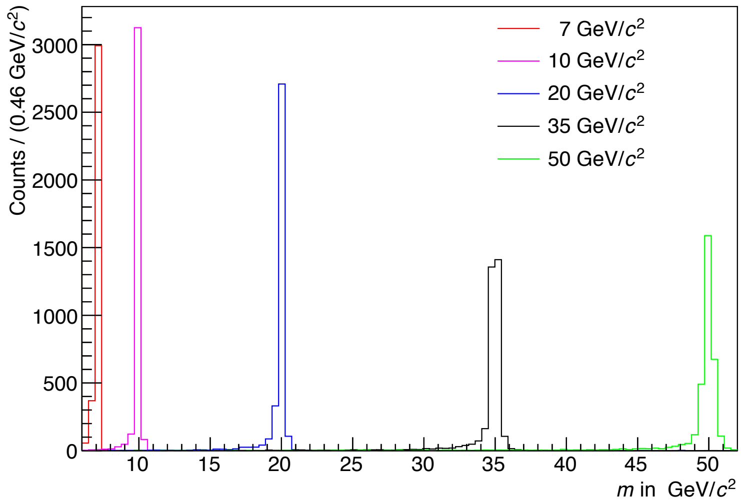


Figure 9: Invariant mass distributions of the 10 ps π_v^0 simulation samples. 2012, magnetic field Up simulation conditions.

is called the generator efficiency. It is listed in Table 2 for the simulation samples with 2012 Up simulation conditions. The full table can be found in Appendix B.

The efficiency is independent of the π_v^0 lifetime, but is lower for larger π_v^0 masses. Muons coming from a heavier π_v^0 have more momentum, because more energy is available when the π_v^0 decays. The forward momentum of particles in the direction of the beam is much larger than their transverse momentum, because of the collision energy. Thus when a π_v^0 decays, the mass that is transformed into momentum, has the largest relative effect on the transverse momentum of the muons. Consequently, the angle between two muons coming from a heavy π_v^0 is larger than when they come from a light π_v^0. This is confirmed by simulation. It makes it more likely that muons coming from a heavier π_v^0 do not fall within the detector acceptance, which explains the lower generator efficiency. This was the motivation for the mass scope of this research, the LHCb detector is not efficient at detecting these signal decays with much larger masses.

The mass distribution of the simulated π_v^0 particles is displayed in Figure 9. A peak is visible for each of the different mass samples, the width of this peak is the resolution with which mass can be measured in the detector.

Table 2: Table of the simulated π_v^0 samples. Listed are the unique sample identification number, the π_v^0 properties and the generator efficieny for the samples with 2012 simulation settings and an upwards magnet field in the detector.

Event ID	π_v^0 mass (GeV/c^2)	π_v^0 lifetime (ps)	Generator efficiency (%)
43114000	7	10	22.12 ± 0.14
43114001	7	100	22.03 ± 0.13
43114002	10	10	20.94 ± 0.13
43114003	10	100	20.80 ± 0.13
43114004	20	10	17.27 ± 0.10
43114005	20	100	17.17 ± 0.11
43114006	35	10	13.04 ± 0.11
43114007	35	100	13.18 ± 0.12
43114008	50	10	9.66 ± 0.09
43114009	50	100	9.61 ± 0.08

4.4 Background samples

The data samples consist of events with a high-quality $\mu^+\mu^-$ pair. When cross-sections $\sigma(\ pp \to X\ )$ and branching fractions $\mathcal{B}\ (X \to \mu^+\mu^-)$ of certain particles X are large, the $\mu^+\mu^-$ pair in the data may have come from this particle. This is the case for Υ mesons and the Drell-Yan (DY) process [20]. The latter is the annihilation of a quark and an anti-quark, creating a virtual photon or Z-boson, which can decay to a muon pair. The J/ψ particle also has a large branching fraction to two muons, but its mass is 3.1 GeV/c^2. This is below the masses in the scope of this research, and this background is discarded in the preselection.

Other possible backgrounds consist of two particles that each decay to one muon in different vertices, for example a heavy flavour $b\bar{b}$ or $c\bar{c}$ pair. Both quarks can decay to a muon, a neutrino and a shower of other particles, called a jet. It will be shown that after discarding the Υ and Drell-Yan background events, the remaining background events have similar properties as $b\bar{b}$ decays.

Diagrams of these two types of background decays can be found in Figure 10. Simulation samples are made for Υ, Z and $b\bar{b}$ decays. Next to that, a sample of these events is extracted from the data.

4.4.1 Υ meson simulation samples

For each of the Υ mesons that have large branching fractions $\mathcal{B}$ to muons ($\Upsilon(1S)$, $\Upsilon(2S)$ and $\Upsilon(3S)$), four simulation samples are generated for 2011 and 2012 simulation conditions and for magnet settings Up and Down. The same triggering, reconstruction and preselection are applied as on the π_v^0 samples and data, in order to make a sample of events with high-quality muon pairs.

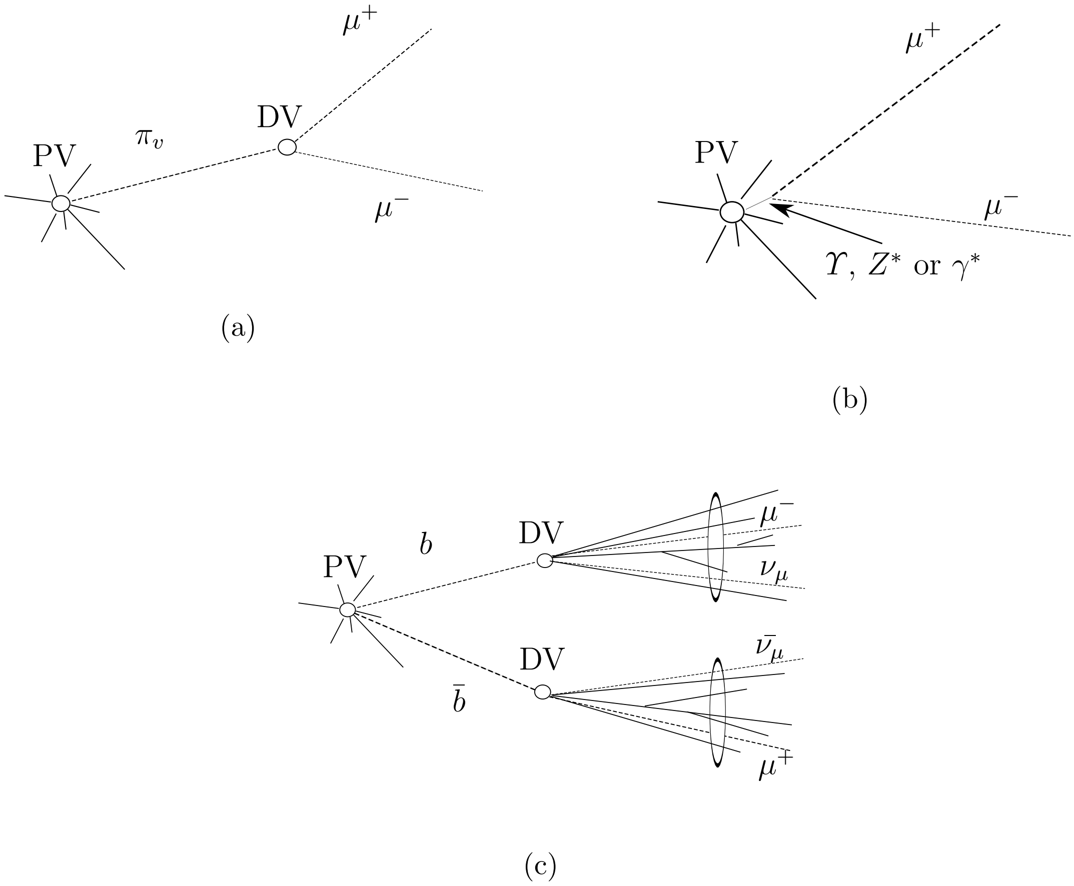
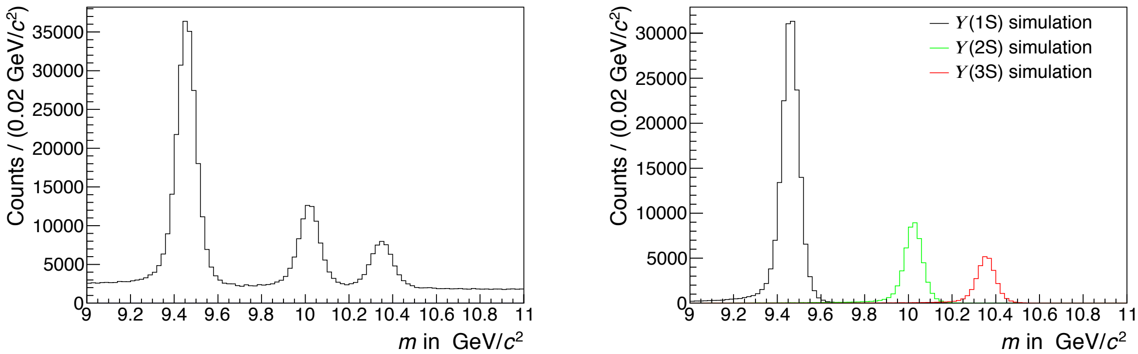


Figure 10: Overview of signal and background decays: (a) π_v^0 signal decay, (b) background decay of short-lived Υ meson and Drell-Yan process (Z and photon), (c) background decay of $b\bar{b}$ to muons and jets.

4.4.2 Υ from data

The masses of the $\Upsilon(1S)$, $\Upsilon(2S)$ and $\Upsilon(3S)$ are 9.46, 10.02 and 10.36 GeV/c^2, respectively [20]. A plot is shown of the invariant mass distribution of the data in this region in Figure 11a. Three peaks are clearly visible and comparing with the simulation samples, shown in Figure 11b, learns that these are indeed Υ particles. Thus the data can be used to obtain a sample with events that most likely contain an Υ meson.

A clear background is also visible in Figure 11a. In order to isolate the events with an Υ particle, three Crystal Ball (CB) functions are fitted to the mass peaks and two exponential functions to the background with the RooFit tools [21]. A CB function is a Gaussian function with a changed tail on the left hand side. Muons are known to emit photons and the loss of energy in this process is translated to a slightly lower reconstructed Υ mass. This brings about so-called radiative tails in the mass distribution, which are



(a) (b)

Figure 11: Invariant mass distributions in the Υ mass window (9 to 11 GeV/c^2), of (a) 2012 Up data and (b) Υ simulation samples with 2012 Up simulation conditions. The ratio's between the simulation samples are calculated in the fit to data.

modelled by the slightly larger left tail in the CB function [22]. A plot of the fit is displayed in Figure 12. Here, the blue line models the backgrounds, the red line the Υ mesons and the black line is the total fit.

So-called 's-weights' are assigned to the events using the RooSplot tools [23]. These weights correspond to the likelihood that the event is part of the Υ peaks or background. By multiplying the data with these weights, 'Υ-data' is created.

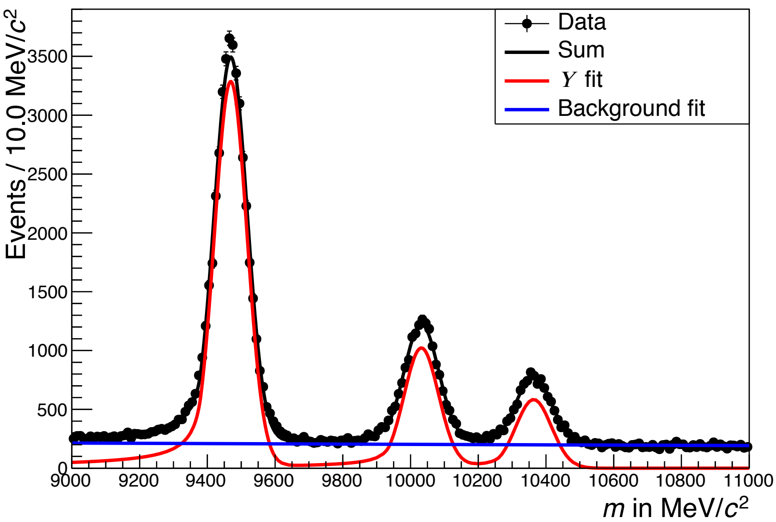


Figure 12: Fit of two exponential backgrounds (blue) and three Crystal Ball functions (red) and their sum (black) to the invariant mass distribution in the window 9 to 11 GeV/c^2 for 2012 Up data.

Properties are calculated for each event in the samples and the distributions are compared for the Υ-data and the $\Upsilon(1S)$ simulation sample in Figure 13. Overall, measured data and simulation results correspond very well, which shows that the procedure worked and indeed selects events containing Υ mesons from data.

Figures 13a and 13b display the transverse momentum (p_{T}) of the $\mu^+\mu^-$ combination and of one of the muons. This is the momentum transverse to the beam axis. The small differences are due to the fact the production of Υ particles in simulation does not perfectly describe all possible production processes in the data. In Section 5.2, the precise preselection criteria will be defined. The preselection constraints on the p_{T} of the $\mu^+\mu^-$ combination ($p_{\mathrm{T}} > 7500$ MeV/c) and on the muons themselves ($p_{\mathrm{T}} > 2500\,\mathrm{MeV}/c$) can clearly be seen.

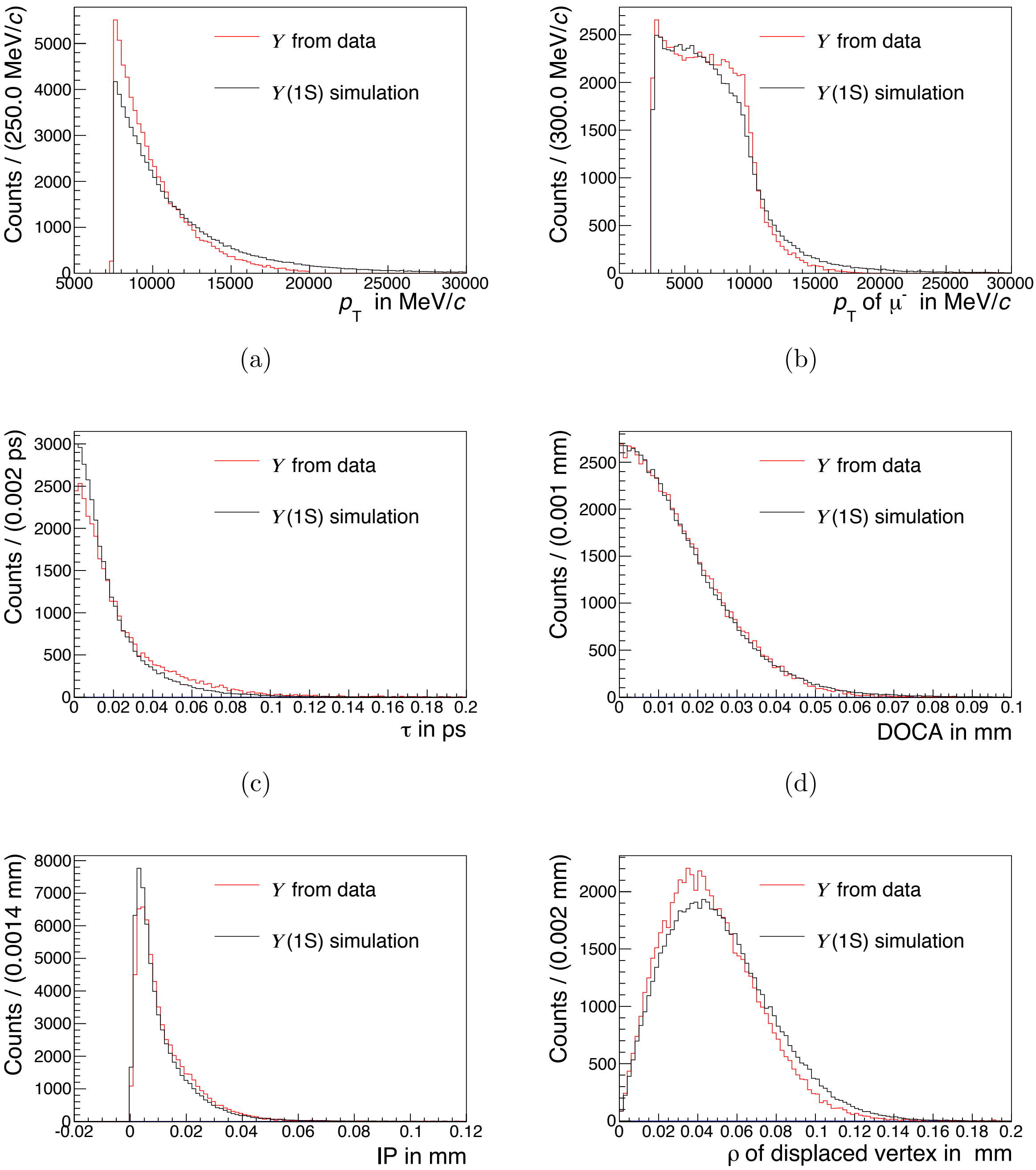
Figure 13c displays the reconstructed proper lifetime τ of the $\mu^+\mu^-$ combination. Υ mesons decay almost instantly after being created and the histogram shows the resolution of the detector.

The DOCA, displayed in Figure 13d, is the Distance Of Closest Approach between the two muons. The Impact Parameter in Figure 13e, is the distance from the $\mu^+\mu^-$ track to the Primary Vertex (PV).

Lastly, ρ in Figure 13f is the radial coordinate of the Displaced Vertex (DV) where the Υ decays to $\mu^+\mu^-$. It is related to the p_{T} of the muon pair by the equation:

$$\rho = \frac{\tau \cdot p_{\mathrm{T}}}{m}.$$

Thus the differences are also caused by the production processes in data and simulation.



(a) (b) (c) (d) (e) (f)

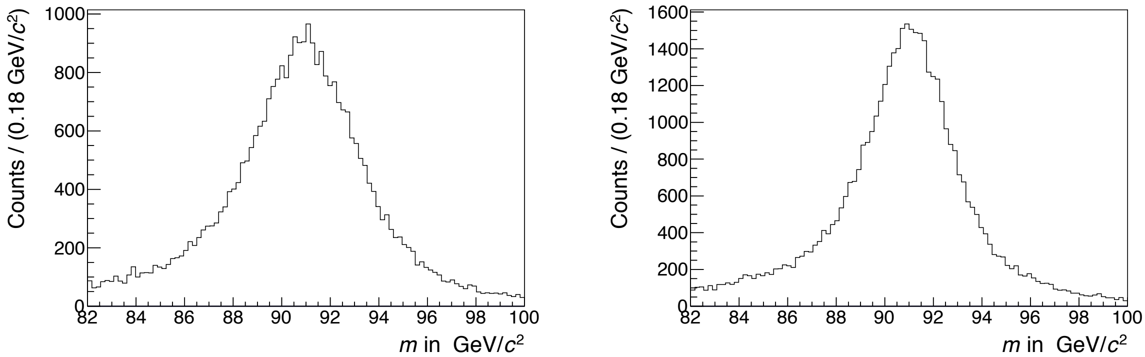
Figure 13: Distributions of some properties of the events for both the $\Upsilon(1S)$ simulation sample (2012 Up simulation conditions) and the Υ events obtained from 2012 Up data. Displayed are (a) the transverse momentum of the $\mu^+\mu^-$ combination, (b) the transverse momentum of the μ^-, (c) the reconstructed proper lifetime τ of the $\mu^+\mu^-$ combination, (d) the Distance of Closest Approach (DOCA) between the two muons, (e) the Impact Parameter (IP) from the $\mu^+\mu^-$ combination to the primary vertex and (f) the radial coordinate ρ of the displaced vertex. The samples are scaled to the area under the data sample in order to compare the distributions.

4.4.3 Z boson simulation samples

Four Z boson samples are generated for 2011 and 2012, Up and Down magnet polarity simulation conditions.

4.4.4 Z boson from data

Like the Υ meson, the Z boson has a large cross-section and branching fraction to two muons, so many of the $\mu^+\mu^-$ pairs in the data are expected to have originated from a Z boson. The mass histogram of the data and Z simulation samples in its mass range are displayed in Figure 14. As the Z boson is very heavy, little to no background is expected in this range, which can be seen in Figure 8. Therefore, a sample of Z boson events can be created by selecting events with a muon pair mass between 85 and 97 GeV/c^2.



(a) (b)

Figure 14: Mass distributions of (a) the 2012 Up data and (b) of the Z boson 2012 Up simulation sample.

In Figure 15, some properties of the events are compared in Z-data and simulation samples. The distributions correspond very well, it can be concluded that a sample with events containing Z bosons has been created. The lifetime of the Z bosons is also very short: they decay instantly and the distribution is a resolution effect from the detector.

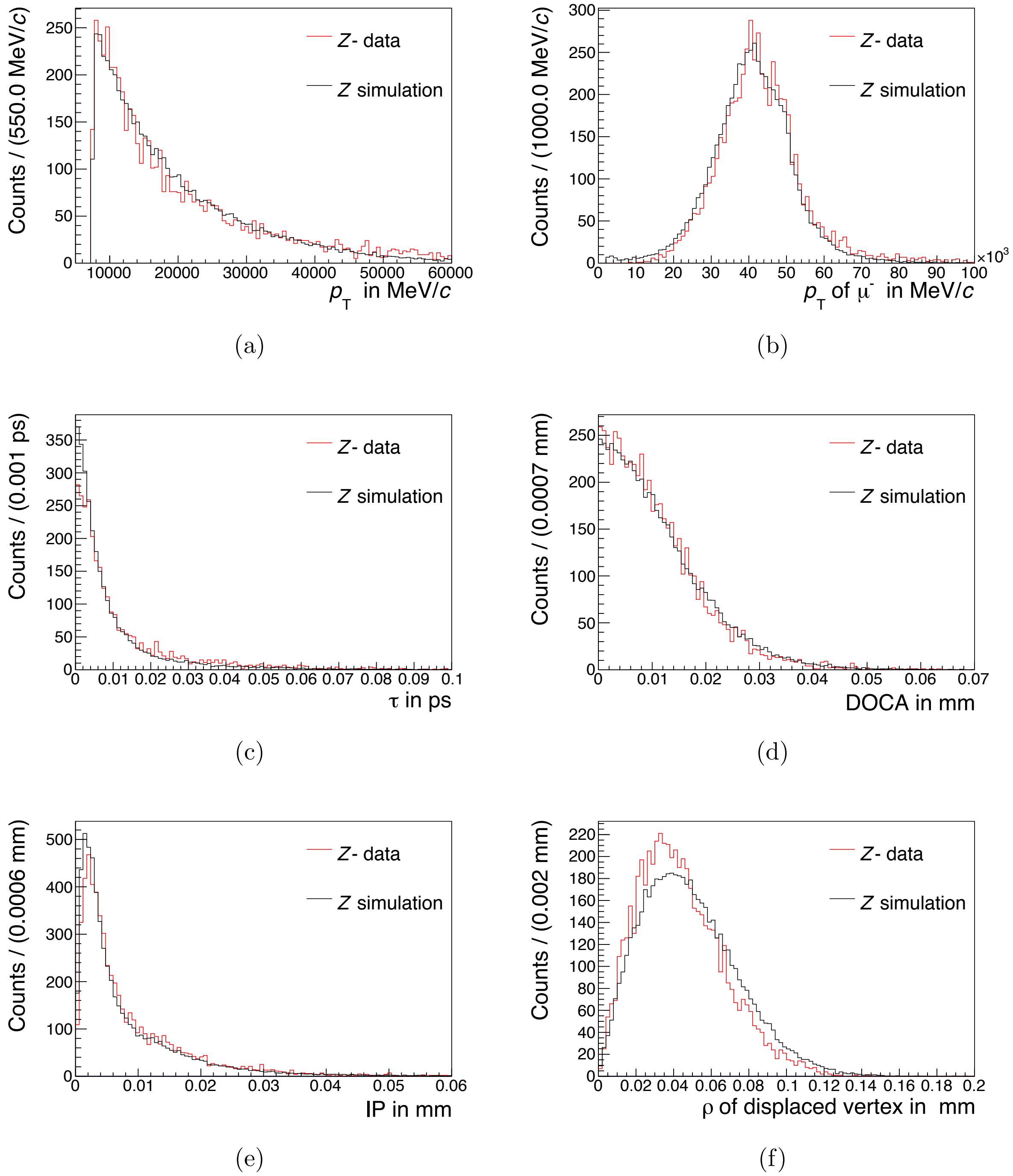


Figure 15: Distributions of some properties of the events for both the Z simulation sample (2012 Up simulation conditions) and the Z events obtained from 2012 Up data. Displayed are (a) the transverse momentum of the $\mu^+\mu^-$ combination, (b) the transverse momentum of the μ^-, (c) the reconstructed proper lifetime τ of the $\mu^+\mu^-$ combination, (d) the Distance of Closest Approach (DOCA) between the two muons, (e) the Impact Parameter (IP) from the $\mu^+\mu^-$ combination to the primary vertex and (f) the radial coordinate ρ of the displaced vertex. The samples are scaled to the area under the data sample in order to compare the distributions.

4.4.5 $b\bar{b}$ simulation samples

In the previous sections, it was demonstrated that Υ mesons and Z bosons decay almost instantly after they are created. The hypothesis is that when these events are suppressed in the data, mostly events with muons originating in two different vertices are left. $b\bar{b}$ decays are an example of such events, and this combination is copiously produced in LHCb as it was designed to these quarks. In order to study properties of these decays, simulation samples are made for $b\bar{b}$ decays and long lifetime data samples are constructed.

The $b\bar{b}$ simulation samples are generated for 2011 and 2012, Up and Down magnet polarity simulation conditions. Unfortunately, no samples are available that have been processed with the same reconstruction and preselection as is used for the other simulation samples and data. To make up for this effect, the preselection momentum and mass requirements have been applied in the analysis.

4.4.6 Long lifetime data

A long lifetime dataset is composed by only selecting events that have a muon pair lifetime longer than the Υ and Z resolution effects. In Figure 16, the distribution of the $\mu^+\mu^-$ pair invariant mass is displayed for the $b\bar{b}$ simulation sample and the data events with the lifetime selection $\tau > 0.2$ ps. This selection is based on Figures 13c and 15c. Most of the events in the Υ and Z peaks are indeed discarded, compared to Figure 8. Data and simulation do not correspond very well, as was expected. The data sample does not only consist of decaying $b\bar{b}$ pairs, e.g. $c\bar{c}$ pairs are also present in the data.

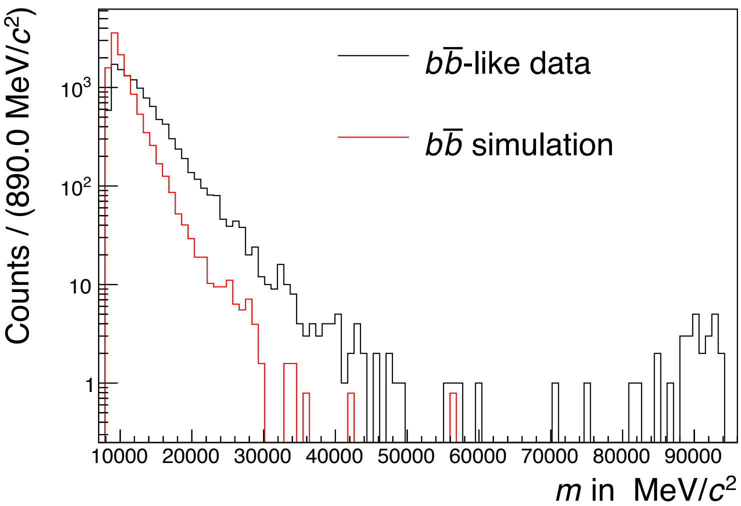


Figure 16: Invariant mass distributions of π_v^0 candidates for long lifetime data and $b\bar{b}$ simulation sample with 2012 Up simulation conditions. The long lifetime data is defined as a random selection corresponding to 200 pb^{-1} from 2012 Up data, with lifetime selection $\tau > 0.2$ ps. The simulation sample is scaled to the long lifetime data.

In Figure 17, some properties of the created long lifetime data and the simulation sample are displayed. The lifetime requirement has also been imposed on the simulation sample, to make a fair comparison.

The simulation and data distributions correspond quite well. The main differences originate from the fact that the data is not all $b\bar{b}$, but it still has similar properties. The different shape of the distribution in Figure 17b is probably due to a preselection requirement that was imposed on the simulation but not on the data. The distributions are different from the Υ and Z distributions, e.g. the DOCA is much larger here. When the muons come from two vertices, the distance between them will be larger than when they come from one vertex. The lifetime is also larger, but not as large as the chosen π_v^0 lifetimes.

It can be concluded that most of the data events after a lifetime selection, have the properties of events like $b\bar{b} \rightarrow \mu^+\mu^-$, i.e. with two muons originating in different vertices. The long lifetime data samples and $b\bar{b}$ simulation samples can be used to study the properties of these background events.

4.4.7 Summary of background decays

Two types of background decays have been identified, muons originating in a short-lived resonance (Υ and DY) and muons from two different vertices like the decay of a b and $\bar{b}$. In Table 3, the properties of these decays and the signal decay are summarised.

Table 3: Overview of properties of π_v^0 signal decays and Υ, DY and $b\bar{b}$- like background decays. Displayed are the lifetime and the displaced vertex (DV) quality.

	π_v^0 signal	Υ and DY background	$b\bar{b}$- like backgrounds
Lifetime	10 - 100 ps	Instantaneous decay	≈ 1 ps
DV quality	Good	Good	Bad

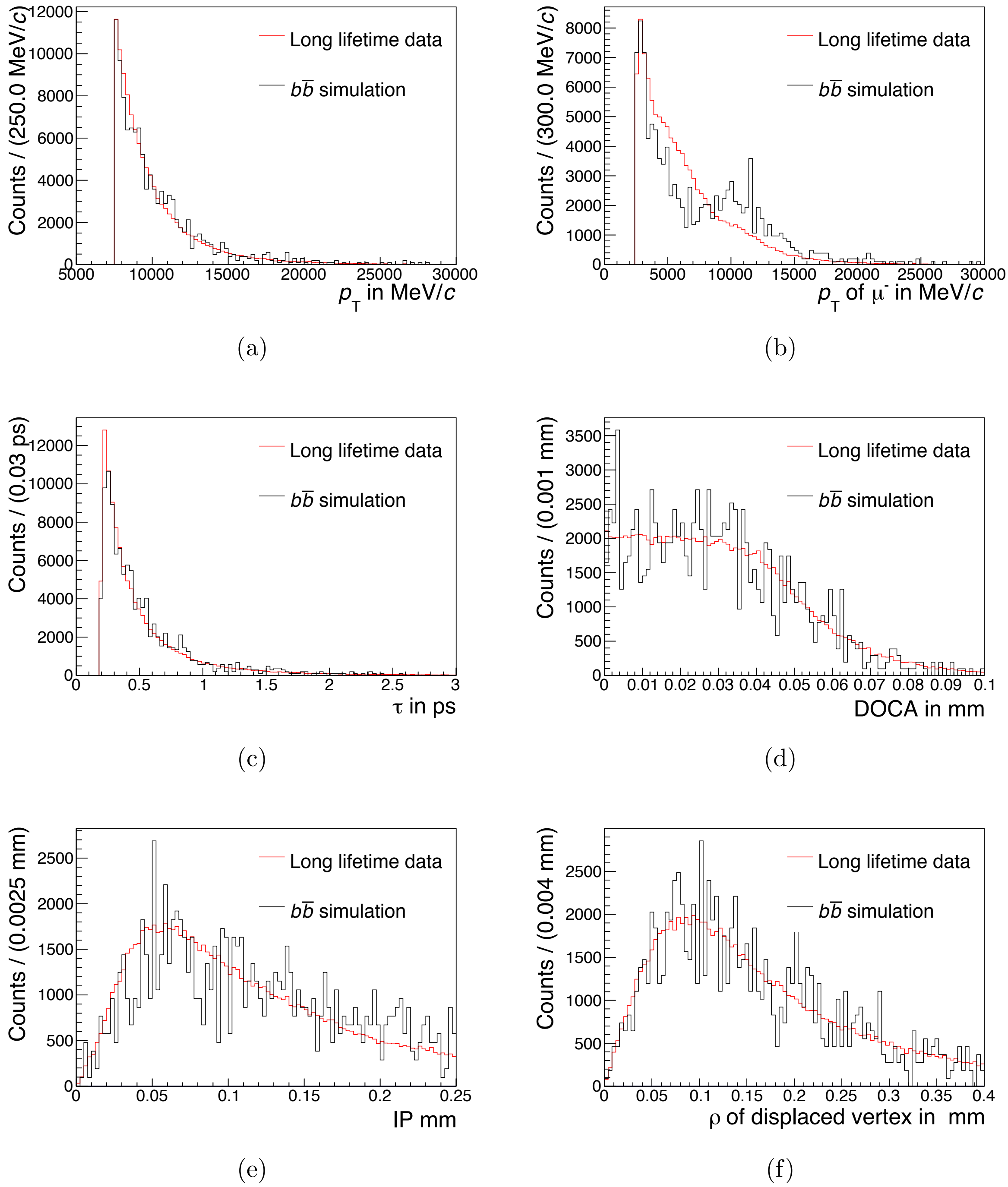


Figure 17: Distributions of some properties of the events for both the $b\bar{b}$ simulation sample (2012 Up simulation conditions) and the events obtained from 2012 Up data with lifetime condition $\tau > 0.2$ ps. Displayed are (a) the transverse momentum of the $\mu^+\mu^-$ combination, (b) the transverse momentum of the μ^-, (c) the proper lifetime τ of the $\mu^+\mu^-$ combination, (d) the Distance of Closest Approach (DOCA) between the two muons, (e) the Impact Parameter (IP) from the $\mu^+\mu^-$ combination to the primary vertex and (f) the radial coordinate ρ of the displaced vertex. The samples are scaled to the area under the data sample in order to compare the distributions.

5 Trigger, reconstruction and preselection

In this Chapter, the process from events in the detector to a sample of events with high-quality muon pairs is described. In the diagram in Figure 7, these are the trigger, reconstruction and preselection steps.

5.1 Trigger

A trigger is a system that decides online (while the data is being taken) whether an event in the detector is stored or not. The LHCb trigger comprises of two stages, the hardware L0 Trigger and the High Level Trigger (HLT) implemented as software running on a CPU farm [24,25]. Data with similar properties is stored in 'trigger lines', which can be selected to use in analyses.

5.1.1 L0 trigger

The L0 trigger is divided into three independent parts: the L0-Calorimeter trigger, L0-PileUp trigger and L0-Muon trigger. The latter, which triggers on high energy muons, is used for this search. It uses the five muon stations displayed in Figure 5. Processors search for hits in the stations that define a straight line and point towards the interaction point of the event. An event is stored in the L0Muon line if the muon track with the largest p_{T}, has a p_{T} over a certain threshold. It is stored in the L0DiMuon line if the muon tracks with largest and second to largest p_{T}, have $p_{\mathrm{T}1} \times p_{\mathrm{T}2}$ over a certain threshold. These thresholds are listed in Table 4.

As the sought after signal has two muons, the L0DiMuon line is used. The L0Muon line is also used, which selects quite a few events that do not get stored in the former line. The muon with the second to largest p_{T} might have so little momentum that the combined p_{T} threshold is not surpassed, while the muon with the largest p_{T} has enough momentum to pass the L0Muon criterion.

The effect of the trigger on π_v^0 decay events is simulated for the π_v^0 simulation samples. The efficiency of a trigger stage is the fraction of π_v^0 events that pass the trigger requirements. It is listed in Tables 5 and 6.

Table 4: L0 thresholds for the p_{T} of muon tracks.

	2011	2012
L0Muon p_{T}^{max}	1.48 GeV/c	1.76 GeV/c
L0DiMuon $p_{\mathrm{T}}^{max} \times p_{\mathrm{T}}^{max-1}$	1.296 (GeV/c)^2	1.6 (GeV/c)^2

5.1.2 High Level trigger

The Hlt trigger comprises two stages: HLT1 and HLT2. HLT1 uses partial event reconstruction and implies a fast muon identification algorithm on all events in the

Table 5: Number of events, relative efficiency and cumulative efficiency on 10 ps π_v^0 simulation samples. Displayed are the generator efficiency and trigger efficiencies. 2012 Up simulation conditions.

	7 GeV/c^2, 10 ps			**50 GeV/c^2, 10 ps**		
	N	ϵ_{rel}	ϵ_{cum}	N	ϵ_{rel}	ϵ_{cum}
After generator	11435	0.2212	0.2212	11325	0.0966	0.0966
L0	10319	0.9024	0.1996	11007	0.9719	0.0939
Hlt1	5573	0.5401	0.1078	9417	0.8555	0.0803
Hlt2	3935	0.7061	0.0761	5779	0.6137	0.0493

Table 6: Number of events, relative efficiency and cumulative efficiency on 100 ps π_v^0 simulation samples. Displayed are the generator efficiency and trigger efficiencies. 2012 Up simulation conditions.

	7 GeV/c^2, 100 ps			**50 GeV/c^2, 100 ps**		
	N	ϵ_{rel}	ϵ_{cum}	N	ϵ_{rel}	ϵ_{cum}
After generator	11435	0.2203	0.2203	11899	0.0961	0.0961
L0	10313	0.9019	0.1987	11562	0.9717	0.0934
Hlt1	1184	0.1148	0.0228	5403	0.4673	0.0436
Hlt2	689	0.5819	0.0133	2250	0.4164	0.0182

L0Muon and L0DiMuon lines. It then applies certain selection criteria and reduces the amount of data coming from the L0 from about 1 MHz to 80 kHz (2012). HLT2 runs a more complete event reconstruction and once again applies selection criteria. This further reduces the data to 5 kHz, which is stored for offline analysis.

The trigger lines impose several criteria on the muons tracks and vertices. The quantity IP-χ^2 is an adaptation of the Impact Parameter, the shortest distance from a track to a vertex. It is calculated in such a way that it no longer depends on the uncertainties in the track and vertex. The χ^2_{vertex} is the sum of the IP-χ^2 for both muons. In is an indication of whether the two muons came from the same vertex. The χ^2_{track} / ndf measures the quality of the muon tracks. Lastly, DOCA is the Distance of Closest Approach between the two muons.

Several Hlt trigger lines are used in this search, they are listed in Tables 7, 8 and 9. The Hlt1TrackMuon line selects events with good quality single muon candidates. Two HLT1 dimuon candidate triggers are used; the Hlt1DiMuonLowMass and Hlt1DiMuonHighMass line both select muons with a good quality vertex. Only HLT2 lines are used that select dimuon candidates, which are grouped in two catagories: ones that select prompt decays and those that select displaced decays. All of these are used in this analysis.

Their combined efficiencies are listed in Tables 5 and 6. It is observed that the HLT1 trigger has a very low efficiency for long lifetimes. This trigger line requires muon hits in the VELO and many π_v^0 particles with a 100 ps lifetime fly out of the VELO before they decay. These events are not triggered and a lot of signal efficiency is lost. That is why the lifetime scope in this search was chosen to be not higher than 100 ps.

Table 7: HLT1 muon lines and their selection criteria.

Hlt1line	`TrackMuon`	`DiMuon HighMass`	`DiMuon LowMass`
Track IP [mm]	$>$ 0.1	-	-
Track IPχ^2	$>$ 16	-	$>$ 3
Track p_{T} [GeV/c]	$>$ 1	$>$ 0.5	$>$ 0.5
Track p [GeV/c]	$>$ 8	$>$ 6	$>$ 6
Track χ^2/ndf	$<$ 2	$<$ 4	$<$ 4
DOCA [mm]	-	$<$ 0.2	$<$ 0.2
χ^2_{vertex}	-	$<$ 25	$<$ 25
Mass [GeV/c^2]	-	$>$ 2.7	$>$ 1

Table 8: HLT2 lines based on two identified muons that select prompt decays.

`Hlt2DiMuon`	`JPsi`	`Psi2S`	`B`	`JPsiHighPT`	`Psi2SHighPT`
Track χ^2/ndf	$<$ 5	$<$ 5	$<$ 5	$<$ 5	$<$ 5
Mass [GeV/c^2]	$\mathrm{M}_{J/\psi} \pm 0.12$	$\mathrm{M}_{\psi(2S)} \pm 0.12$	$>$ 4.7	$\mathrm{M}_{J/\psi} \pm 0.12$	$\mathrm{M}_{\psi(2S)} \pm 0.12$
χ^2_{vertex}	$<$ 25	$<$ 25	$<$ 10	$<$ 25	$<$ 25
${p_{\mathrm{T}}}^{\mu\mu}$ [GeV/c]	-	-	-	$>$ 2	$>$ 3.5

Table 9: HLT2 lines based on two identified muons that select displaced decays.

`Hlt2DiMuon`	`Detached`	`DetachedHeavy`	`DetachedJPsi`
Track χ^2/ndf	$<$ 5	$<$ 5	$<$ 5
Track IPχ^2	$>$ 9	-	-
Mass [GeV/c^2]	$>$ 1	$>$ 2.95	$\mathrm{M}_{J/\psi} \pm 0.12$
FDχ^2	$>$ 49	$>$ 25	$>$ 9
χ^2_{vertex}	$<$ 25	$<$ 25	$<$ 25
${p_{\mathrm{T}}}^{\mu\mu}$ [GeV/c]	$>$ 1.5	-	-

5.2 Reconstruction and preselection

The data stored on disk by the trigger is called raw data. It consists stored events with information from the detector. The events are reconstructed again, more precisely this time because there is less time limitation. Particle tracks and reconstructed neutral particles are the output.

After the reconstruction, the preselection is applied, which further reduces the amount of data by imposing more requirements on the events. It also runs a particle identification algorithm. The data is stored in different 'stripping lines', each with their own criteria. The A1MuMu stripping line is used. This line requires a positively and negatively charged muon, both with a transverse momentum $p_T > 2.5\,\mathrm{GeV}/c$ with tracks of high quality. The reconstructed mother particle of the two muons must have a mass higher than 5 GeV/c^2 and a $p_T > 7.5\,\mathrm{GeV}/c$.

Lastly, the displaced vertex (DV) must also be of high quality. The preselection requirement is IP-$\chi^2 < 12$, for larger values it is no longer credible that the two muons came from one vertex.

The events with $\mu^+\mu^-$ pairs (π_v^0 candidates) that meet all of these requirements, are saved and are the starting point of this analysis.

6 Selection

In this Chapter, the selection is described that suppresses events with properties of known backgrounds and keeps events that could contain a π_v^0 particle. Firstly, the analysis of the samples is described, after which different selection criteria are evaluated. A summary is given and the results of the selection are displayed on both data and π_v^0 simulation samples.

6.1 Analysis

The data and simulation samples consist of events with a π_v^0 candidate particle (a high-quality $\mu^+\mu^-$ pair that could have been a π_v^0). Different properties of this candidate and the muons are calculated, like the smallest distance between the two muons and the number of particles in a small cone around a muon. Based on these properties, it is decided whether the decay was a background or possible signal ($\pi_v^0 \to \mu^+\mu^-$) decay.

The momenta of the muons were calculated during the reconstruction of their tracks. This calculation is not equally accurate in all parts of the detector. To compensate for this, a scaling is added to more accurately measure the momenta [26].

Particles with colour, like quarks, cannot exist on their own right and therefore fragment into a cone of particles, called a jet. A jet reconstruction algorithm is run on the events, in order to find these jets and calculate their properties.

An event can have multiple primary vertices (PV's) from other colliding protons. They are determined by looking at clusters of tracks. A PV-refinding algorithm is applied, which recalculates the PV locations. This time it does not use the two muon tracks, because it is now assumed that these come from a displaced vertex and should no longer be taken in to account when calculating the PV.

The 'best PV' of the event is determined by running over all resulting PV candidates and calculating the quantity IP-χ^2 from the π_v^0 candidate to the PV (see Section 5.2). It is a measure of whether or not the π_v^0 candidate came from that PV. The PV candidate for which the IP-χ^2 is smallest, is called the 'best PV' of the event. Calculated quantities involving a PV, use the 'best PV', unless otherwise stated.

6.2 Outline

In the next sections, different criteria are discussed that can separate the π_v^0 signal events from background events. It is first determined which variables separate the best, and a loose criterion for these variables is determined. This does not use the full separating potential of the variable yet, in order to prevent making a too stringent selection when the criteria are combined. They can be made more strict after they have been consecutively applied.

6.3 Lifetime selection

In Chapter 4, it was shown that the backgrounds Υ and Z decay instantaneously, and $\mu^+\mu^-$ pairs from $b\bar{b}$-like backgrounds also have smaller lifetimes than π_v^0 particles. In order to suppress the first two backgrounds and part of the $b\bar{b}$-like backgrounds, particles are selected according to how long they live.

Different quantities are considered as selection criteria. The first is the displacement d, the length of the displacement vector from the best PV to the DV. The distribution of this quantity is displayed in Figure 18a for the 7 and 50 GeV/c^2 π_v^0 simulation samples and the Z and Υ samples from data. It can be observed that the Z and Υ particles decay much faster than the simulated π_v^0 particles.

Other possible criteria are the spherical coördinate ρ of the DV and the lifetime τ of the π_v^0 candidate. The distribution of τ is displayed in Figure 18b.

The lifetime is expected to depend the least on the mass of the π_v^0, as it is predefined in the simulation. ρ and d depend on the momentum of the particles, which depends on their mass. It is preferable to impose the same requirements on all π_v^0 mass samples, thus the lifetime is the preferred variable when its separation power is good.

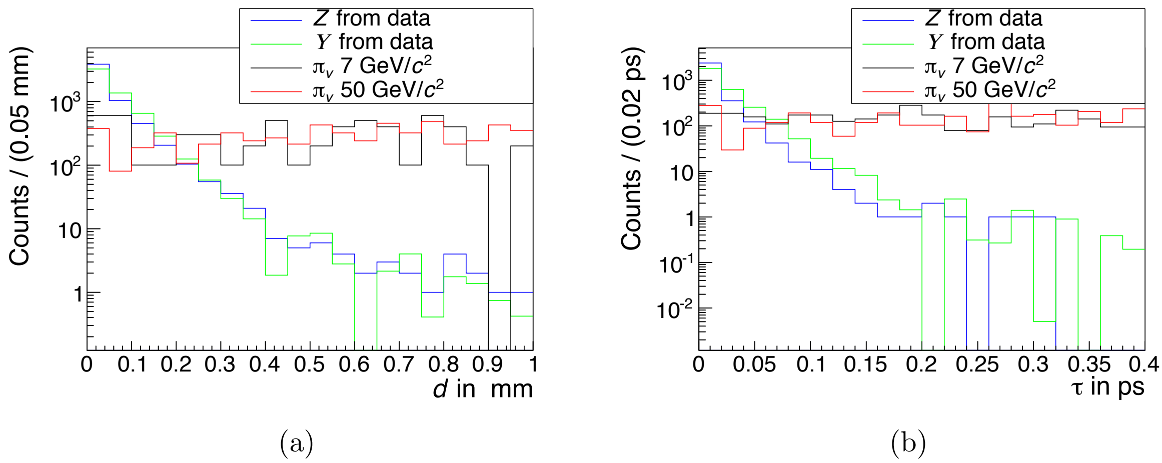


Figure 18: Distributions of (a) the displacement d (distance from best PV to DV) and (b) lifetime τ of the π_v^0 candidate on a log logarithmic scale. Displayed are the Υ and Z samples from data and 7 and 50 GeV/c^2, 10 ps π_v^0 simulation samples, scaled to the Z data sample for comparison. 2012 Up (simulation) conditions.

6.3.1 ROC curves

In order to determine how good these variables are at separating signal from background, so-called Receiver Operating Characteristic (ROC) curves are compared. These are constructed by calculating the efficiency on the signal and rejection of the background for different values of the criterion. The efficiency on the π_v^0 signal for a selection on d is calculated as:

$$\epsilon_{\pi_v^0}(d < 0.5) = \frac{N_{\pi_v^0, d<0.5}}{N_{\pi_v^0,\mathrm{tot}}}.$$

Here, $\epsilon_{\pi_v^0}(d < 0.5)$ is the efficiency of the $d < 0.5$ requirement on the π_v^0 sample, $N_{\pi_v^0, d<0.5}$ is the number of events in which the π_v^0 has a displacement smaller than 0.5 and $N_{\pi_v^0,\mathrm{tot}}$ is the number of events without a displacement requirement. These efficiencies are a measure of the performance of the selection criterion, it is the fraction of π_v^0 particles that survives the selection. It is calculated for different values of d, τ and ρ. The rejection of background events is defined as:

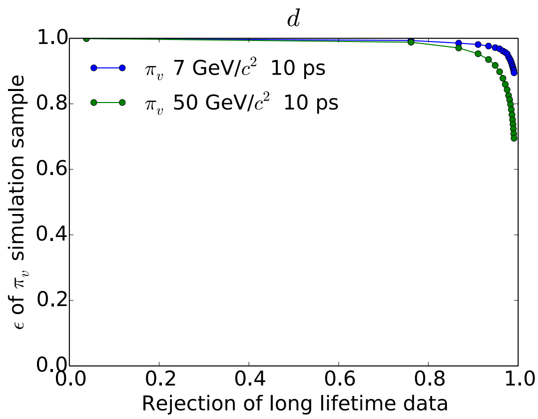
$$\mathrm{Rejection}_{\mathrm{BG}}(d < 0.5) = 1 - \epsilon_{\mathrm{BG}}(d < 0.5) = 1 - \frac{N_{\mathrm{BG}, d<0.5}}{N_{\mathrm{BG,tot}}}.$$

The ROC curve is constructed by plotting the efficiency on signal on the y-axis and the rejection of background on the x-axis for different values of the variable. The ROC curve of an ideally separating variable would be a stepfunction that steps down from one to zero at x=1. This indicates that no signal efficiency is lost when the background rejection is increased.

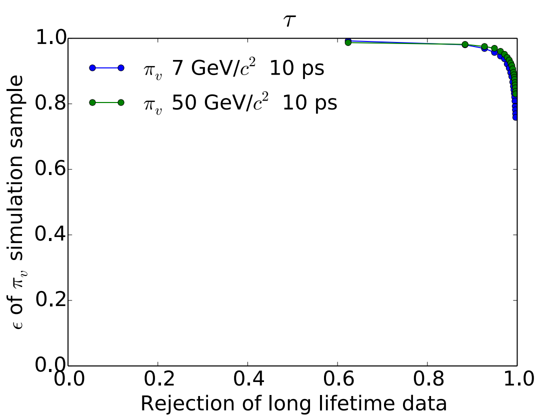
The ROC curves for the variables described in this section, are displayed in Figure 19. The rejection is calculated on the data sample, as most events are expected to be background events. A random selection is made out of the data events in order to do a 'blind search'. This prevents an over-optimisation on the data, which would never lead to a discovery. The amount of selected data corresponds to an integrated luminosity of 200 pb^{-1} (a measure of the number of proton proton collisions).

Both the efficiency on signal and rejection of background are high for all three variables, so they are suitable selection criteria. This was expected, as it was shown in Chapter 4 that a large fraction of the π_v^0 candidates are Z bosons or Υ particles. d and ρ are indeed mass dependent and the efficiency on signal is less good for the 50 GeV/c^2 sample. Therefore τ is chosen as a selection variable.

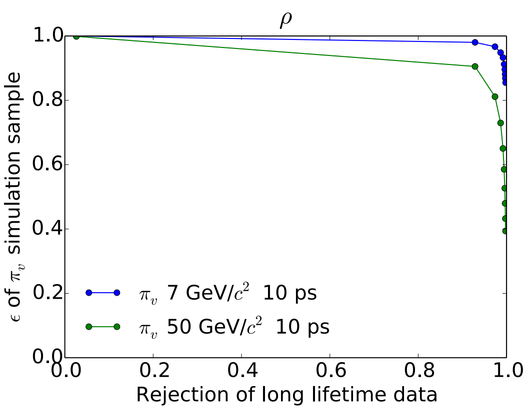
Figure 20b shows the distribution of τ on a normal scale, from which can be concluded that the Υ and Z backgrounds will be mostly suppressed when $\tau > 0.2$ ps is required. The efficiency of the lifetime selection criterion on different samples is plotted in Figure 20a. For $c\tau > 0.2$ ps, the efficiency on the π_v^0 samples is about 97% , whilst the efficiency on data is less than 10 %. It is a loose selection criterion and can later be made more stringent.



(a)

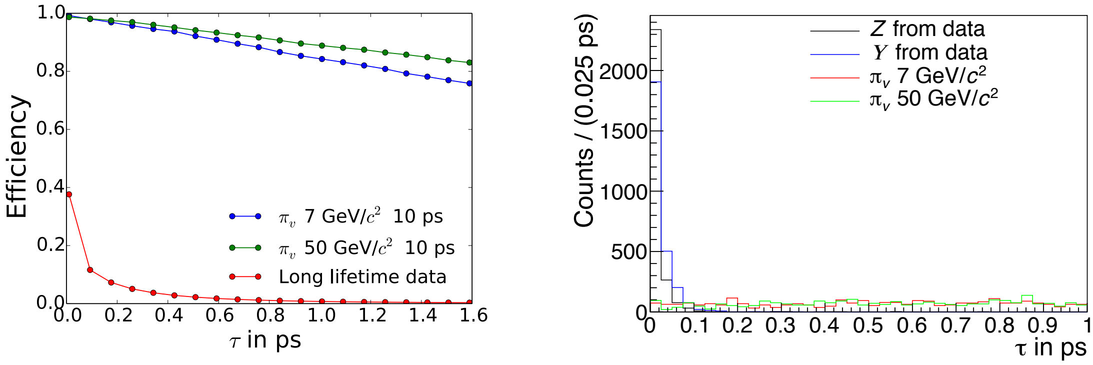


(b)



(c)

Figure 19: ROC curves of (a) displacement d (distance from best PV to DV), (b) lifetime τ of the π_v^0 candidate and (c) radial coördinate ρ of the DV. The efficiency on the 10 ps π_v^0 simulation samples is on the y-axis, the rejection of data corresponding to 200 pb^{-1} on the x-axis. 2012 Up (simulation) conditions.



(a)

(b)

Figure 20: (a) Efficiencies of the requirement $\tau >$ on data (200 pb^{-1}) and 10 ps π_v^0 simulation samples. (b) Lifetime distribution of Υ and Z in data (200 pb^{-1}) and 10 ps π_v^0 simulation samples, on a linear scale. 2012 Up (simulation) conditions.

6.4 Additional selection

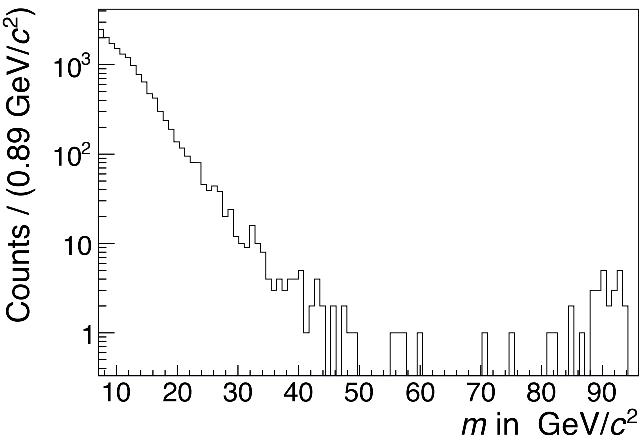


Figure 21: Invariant mass distribution of the data (200 pb^{-1}) with lifetime requirement $\tau >$ 0.2 ps, 2012 Up sample.

In Figure 21, the data is displayed with the above described loose requirement on the lifetime. The Υ and Z backgrounds are clearly almost all suppressed. In Chapter 4, it was shown that the remaining data mostly has the properties of a background with muons coming from two different vertices. There are many differences between these decays and signal decays. For example, the muons from these background decays will have more particles in a cone around them, because of the jets that accompany these muons. In the next sections, various criteria are discussed. They are grouped in categories, the displaced vertex, lifetime, pointing, isolation and jets selection criteria.

Different separating variables are discussed for each catagory and ROC curves are used to show which variable performs the best. The rejection of backgrounds on the x-axis is evaluated on data (200 pb^{-1}) with the loose lifetime selection $\tau > 0.2\,\mathrm{ps}$. This allows a study of $b\bar{b}$-like backgrounds with enough statistics to make a well-informed selection. The efficiency on signal on the y-axis of the ROC curves is evaluated on the 7 and 50 GeV/c^2 π_v^0 simulation samples with the same lifetime selection. Both samples are used to check the mass dependency.

The software TMVA (Tool for MultiVariate Analyses) is used as a complementary method to determine the best separating variable in each catagory [27]. It uses a π_v^0 simulation sample, a background sample and possibly separating variables as input and returns the (combination of) variables that are the best selection criteria. Different methods can be called, this analysis uses the Boosted Decision Tree (BDT) method. The background input is data (200 pb^{-1}) with the loose lifetime selection. The signal input is the π_v^0 simulation sample with mass 10 GeV/c^2, a lifetime of 10 ps and the same lifetime criterion as on data.

6.4.1 Displaced vertex quality

The muons from the $b\bar{b}$-like background decays originate in two different vertices. This implies that the quality of the reconstructed DV is not good.

Different variables can be used to distinguish between this scenario and the π_v^0 signal, the first being the distance of closest approach (DOCA) between the two muons. This is the smallest distance between the two reconstructed muon tracks and is expected to be larger for $b\bar{b}$-like data. A graph of the DOCA for π_v^0 samples and data with a loose lifetime selection is displayed in Figure 22. It is clear that it is larger for the long lifetime data.

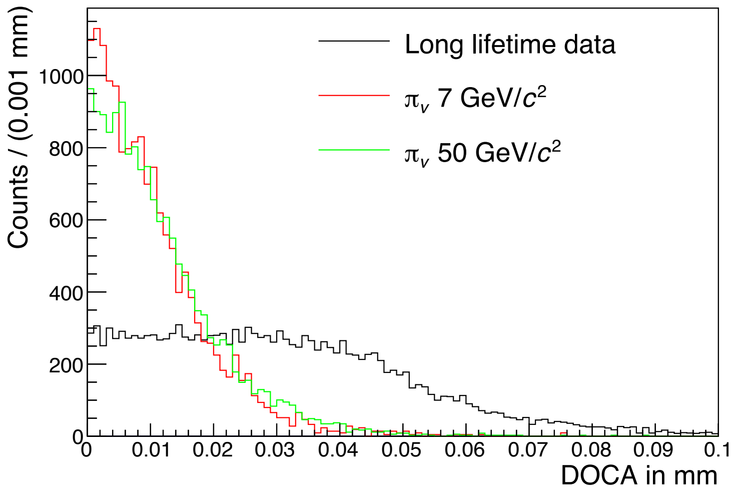


Figure 22: Distance of closest approach (DOCA) between the two muons, for two 10 ps π_v^0 samples and 200 pb^{-1} data with loose lifetime selection. 2012 Up (simulation) conditions, samples scaled to data.

The IP-χ^2 from the muons to the DV is also a possible selection criterion. It was introduced in Section 5.2 and is a measure of the quality of the DV, that does not depend on the errors in calculating it.

Events with muons originating in one vertex can still have a large DOCA when the error in evaluating the DOCA is large. IP-χ^2 does not discard such events. For a smaller mass, the error on calculating the DOCA is larger. This has been checked on π_v^0 simulation samples. By using IP-χ^2 as a criterion, events that only would have been discarded because of their large error, still pass the selection. The variable gets less mass dependent, which is preferable for this analysis.

There are also downsides to using IP-χ^2 as a criterion. The errors for the π_v^0 samples are all very similar and a large error indicates that the event most likely did not contain a π_v^0 particle. Then a correction for for the error by using IP-χ^2 is unwanted.

In Figure 23, the ROC curves can be found for both variables. It can be concluded that the DOCA is the best separating variable as it rejects more data for the same signal efficiency. It should be noted that an IP-χ^2 > 12 selection was already made in the preselection. This already discards of a lot of the long lifetime data and influences which variable is now better at separating signal and background. It does not influence the conclusion that DOCA is a better separating variable to apply after the preselection.

In Figure 24a, the efficiencies of the different π_v^0 samples and the long lifetime data sample can be found for the different values of DOCA. A loose selection criterion is determined to be DOCA < 0.05 mm. The efficiency on π_v^0 simulation samples is 99% for this value, the rejection of long lifetime data is 10 %. In Figure 24b, the 'S over B' curve is displayed for the DOCA. This is a graph of the quantity:

$$\frac{N_{\mathrm{S}}}{\sqrt{N_{\mathrm{BG}}+1}} = \frac{N_{\pi_v^0}(\mathrm{DOCA} < x)}{\sqrt{N_{\mathrm{data}}(\mathrm{DOCA} < x)+1}},$$

where $N_{\pi_v^0}(\mathrm{DOCA} < x)$ is the number of π_v^0 particles that have DOCA $< x$ and $N_{\mathrm{data}}(\mathrm{DOCA} < x)$ is the number of long lifetime data events with DOCA $< x$. It is a measure of the performance of the selection variable DOCA. If it separates well, the number of remaining signal events is large and number of background events low, so the quantity is large.

TMVA also returned DOCA as the best separating variable.

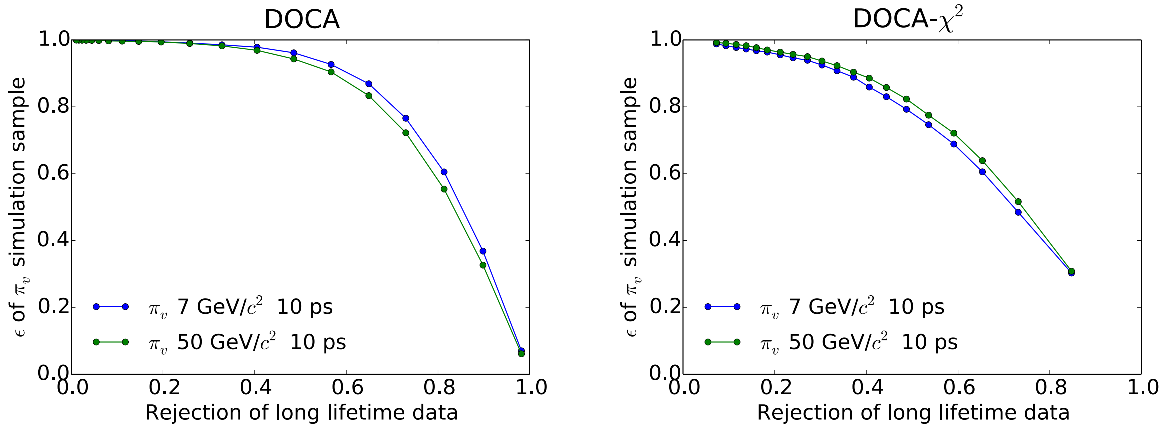
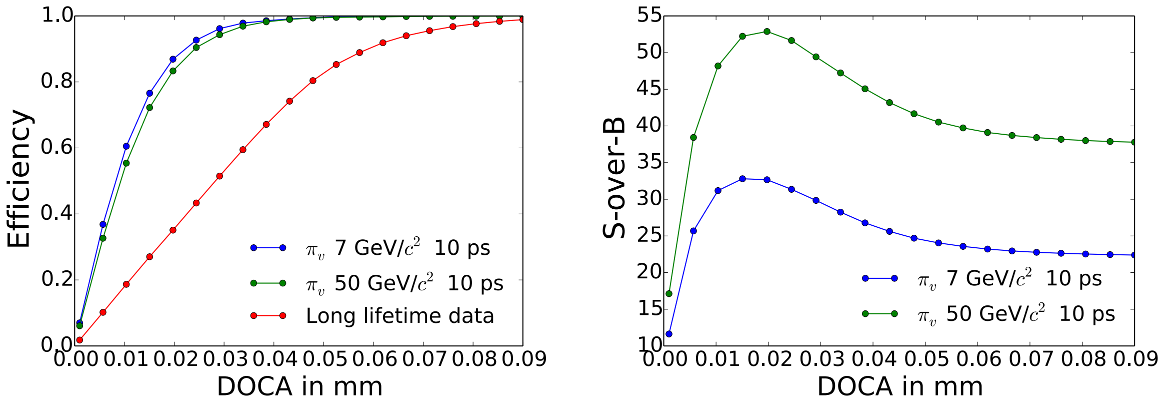


Figure 23: ROC curves for the DOCA and DOCA$-\chi^2$ with efficiency on two 10 ps π_v^0 simulation samples on the y-axis and rejection on 200 pb^{-1} 2012 Up data on the x-axis, both with loose lifetime selection $\tau >$ 0.2 ps.



(a) (b)

Figure 24: Efficiencies and 'S over B' curve for the DOCA for two 10 ps π_v^0 simulation samples and the 200 pb^{-1} 2012 Up data sample, all with loose lifetime selection τ >0.2 ps .

6.4.2 Lifetime

As the $b\bar{b}$-like background decays have other properties than Υ and Z decays, another lifetime variable might be more suitable to suppress them. After applying the requirement for τ, the separating lifetime variables discussed in Section 6.3 are applied again.

An additional separating variable that depends strongly on the lifetime of the π_v^0 candidate is θ_{DIRA}. This is the angle between the displacement vector $\vec{\mathbf{d}}$ (the vector between the PV and DV) and the momentum $\vec{\mathbf{p}}$ of the π_v^0 candidate. From Figure 25, it is observed that this angle is much larger for the long lifetime data than for the π_v^0 simulation samples.

The ROC curves of all of these variables are displayed in Figures 26 and 27. These show that θ_{DIRA} separates the best, although the τ can also be made more stringent without losing much signal efficiency.

The graph of the efficiency of selecting candidates with different θ_{DIRA} values, and the 'S over B' curve for this variable, are displayed in Figure 28. Thus a loose selection on this variable would be to require $\theta_{\text{DIRA}} < 0.05$. The efficiency on π_v^0 samples is then higher than 0.95, while the efficiency on data is less than 0.40. It should not be made smaller than about 0.005, the π_v^0 efficiency drops significantly there, as expected from the θ_{DIRA} distribution in Figure 25. From the 'S-over-B' curve, it can be observed that lower values of $\theta_{\text{DIRA}} < 0.05$ will perform even better, but this is an initial loose criterion.

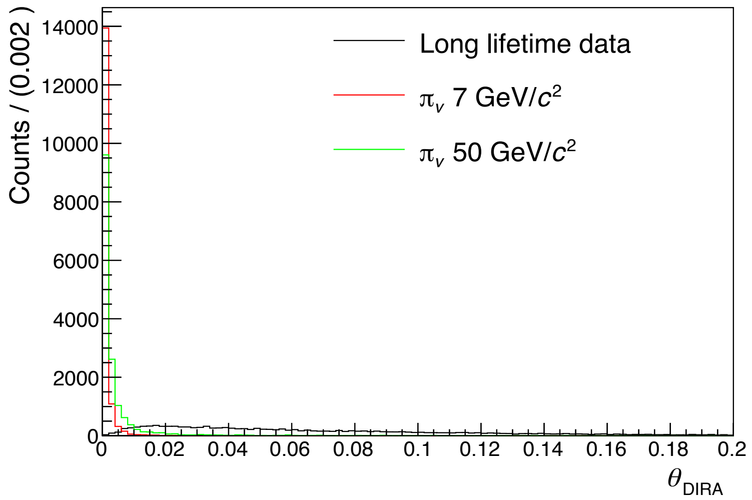
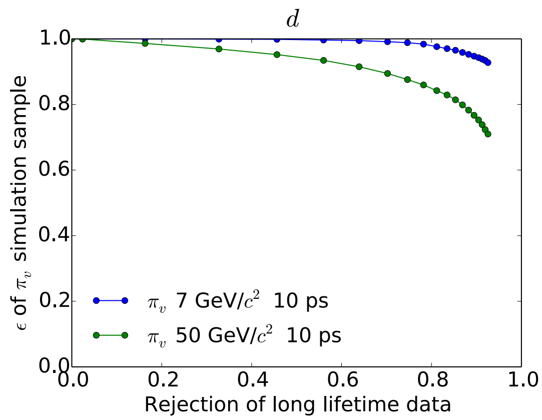
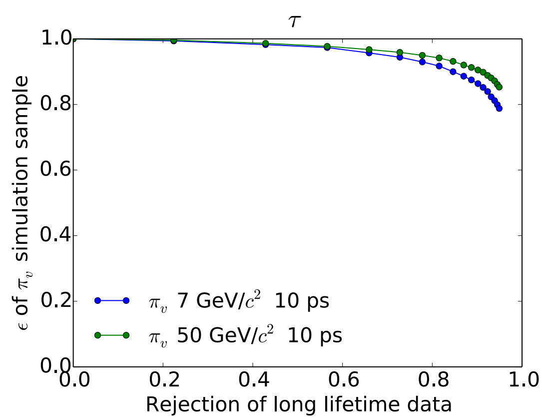


Figure 25: θ_{DIRA} for two 10 ps π_v^0 samples and 200 pb^{-1} data with loose lifetime selection. 2012 Up (simulation) conditions, samples are scaled to data.

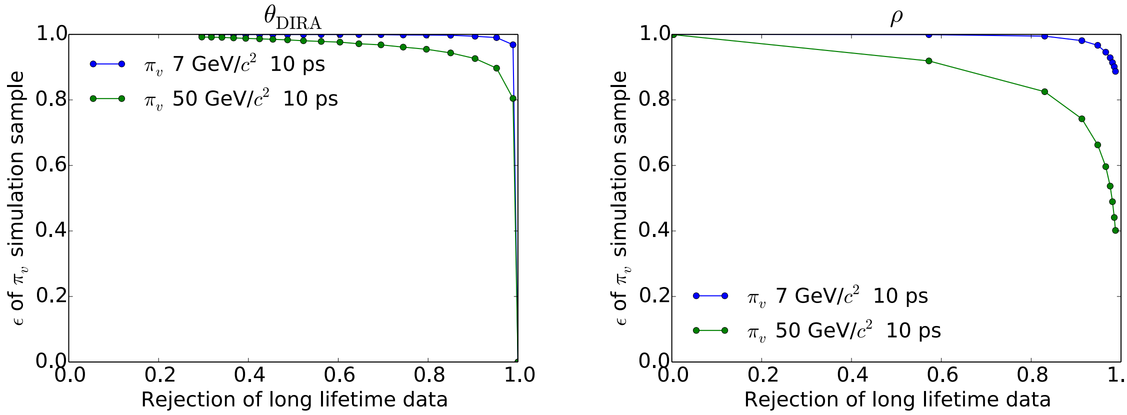


(a)



(b)

Figure 26: ROC curves for (a) the displacement d (distance from best PV to DV) and (b) lifetime τ of the π_v^0 candidate. The efficiency on two 10 ps π_v^0 simulation samples on the y-axis and rejection on 200 pb^{-1} 2012 Up data on the x-axis, both with loose lifetime selection $\tau > 0.2\,\mathrm{ps}$.



(a) (b)

Figure 27: ROC curves for (a) θ_{DIRA} and (b) radial coördinate ρ of the DV. The efficiency on two 10 ps π_v^0 simulation samples on the y-axis and rejection on 200 pb^{-1} 2012 Up data on the x-axis, both with loose lifetime selection $\tau > 0.2\,\mathrm{ps}$.

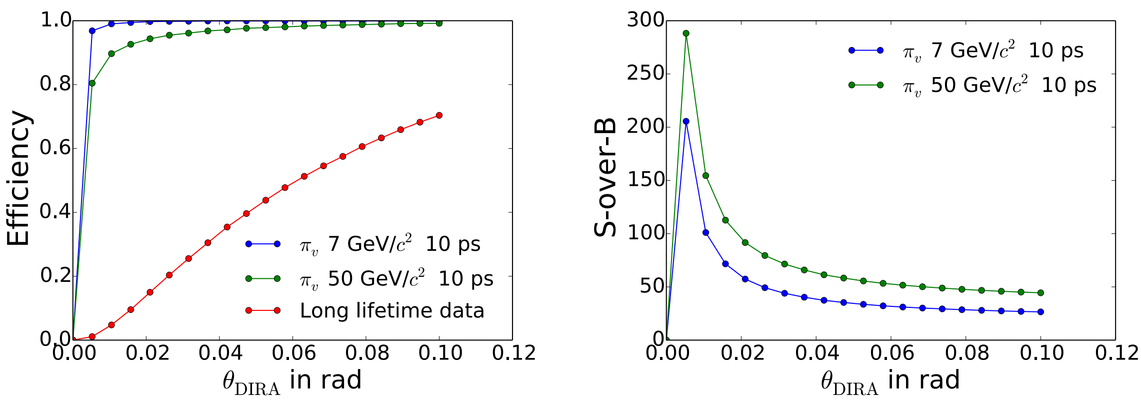


Figure 28: Efficiencies and 'S over B' curve for the θ_{DIRA} for two 10 ps π_v^0 simulation samples and the 200 pb^{-1} 2012 Up data sample, all with loose lifetime selection $\tau > 0.2\,\mathrm{ps}$. 2012 Up simulation conditions.

6.4.3 Pointing

Another difference between the $b\bar{b}$-like background and signal, is that a π_v^0 candidate that was reconstructed from two muons from different vertices, will not 'point back' to the PV very well. The Impact Parameter (IP) is a possible variable, it is the distance from the π_v^0 candidate track to the PV. As can be seen from Figure 29, this is large for the $b\bar{b}$-like background and small for the π_v^0 sample. The IP-χ^2 of the PV can be used to make the variable independent of the error in calculating it. Next to that, the θ_{DIRA} is also a measure of pointing.

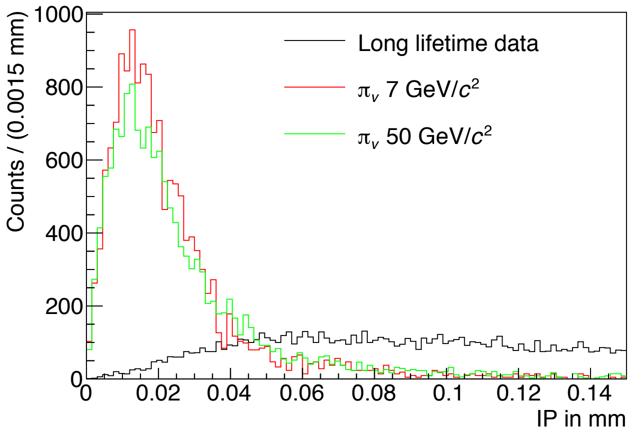


Figure 29: Impact Parameter (IP) of the π_v^0 candidate to the best PV for two 10 ps π_v^0 samples and 200 pb^{-1} data with loose lifetime selection. 2012 Up (simulation) conditions, samples are scaled to the long lifetime data.

In Figure 30, the ROC curves for IP and IP-χ^2 are displayed. Comparing to 27 shows that θ_{DIRA} is the variable that separates the best. A loose selection criteria was already determined in the previous section. The IP and IP-χ^2 do not have a very good efficiency on the π_v^0 simulation samples and are not used as selection criteria.

These variables are used as input for TMVA, which returned θ_{DIRA} as the best separating variable.

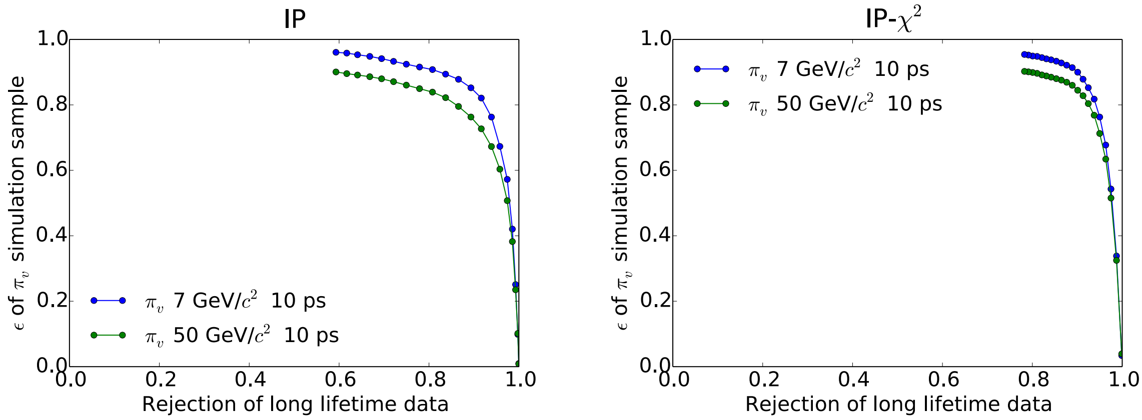


Figure 30: ROC curves for the IP and IP-χ^2 with efficiency on two 10 ps π_v^0 simulation samples on the y-axis and rejection on 200 pb^{-1} 2012 Up data on the x-axis, both with loose lifetime selection $\tau > 0.2$ ps.

6.4.4 Isolation

A b quark hadronises to a jet, so on average more particles are found in the neighbourhood of the muon when the event was in fact a $b\bar{b}$ decay. By looking at the particles in a cone around the muon, separating variables can be defined. It is decided that a particle track is part of the cone by calculating the $\Delta\phi$ and $\Delta\eta$ between the particle track and muon track. ϕ is the angle in the plane transverse to the beamline and the pseudorapidity $\eta = -\ln[\tan(\theta/2)]$, with θ the angle between the momentum and the positive z-axis. The cone is then described by tracks with $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.7$.

This is performed in two different ways, the first method uses all particles in the cone. The second uses the input for the jet algorithm, which is slightly different from all particles, e.g. it disregards particles if their quality is not good enough and particles that are erroneously reconstructed twice.

A possible variable is the number of particles in the cones. However, this is not a very well defined quantity, low energy gluons and photons are produced often and it is hard to predict in what quantities. Therefore, a better motivated variable is the sum of the momenta of all the particles in the cone. This way, the influence of low energy particles is reduced. Events can be selected for which this quantity is low. As the p_T of the muon is in the sum, this would mean that events would sooner be suppressed for high p_T muons. This is an unwanted effect, so the separating variable is defined by dividing by the p_T:

$$\sum_{p_\mathrm{T}} = \frac{\sum_i (p_\mathrm{T})_i}{(p_\mathrm{T})_\mu}.$$

This variable can be computed in a number of different ways. The vectorial or scalar sum of the momenta can be taken, p_T or p of the muons can be used and all particles or just the neutral particles can be included. The p_{T} is expected to separate better, as this is not biased by the forward momentum.

In Figure 31, the number of neutral particles N^0 in a cone around the μ^- is displayed, as well as $\sum_{(p_{\mathrm{T}})^0}$, the above defined scalar sum for only the neutral particles. It is clear that the sum of momenta distinguishes better between signal and background, the distributions of N^0 are too similar to work as a separating variable.

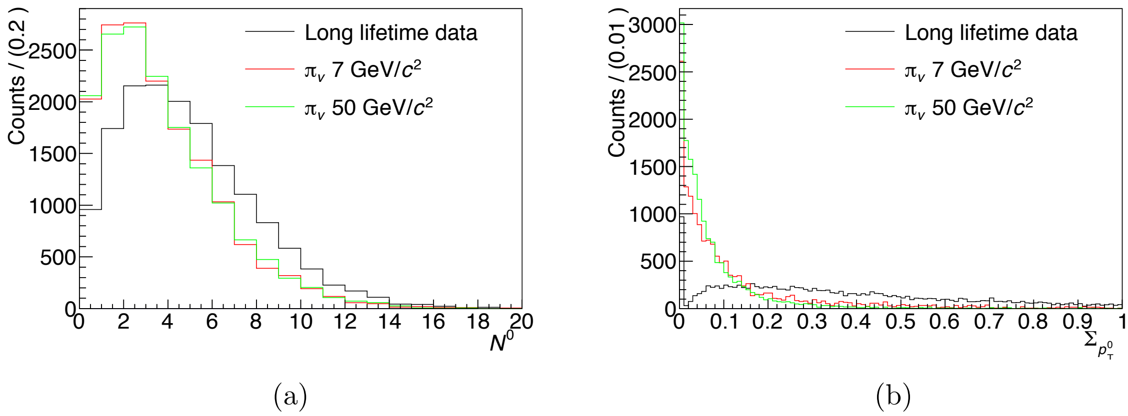
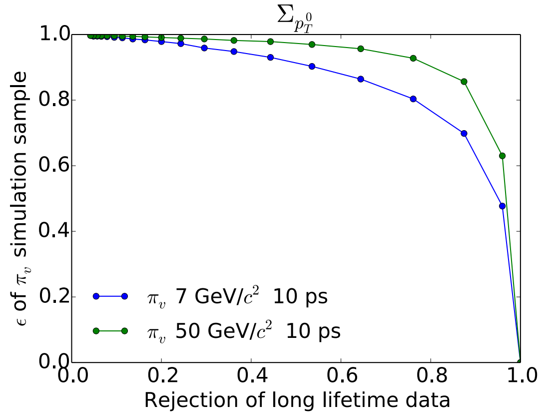


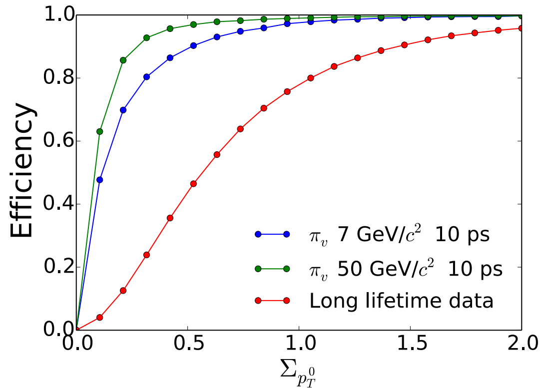
Figure 31: Number of neutral particles N^0 and scalar sum of transverse momentum p_{T} of neutral particles, scaled by the p_{T} of the muon, in a cone of $\Delta R = 0.7$ around the μ^-. For two 10 ps π_v^0 samples and 200 pb^{-1} data with loose lifetime selection. The input from the jet algorithm is used and the samples are scaled to the long lifetime data.

Each variable is calculated for both muons and a requirement is put on the maximum of these values. ROC curves proved that the particles that are the input for the jet algorithm, provide better distinguishing variables. As predicted, the sum over p_{T} distinguishes better than p. The scalar sum proved to separate the best. The ROC curve for $\sum_{(p_{\mathrm{T}})^0}$ and the efficiencies of the different samples for different requirements on this variable are displayed in Figure 32. From the latter graph, it can be concluded that close attention should be paid when this variable is used as a selection criterion. As there are certainly also particles around the muons originating in π_v^0 particles, this variable quickly reduces the efficiency on the signal sample. A loose selection criterion would be $\sum_{(p_{\mathrm{T}})^0} < 1.0$.

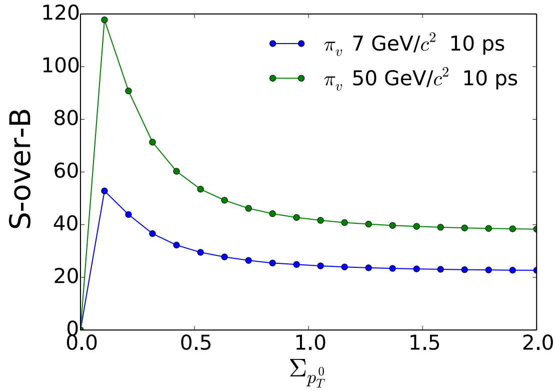
TMVA with all of these variables as input, returns $\sum_{p^0}$ as the best separating variable. The difference with $\sum_{(p_{\mathrm{T}})^0}$ is small and this variable is used as it has been proved with ROC curves that it performs well.



(a)



(b)



(c)

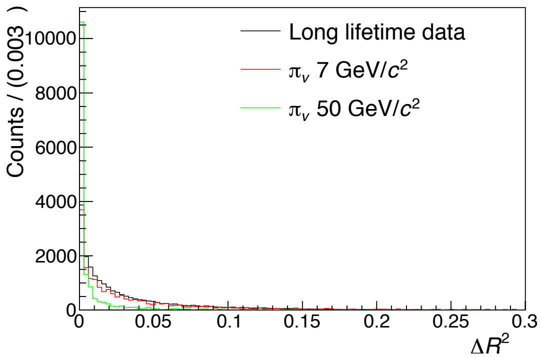
Figure 32: (a) ROC curve for scalar sum of transverse momentum p_T of neutral particles, scaled by p_T of muon, in a cone of $\Delta R = 0.7$ with the efficiency on two π_v^0 simulation samples on the y-axis and rejection on 200 pb^{-1} 2012 Up data on the x-axis, both with loose lifetime selection $\tau > 0.2$ ps. (b) Efficiencies on these samples as a function of selection values for this variable. (c) 'S over B' curve.

6.4.5 Jets

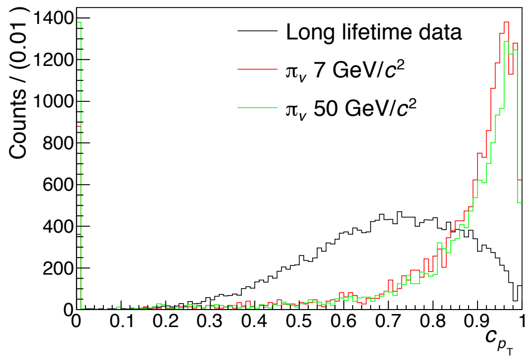
Lastly, the jet from the jet reconstruction can be identified that is matched with one of the muons. Its properties could provide separating variables. The distance between the jet and the muon is determined by $(\Delta R)^2 = (\Delta\phi)^2 + (\Delta\eta)^2$ (calculated between their momenta) and the number of tracks in the jet are counted.

The $c p_\mathrm{T}$ could also be separating, which is the fraction of charged p_T in the jet with respect to the total p_T of both charged and neutral particles. It is reasoned that as there are less particles around the muon coming from a π_v^0, most of the p_T in the associated jet will come from the muon itself (which is charged). However, the jet algorithm is complicated and it is hard to predict what the distributions will look like. The distributions of c_{p_T} and $(\Delta R)^2$ are displayed in Figure 33. A large fraction of the π_v^0 events do not have identified jets, these have $c_{p_\mathrm{T}} = 0$ and a non-defined $(\Delta R)^2$, the former can be seen in the graph. It is observed that the distribution of $(\Delta R)^2$ for long lifetime data and π_v^0 simulation are not very different and the number of tracks (not displayed here) also does not have good separating power.

By requiring that the event either has no jet at all, or has a large c_{p_T}, ROC curves and efficiency graphs are constructed. They are displayed in 34. It can be concluded that $c_{p_\mathrm{T}} > 0.5$ is a loose selection criterion, with a signal efficiency of nearly one and a background rejection of less than 0.9.

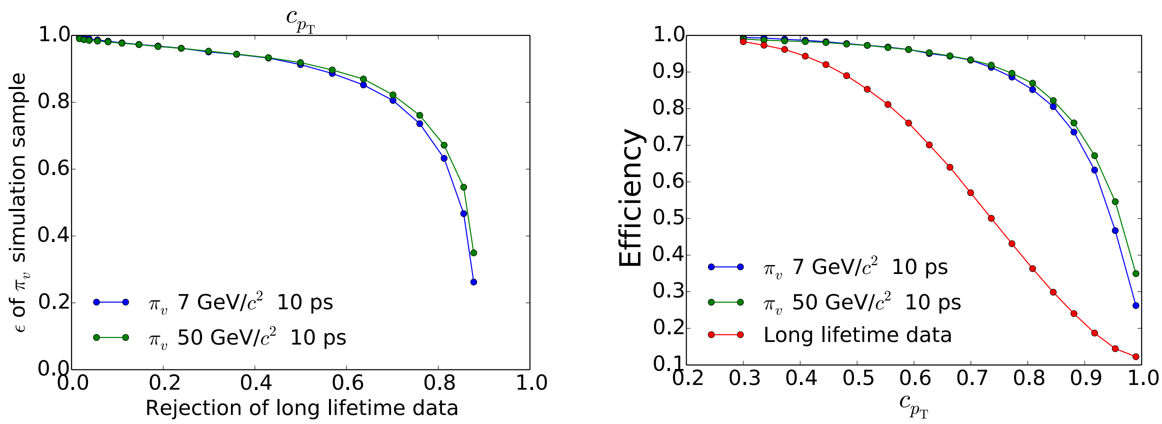


(a)



(b)

Figure 33: (a) Distance $(\Delta R)^2$ between the momenta of muon and matching jet and (b) the fraction c_{p_T} of charged p_T in the matching jet for two 10 ps π_v^0 samples and 200 pb^{-1} data with loose lifetime selection. Scaled to long lifetime data.



(a)

(b)

Figure 34: (a) ROC curve for the fraction $c_{p_{\mathrm{T}}}$ of charged p_{T} in the matching jet. The efficiency on two 10 ps π_v^0 simulation samples on the y-axis and rejection on 200 pb^{-1} 2012 Up data on the x-axis, both with loose lifetime selection $\tau > 0.2$ ps. (b) Efficiencies on these samples as a function of selection values for $c_{p_{\mathrm{T}}}$.

6.5 Overview

The variables that were determined to be best separating in the above sections and their loose selection values are listed in Table 10. In Figure 35, the effect of these selections is shown by consecutively applying them on the data. Especially the lifetime selection discards of a lot of background, which was expected as a large fraction of the data was shown to be short living Υ and Z background decays. Many background events remain; the selection needs to be optimised.

Table 10: Loose selection criteria.

Category	**Variable**	**Selection value**
Displaced vertex	DOCA	< 0.05 mm
Isolation	$\sum_i (p_{\mathrm{T}})_i^0/(p_{\mathrm{T}})_\mu$	< 1.0
Lifetime	τ	> 0.2 ps
Pointing	θ_{DIRA}	< 0.05 rad
Jets	$c_{p_{\mathrm{T}}}$	> 0.5

It has to be checked that the best separating variable in each category does not depend on whether or not a selection has already been made in the other categories, as the variables might be correlated. The above described loose selection criteria are applied except for one category, and in this category it is checked if the same variable still separates the best. This is done for all categories and the found ROC curves are very similar to the ones without additional selection criteria. This indicates that the categories are reasonably

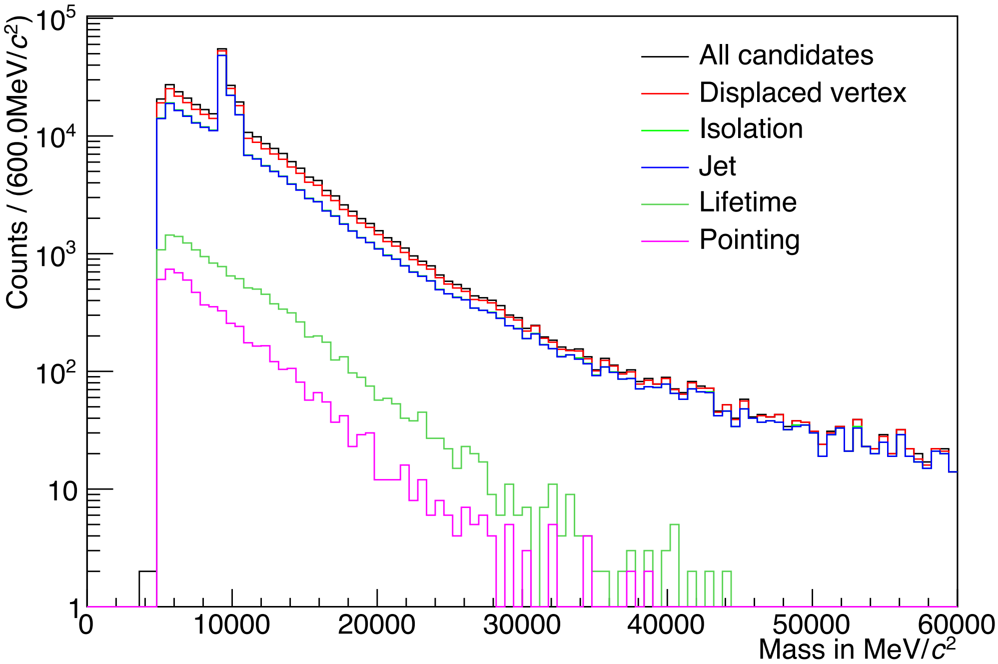


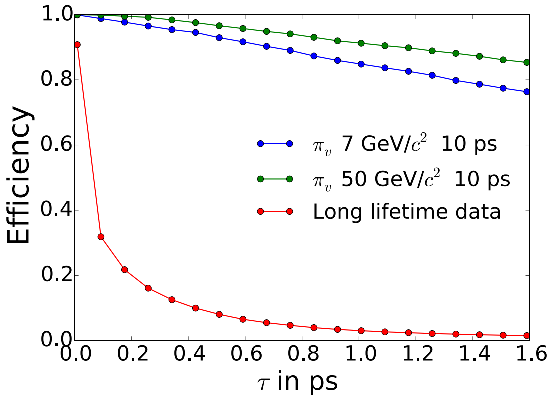
Figure 35: Invariant mass distribution of the data (200 pb^{-1}) after consecutive loose selection criteria are applied, see Table 10. Starting with all π_v^0 candidates after trigger and preselection. 2012 Up conditions.

uncorrelated. It also proves that the procedure of the above section provides good selection variables.

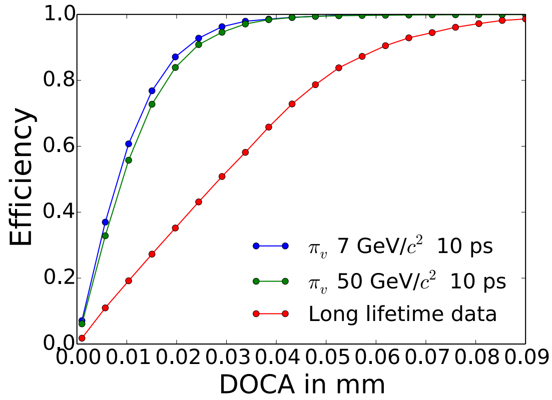
The efficiency and 'S-over-B' curves with all loose selection criteria applied (except in the catagory under investigation) can be found in Figure 36 and Figure 37. DOCA and θ_{DIRA} appear to separate the best, they are quite mass independent, discard a lot of background and retain a high signal efficiency. From these graphs, the final selection is determined, listed in Table 11.

Table 11: Final selection criteria.

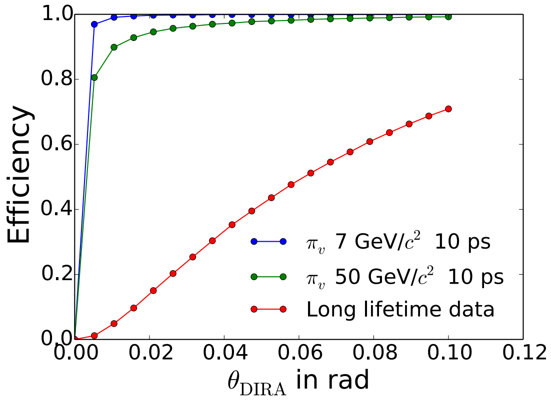
Catagory	**Variable**	**Selection value**
Displaced vertex	DOCA	< 0.03 mm
Isolation	$\sum_i (p_{\mathrm{T}})_i^0/(p_{\mathrm{T}})_\mu$	< 0.8
Lifetime	τ	> 2.7 ps
Pointing	θ_{DIRA}	$< 7 \times 10^{-3}$ rad
Jets	$c_{p_{\mathrm{T}}}$	> 0.7

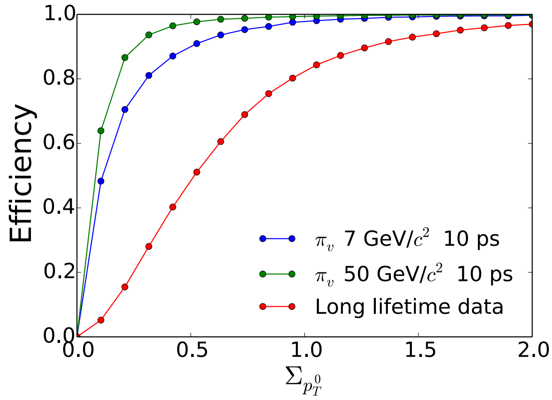


(a)



(b)





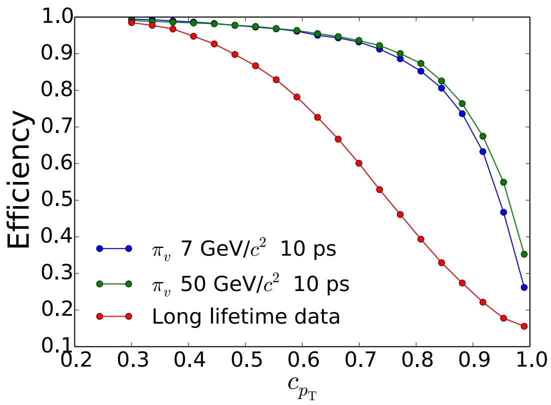


Figure 36: Efficiencies for the best separating variable in each catagory for two 10 ps π_v^0 simulation samples and the 200 pb^{-1} 2012 Up data sample, all with the loose selection criteria listed in Table 10.

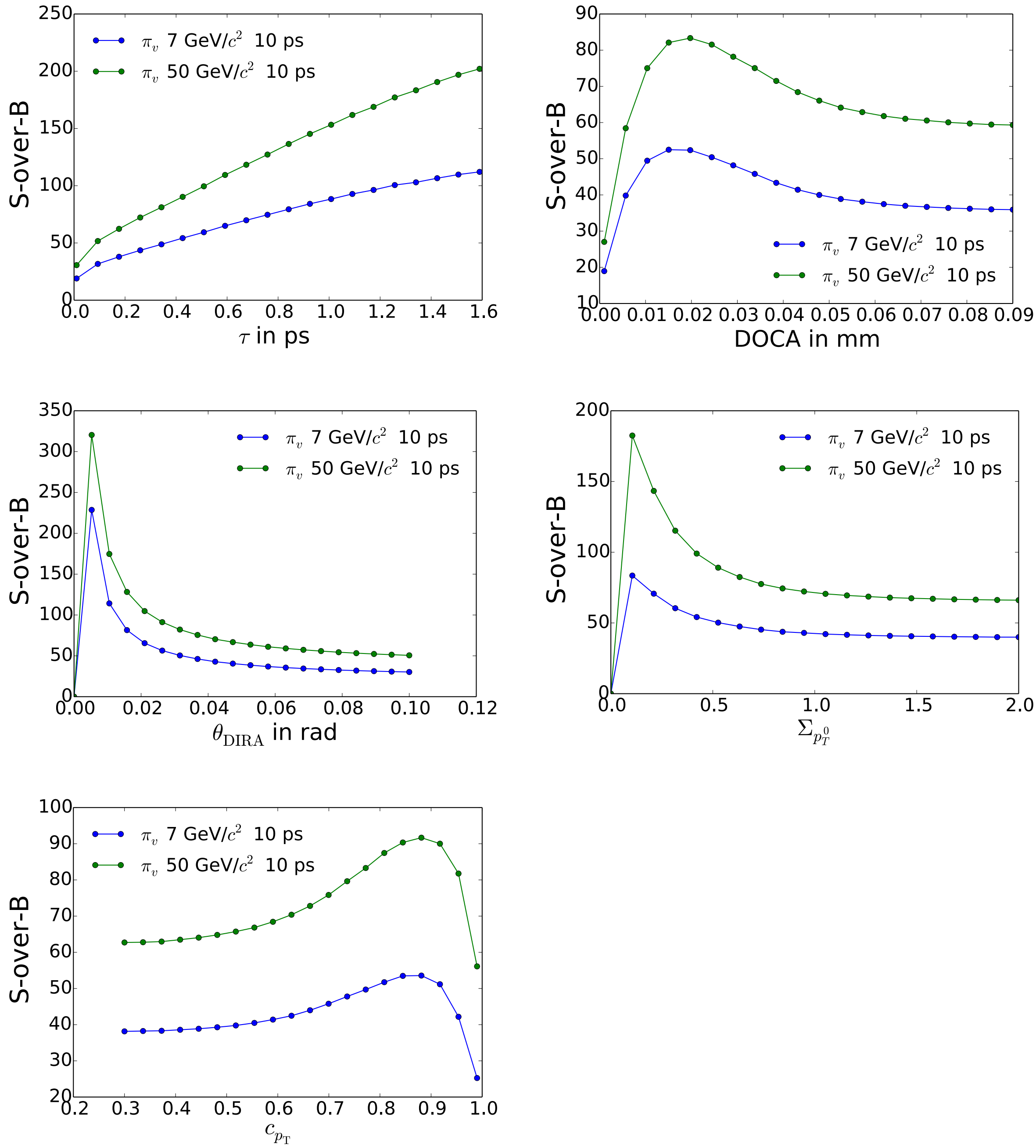


Figure 37: 'S over B' curves for the best separating variable in each catagory for two 10 ps π_v^0 simulation samples and the 200 pb^{-1} 2012 Up data sample, all with the loose selection criteria listed in Table 10.

6.6 Result of selection

The selection criteria have been determined on a subset of the data. In Figure 38, the effect of the selection on all unblinded data is displayed.

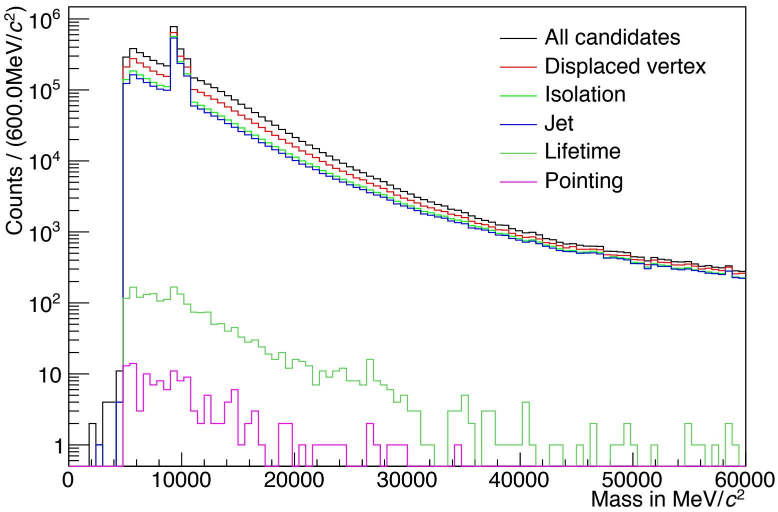


Figure 38: Invariant mass distribution of all data (2011 and 2012, Up and Down) after the final selection criteria are consecutively applied, see Table 11. Starting with all π_v^0 candidates after trigger and preselection, on a logarithmic scale.

The selection discarded of almost all data, the remaining number of π_v^0 candidates is 137. No evidence of the π_v^0 signal decay is found. Therefore, a limit will be set on the production rate.

Next to that, it can be concluded from the graph that the selection criterion on jets does not remove a significant part of the data. The selection criteria motivated by properties of muons from two separate DV's were applied in different orders. It was found that the last applied criterion did not suppress a significant amount of background decays. The criteria are complementary and in future research one of them can be omitted to reduce the systematic error.

The efficiency of the selection on the π_v^0 simulation samples is listed in Tables 12 and 13. The cumulative efficiency is calculated by dividing the number of candidates after a selection step by the number total number of events (generated events divided by generator efficiency). For completeness, the generator, trigger and preselection efficiencies are also listed. The cumulative efficiency of the selection described in this chapter lies between 0.53 and 0.74, this is necessary in order to suppress all backgrounds. It is higher for 100 ps lifetimes and higher masses, because the pointing criterion performs better in those cases.

The overall cumulative efficiency is worse for long lifetimes. As pointed out in Section 5.1, this is due to π_v^0 particles that fly out of the VELO, which results in a low HLT 1 trigger efficiency.

In Appendix C, the efficiencies on the π_v^0 simulation samples are compared for 2011 and 2012, Up and Down simulation conditions. The differences are small, because the simulation conditions are not very different. The effect on the efficiency still needed to be checked.

6.6.1 Multiple candidates

It is possible that multiple $\mu^+\mu^-$ combinations pass the preselection requirements for one event. All passing combinations are stored, so some events are stored multiple times. The Tables 12 and 13 list the number of events in the upper half of the table and the number of candidates in the lower half. The cumulative efficiency is evaluated by dividing the number of candidates by the initial number of events. Equation 1 then correctly returns the number of signal decays after selection. The fraction of the events that survive the selection and have two candidates, ranges from about 8% for 10 ps 50 GeV/c^2 samples to about 0.5% for 100 ps 7 GeV/c^2 samples.

Table 12: Number of events (upper half) or candidates (lower half), relative efficiency and cumulative efficiency on π_v^0 simulation samples. Displayed are different steps: the generator efficiency, trigger efficiencies, preselection and the final selection. 2012 Up simulation conditions.

	7 GeV/c², 10 ps			**50 GeV/c², 10 ps**		
	N	ϵ_{rel}	ϵ_{cum}	N	ϵ_{rel}	ϵ_{cum}
After generator	11435	0.2212	0.2212	11325	0.0966	0.0966
L0	10319	0.9024	0.1996	11007	0.9719	0.0939
Hlt1	5573	0.5401	0.1078	9417	0.8555	0.0803
Hlt2	3935	0.7061	0.0761	5779	0.6137	0.0493
After preselection	3485	0.8856	0.0674	5819	1.0069	0.0496
Displaced vertex	3362	0.9647	0.0650	5509	0.9467	0.0470
Isolation	3213	0.9557	0.0622	5428	0.9853	0.0463
Jets	3096	0.9636	0.0599	4634	0.8537	0.0395
Lifetime	3028	0.9780	0.0586	4506	0.9724	0.0384
Pointing	1850	0.6110	0.0358	3634	0.8065	0.0310
$\epsilon_{\mathbf{cum}}$ **Selection**			0.5308			0.6183

Table 13: Number of events (upper half) or candidates (lower half), relative efficiency and cumulative efficiency on π_v^0 simulation samples. Displayed are different steps: the generator efficiency, trigger efficiencies, preselection and the final selection. 2012 Up simulation conditions.

	7 GeV/c^2, 100 ps			**50 GeV/c^2, 100 ps**		
	N	ϵ_{rel}	ϵ_{cum}	N	ϵ_{rel}	ϵ_{cum}
After generator	11435	0.2203	0.2203	11899	0.0961	0.0961
L0	10313	0.9019	0.1987	11562	0.9717	0.0934
Hlt1	1184	0.1148	0.0228	5403	0.4673	0.0436
Hlt2	689	0.5819	0.0133	2250	0.4164	0.0182
After preselection	552	0.8012	0.0106	1985	0.8822	0.0160
Displaced vertex	529	0.9583	0.0102	1846	0.9300	0.0149
Isolation	509	0.9622	0.0098	1809	0.9800	0.0146
Jets	493	0.9686	0.0095	1654	0.9143	0.0134
Lifetime	482	0.9777	0.0093	1594	0.9637	0.0129
Pointing	379	0.7863	0.0073	1479	0.9279	0.0119
$\epsilon_{\mathbf{cum}}$ **Selection**			0.6866			0.7390

7 Systematic uncertainties

Multiple sources of systematic uncertainties arise in the limit setting procedure. In this Chapter, an overview is given and the uncertainties are either neglected or an evaluation method is suggested. Due to time constraints, the errors have not been calculated and an estimation of the total systematic uncertainty is given.

In Chapter 8, it is described how the number of candidates after the selection translates to a limit on the number of π_v^0 decays. For this procedure, an estimate of the number of background decays is needed. The limit translates to a limit on the production rate by Equation 1, for which the luminosity and the efficiency of the selection on π_v^0 events are required.

This procedure brings about three main sources of systematic uncertainties. Firstly, the efficiency of the selection on events containing a π_v^0 is estimated using simulation samples. The efficiency on π_v^0 decays in data can be different, which results in a systematic error. Next to that, systematic uncertainties arise in the estimation of the luminosity and estimation of the number of background events.

7.1 Efficiency of data processing

Each of the selection steps listed in Table 12, brings about a systematic error. In order to estimate the uncertainty in calculating the efficiency on simulated signal samples, decays with similar properties as the π_v^0 signal can be investigated. The Υ and Z decays resemble π_v^0 decays, as the muons from these decays also come from the same vertex. In Chapter 4, it was shown that their simulation corresponds very well with data for different variables. This indicates that the systematic error will be small for evaluating the π_v^0 efficiency on simulation.

7.1.1 Trigger and preselection

The preselection imposes restrictions on the transverse momentum of the muons and π_v^0 candidate. In Figures 13 and 15, it is shown that the momentum distributions for the Z simulation samples agree very well with the Z data sample. The momentum distributions agree less well for Υ data and simulation, but this was predicted as not all Υ production processes are taken into account in the simulation. In [28] it is shown that momentum distributions are modelled well in simulation. Therefore, the systematic uncertainties for the momentum requirements are neglected.

The preselection also requires a good displaced vertex: IP-χ^2 < 12 (see Section 5.2 for a definition). It has been checked that this is a very loose requirement that does not suppress a significant amount of events of the π_v^0 simulation samples. Next to that, the requirement in the final selection on the DOCA is a much more stringent constraint on the quality of the vertex. Therefore, the systematic error of the IP-χ^2 requirement is small (in comparison) and thus neglected.

Lastly, the preselection requires both muons to be identified as such by the particle identification algorithm. Methods exist in LHCb to calculate the associated systematic error, this remains to be evaluated in future research.

The systematic errors originating in the calculation of trigger efficiencies on simulation samples also remain to be evaluated.

7.1.2 Selection

In order to estimate the systematic uncertainty of the selection, the efficiency of each selection step can be calculated for Υ or Z data and simulation samples. The difference between the samples is a measure of the systematic uncertainty on the π_v^0 efficiency. The main difference between the Υ, Z and signal are the kinematic variables. In order correct for this when evaluating the systematic uncertainty on these decays, the samples can be divided in intervals (bins) of the transverse momentum p_{T}. The systematic uncertainty can be determined by adding the difference between data and simulation in each momentum region, weighted by the number of signal events in that momentum region. This is represented by the following formula:

$$\sigma = \sum_{p_{\mathrm{T}}\ \mathrm{bin}} |\epsilon_{\mathrm{Sim}}(p_{\mathrm{T}}\ \mathrm{bin}) - \epsilon_{\mathrm{Data}}(p_{\mathrm{T}}\ \mathrm{bin})| \frac{N_{\mathrm{Sig}}(p_{\mathrm{T}}\ \mathrm{bin})}{N_{\mathrm{Sig}}(\mathrm{total})}. \tag{2}$$

Here, σ is the systematic uncertainty on the efficiency of the selection criterion on signal. $\epsilon_{\mathrm{MC}}(p_{\mathrm{T}}\ \mathrm{bin})$ and $\epsilon_{\mathrm{Data}}(p_{\mathrm{T}}\ \mathrm{bin})$ are the efficiencies of this selection criterion on Υ or Z simulation and data in a certain p_{T} interval. $N_{\mathrm{Sig}}(p_{\mathrm{T}}\ \mathrm{bin})$ and $N_{\mathrm{Sig}}(\mathrm{total})$ are the number of signal events in the corresponding p_{T} interval and in total.

This is not a trivial calculation. It has been observed that the Υ does not have enough momentum to compare data and simulation for all π_v^0 momentum intervals. Z simulation samples contain less events than Υ samples, which could introduce a significant statistical error in the evaluation of 2. This remains to be studied in future research.

7.2 Background estimation

In order to set a limit, the number of remaining background decays needs to be estimated. In the next chapter, it is shown that this is done by looking at π_v^0 candidates that survive the selection in mass sidebands around the hypothesised π_v^0 mass. The background is assumed to be uniform, the π_v^0 candidates in the sidebands are summed and scaled to the window size to provide an estimate for the background around the hypothesised π_v^0 mass. From Figure 38 it is concluded that this is a reasonable assumption for small mass windows, as there is very little background left. The background decreases for larger masses, but this effect can only lead to an overestimation of the backgrounds. The error in the background estimation is therefore neglected.

7.3 Luminosity

The systematic error on the integrated luminosity measurement at LHCb is estimated to be 1.71% on data recorded in 2011 and 1.16% in 2012 [29].

7.4 Overview

The agreement between Υ and Z data and simulation indicates that the systematic error in data processing is small. The error on the integrated luminosity is at most 2% and the error on the background estimation is neglected. The total systematic error is estimated to be 10%.

8 Setting a limit

No evidence for a π_v^0 signal decay was found and this analysis now aims to put a limit on the production rate of π_v^0 particles that decay to two muons. In this section, the statistical procedure used to set this limit is described. The systematic error is taken into account and the expected limit for a no signal hypothesis is calculated. Graphs of the limit are presented as a function of the mass of the π_v^0.

8.1 Strategy

A limit is set on the production rate of a Higgs, decaying to two π_v^0 particles, of which at least one decays to two muons. Rewriting formula 1 for the production rate gives:

$$\sigma(pp \to H) \times \mathcal{B}(H \to \pi_v^0\pi_v^0) \times \mathcal{B}(\pi_v^0 \to \mu^+\mu^-) = \frac{\mu_{\text{Sig}}}{L \times \epsilon}. \tag{3}$$

It is assumed that $\mathcal{B}(\pi_v^0 \to \mu^+\mu^-) = 1$. L is the total integrated luminosity in fb^{-1}, which represents the total number of pp collisions and is listed in Table 1. μ_{Sig} is the number of signal decays ($\pi_v^0 \to \mu^+\mu^-$) after the selection and ϵ is the efficiency of the selection on π_v^0 decays. A limit is set on μ_{Sig}, which directly translates to a limit on the production rate by using L and ϵ. This procedure is first described without taking the systematic error into account, its implementation is discussed in Section 8.5.

In order to set a limit on μ_{Sig}, an assumption is made about the mass of the π_v^0. The mass resolution Δm_{Res} is determined from the π_v^0 simulation samples. It is illustrated in Figure 39. This is the precision with which the mass can be determined in the detector and thus a measure of the mass range in which signal events are expected. The signal region is defined as the region around the hypothesised mass of the π_v^0 with width $\Delta m_{\text{SigRegion}} = 6 \cdot \Delta m_{\text{Res}}$. The mass resolution depends on the kinematics and thus on the mass of the π_v^0, it is determined as a function of the mass in the next section.

Two ranges around the signal region in the mass histogram are defined as sidebands. These have width $\tau \cdot \Delta m_{\text{SigRegion}}$, with τ a to be determined transfer factor. In Figure 39, a hypothesised π_v^0 mass, the signal region around it and two sidebands are drawn for illustration. Outside the signal region, no signal events are expected, and all observed events are assumed to be background events. The number of background events in the sidebands is used to estimate the number of background events in the signal region. The number of observed events in the signal region is then used to put a limit on μ_{Sig}. The exact procedure for this is described in Section 8.3.

The assumption about the mass of the π_v^0 is made for several points over the region from 7 to 50 GeV/c^2. By following the above procedure for each point, the limit is determined for the entire mass range.

8.2 Mass resolution

The mass resolution is determined by fitting a Gaussian function to the mass histograms of the π_v^0 simulation samples. An example is displayed in Figure 40a. It is not a perfect

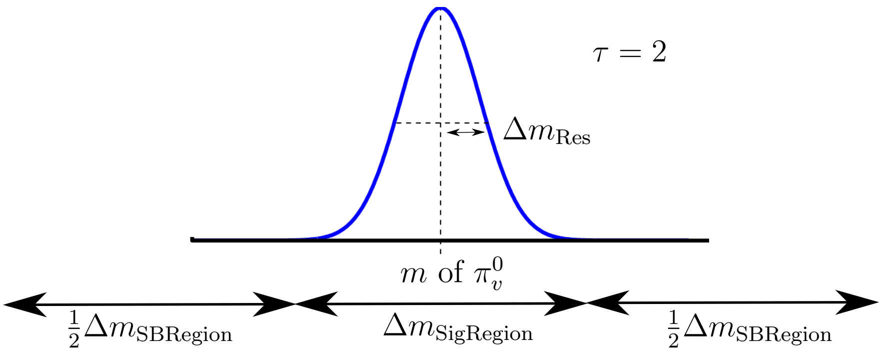
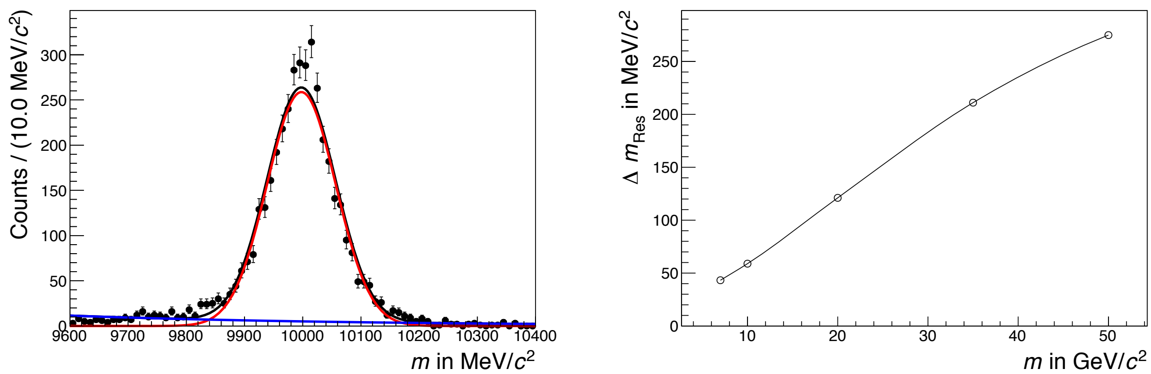


Figure 39: Illustration of the signal region and sideband regions around the hypothesised mass of the π_v^0. The transfer factor is $\tau = 2$ and the signal region is defined as $\Delta m_{\mathrm{SigRegion}} = 6 \cdot \Delta m_{\mathrm{Res}}$.

fit, but good enough to estimate the range in which signal events are expected.

The σ parameter of the fit is the mass resolution. It is plotted as a function of the mass of the π_v^0 in Figure 40b. The area under a normalised Gaussian function from $\mu - 3\sigma$ to $\mu + 3\sigma$ is 0.997. Therefore, most of the signal decays are expected to have masses in the signal region with width $6 \cdot \Delta m_{\mathrm{Res}}$. This has been checked for the sample in Figure 40a; a fraction of 0.998 ± 0.0007 of the events has a π_v^0 mass in the signal region.



(a) (b)

Figure 40: (a) Gaussian function plus exponential background fit to the mass histogram of the 10 GeV/c^2, 10 ps π_v^0 sample, and (b) the $\sigma = \Delta m_{\mathrm{Res}}$ parameter versus the mass of the π_v^0 for the 10 ps samples, linearly interpolated between the points. Both with 2012 Up simulation conditions.

8.3 Observed limit on number of signal events

In order to set a limit, a statistical test is performed. The hypothesis H_0 is defined as the presence of a signal. This is tested against the alternative hypothesis H_1 that there is no signal:

$$\begin{aligned} H_0 &: \mu_{\text{Sig}} > 0, \\ H_1 &: \mu_{\text{Sig}} = 0. \end{aligned} \tag{4}$$

The quantity $O_{\text{SigRegion}}$ is measured, the number of observed events in the signal region, which is the so-called 'test statistic' t. The actually measured value of a quantity is indicated with a prime: $O'_{\text{SigRegion}}$. For different values of the hypothesised μ_{Sig}, the probability is calculated that the measured value $O'_{\text{SigRegion}}$ or an even lower value is found. This probability is called the p-value. For a p-value smaller than 0.05, the hypothesis H_0 is rejected and it is concluded that there was no signal. The smallest μ_{Sig} for which H_0 can be rejected, is the limit on the μ_{Sig}.

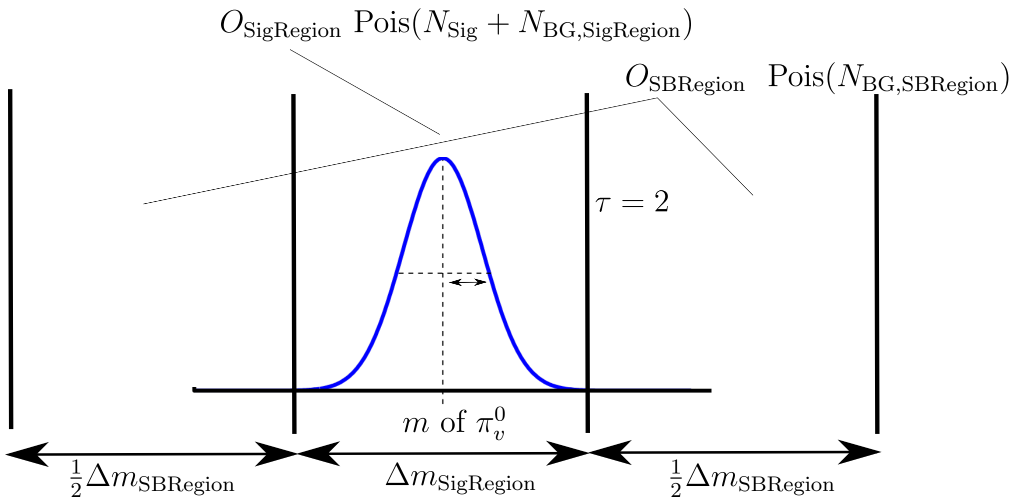


Figure 41: Illustration of the observed values and their distributions. The transfer factor is $\tau = 2$.

$O_{\text{SigRegion}}$ is Poisson distributed with mean $\mu_{\text{Sig}} + N_{\text{BG,SigWindow}}$, where the latter variable is the number of background events in the signal region, see Figure 41 for illustration. The number of observed events in the sidebands O_{SBRegion} is also measured, which is Poisson distributed with mean $N_{\text{BG,SBRegion}}$. As the width of the sidebands is $\tau \cdot \Delta m_{\text{SigRegion}}$, the number of events in the sidebands is $N_{\text{BG,SBRegion}} = \tau \cdot N_{\text{BG,SigRegion}}$. Here, it is assumed that the background events are uniformly distributed within the small mass range. Then the likelihood of finding the observed values $O'_{\text{SigRegion}}$ and O'_{SBRegion} is defined as [30]:

$$
\begin{aligned}
L(\mu_{\text{Sig}}, N_{\text{BG,SigRegion}}) &= \frac{(\mu_{\text{Sig}} + N_{\text{BG,SigRegion}})^{O'_{\text{SigRegion}}}}{O'_{\text{SigRegion}}!} e^{-(\mu_{\text{Sig}} + N_{\text{BG,SigRegion}})} \\
&\times \quad \frac{(\tau N_{\text{BG,SigRegion}})^{O'_{\text{SBRegion}}}}{O'_{\text{SBRegion}}!} \quad e^{-\tau N_{\text{BG,SigRegion}}} \ . \qquad (5)
\end{aligned}
$$

In order to calculate the p-value for a value of test statistic $t = O'_{\text{SigRegion}}$, the distribution of $O_{\text{SigRegion}}$ is needed. As it is Poisson distributed with the mean $\mu_{\text{Sig}} + N_{\text{BG,SigRegion}}$, $N_{\text{BG,SigRegion}}$ needs to be estimated. The conditional maximum likelihood estimator for this variable is obtained by maximizing Equation 5 for $N_{\text{BG,SigRegion}}$ under the μ_{Sig} hypothesis:

$$
\frac{\partial L}{\partial N_{\text{BG,SigRegion}}}(\mu_{\text{Sig}}, N_{\text{BG,SigRegion}}) = 0,
$$

which gives:

$$
\begin{aligned}
\hat{\hat{N}}_{\text{BG,SigRegion}} &= \frac{O_{\text{SigRegion}} + O_{\text{SBRegion}} - (1+\tau)\mu_{\text{Sig}}}{2(1+\tau)} \\
&+ \left[\frac{(O_{\text{SigRegion}} + O_{\text{SBRegion}} - (1+\tau)\mu_{\text{Sig}})^2 + 4(1+\tau)O_{\text{SBRegion}}\mu_{\text{Sig}}}{4(1+\tau)^2}\right]^{1/2} (6)
\end{aligned}
$$

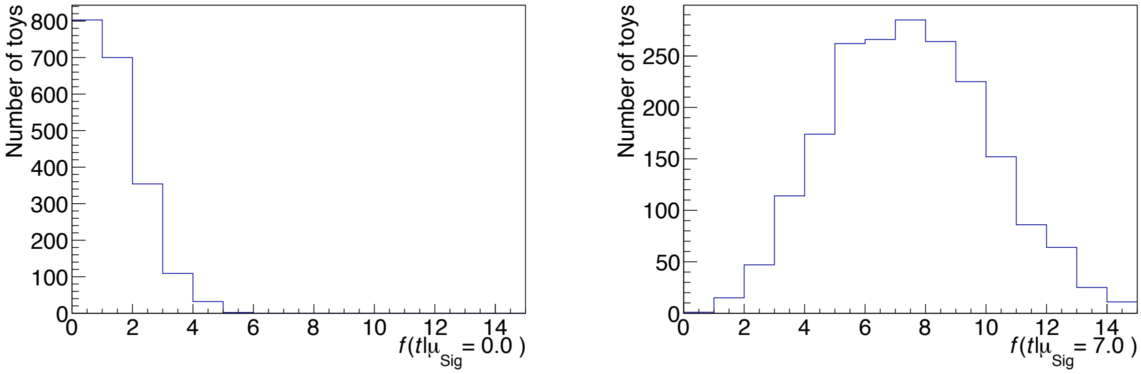
Several ROOSTATS and other tools are available for calculating limits, e.g. the Frequentist and Asymptotic calculators and the TROLKE method [31]. This analysis has very few background events, and these methods require the number of events to be large in the region used for estimating the background. Therefore, a new method has been developed.

The method uses $O'_{\text{SigRegion}}$, O'_{SBRegion} and τ as input, and returns a limit. It makes the following calculations:

1. For multiple hypothesised values of μ_{Sig}, calculate $\hat{\hat{N}}_{\text{BG,SigRegion}}$ with Equation 6;
2. For each μ_{Sig}, calculate distribution of test statistic t, $f(t|\mu_{\text{Sig}})$, by throwing toys (picking numbers) with a Pois($\mu_{\text{Sig}} + \hat{\hat{N}}_{\text{BG,SigWindow}}$) distribution;
3. For each μ_{Sig}, calculate p-value by integrating the lower tail of this distribution up to the measured $O'_{\text{SigRegion}}$:

$$
p = \int_0^{O'_{\text{SigRegion}}} f(t|\mu_{\text{Sig}})\text{d}t;
$$

4. Calculate the intersection of the μ_{Sig} versus p-value graph with the line p=0.05. This value of μ_{Sig} is the limit.



(a)

(b)

Figure 42: Distributions of test statistic $t = O_{\text{SigRegion}}$ for hypotheses $\mu_{\text{Sig}} = 0$ and $\mu_{\text{Sig}} = 7$. Here, $O_{\text{SigRegion}} = 2$, $O_{\text{SBRegion}} = 4$ and $\tau = 4$.

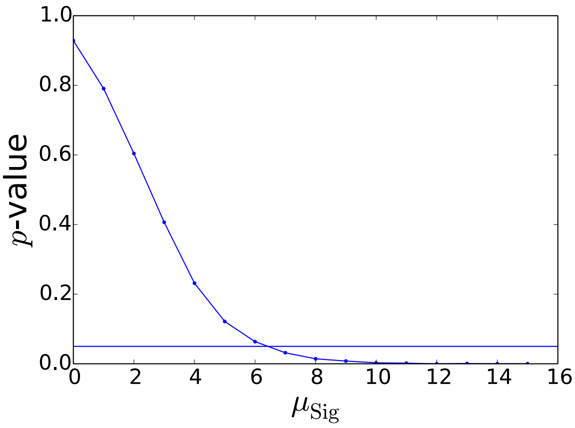


Figure 43: μ_{Sig} hypothesis versus p-value for the observed $O_{\text{SigRegion}} = 2$, $O_{\text{SBRegion}} = 4$. The hypothesis is rejected when the probability p to observe these values is smaller than 0.05.

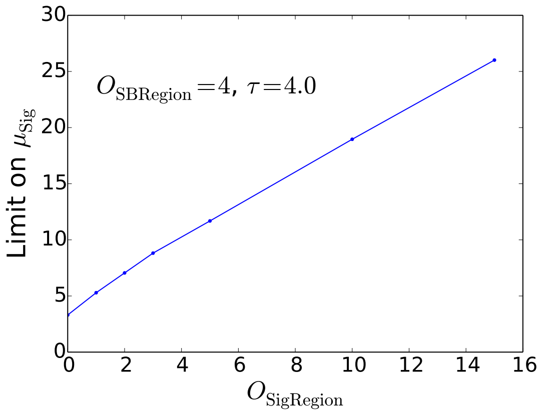
The distributions of test statistic $t = O_{\text{SigRegion}}$ for the hypotheses $\mu_{\text{Sig}} = 0$ and $\mu_{\text{Sig}} = 7$ are plotted in Figure 42. Here, $O'_{\text{SigRegion}} = 2$, $O'_{\text{SBRegion}} = 4$ and $\tau = 4$. It can be concluded that it is quite likely that $O_{\text{SigRegion}} \leqslant 2$ when no signal is expected ($p \approx 0.95$), but that it is very unlikely when 7 signal events are expected ($p \approx 0.03$). In Figure 43, the graph of the hypothesis for μ_{Sig} versus the p-value is displayed. The limit is $\mu_{\text{Sig}} \approx 6.33$, there is only a 5% probability that $O'_{\text{SigRegion}} \leqslant 2$ when the real number of signal events was 6.33.

Some consistency checks are performed and displayed in Figure 44. Figure 44a shows the limit as a function of $O'_{\text{SigRegion}}$, it is expected to be less good (larger) when more events are observed in the signal region. This is indeed the case.

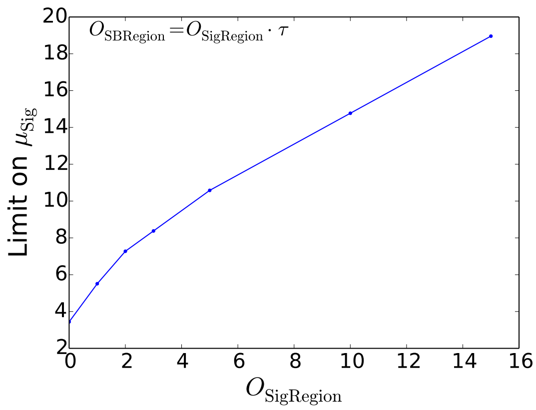
Figure 44b is used to check that when more events are observed, in both the signal region and the sidebands, the limit gets worse. A uniform observation is defined, where

$O'_{\text{SBRegion}} = \tau \cdot O'_{\text{SigRegion}}$. $O_{\text{SigRegion}}$ is plotted versus the limit. The graph shows that a less stringent limit can be set when more events are observed. The dependency is strong, so it is crucial for the selection to exclude as many background events as possible.

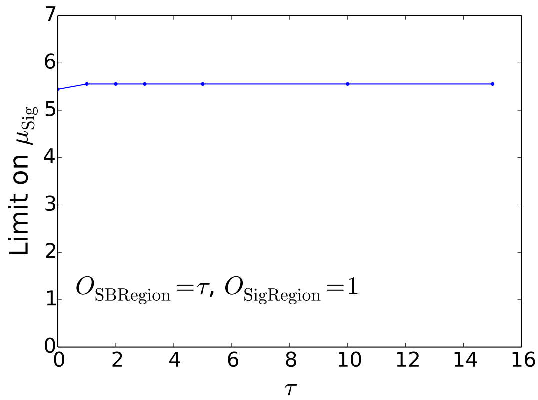
Figure 44c displays the limit as a function of τ. Here, $O'_{\text{SigRegion}} = 1$ and $O'_{\text{SBRegion}} = \tau$ in order to take into account that when the sidebands are larger, more events are used for the background estimation. Ideally, the limit should not depend too much on the width of the sidebands. From the graph, it can be concluded that the limit does not significantly depend on the choice of τ. It is chosen to be $\tau = 6$.



(a)



(b)



(c)

Figure 44: Consistency checks of the method. (a) $O_{\text{SigRegion}}$ versus the limit on μ_{Sig}. Here, $O_{\text{SBRegion}} = 4$ and $\tau = 4$. (b) $O_{\text{SigRegion}}$ versus the limit on μ_{Sig} with $O_{\text{SBRegion}} = \tau \cdot O_{\text{SigRegion}}$ and $\tau = 4$. In (c), τ versus the limit on μ_{Sig}, with $O_{\text{SigRegion}} = 1$ and $O_{\text{SBRegion}} = \tau$.

8.4 Expected limit on number of signal events

It is customary to quantify the expected limit, based on the H_1 hypothesis: there is no signal. Based on this hypothesis, the distribution of the test statistic $t = O_{\text{SigRegion}}$ is studied. The median of this distribution is the value of $O_{\text{SigRegion}}$ for which the area under the normalised distribution is 0.5:

$$\frac{\int_0^{E_{\text{SigRegion}}} f(t|\mu_{\text{Sig}} = 0)\mathrm{d}t}{\int_0^{\infty} f(t|\mu_{\text{Sig}} = 0)\mathrm{d}t} = 0.5.$$

The notation $E_{\text{SigRegion}}$ is adopted for the median, it is the expected number of events in the signal region under the no signal hypothesis. The limit setting procedure is repeated with $E_{\text{SigRegion}}$ instead of $O_{\text{SigRegion}}$ (number 3 and 4 in the list of calculations above). This way, the expected limit is calculated on μ_{Sig}. For example, for $O_{\text{SigRegion}} = 2$, $O_{\text{SBRegion}} = 4$, $\tau = 4$, $E_{\text{SigRegion}} \approx 1.25$ is found from integrating up to 50% of Figure 42a.

A confidence interval for the observed limit is obtained by respectively calculating $E_{0.975}$, $E_{0.84}$, $E_{0.16}$ and $E_{0.025}$ in:

$$\frac{\int_0^{E_{\text{x}}} f(t|\mu_{\text{Sig}} = 0)\mathrm{d}t}{\int_0^{\infty} f(t|\mu_{\text{Sig}} = 0)\mathrm{d}t} = x.$$

Under the no-signal hypothesis, there is a 68% probability that the observed number of events in the signal region will be between $E_{0.16}$ and $E_{0.84}$ and 95% that they are between $E_{0.975}$ and $E_{0.025}$. This translates to confidence intervals on the limit and is used as a check of the method; when the observed limit does not fall within these confidence intervals, it is an indication that the no-signal hypothesis is not true. The hypotheses in Equation 5 should be reversed in order to prove a discovery.

For the above example, $E_{0.975} \approx 3.76$, $E_{0.84} \approx 2.44$, $E_{0.16} \approx 0.39$ and $E_{0.025} \approx 0.061$. When these values are used to calculate the limit, the expected limit is found to be 4.98. The 68% confidence interval is between 3.06 and 6.46 and the 95% confidence interval between 3.06 and 7.87. The lower bounds are the same, because the test statistic distributions are discrete and thus $E_{0.16}$ and $E_{0.025}$ can fall in the same bin and have the same p-value.

8.5 Systematic error

In Chapter 7, it was discussed that the main systematic error in this analysis is due to the evaluation of the efficiency on π_v^0 simulation samples. It can over- or underestimated. The systematic error that arises due to these differences, was estimated to be about 10%. When a hypothesis is made about a number of signal decays in the limit program, this could have been 10 % higher or lower because of this systematic error. This influences the μ_{Sig} versus p-value graph by smearing it as illustrated in Figure 45. The conservative limit is the intersection of the upper error band with p=0.05. The limit on μ_{Sig} becomes 10% larger and so does the limit on the production rate.

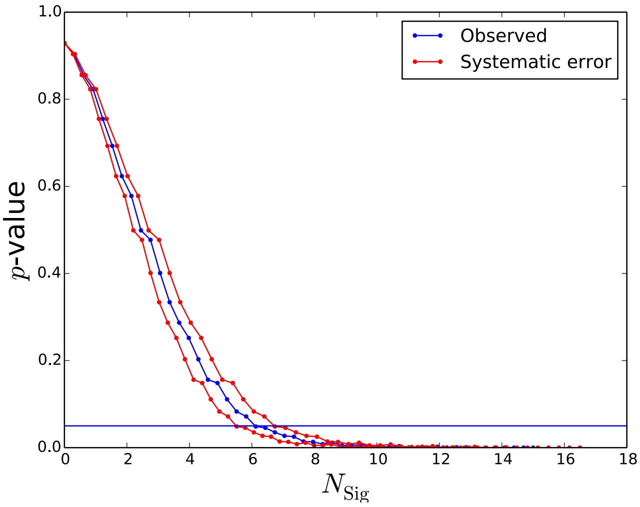


Figure 45: μ_{Sig} hypothesis versus p-value for the observed $O_{\text{SigRegion}} = 2$, $O_{\text{SBRegion}} = 4$ and $\tau = 2$, including the systematic error. The hypothesis is rejected when the upper band probability p to observe these values is smaller than 0.05.

8.6 Limit on production

Figure 46 displays $O'_{\text{SigRegion}}$ and O'_{SBRegion} for several mass points. A limit can now be set on μ_{Sig} by repeating the above procedure for these mass points. It is translated to a limit on the production rate by Equation 3. The efficiency for each mass point is calculated by applying the final selection on the π_v^0 simulation samples (see Table 11) and dividing the number of candidates by the initial number of events. A linear interpolation is applied to estimate the efficiencies for mass other mass points, which is displayed in Figure 47. The interpolation is performed for all production years, polarities and both lifetimes. The quantity $L \times \epsilon$ in Equation 3 is then evaluated as:

$$L_{\text{2011,Up}} \times \epsilon_{\text{2011,Up}} + L_{\text{2011,Down}} \times \epsilon_{\text{2011,Down}} + L_{\text{2012,Up}} \times \epsilon_{\text{2012,Up}} + L_{\text{2012,Down}} \times \epsilon_{\text{2012,Down}}.$$

The graphs of the observed and expected limits for π_v^0 particles with lifetimes 10 and 100 ps are displayed in Figure 48. The observed limit is consistent with the confidence interval for a no signal hypothesis, no evidence for the signal is found. The graph shows discontinuities that are consistent with the mass distribution of the remaining candidates after selection, see Figure 38. Due to the discreteness of Poisson statistics, the confidence interval boundaries often overlap. The difference between 10 and 100 ps lifetimes is a scaling factor consistent with the difference in efficiency listed in Tables 12 and 13.

More background events were present for masses below 30 GeV/c^2, which makes it harder to exclude the possibility of a $\pi_v^0 \to \mu^+\mu^-$ decay and results in a larger limit. In Chapter 4 it was shown that π_v^0 particles with a long lifetime have a low trigger efficiency. They fly out of the VELO and muon hits in the VELO are required to trigger the event.

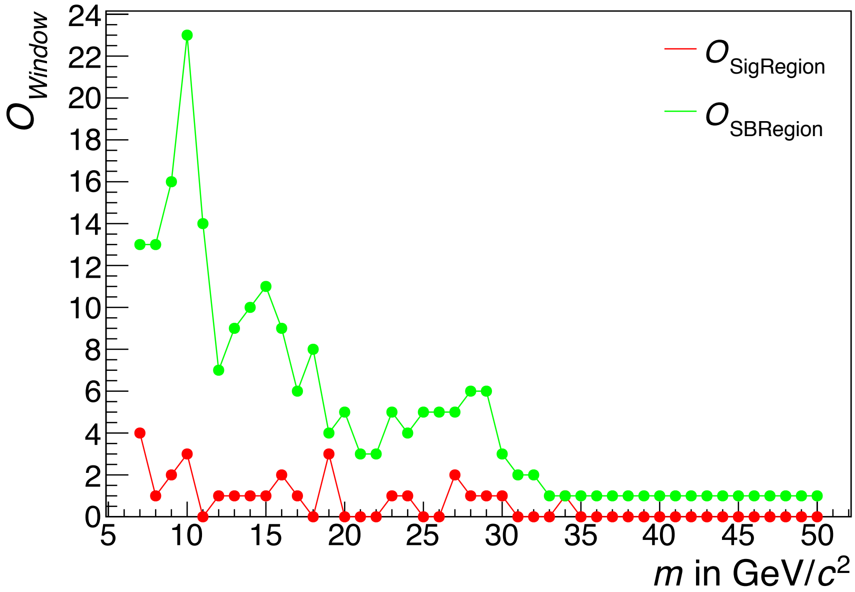


Figure 46: Observed number of events in signal region ($O_{\mathrm{SigRegion}}$) and sideband regions (O_{SBRegion}) for several mass points in all data.

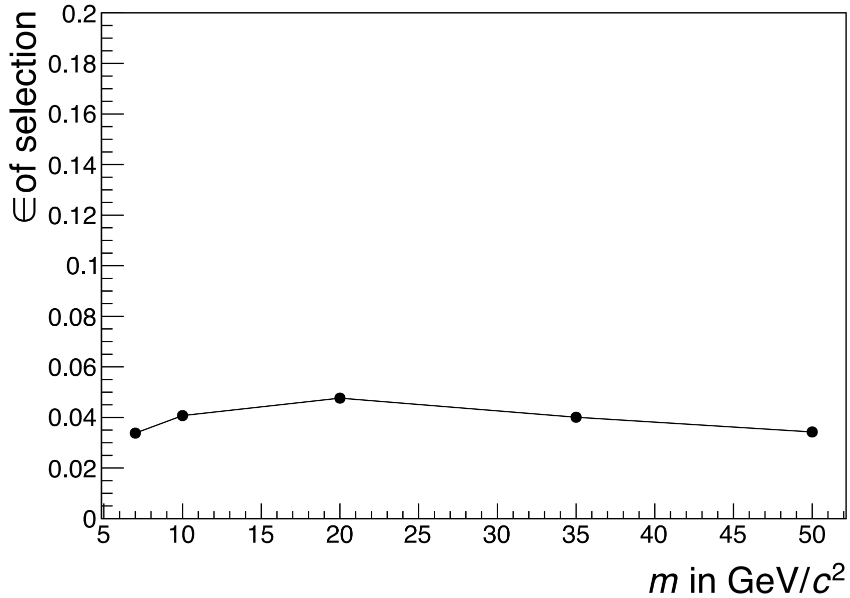
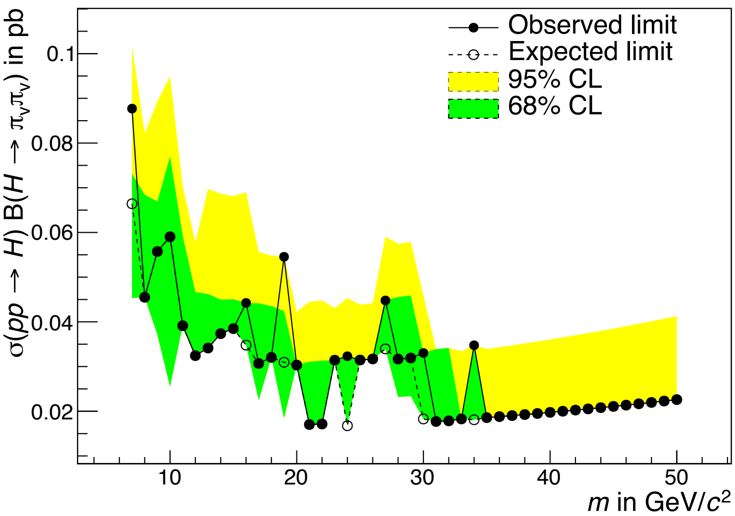


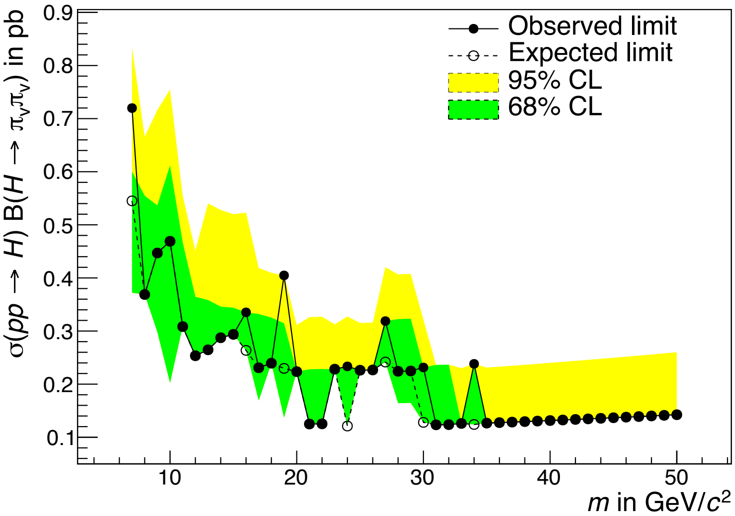
Figure 47: Mass m versus efficiency ϵ of the selection on the 10 ps π_v^0 simulation sample with 2012 Up simulation conditions. Linearly interpolated between the mass samples.

Missing produced π_v^0 particles makes it harder to impose a limit on the production, therefore the limits are higher for larger lifetimes.

The limit of CMS for 20 and 50 GeV/c^2 and 10 and 100 ps is stable and about 9×10^{-3} pb. Unfortunately, this search is thus not yet competitive with CMS for these masses. This limits on π_v^0 production with 7 to 20 GeV/c^2 masses are complementary results.



(a)



(b)

Figure 48: Observed and expected limits on the signal production rate and 95% and 68% Confidence Levels in 3 fb^{-1} data from 2011 and 2012, as a function of π_v^0 mass, for (a) a π_v^0 lifetime of 10 ps and (b) a π_v^0 lifetime of 100 ps.

9 Conclusion

In this analysis, a search was performed for long-lived exotic particles decaying to oppositely charged muons. The Hidden Valley $\pi_v^0 \to \mu^+\mu^-$ decay was used as a benchmark signal. The search was confined to particles with lifetimes 10 and 100 ps and with masses ranging from 7 to 50 GeV/c^2.

The used data consisted of events with muon pairs; the π_v^0 candidates. Two types of background processes were identified that could also be the source of these muon pairs. The first type consisted of two muons from instantaneously decaying resonances and the second was the topology of muons coming from two particles decaying in separate vertices. It was found that after suppressing the first type of background by selecting on high lifetime, the remaining backgrounds were all of the second type.

Based on differences between π_v^0 decays and muons coming from backgrounds with two seperate vertices, selection criteria were defined. About 60% of the simulated π_v^0 decays passed these criteria. Almost all of the background decays were suppressed, the remaining number of π_v^0 candidates in data was 137. No evidence for the π_v^0 signal was found.

A limit was set on the signal production rate as a function of the mass of the π_v^0. For 10 ps lifetimes the limit ranged from 0.09 pb^{-1} for low masses to 0.02 pb^{-1} for high masses. Most of the remaining background decays had low masses, which makes it harder to exclude the possibility of a $\pi_v^0 \to \mu^+\mu^-$ decay for these masses and leads to a higher limit. The limit for 100 ps lifetimes ranged from 0.7 pb^{-1} to 0.16 pb^{-1}. It is higher than for low lifetimes, as long-lived particles often fly out of the detector range before they decay. When a large fraction of the decays remain undetected, it is harder to set a limit on their production.

10 Discussion and outlook

In this analysis, a search for a 'Beyond the Standard Model' (BSM) particle is performed. Many physical phenomena remain unexplained by the Standard Model (SM) and BSM searches can provide useful information needed to extend our understanding of the universe. It is a challenge to decide where to look for new physics. Numerous theories are available that predict a wide range of phenomena that are often hard to verify by experimental results. Therefore, searches often combine theory with a topology that provides a clear signature for a search. Well motivated searches have potential to discover or exclude BSM theories.

This search provides a contribution to this effort. Several theories predict the existence of long-lived exotic particles that decay in vertices away from the main interaction point. Combined with a $\mu^+\mu^-$ final state, this provides a signature for which SM backgrounds are quite easily suppressed, as shown in this search. The scope of this research comprises exotic particle masses of 7 to 50 GeV/c^2 and lifetimes of 10 and 100 ps.

The fact that no evidence for this specific signal was found, does not mean that it was not there. The conclusion is that the signal does not occur with a sufficiently high rate to detect it with the current analysis and detector constraints.

In March 2015, the LHC will be restarted and will run at a higher center of mass energy (14 TeV). This will increase the likelihood that exotic particles are produced and will provide more data. This search can be repeated with this data, it can be added to the current search to improve the chance of finding significant evidence for the signal.

The CMS collaboration also performed a search for these signatures. For the exotic particle masses 20 and 50 GeV/c^2, this search is not yet competitive. The higher luminosity and different acceptance of CMS has so far proven more successful for this signal. However, this search can still be improved upon. Next to that, it it complementary to the CMS results as it reaches lower masses.

It would be interesting to widen the scope and search at more masses. For low masses, the SM background is larger and less easy to suppress than for the current search. It would require a new search and strategy. Searching for higher masses is challenging at the LHCb experiment. Muons originating in a heavier particle have a larger angle between them, as shown in this analysis. They then fly out of the acceptance of the forward LHCb detector.

The limits in this search have been set for π_v^0 lifetimes of 10 and 100 ps. It is possible to extrapolate these results to different lifetimes, which remains to be done in future research.

In this analysis, the branching fraction of the exotic particles to $\mu^+\mu^-$ is assumed to be one, i.e. they always decay to a muon pair. It may be possible to correct for different branching fractions, but this is complicated by the possibility of two π_v^0 decays in one collision. This can be further investigated.

10.1 Outlook current analysis

The current analysis is not entirely finished, it remains to be shown that the systematic error is of the predicted order of magnitude.

The search can also be improved on several points. Firstly, the selection of good π_v^0 candidates can be made mass dependent. It was observed that most background was present for masses below 30 GeV/c^2. Thus for higher masses it is not necessary to make a selection as stringent as in the present analysis. This will improve the chance of finding the π_v^0 particle and set a better limit in that range.

Next to that, the selection can be further improved by looking at more separating variables. For example, the possibilities of jets as identified by the particle flow algorithm have not been fully explored. The 'Tool for Multivariate Analysis' software has so far only been used to check if the chosen variables are separating well. Its possibilities can be further applied to this analysis.

The systematic error can be determined for multiple separating variables and the selection can be re-evaluated based on this. Variables with lower systematic errors are preferable as this results in more stringent limits.

Lastly, the systematic error is conservatively incorporated in calculating the limit. It can also be implemented as a nuisance parameter when the likelihood is calculated that the signal hypothesis is true for the number of observed events.

This thesis will be modified to an LHCb analysis note, including an evaluation of the systematic errors. It will then head for publication.

Appendices

A Software

All LHCb software is built on a common framework named GAUDI [32]. Physics simulations are done in the GAUSS [33] application, which uses PYTHIA [28] for the event generation and EVTGEN to simulate the time evolution and decay of the particles [34]. GAUSS uses the GEANT4 toolkit to simulate the detector response.

Not all data collected by the detector can be stored on disk. The trigger software in the MOORE project decides which events will be stored based on criteria described in Section 5.1. The same criteria are applied to the simulation samples.

'Raw data' coming from the detector (data samples) or the detector simulation (simulation samples), is further processed in the same way. Particle tracks are reconstructed using the software BRUNEL, version Reco14a. Then the preselection is applied, a procedure to further reduce the amount of events. The algorithms are included in the DAVINCI package. For 2012 samples, version Stripping20 is used, for 2011, Stripping20r1. The samples are further analysed with physics analysis tools from the DAVINCI package.

B Table of π_v^0 simulation samples

Table 14: Table of the simulated π_v^0 samples. Listed are the unique sample identification number, the π_v^0 simulation properties and the generator efficieny.

Event ID	Year	Magnet	Mass (GeV/c^2)	Lifetime (ps)	Generator efficiency (%)
43114000	2012	Up	7	10	22.12 ± 0.14
43114000	2012	Down	7	10	21.93 ± 0.12
43114000	2011	Up	7	10	21.28 ± 0.16
43114000	2011	Down	7	10	21.22 ± 0.14
43114001	2012	Up	7	100	22.03 ± 0.13
43114001	2012	Down	7	100	21.80 ± 0.14
43114001	2011	Up	7	100	21.28 ± 0.14
43114001	2011	Down	7	100	21.33 ± 0.16
43114002	2012	Up	10	10	20.94 ± 0.13
43114002	2012	Down	10	10	20.86 ± 0.13
43114002	2011	Up	10	10	20.06 ± 0.14
43114002	2011	Down	10	10	20.16 ± 0.15
43114003	2012	Up	10	100	20.80 ± 0.13
43114003	2012	Down	10	100	20.82 ± 0.12
43114003	2011	Up	10	100	19.92 ± 0.16
43114003	2011	Down	10	100	20.19 ± 0.15
43114004	2012	Up	20	10	17.27 ± 0.10
43114004	2012	Down	20	10	16.98 ± 0.10
43114004	2011	Up	20	10	16.30 ± 0.12
43114004	2011	Down	20	10	16.27 ± 0.12
43114005	2012	Up	20	100	17.17 ± 0.11
43114005	2012	Down	20	100	17.33 ± 0.11
43114005	2011	Up	20	100	16.32 ± 0.11
43114005	2011	Down	20	100	16.35 ± 0.12
43114006	2012	Up	35	10	13.04 ± 0.11
43114006	2012	Down	35	10	12.89 ± 0.11
43114006	2011	Up	35	10	12.11 ± 0.11
43114006	2011	Down	35	10	12.15 ± 0.11
43114007	2012	Up	35	100	13.18 ± 0.12
43114007	2012	Down	35	100	12.90 ± 0.11
43114007	2011	Up	35	100	12.22 ± 0.11
43114007	2011	Down	35	100	12.13 ± 0.10

Table 15: Table of the simulated π_v^0 samples (continuation). Listed are the unique sample identification number, the π_v^0 simulation properties and the generator efficieny.

Event ID	Year	Magnet	Mass (GeV/c^2)	Lifetime (ps)	Generator efficiency (%)
43114008	2012	Up	50	10	9.66 ± 0.09
43114008	2012	Down	50	10	9.43 ± 0.09
43114008	2011	Up	50	10	8.73 ± 0.07
43114008	2011	Down	50	10	8.67 ± 0.08
43114009	2012	Up	50	100	9.61 ± 0.08
43114009	2012	Down	50	100	9.49 ± 0.08
43114009	2011	Up	50	100	8.81 ± 0.08
43114009	2011	Down	50	100	8.75 ± 0.08

C Efficiencies on π_v^0 simulation samples

Table 16: Number of events (upper half) or candidates (lower half), relative efficiency and cumulative efficiency on 50 GeV/c^2 10 ps π_v^0 simulation samples. Displayed are different steps: the generator efficiency, trigger efficiencies, preselection and the final selection.

	2011, Up			**2012, Up**		
	N	ϵ_{rel}	ϵ_{cum}	N	ϵ_{rel}	ϵ_{cum}
After generator	12059	0.0881	0.0881	11899	0.0961	0.0961
L0	11674	0.9681	0.0853	11562	0.9717	0.0934
Hlt1	5992	0.5133	0.0438	5403	0.4673	0.0436
Hlt2	2643	0.4411	0.0193	2250	0.4164	0.0182
After preselection	2361	0.8933	0.0172	1985	0.8822	0.0160
Displayed vertex	2241	0.9492	0.0164	1846	0.9300	0.0149
Isolation	2210	0.9862	0.0161	1809	0.9800	0.0146
Jets	1995	0.9027	0.0146	1654	0.9143	0.0134
Lifetime	1948	0.9764	0.0142	1594	0.9637	0.0129
Pointing	1833	0.9410	0.0134	1479	0.9279	0.0119
$\epsilon_{\mathbf{cum}}$ **Selection**			0.7717			0.7390

Table 17: Number of events (upper half) or candidates (lower half), relative efficiency and cumulative efficiency on 50 GeV/c^2 10 ps π_v^0 simulation samples. Displayed are different steps: the generator efficiency, trigger efficiencies, preselection and the final selection.

	2011, Down			**2012, Down**		
	N	ϵ_{rel}	ϵ_{cum}	N	ϵ_{rel}	ϵ_{cum}
After generator	11232	0.0875	0.0875	11913	0.0949	0.0949
L0	10878	0.9685	0.0847	11582	0.9722	0.0923
Hlt1	5626	0.5172	0.0438	5533	0.4777	0.0441
Hlt2	2524	0.4486	0.0197	2319	0.4191	0.0185
After preselection	2237	0.8863	0.0174	2040	0.8797	0.0163
Displayed vertex	2105	0.9410	0.0164	1892	0.9275	0.0151
Isolation	2065	0.9810	0.0161	1865	0.9857	0.0149
Jets	1887	0.9138	0.0147	1715	0.9196	0.0137
Lifetime	1844	0.9772	0.0144	1649	0.9615	0.0131
Pointing	1735	0.9409	0.0135	1550	0.9400	0.0123
$\boldsymbol{\epsilon_{cum}}$ **Selection**			0.7711			0.7544

References

[1] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961), no. 4 579 .

[2] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[3] J. SolÃ , *Dark matter, dark energy and the time evolution of masses in the Universe*, Int. J. Mod. Phys. **A29** (2014), no. 2 1444016, `arXiv:1408.4427`.

[4] R. Barbier *et al.*, *R-parity violating supersymmetry*, Phys. Rept. **420** (2005) 1, `arXiv:hep-ph/0406039`.

[5] J. L. Hewett, B. Lillie, M. Masip, and T. G. Rizzo, *Signatures of long-lived gluinos in split supersymmetry*, Journal of High Energy Physics **2004** (2004), no. 09 070.

[6] M. J. Strassler and K. M. Zurek, *Echoes of a hidden valley at hadron colliders*, Phys. Lett. **B651** (2007) 374, `arXiv:hep-ph/0604261`.

[7] T. Han, Z. Si, K. M. Zurek, and M. J. Strassler, *Phenomenology of hidden valleys at hadron colliders*, Journal of High Energy Physics **2008** (2008), no. 07 008.

[8] Planck, P. A. R. Ade *et al.*, *Planck 2013 results. I. Overview of products and scientific results*, Astron. Astrophys. **571** (2014) A1, `arXiv:1303.5062`.

[9] M. Cvetic, J. Halverson, and H. Piragua, *Stringy Hidden Valleys*, JHEP **1302** (2013) 005, `arXiv:1210.5245`.

[10] T. Han, Z. Si, K. M. Zurek, and M. J. Strassler, *Phenomenology of hidden valleys at hadron colliders*, JHEP **0807** (2008) 008, `arXiv:0712.2041`.

[11] M. J. Strassler, *Possible effects of a hidden valley on supersymmetric phenomenology*, `arXiv:hep-ph/0607160`.

[12] CMS, S. Chatrchyan *et al.*, *Search in leptonic channels for heavy resonances decaying to long-lived neutral particles*, JHEP **1302** (2013) 085, `arXiv:1211.2472`.

[13] CMS, V. Khachatryan *et al.*, *Search for long-lived particles that decay into final states containing two electrons or two muons in proton-proton collisions at $\sqrt{s} = 8$ TeV*, `arXiv:1411.6977`.

[14] CMS and LHCb collaborations, R. Aaij *et al.*, *Discovery of $B_s^0 \to \mu^+\mu^-$ and evidence for $B^0 \to \mu^+\mu^-$ by the CMS and LHCb experiments*, LHCb-PAPER-2014-049, in preparation.

[15] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030. LHCB-TDR-009.

[16] A. A. Alves Jr *et al.*, *The lhcb detector at the lhc*, Journal of instrumentation **3** (2008), no. 08 S08005.

[17] T. L. Collaboration, *The lhcb detector at the lhc*, Journal of Instrumentation **3** (2008), no. 08 S08005.

[18] LHCb collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011. LHCB-TDR-005.

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

[20] Particle Data Group, J. Beringer *et al.*, *Review of particle physics*, Phys. Rev. D **86** (2012) 010001.

[21] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, ArXiv Physics e-prints (2003) `arXiv:physics/0306116`.

[22] LHCb collaboration, R. Aaij *et al.*, *Measurement of Υ production in pp collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. **C72** (2012) 2025, `arXiv:1202.6579`.

[23] M. Pivk and F. R. L. Diberder, *Splot: A statistical tool to unfold data distributions*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **555** (2005), no. 1-2 356 .

[24] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, `arXiv:1211.3055`.

[25] J. Albrecht *et al.*, *Performance of the lhcb high level trigger in 2012*, Journal of Physics: Conference Series **513** (2014), no. 1 012001.

[26] LHCb collaboration, R. Aaij *et al.*, *Measurements of the Λ_b^0, Ξ_b^-, and Ω_b^- baryon masses*, Phys. Rev. Lett. **110** (2013) 182001, `arXiv:1302.1072`.

[27] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, `arXiv:physics/0703039`.

[28] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, `arXiv:hep-ph/0603175`.

[29] LHCb collaboration, *Precision luminosity measurements at the LHC*, LHCb-PAPER-2014-047, in preparation.

[30] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, European Physical Journal C **71** (2011) 1554, `arXiv:1007.1727`.

[31] L. Moneta, K. Cranmer, G. Schott, and W. Verkerke, *The RooStats project*, in *Proceedings of the 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research.*, p. 57, 2010. `arXiv:1009.1003`.

[32] G. Barrand *et al.*, *GAUDI - A software architecture and framework for building HEP data processing applications*, Comput. Phys. Commun. **140** (2001) 45.

[33] T. L. Collaboration, *The lhcb simulation application, gauss: Design, evolution and experience*, Journal of Physics: Conference Series **331** (2011), no. 3 032023.

[34] LHCb collaboration, *LHCb computing: Technical Design Report*, CERN-LHCC-2005-019. LHCB-TDR-011.